Target the Right Lead: A "Smart Sales" Predictive Lead Scoring Model for Business-to-Business Inside Sales Success

by

Migao Wu

A Thesis by Articles Submitted to the University of Ottawa
in Partial Fulfillment of the Requirements for the
Doctor of Philosophy (Ph.D.) degree in
Digital Transformation & Innovation



School of Electrical and Computer Science
Faculty of Engineering
University of Ottawa

© Migao Wu, Ottawa, Canada, 2024

Abstract

The way of selling has evolved over the last decade, mostly driven by the swift profession of digital technologies and social changes. Inside sales have grown significantly due to the high costs of traditional field sales and advances in information and communication technologies (ICTs). The COVID-19 pandemic has further accelerated this shift, compelling selling organizations to maximize opportunities from both new prospects and existing customers through sales that are conducted remotely or virtually. As a fundamental task in the initial prospecting stage of the inside sales process, lead scoring evaluates and ranks leads based on their conversion likelihood to improve sales success. Despite the importance of lead scoring tasks in inside sales, there is a lack of research studying the current state of lead scoring models and how to improve overall sales performance by enhancing the efficacy of lead scoring models. This dissertation investigates the evolution and optimization of lead scoring tasks in the business-to-business (B2B) inside sales context, focusing on transitioning from traditional to predictive lead scoring models.

This doctoral dissertation aims to: (1) Summarize critical gaps in the current understanding of lead scoring models, including their advantages, disadvantages, and impact on sales performance. (2) Identify and validate the key factors for a potential buyer's purchase decision-making process. (3) Optimize lead scoring outcomes through predictive modeling. By employing a systematic literature review, empirical validation, and the development and implementation of a novel predictive lead scoring model, this dissertation demonstrates the potential of data-driven predictive analytics using datasets from online surveys and service industries and offers a comprehensive framework for improving inside sales performance.

The key contributions of this doctoral dissertation include the integration of a seller-centric marketing sales funnel with a lead-centric purchase decision-making process, providing a holistic approach to predictive lead scoring. The findings demonstrate that incorporating insights from leads' profiles into predictive models can significantly enhance sales performance. This research not only aids industry practitioners in refining their lead management strategies but also advances academic understanding of the role of predictive analytics in B2B inside sales success.

Acknowledgment

I would like to begin by extending my sincere gratitude to the individuals and organizations who have supported and inspired me throughout this Ph.D. journey.

I am grateful to my exceptional supervisors, Professor Morad Benyoucef and Professor Pavel Andreev for their invaluable time, selfless knowledge sharing, patience, academic guidance, and personal and professional support during my Ph.D. study.

I am also thankful to VanillaSoft, Mitacs, Faculty of Engineering, and Telfer School of Management for their financial support of my research. Especially, I would like to express my gratitude to David Hood and Daniel Sims for their support in retrieving the data and practical insights that made this research possible. Additionally, I appreciate the technical support and resources provided by the staff at VanillaSoft.

Furthermore, my appreciation also extends to my thesis advisory committee members, Professor Ajax Persaud and Professor Herna Viktor, whose insightful feedback and constructive criticism in the comprehensive exam, thesis proposal, and thesis defenses have greatly enhanced the technical and theoretical aspects of my research.

Lastly, my deepest gratitude goes to my family and friends, whose financial and mental support have been fundamental to the success of this research. A very special thanks to my wife who has provided unwavering support in every possible way during the most challenging moments of my Ph.D. research journey.

Table of Contents

Abstract ii

Acknowledgment iii

Table of Contents iv

Chapter 1: Introduction 1

1.1 Research problem and motivation 4

1.2 Research questions and objectives 6

1.3 Research contributions 6

1.4 Related publications 7

1.5 Thesis outline 8

Chapter 2. The state of lead scoring models and their impact on sales performance 9

2.1 Introduction 10

2.1.1 Inside sales and lead scoring modeling 10

2.1.2 Problem and motivation 11

2.1.3 Scope and contribution 12

2.2 Methodology 13

2.2.1 Planning the review 13

2.2.2 Conducting the review 14

2.2.3 Reporting on the review 16

2.3 Search results 16

2.4 State of lead scoring models 18

2.4.1 Traditional lead scoring models 18

2.4.2 Predictive lead scoring models 19

2.4.3 Lead scoring models in grey literature 24

2.4.4 The preferred model: supervised vs. unsupervised 24

2.4.5 Impact of lead scoring models on sales performance 26

2.4.6 Limitations 31

2.5 Discussion and conclusions 31

Appendix: table of extracted literature data 39

References 44

Chapter 3. Unlocking B2B buyer intentions to purchase: Conceptualizing and validating inside sales purchases 50

3.1 Introduction 51

3.2 Background 52
3.2.1 B2B purchase decision-making process 52
3.2.2 B2B buying behavior 53
3.3 B2B buyer's intention to purchase (BIP) decision-making model 54
3.4 Methodology 57
3.5 Results 58
3.5.1 Data collection 58
3.5.2 BIP model evaluation 59
3.5.3 Multigroup analyses 62
3.6 Discussion and implications 64
3.6.1 Theoretical implications 65
3.6.2 Practical implications 68
3.6.3 Limitations and future work 70
3.6.4 Conclusion 70
References 81
Chapter 4. Smart Sales: Amplifying the Power of Predictive Lead Scoring in B2B Sales 87
4.1 Introduction 87
4.2 Background 90
4.2.1 From traditional to predictive lead scoring models 90
4.3 Methodology 92
4.3.1 Data collection 94
4.3.2 Machine learning (ML) algorithms 95
4.3.3 Model evaluation and imbalanced dataset 97
4.4 Results 98
4.4.1 Modeling procedure 98
4.4.2 Modeling results 98
4.5 Conclusion 100
4.5.1 Implications 100
4.5.2 Research contributions 102
4.5.3 Limitations and future research 103
References 106
Chapter 5. Data-Driven Predictive Lead Scoring Model With Lead Profiling 114

5.1 Introduction 114
5.2 Background 118
5.2.1 Marketing-sales funnel 118
5.2.2 Buyer purchase decision-making process 119
5.2.3 Application of machine learning (ML) and data mining (DM) in lead scoring 119
5.3 Methodology 120
5.3.1 Data collection and preparation 121
5.3.2 Machine learning algorithms 123
5.3.3 Model evaluation and imbalanced dataset 124
5.4 Results 126
5.4.1 Modeling procedure 126
5.4.2 Modeling results 127
5.4.3 Profiling the clusters 130
5.5 Discussion and conclusion 131
5.5.1 Implications 132
5.5.2 Research contributions 134
5.5.3 Limitations and future research 135
References 144
Chapter 6. Discussion 151
6.1 Conclusion and implications 151
6.2 Limitation and future work 153
6.2.1 Potential biases in data 153
6.2.2 Model generalization 154
6.2.3 Real-world case study 154
6.2.4 Cutting-edge alternative model and learning method 155
References 157

Chapter 1: Introduction

Companies invest heavily in advertisements, trade shows, database purchases, web campaigns, and telemarketing to generate leads and dedicate enormous effort and resources to nurture and convert these leads into customers (Sabnis et al., 2013; Smith et al., 2006). A *Lead* is a member of a target market segment who has signaled an interest in the company's product(s)/service(s) or requested information, regardless of whether that particular interest comes from a new prospect or an existing customer (Carroll, 2006; Monat, 2011). Conventional, outside sales (also called field sales) have been primarily based on in-person interactions with leads. In contrast, *inside sales* mainly rely on remote sales done with the help of Information and Communication Technologies (ICT) (e.g., phone, Internet) (Ohiomah et al., 2016, 2019, 2020). *Lead Management Systems* (LMSs), being a "driving force" of inside sales, use various Information Technology (IT) tools to streamline and automate complicated lead management processes (Ohiomah et al., 2019), for example, lead generation, lead nurturing, lead distribution, and lead scoring (Järvinen & Taiminen, 2016; Michiels, 2008, 2009; Syam & Sharma, 2018, Velocity, 2003).

The modern sales function has ushered in an era marked by innovative interactions between buyers and sellers (Pullins et al., 2017). The increasing cost of conventional field sales, advances in technology tools, and buyers' higher demands and expectations have contributed to the rapid growth of inside sales (Rutherford et al., 2014; Thaichon et al., 2018). For the last two decades, we have observed a significant shift from conventional/field sales to the dominating inside sales enabled by ICTs. The current COVID pandemic forced many organizations to reduce costs and eliminate unnecessary spending (Patma et al., 2021). For this reason, it has become increasingly essential for organizations to maximize opportunities from new prospects and existing customers by taking advantage of inside sales. While the early adoption of inside sales involved supporting outside salespeople, many contemporary organizations have completely embraced inside sales covering business activities from prospecting to exclusive account management (Harmon & Funk, 2014). Inside sales became the only way to sell products and services for some industries. However, not only the practice of selling (i.e., traditional vs. ICT-enabled inside sales) has evolved, but also inside sales have further benefited from the switch towards queue-based LMSs. A lead management system can be either list-based or queue-based (Ohiomah et al., 2016). A list-based LMS gives a salesperson freedom to choose which leads to contact given a long unfiltered list of leads. On the other hand, a queue-based LMS applies pre-defined business and other rules to automatically filter and assign the next best lead to a salesperson to contact. Queue-based LMSs proved to be more efficient and, as a result, became the leading LMSs for inside sales. Increased productivity, deeper list penetration, more efficient management control, and quicker lead response have made queue-based LMSs the most promising solution for managing inside sales (VanillaSoft, 2014).

For most inside sales organizations, practitioners are assigned with customer acquisition duties. These include sales lead generation and scoring/qualification, contacting leads through various ICTs, organizing meetings/appointments with leads, and eventually converting to customers/sales (Harmon & Funk, 2014; Zoltners et al., 2013). Lead generation and scoring are the tasks in the prospecting stage of the business-to-business (B2B) inside sales process (Paschen et al., 2020). In prospecting stage, the goal is to find potential customers (leads) and identify high-quality leads (who have a higher chance of conversion into customers). In the past, marketers and salespeople consider different factors in deciding what leads to pursue out of a large pool of leads (D'Haen et al., 2013). They usually rely heavily on personal experience, intuition, and gut feeling to make such a pursuing decision. This traditional approach could result in company resources being wasted by dedicating them to the wrong leads (D'Haen & Van den Poel, 2013; McDonnell, 2014; Monat, 2011). A better approach is *lead scoring* which is an effective and efficient way of qualifying the quality of a pool of leads in queue-based LMSs (D'Haen & Van den Poel, 2013; Duncan & Elkan, 2015). Lead scoring prioritizes marketing and sales efforts towards leads that are more likely to be converted to customers (Benhaddou & Leray, 2017; Duncan & Elkan, 2015; Nygård & Mezei, 2020), which is a critical task at the beginning stage of a successful inside sales process.

A *lead scoring model* is an IT tool that evaluates the leads' likelihood of making a purchase by ranking them against a scale to differentiate and prioritize them for sales (Boogar, 2019; Duncan & Elkan, 2015; McDonnell, 2014). A decent lead scoring model usually requires marketing and sales functions to collaborate closely by defining what constitutes a good lead to pursue. There are two categories of lead scoring models: traditional and predictive (Duncan & Elkan, 2015; McDonnell, 2014). Traditional lead scoring is a strategy employed by the marketing team, which endeavors to quantify the quality of a lead (McDonnell, 2014; Michiels, 2008; Monat, 2011). Traditional lead scoring is mostly based on the subjective judgment and experience of marketers and salespeople, while the predictive approach promises to be more objective and effective.

The profound transformation of business operations is being driven by the swift profession of digital technologies, including artificial intelligence, automation, cloud computing, blockchain, the Internet of Things, and etc. (Paschen et al., 2019, 2020; Syam & Sharma, 2018). This evolution is fundamentally altering the ways in which organizations generate and provide enhanced value to their decision-makers. As the major branches of artificial intelligence, various machine learning (ML) algorithms and data mining (DM) models have been applied to deal with tasks in the prospecting stage (Chorianopoulos, 2016; Ngai et al., 2009; Rodrigues Chagas et al., 2018), predictive lead scoring models use advanced data-driven predictive analytics to discover insights within "cold" leads/prospects data, uncover hidden/non-linear relationships between predictors and target/outcome event, and finally estimate a propensity score for each

new prospect (Boogar, 2019; Chorianopoulos, 2016; McDonnell, 2014; Nygård & Mezei, 2020). In predictive lead scoring models, DM and ML techniques could be leveraged to identify hidden patterns within data, filter out the most influencing attributes, and generate models with accurate predictions based on historical data. The utilization of these predictive analytics have the potential to optimize inside sales activities, improve lead conversion rates, and enhance the overall sales performance, while minimizing the reliance on arbitrary rules employed by sales practitioners.

Some results generated by traditional lead scoring can be inaccurate, which exerts minimal impact on sales performance or even misleads the direction of sales. For instance, under the traditional lead scoring, salespeople usually spend too much time dealing with a large volume of low-quality market-qualified leads (MQLs) that most unlikely become sales-qualified leads (SQLs) or genuine sales at the end (McDonnell, 2014; Nygård & Mezei, 2020). Instead of hiring more salespeople to make contact with leads, which is expensive, more efforts should be put into building predictive lead scoring models that can produce high-quality lists of MQLs during the prospecting stage of the sales process (Buttle, 2009; Duncan & Elkan, 2015). Predictive lead scoring models can produce much-refined lists of leads in the prospecting stage for salespeople to contact in the following pre-approach and approaching stages.

Furthermore, the capability of processing a large amount of data is another reason why predictive lead scoring is better than traditional lead scoring (McDonnell, 2014). A successful prediction of lead conversion relies on the analysis of vast amounts of data. Traditional lead scoring does not have the ability to analyze vast amounts of potential customers' data due to the lack of computational power. Predictive lead scoring is the right solution for forecasting the likelihood of prospects converting to customers, given its high computational and analytical ability. Meanwhile, traditional lead scoring models may fail to capture nonlinear effects between independent and dependent variables or complicated interactions between features (Duncan & Elkan, 2015). These disadvantages indicate spending resources on converting low-quality prospects who are unlikely to convert at the end, which may degrade sales performance. Additionally, predictive lead scoring can help find complex relationships between predictors and target event/outcome as well as identify trends and the most determinant features in the prospects data. In terms of the lead conversion rate which is a common metric used in assessing the performance of lead scoring models, research shows that the lead conversion rate (i.e., from prospects to customers) increases by 5% on average as a result of using predictive lead scoring models, compared to an average 5% conversion rate in the traditional approach. (D'Haen & Van den Poel, 2013; Duncan & Elkan, 2015).

Despite the evident promise of predictive lead scoring, there remains a dearth of understanding regarding the factors and practices conducive to enhancing the efficacy of predictive lead scoring (Başarslan & Argun, 2020; Eitle & Buxmann, 2019; Nygård & Mezei, 2020). This knowledge gap is particularly

significant given predictive lead scoring's pivotal role in the digital transformation of B2B inside sales context.

1.1 Research problem and motivation

Lead scoring and conversion to sales are the core components of the inside sales process (Pullins et al., 2017). Qualified leads that do not result in short-term sales often slip through and become lost revenue opportunities without an appropriate inside sales lead management strategy (Michiels, 2008, 2009). The average conversion rate of prospects to qualified leads is approximately 10% and merely 1-6% of leads will ultimately become customers (Coe, 2004; D'Haen & Van den Poel, 2013; Duncan & Elkan, 2015; Eitle & Buxmann, 2019). In this thesis, we rely on Churchill et al.'s description (1985, p.116) of sales effectiveness to define sales performance as "*a summary index of organizational outcomes for which an individual is at least partly responsible.*" The likelihood of conversion directly influences sales performance (Ohiomah et al., 2020), and both represent significant challenges for inside sales. There is a need to find a better way to increase the efficiency of the sales team and improve conversion rate in LMS-enabled inside sales (Duncan & Elkan, 2015; McDonnell, 2014). In addition, some determinants of sales success are more robust when selling remotely (i.e., when engaging in inside sales) (Ohiomah et al., 2020). Thus, this thesis's primary challenge and motivation are finding a better way to increase sales performance in B2B inside sales.

Research has shown that effective lead scoring and qualification in inside sales can reduce budgets and maximize revenue by focusing on the quality and not the number of leads (Michiels, 2008, 2009). Lead scoring is widely acknowledged as a promising way to qualify leads and assist with the low conversion challenge (D'Haen & Van den Poel, 2013; Duncan & Elkan, 2015; Lindahl, 2017; McDonnell, 2014; Monat, 2011). Companies that employ lead scoring in the LMSs can potentially benefit from up to a 70% increase in lead generation return on investment compared to companies that do not use lead scoring (McDonnell, 2014). According to a 2018 report from Global Entrepreneurship Monitor (GEM), on average, three new companies are created every second (Jaskaran, 2018). This means that at least 0.8 million new companies are created overnight. A salesperson needs to contact approximately 800 leads per day even if only 0.1% of the 0.8 million prospects are relevant to the company. Despite the number of leads, being able to differentiate between high-quality and low-quality leads in such a rapidly growing market is challenging.

Lead scoring models are emerging as a solution to that challenge, yet there remains limited understanding of the methods and criteria suitable for inside sales. The existing literature is characterized by three major limitations. First, despite the recognized importance of leads scoring models in inside sales and the imperative to explore how these models could effectively address the challenges of inside sales performance, there is a notable absence of studies summarizing the knowledge about existing lead scoring models and their impact on sales performance. Although a few studies have been dedicated to this subject

(Başarslan & Argun, 2020; Eitle & Buxmann, 2019; Nygård & Mezei, 2020), there is still a gap in understanding existing lead scoring models, types, advantages, disadvantages, the algorithms employed in building them, their appropriateness and efficiency under specific conditions (such as data sources), and how these models influence sales performance.

Second, in the dynamic and rapidly evolving business environment, grasping the needs and preferences of leads is paramount in the prospecting stage and ultimately for attaining inside sales success (Homburg et al., 2011; Rapp et al., 2012). Yet, existing sales research efforts have predominantly concentrated on the organizational and salesperson's viewpoint (e.g., Ohiomah et al., 2019; Pöyry et al., 2017; Sabnis et al., 2013; Smith et al., 2006), neglecting a crucial aspect --- the perspective of leads. More specifically, there is a notable gap in understanding how to model buying behavior and the decision-making process in B2B inside sales from the lead's perspective. Therefore, we are motivated by the desire to uncover profound insights into the phased purchase decision-making process and critical factors that influence inside sales success from the lead's perspective. By understanding what different leads truly need in various phases, we aim to integrate these insights into the implementation of predictive lead scoring models and anticipate improved performance.

Third, predictive lead scoring models in existing studies are built on the foundation of the marketing-sales funnel (D'Haen & Van den Poel, 2013; Duncan & Elkan, 2015; Eitle & Buxmann, 2019), which was transformed from the famous seven steps of selling (D'Haen & Van den Poel, 2013; Dubinsky, 1981). A B2B selling organization could model the complex sales process by using the marketing-sales funnel. However, the marketing-sales funnel is seller-centric. The recognition of leads as the initiating point of business activities is strong evidence that the sales process has become increasingly lead-centric (e.g., Shannahan et al., 2013; Sheth et al., 2000, 2009). Sellers have begun to consider how to better respond to leads' various demands, purchase intentions, and behaviors in each phase of the sales process, instead of actions to influence leads. Nevertheless, there is a lack of research exploring how to implement predictive lead scoring models that are both seller and lead-centric. This research aims to do just that. Much literature has studied industrial customer buying behavior and suggested that purchase decision-making is influenced by buyers' demographic and behavioral factors (Dickson et al., 2009; Sharma, 1997; Sheth, 1973). The literature suggests that the B2B buying decision-making process could be explained through a buying funnel, which is referred to as a staged process that a lead takes in order to purchase a product or service for the organization (Jansen & Schuster, 2011). Thus, we leverage the phased buying funnel from the lead's perspective to study the factors that influence a lead's decision-making. Eventually, we integrate the findings into the implementation of the predictive lead scoring model. The data-driven predictive analytics can help provide insights into lead scoring and qualification activities and improve lead conversion rates.

1.2 Research questions and objectives

To address the aforementioned limitations in the inside sales and lead scoring literature, this doctoral research aims to explore the following research questions:

Research Question 1 (RQ1): *What are the advantages and disadvantages of the existing lead scoring models and how do they influence sales performance?*

Research Objective 1 (RO1): reviewing and analyzing the existing literature and industrial white papers on lead scoring models as well as their impact on sales performance using a systematic literature review. Proposing a classification framework to classify lead scoring models and suggests ways of improving sales performance in lead scoring models. Furthermore, suggesting which learning methods (i.e., supervised and/or unsupervised) should be used when building predictive lead scoring models given the availability of data sources.

Research Question 2 (RQ2): *How do drivers and enablers influence a buyer's purchase decision-making process in a B2B inside sales context?*

Research Objective 2 (RO2): conceptualizing and empirically validating critical factors that influence a B2B inside sales buyer's purchase intention during the purchase decision-making process. Additionally, conducting multigroup analyses to reveal buyers with various demographics need to be approached and treated with different selling strategies in the sales process.

Research Question 3 (RQ3): *How can predictive analytics optimize lead scoring outcomes for B2B inside sales?*

Research Objective 3 (RO3): developing and implementing a predictive lead scoring model considering the perspectives of both sellers and leads. Applying both unsupervised and supervised machine learning methods to provide concise and high-quality lists of leads to improve leads conversion rates and reduce cost of acquisition.

Each research question is motivated by the abovementioned limitation respectively.

1.3 Research contributions

The major contribution of this doctoral research is the theoretical design and implementation of a "smart sales" predictive lead scoring model and the recommendation of effective lists of sales leads that improve B2B inside sale success. To the best of our knowledge, this is the first study that uses two theoretical foundations for modeling lead scoring process, the integrated marketing-sales funnel from a seller's perspective (D'Haen & Van den Poel, 2013; Järvinen & Taiminen, 2016), and the B2B buyer purchase decision-making process from a lead's point of view (Jansen & Schuster, 2011). The existing studies on lead processing and scoring models only focus on the conversion steps in the marketing-sale funnel and consider from sellers' points of view (Benhaddou & Leray, 2017; D'Haen & Van den Poel, 2013; Duncan

& Elkan, 2015; Nygård & Mezei, 2020). However, we showed that the sales performance (e.g., lead conversion rate) could be further improved by integrating insights discovered from the lead purchase decision-making process into the modeling process. Specifically, we

- Identified, evaluated, and analyzed various lead scoring models. In particular, we conducted a systematic literature review of the existing body of knowledge focusing on summarizing conventional methods, data mining models, and machine learning algorithms applied to lead scoring to uncover future research avenues. Furthermore, we proposed a classification framework and used it to classify lead scoring models, summarized modeling processes, and examined the models' impact on sales performance. Additionally, we suggested ways of improving sales performance in lead scoring models. Most importantly, we recommended which learning methods should be used when building predictive lead scoring models, given the availability of data sources.
- Proposed and empirically validated a B2B buyer's intention to purchase decision-making model based on theories of the B2B purchase decision-making process (Jansen & Schuster, 2011; Puccinelli et al., 2009; Sirakaya & Woodside, 2005) and buying behavior (Johnston & Lewin, 1996). We captured key drivers and enablers of a B2B inside sales buyer's purchase intention.
- Developed and implemented predictive lead scoring models using a compositive approach (i.e., clustering then classification) that considers leads' profiles (demographics of leads) and the interactions between salespeople and leads. The proposed "smart sales" predictive lead scoring model assists industry practitioners in gaining a better understanding of challenges and opportunities in lead scoring process, while helping academic researchers in the data science field understand the implications of advanced predictive analytics in B2B inside sales success.

1.4 Related publications

- Wu, M., Andreev, P., & Benyoucef, M. (2024). Smart Sales: Amplifying the Power of Predictive Lead Scoring in B2B Sales. *The Pacific Asia Conference on Information Systems (PACIS) 2024 Proceedings. 3.* https://aisel.aisnet.org/pacis2024/track09_digittrans/track09_digittrans/3
- Wu, M., Andreev, P., Benyoucef, M., & Hood, D. (2024). Unlocking B2B buyer intentions to purchase: Conceptualizing and validating inside sales purchases. *Decision Support Systems, 179*. 114165. https://doi.org/10.1016/j.dss.2023.114165
- Wu, M., Andreev, P., & Benyoucef, M. (2023). The state of lead scoring models and their impact on sales performance. *Information Technology and Management*, *25*(1), 69–98. https://doi.org/10.1007/s10799-023-00388-w
- Wu, M., Andreev, P., & Benyoucef, M. (2024). Data-driven predictive lead scoring model with lead profiling. *Decision Support Systems.* (*working paper*)

1.5 Thesis outline

The remainder of this doctoral thesis is organized as follows. In the second chapter, we examined lead scoring models and identified metrics that have been used to measure the impact of lead scoring models on sales performance. The third chapter elaborates how drivers and enablers influence a B2B inside sales buyer's intention to purchase during various phases in the purchase decision-making process. The fourth chapter proposes a predictive lead scoring model and empirically validates the effectiveness of it using a small B2B dataset. As a result, the model generates much-refined prioritized lists of sales leads and provides insights on effective lead scoring process practices in the sales pipeline with the consideration of both sellers and leads. On the basis of chapter four, the fifth chapter further validates the proposed "smart sales" lead scoring model by using larger datasets collected from three service industries and elaborates on the analysis of leads' profiling. The study highlights the important role of data-driven predictive modeling in shifting from human decision-making to ML model-based approach in lead scoring task. In the conclusion chapter, this doctoral thesis by articles concludes with a brief summary of the implications arising from our research findings for the inside sales literature and the sales practices.

Chapter 2. The state of lead scoring models and their impact on sales performance



Article Reference: Wu, M., Andreev, P., & Benyoucef, M. (2023). The state of lead scoring models and their impact on sales performance. *Information Technology and Management*, *25*(1), 69–98. https://doi.org/10.1007/s10799-023-00388-w

Article Abstract: Although lead scoring is an essential component of lead management, there is a lack of a comprehensive literature review and a classification framework dedicated to it. Lead scoring is an effective and efficient way of measuring the quality of leads. In addition, as a critical Information Technology tool, a proper lead scoring model acts as an alleviator to weaken the conflicts between sales and marketing functions. Yet, little is known regarding lead scoring models and their impact on sales performance. Lead scoring models are commonly categorized into two classes: traditional and predictive. While the former primarily relies on the experience and knowledge of salespeople and marketers, the latter utilizes data mining models and machine learning algorithms to support the scoring process. This study aims to review and analyze the existing literature on lead scoring models and their impact on sales performance. A systematic literature review was conducted to examine lead scoring models. A total of 44 studies have met the criteria and were included for analysis. Fourteen metrics were identified to measure the impact of lead scoring models on sales performance. With the increased use of data mining and machine learning techniques in the fourth industrial revolution, predictive lead scoring models are expected to replace traditional lead scoring models as they positively impact sales performance. Despite the relative cost of implementing and maintaining predictive lead scoring models, it is still beneficial to supersede traditional lead scoring models, given the higher effectiveness and efficiency of predictive lead scoring models. This study reveals that classification is the most popular data mining model, while decision tree and logistic regression are the most applied algorithms among all the predictive lead scoring models. This study contributes by systematizing and recommending which machine learning method (i.e., supervised

and/or unsupervised) shall be used to build predictive lead scoring models based on the integrity of different types of data sources. Additionally, this study offers both theoretical and practical research directions in the lead scoring field.

Keywords: Lead scoring model; Sales performance; Data mining model; Machine learning algorithm; Systematic literature review

Note: This chapter is an author accepted manuscript submitted to the journal referred to above.

2.1 Introduction

2.1.1 Inside sales and lead scoring modeling

A *lead* is an essential raw material for sales organizations. Leads, being members of a target market segment, intentionally or unintentionally signal an interest in a company's product(s)/service(s), regardless of whether that particular interest comes from a new prospect or an existing customer [14, 44]. Companies invest significantly in advertisements, web campaigns, and marketing to generate new leads and allocate enormous resources to nurture and convert these leads into customers [56, 59]. Conventional, outside sales (also called field sales) that are primarily based on in-person interactions with leads have been giving up the leading role to *inside sales* that mainly rely on remote sales conducted with the help of information and communication technologies (ICT) (e.g., phone, Internet) [49–51]. For some industries, inside sales became dominant and sometimes the only way to sell their products and services. The increasing cost of conventional sales, as well as advances in information technology (IT) tools and buyers' higher demands and expectations, have contributed to the rapid growth of inside sales [55, 62]. For the last two decades, we have observed a significant shift from conventional field sales to the dominating inside sales enabled by ICT. The current COVID-19 pandemic forced many organizations to reduce costs and eliminate unnecessary spending [75]. For this reason, it has become increasingly essential for organizations to maximize opportunities from new prospects and existing customers by taking advantage of inside sales.

Lead Management System (LMS), an integrated information system of inside sales, became the "driving force" for operations with leads. LMS uses various IT tools to streamline and automate complicated lead management processes [49], for example, lead generation, lead nurturing, lead distribution, and lead scoring [28, 42, 43, 61, 66]. However, not only the way of selling (i.e., traditional vs. ICT-enabled inside sales) has evolved during the last decades, but inside sales have further benefited by shifting from list-based (manually prioritizing and filtering of leads based on sales representatives' knowledge and experience) to queue-based LMSs (an approach for prioritizing leads when the most promising leads are served first) [49, 50]. The increased productivity, more efficient management control, and quicker response to leads have made queue-based LMSs the best solution for managing leads in inside sales [65].

Lead scoring has been widely acknowledged as the most effective and efficient way of qualifying the quality of a large number of leads for queue-based LMSs [11, 17, 20, 37, 39, 44]. *Lead scoring modeling* is at the core of lead scoring, a qualification approach that assesses the leads' likelihood of making a purchase by ranking them against a scale to differentiate and prioritize them by generating a queue-based list for sales [7, 20, 48]. A high-quality lead scoring model with superior predictive power could convince salespeople to contact more market-qualified leads (MQLs) and convert those "ready-to-buy" leads to customers in a short time [25, 56]. From a long-term perspective, having a high-quality lead scoring model can also improve the internal collaboration between the marketing and sales functions [56].

2.1.2 Problem and motivation

Lead qualification and conversion to sales are the most critical success components of the inside sales process [54]. Without an appropriate inside sales lead management strategy, qualified leads that do not result in short-term sales often slip through and become lost revenue opportunities [42, 43]. The average conversion rate of prospects to qualified leads is approximately 10%, and only 1-6% of leads ultimately become customers [16, 17, 21]. Such low conversion rates of leads to customers are mostly associated with the low quality of leads in queues that sales teams work with [39, 48]. Sales teams spend valuable and limited time resources on low-quality leads that may not be converted [20]. The likelihood of conversion directly influences sales performance [51]. There is an overall challenge to find a better way to increase sales performance and improve conversion rates in inside sales [20, 39]. In addition, some determinants of sales success are stronger when selling remotely (i.e., when engaging in inside sales) [51].

Effective lead management in inside sales can reduce budgets and maximize revenue by focusing on the quality and not the number of leads [42, 43]. Lead scoring has been widely acknowledged as a promising way to assist with the low conversion challenge [17, 20, 37, 39, 44]. Companies that employ lead scoring in their LMSs can potentially benefit from up to 70% increase in lead generation return on investment compared to companies that do not use lead scoring [39]. In addition, the conversion rate from prospects to qualified leads increases to 15-20% which means that eventually, more leads will convert to sales [17, 20]. According to a 2018 report from GEM (Global Entrepreneurship Monitor), on average, three new companies are created every second [29]. This means that at least 0.8 million new companies are created over one night. A salesperson needs to contact approximately 800 leads per day, even if only 1% of the 0.8 million prospects are relevant to the company. Despite the number of leads, being able to differentiate between high-quality and low-quality leads in such a rapidly growing market is challenging.

Lead scoring models are emerging as a solution to that challenge but still little is known about how and what lead scoring models need to be employed for inside sales. Despite the importance of lead scoring models in inside sales and the call to find out how these models can address the challenge of inside sales

performance; no study summarized the knowledge about existing lead scoring models and their impact on sales performance. Although a few studies have been dedicated to the subject [6, 21, 48], little is known regarding existing lead scoring models, their types, advantages and disadvantages; what algorithms have been used in building lead scoring models; which of them are more appropriate and efficient for specific conditions (i.e., data sources), and how lead scoring models influence sales performance. This paper aims to address this challenge. From a theoretical perspective, our motivation is to conduct a systematic literature review in the field of lead scoring models to identify research areas that require further investigation.

By knowing what lead scoring models exist and their corresponding suitability, we could choose which one to use, given the availability of data sources. As more efficient and effective machine learning (ML) algorithms are introduced, marketing teams can implement more sophisticated lead scoring models by integrating the more advanced algorithms to handle datasets with higher degrees of complexity. Because large datasets with higher degrees of complexity normally contain more hidden signals of good potential customers [15, 29]. Therefore, it is imperative to grasp the knowledge of algorithms that have been applied to build lead scoring models. With more hidden signals extracted from datasets, more profitable leads could be identified, thus, sales performance could eventually increase.

2.1.3 Scope and contribution

To foster our understanding of how to improve inside sales performance, what role lead scoring models play in this improvement, and how different types of such models can influence sales performance, it is essential to summarize the existing knowledge on the domain. Because it is important to investigate the existing types of lead scoring models and their impact on sales performance, we argue that a framework is needed to classify existing lead scoring models. A systematic literature review (SLR) on lead scoring models should help fill the above-mentioned gaps. Hence this study proposed and addressed the following research questions (RQs):

- RQ1. What are the advantages and disadvantages of the existing lead scoring models?
- RQ2. What is the preferred model for which data source?
- RQ3. How do lead scoring models influence sales performance?

The main contributions of this study to academic research and sales practice are:

1. This study identifies, evaluates, and analyzes various lead scoring models. In particular, it focuses on summarizing conventional methods, data mining (DM) models, and ML algorithms applied to lead scoring to uncover future research avenues.
2. Furthermore, this study proposes a classification framework and uses it to classify all the identified lead scoring models, summarize modeling processes, examine the models' impact on sales

performance, and compare models' impact to suggest lead scoring models. Additionally, this study suggests ways of improving sales performance in lead scoring models.

3. Moreover, since predictive lead scoring has become the trend, this study investigates the reason why the predictive approach is better than the traditional approach.
4. Most importantly, this study recommends which learning methods (i.e., supervised and/or unsupervised) should be used when building predictive lead scoring models, given the availability of data sources.

The rest of the paper is structured as follows. Section 2.2 Methodology describes the methodology employed in this study, followed by Section 2.3 Search results, which discusses the proposed classification framework and shows the results of the literature review. Section 2.4 State of lead scoring models expatiates answers to the research questions and the limitations of this study. Finally, Section 2.5 Discussion and conclusions presents the discussion, conclusions, and implications of this study.

2.2 Methodology

In this review, we followed Kitchenham's SLR approach [34], which consists of three main steps, namely planning the review, conducting the review, and reporting on the review. We defined the review's objectives and developed a review protocol in the planning step. In the conducting review step, we executed search queries, selected studies, and assessed their quality. Finally, we extracted and synthesized the data in the reporting step. In addition, we validated, analyzed, and described the results and tabulated them in quantitative summaries. The entire process is detailed below.

2.2.1 Planning the review

Search strategy: After defining research questions, we identified the concepts in the two research questions involving two disciplines: (1) computer science and (2) business management. Therefore, an interdisciplinary literature search needed to be carried out for this SLR. Hence, we used synthetic databases such as Scopus and Web of Science to account for the interdisciplinary nature of the study. Furthermore, the technology and science-focused Institute of Electrical and Electronics Engineers (IEEE) Xplore library and business-focused databases such as Business Source Complete and ABI/INFORM Global were scanned for relevant studies. Moreover, since lead scoring systems have been employed in the industry, we searched grey literature to ensure complete coverage of industrial technical reports, research papers, project reports, and white papers. Typically, a grey literature scan is necessary to address publication bias in SLRs. For the grey literature search, we searched the OpenGrey database the same way we searched traditional literature databases.

To define the search queries, first, the research questions were decomposed into four concepts: (1) Lead Scoring, (2) Modeling, (3) Sales, and (4) Performance. Second, keywords were generated for each concept

by using relevant background knowledge in the fields, the pearl growing technique [52], and brainstorming (see **Table 2.1**).

Table 2.1 appears here

Third, we conducted a keyword search on the aforementioned databases. The final search queries were applied to the article title, abstract, and keywords fields. There was no need to include "full-text" in the search field since this would have led to many false positives at this search stage.

Inclusion/Exclusion criteria: In the evaluation stage, the retrieved results were evaluated against the following inclusion criteria:

- Peer-reviewed journal articles and conference proceeding papers that focus on lead scoring models, or applying DM models or ML algorithms to lead scoring;
- Relevant papers on lead scoring models that are identified by the snowballing technique [12];
- Grey literature: industrial technical reports, project reports, and white papers on lead scoring models.

Next, the included studies were evaluated against the following exclusion criteria:

- Language: not in English;
- Subject area: not in the business management domain;
- Primary focus: not related to lead scoring;
- Study form: studies in the form of abstracts or posters.

Quality criteria: Each included study must meet all the following quality assessment criteria:

- A study focuses on a lead scoring technique within the business scope;
- A study addresses the impact of the proposed lead scoring model(s) on sales;
- A study includes a performance evaluation scheme for evaluating the alleged lead scoring model(s);
- Grey literature articles must address both traditional and predictive lead scoring for comparison purposes. Evaluation metrics can be left out in grey literature. Because it is an uncommon practice to include model performance evaluation schemes in grey literature.

2.2.2 Conducting the review

Identification: After locating a few key papers by entering the preliminary search queries into the databases [7, 17, 20, 48, 60], a pearl growing technique (i.e., using keywords and index terms of key papers) [52] was executed to optimize search terms in the initial search queries. The subject headings and keywords of these key papers were used to optimize search terms and refine the preliminary search queries. Eventually, the final search queries were run in six databases. Additionally, a backward snowballing technique [3, 12] was

adopted to complement literature database searches. **Table 2.2** shows the number of articles retrieved from each source.

Table 2.2 appears here

Screening: Then, we removed duplicates from a total of 1150 records. After excluding 345 duplicated records, we ended up with 805 papers. **Figure 2.1** shows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram for visualizing each stage's selection process and results [36].

Figure 2.1 appears here

A total of 685 records were excluded, leaving 120 articles for full-text quality assessment (see **Figure 2.1**). We used the inclusion/exclusion criteria, so articles fulfilling any of the exclusion criteria were disqualified. Specifically, we read each study's title, abstract, and keywords.

Eligibility and inclusion: After a full read of the 120 articles, a total of 44 studies were retained for qualitative synthesis using the quality criteria. The reasons for excluding articles during the full-text assessment phase were ineligible areas of focus, not addressing the impact on sales, retracted articles, and missing model evaluation scheme. For example, Hou and Yang [27] presented a classification model for potential customer identification and prioritization via the 80-20 principle. However, they omitted to address the evaluation metrics for assessing the proposed model. Thus, this study was not included in the final list of qualified studies. Xiaowen et al. [70] established a traditional lead scoring system of a three-layer value assessment structure via the analytic hierarchy process to locate potential government and corporate customers. However, this study lacked an evaluation of the proposed system and measurement of sales performance.

Some studies failed to meet the focused area criterion. For instance, Nguyen et al. [47] introduced a new model to identify the characteristics of customers using the rough set theory. However, their study mainly focused on distinguishing customers' characteristics while staying away from applying the lead scoring process. Therefore, this study was excluded from the final list. Furthermore, studies by Baecke and Van Den Poel [4, 5] focused on incorporating spatial interdependence feature with auto-correlation and regression techniques to improve existing customer acquisition models instead of building a lead scoring model and analyzing its impact on sales. Hence these studies were excluded. Moreover, a qualitative case study by Jarvinen and Taiminen [28] demonstrated the use of marketing automation tools to generate high-quality sales leads through behavioral targeting and personalizing content without discussing lead scoring. Thus, this study was excluded.

2.2.3 Reporting on the review

All lead scoring models in the qualified studies were extracted and categorized to make the review unambiguous and comprehensive to address the research questions. Each study was thoroughly reviewed and classified according to the two lead scoring approaches and six classes of models (see **Figure 2.2**). We validated the result and performance of each lead scoring model by assessing the proposed evaluation metrics and sales performance measurements. Additionally, we extracted all impact/influences of lead scoring models on sales performance. In the end, we analyzed how lead scoring models influence/impact sales performance.

The data extracted from the qualified studies (see **Appendix: table of extracted literature data**) includes:

- Title, authors, and year of the paper.
- Suggested methods/models/algorithms.
- Approach (traditional or predictive) of the proposed lead scoring models.
- Evaluation metrics applied to the proposed lead scoring models.
- Metrics used to measure sales performance.
- A summary of lead scoring models' impact on sales.

The results of the literature analysis are presented in the following section.

Figure 2.2 appears here

2.3 Search results

We developed a classification framework for traditional and predictive lead scoring models to summarize the existing models (**Figure 2.2**). This framework is based on a comprehensive review of the academic and grey literature on lead scoring models in LMSs. According to the studies conducted by Duncan and Elkan [20] and McDonnell [39], lead scoring can be split into two approaches: traditional and predictive. Additionally, Chorianopoulos [15] and Ahmed [2] described the major types of DM models in Customer Relationship Management (CRM), namely classification, clustering, and regression. Given the flexibility of this classification framework, it can be expanded as more traditional techniques or DM models are applied to lead scoring. In the following paragraphs, a short description of two lead scoring approaches, as well as traditional techniques and DM models are provided with some references for more details.

According to Duncan and Elkan [20] and McDonnell [39], lead scoring consists of two approaches, traditional and predictive. These two approaches have been studied by academics and practitioners. They share a common goal of scoring and prioritizing leads according to their likelihood of purchase. In traditional lead scoring, marketers attempt to quantify the quality of a lead to determine when it should be passed to salespeople [39].

In traditional lead scoring, marketers usually analyze explicit (e.g., industry type, job role, company size, and revenue) and implicit (e.g., website visits, email opens, clicks, form completions, and online behaviors) information on leads. They assign scores to leads based on criteria/ rules and track them. Traditional lead scoring is usually supported by marketing automation software, such as Oracle Eloqua [37]. There are three main models in the traditional lead scoring segment, which are Lamb or Spam [11], rules/points-based [11, 29, 39], and scorecard [44, 67]. The Lamb or Spam model filters out low-quality leads and surfaces relatively high-quality ones by assessing their attributes (e.g., email domain, company size) [11]. Rules-based or points-based lead scoring assigns points to leads' demographic and behavioral characteristics according to specific rules. These rules are stipulated based on human experience and intuition [29]. As a result, the lead score is the outcome of a weighted function of these attributes. The scorecard model is similar to the rules/points based model, except for one major difference: the score is calculated by statistical and mathematical approaches based on the different factors' importance levels [67].

Predictive lead scoring uses advanced data-driven predictive analytics to discover insights within "cold" leads/prospects data, uncover hidden/non-linear relationships between various predictors and target/outcome events, and finally estimate a propensity score for each new prospect [11, 15, 39, 48]. Within the context of predictive lead scoring, DM and ML techniques can be leveraged to identify various data patterns, filter out the most influential attributes, and generate predictive models based on historical data [15]. They can be leveraged to guide the decision-making process and predict decisions' effects [46]. Predictive lead scoring is supported by different DM models, including classification, clustering, and regression.

Classification is commonly used in DM [2, 15, 46]. Its goal is to build a model to predict the outcome of an event by classifying new records to the predefined classes [2, 15]. The algorithms commonly used for classification are decision trees, logistic regression, and neural networks. Clustering segments a heterogeneous population into a few homogeneous clusters [2, 46]. The major difference between clustering and classification is that the number of clusters is unknown in clustering. The commonly applied clustering algorithm is k-means. Regression is a frequently used statistical estimation technique for predicting the value of a continuous output based on the inputs [15, 46]. Regression has been applied to test the significance of relationships between variables, fit curves, and predict continuous outcomes. Linear regression is the most common technique.

Table 2.3 shows the 44 qualified studies according to the proposed classification framework. The number of lead scoring studies has increased in the last few years. **Figure 2.3** shows a trend against the timeline of the qualified lead scoring studies covered in this SLR. Out of the 44 selected studies, 39 are journal or conference papers, while 5 are grey literature (**Table 2.4**). We can observe that the selected

studies emphasize predictive lead scoring more than traditional lead scoring, showing the current trend in the research field. We analyzed predictive lead scoring models in all the qualified predictive studies. The literature analysis reveals that there are 18 different predictive lead scoring models. **Table 2.5** shows that the most popular models are decision tree classification and logistic regression.

Table 2.3 appears here

Figure 2.3 appears here

Table 2.4 appears here

Table 2.5 appears here

Table 2.6 shows the number of traditional lead scoring models in the qualified studies. The most popular one is the rules/points-based model. **Table 2.7** shows various metrics used to measure sales performance after applying the proposed lead scoring model(s) in all the qualified studies. The lead conversion rate is the most popular metric.

Table 2.6 appears here

Table 2.7 appears here

2.4 State of lead scoring models

As an IT tool in LMSs, a lead scoring model prioritizes sales and marketing efforts towards leads that are more likely to convert into customers [7, 20, 48]. Marketing and sales usually collaborate to build lead scoring models by defining what constitutes a good lead to pursue. There are two approaches to lead scoring: the traditional lead ranking method and the more advanced data-driven predictive approach [20, 39]. Traditional lead scoring endeavors to quantify the quality of a lead [39, 42, 44]. Its goals are to prioritize leads to sales and develop scalable approaches if leads meet the minimum qualification requirements. Traditional lead scoring is based on salespeople's experience and judgment, while the data-driven predictive approach promises to be more objective and efficient. Various DM models and ML algorithms have been applied to build predictive lead scoring models that assess the likelihood of converting leads to customers [11, 15, 39, 48]. Recently, the sales industry has been leaning more towards predictive lead scoring approaches [11, 29]. Moreover, with an increasing number of studies published on lead scoring models in the last few years, one can claim that this field has been gaining attention from academic research, especially predictive lead scoring. We first discuss lead scoring models in academic literature. As mentioned earlier, lead scoring can be divided into two categories, namely traditional and predictive.

2.4.1 Traditional lead scoring models

In traditional lead scoring, rules/points-based and score- card models were frequently used a decade ago. Monat [44] proposed a practical qualitative modeling tool that predicts the probability of an industrial sales lead converting to a customer based on observable lead characteristics. The author claimed that this is the

first lead characterization model that is theoretically based. A rules/points-based sales lead evaluation scorecard was provided to assess leads from eight determinants of sixteen manifest characteristics. In addition, details of points assignment, scoring procedure, and accuracy measures on a real company dataset were provided in this study. Furthermore, a couple of studies developed simple linear lead scoring models by combining points from several factors [24, 26]. However, none of these studies have been field-validated.

Another conventional way to calculate a lead score is to use a scorecard model. An analytical hierarchy process- based (AHP-based) framework helps companies rank and prioritize prospective leads based on the different factors' importance levels [67]. More specifically, input statistics are normalized on a relative scale for each criterion. On the first row of the scorecard are the weights of all criteria. The remaining rows are leads. The total score of a lead is the sum of all weighted factors in that row [67]. The lead scorecard model enables companies to assess the factors for acquiring potential leads. For instance, Lindahl [37] qualitatively examined traditional lead scoring in the Business-to-Business (B2B) marketing automation domain by exploring how lead scoring contributes to a more efficient and effective marketing process for a B2B service company. The results of this study indicated that the examined scorecard lead scoring model can be used in multiple ways in marketing automation. Indeed, Lindahl [37] presented a complete lead profile scorecard and an intact lead engagement scorecard with a lead score matrix that specifies lead score value thresholds. Additionally, this study recommended corresponding marketing actions for lead score values.

In traditional lead scoring systems, the decision to pursue a lead typically relies on personal experience, intuition, and cognitive capability [23, 29]. These approaches can result in company resources being used inefficiently by dedicating them to the wrong leads [17]. Also, traditional lead scoring models could be error prone due to the manual selection of values and human intervention [11, 20, 29]. Therefore, some of the results generated by traditional lead scoring could be inaccurate and biased. Moreover, traditional lead scoring models may fail to capture nonlinear effects and heavily rely on behavioral data [20].

2.4.2 Predictive lead scoring models

As various DM and ML techniques started to re-emerge, using advanced data-driven predictive analytics to discover insights within leads data and predicting lead scores have become the solution of choice in lead scoring [11, 20, 29, 39]. These techniques can be leveraged to generate predictive models based on historical data to identify various data patterns, filter out the most influencing attributes of leads, and calculate lead scores [15, 29]. As mentioned earlier, we adopted Chorianopoulos [15], Ahmed [2], and Ngai et al.'s [46] classification of DM models. DM models used to build predictive lead scoring models can be categorized into *classification*, *clustering*, and *regression*.

2.4.2.1 Classification

Under the classification category, we unidentified the following common algorithms used for lead scoring models: decision tree, random forest, neural network, and logistic regression.

Decision tree: aims to identify and classify the factors for turning potential customers into "real" customers [31]. Specifically, decision trees can deal with both continuous and discrete attributes for extracting valuable hidden knowledge from leads data. The decision tree consists of rules which can be automatically employed to predict the conversion of leads into customers [38]. Peng and Xu [53] proposed a predictive lead scoring model to identify potential and sustain- able leads. This model adopts correlation analysis to detect relationships between variables and decision trees to find rules for identifying leads. As a supervised learning model [15], decision trees have been used to optimize and stabilize the predictive lead scoring model based on feedback information from the previous phase in the sales process [17]. With regards to the application of decision trees in the sales industry, the same success has been observed. GE Capital built a financing lead triggers system to automate the col- lection and aggregation of information on companies, which was then mined to identify actionable sales leads by using an embedded decision tree algorithm [1]. More specifically, a two-class decision tree was used to identify combinations of financial metrics and values over time that depict patterns common across the positive cases while not present in the negative cases. GE Capital announced that the productivity of salespeople had been improved by 30–50% in terms of phone calls and meetings after deploying the new system [1]. The salesforce's increased productivity and effectiveness have led to a growth in the total volume of deals as approved by GE Capital.

As an ensemble method based on decision trees, bagged decision trees have been used in predicting potential customers during the acquisition process and have shown a decent prediction accuracy [18]. Moreover, as an improvement over the bagging technique, the gradient boosted trees algorithm has been used to prioritize leads based on the probability of conversion to sales opportunities [20]. Duncan and Elkan [20] used the three-class gradient boosted trees algorithm to classify leads with different characteristics into three classes in the lead qualification model. Gokhale and Joshi [23] showed that the two-class boosted decision tree has the best performance in an experiment of a set of ML models when modeling a lead identification and qualification process. A group of Microsoft researchers presented a generic automated lead ranking system based on a boosted decision tree with Bayesian optimization on hyperparameter tuning [30]. A dynamic CRM system integrates a special feature which enables human inputs into the loop for feature engineering and selection. Furthermore, a data mashup approach combining high-scale mobile consumer data with online food company data was introduced to acquire high-value potential customers [58]. A gradient boosted tree was used as a prediction classifier with an RFM (recency, frequency, and monetary value) model to label customers.

Random forest: is considered an all-sided classification algorithm built on the concept of decision trees. As an ensemble learning technique, random forest builds multiple decision trees on different bootstrap data samples [9]. It has been applied to perform classification on textual data. Meire et al. [41] used the random forest model to classify prospects with social media data as input. The random forest has also been used with explanation models for qualifying and classifying prospects based on a set of predefined features and historical data about existing customers [9, 10]. The random forest is considered a "black box" algorithm, making it challenging to interpret the generated results and their implications [9]. Despite the superior capability of analyzing large datasets with complicated relationships between variables, "black box" algorithms cannot generally provide business practitioners with understandable insights that can help decision-making [17]. A couple of explanation methods, namely EXPLAIN and the Interactions-based Method for Explanation (IME) methods, were used to help comprehend how the output was achieved by a given input in "black box" algorithms [10]. A social CRM analytics framework was introduced to improve customer acquisition, conversion, and retention [35]. In this framework, customer acquisition is an optimization task relying on a linear optimization model with random forest for lead classification, and a Latent Dirichlet Allocation (LDA) [8] to uncover the topics mentioned by customers on social media.

Neural network: In a case study conducted in a telecommunications company, a lead qualification model based on ensemble neural networks was implemented to estimate the conversion probability of each lead [22]. The model integrates regression and principal component analysis to select significant variables before building the propensity model. In data mining, a common challenge is the imbalanced number of outcomes in each target class. Imbalanced distribution of class labels within a dataset can have a negative impact on the prediction results of lead scoring models (i.e., a high rate of false positives since many minority class instances are misclassified as belonging to the majority class) [15]. Neural networks do not have an inherent ability to handle imbalanced data effectively. In fact, they can be particularly sensitive to class imbalances because of how they optimize based on the overall dataset. However, some approaches could be applied to help neural networks handle imbalanced data more effectively. One approach involves applying resampling techniques to balance the data before training neural networks. Another method is adjusting the loss function by assigning higher weights to the minority class and lower weights to the majority class, which encourages the network to prioritize the minority class during training, thus reducing bias. Additionally, data augmentation techniques such as rotation transformation, cropping, or noise injection could be used to create new samples for the minority class, particularly in image or text data, to further address class imbalance.

Logistic regression: was found to be a popular classification modeling algorithm for scoring and prioritizing leads. Logistic regression was used to predict the lead conversion probability after an initial list

of features was extracted from a given customer dataset using a forward stepwise regression algorithm [60]. Moreover, Yan et al. [72] proposed a predictive lead scoring model to forecast the win propensity of sales leads over a period. They applied logistic regression to capture and estimate the impact of a salesperson's daily activities and lead information on the sales outcomes (i.e., won or lost). Additionally, a logistic regression model was built based on the concepts extracted from existing customers' websites to predict the probabilities of new profitable leads [63]. Furthermore, a predictive lead scoring model was trained by using logistic regression to dis- cover which concepts (i.e., words concurrently appearing across leads' websites are grouped into "concepts") are more related to converted than unconverted leads [19]. The results of spherical clustering, latent semantic analysis, and expert knowledge are the input sources of the proposed model. As a common supervised learning algorithm, D'Haen and Van den Poel [17] used logistic regression to optimize a predictive lead scoring model by applying a stepwise selection to avoid possible model overfitting. In a case study on targeting potential customers of an energy service, a classification prediction model based on logistic regression was presented to accurately identify and prioritize target customers [68]. The features are selected from four dimensions by using a customer evaluation index system. In another case study on recruiting businesses for a building retrofit project, logistic regression was applied to find prospective leads, then screen and prioritize them for targeting [57].

In addition, companies can use the nearest neighbor algorithm to find similarities among prospects and construct a profiling model to group prospects of similar characteristics into the same group [17]. The nearest neighbor algorithm can be run when there is only available data on the current customer base and a list of prospects. The most significant advantage of the nearest neighbor algorithm is that it does not require prior knowledge of the distribution [17]. Benhaddou and Leray [7] applied the Bayesian network algorithm, a supervised learning that focuses on building probabilistic models to estimate the probabilities of leads belonging to target classes.

Furthermore, an empirical study was conducted to evaluate the feasibility and performance of four algorithms for automating lead scoring by using several assessment metrics [48]. The logistic regression model achieved the highest sensitivity but the lowest specificity. In other words, this model was more capable of identifying a positive class than a negative class. Overall, the random forest model was selected as the best-performing model. Nygård and Mezei [48] showed that automated lead scoring could improve the sales process by revealing insights into sales. A couple of studies [21, 45] agreed on the performance of random forest by demonstrating its superior ability in predicting the probability of a lead conversion or a sales deal. Eitle and Buxmann [21] also extolled the predictive performance of CatBoost after comparing it with decision trees, support vector machine, and XGBoost. In a case study on bank customer acquisition, Başarslan and Argun [6] built multiple classifiers to estimate potential customers. Using k-fold cross-

validation and holdout methods, they found that the best classifier to be random forest with an overall balanced performance among all evaluation metrics.

2.4.2.2 Clustering

With respect to the clustering DM model, k-means has been used to cluster potential customers into various groups for scoring purposes and to develop different marketing strategies accordingly [69]. Additionally, the application of the k-means clustering algorithm helps segment large amounts of customer data into groups and extracts hidden relation- ships from them [38]. The results of k-means clustering provide subjective segmentation, making the data more applicable and informative for further analysis. Self-organizing maps form another clustering approach identified by our review. They have been applied to customers' data to determine the number of clusters prior to cluster analysis [69]. Each cluster contains potential customers with similar behavioral and demographic characteristics.

An Expectation-Maximization (EM) clustering algorithm was used to cluster potential prospects' websites based on prevalent terminologies from the concepts that mainly occur on the websites of profitable business prospects and that seldom occur on the websites of non-profitable customers [63]. Consequently, the results of this clustering analysis can help companies identify profitable leads. D'Haen et al. [19] developed a lead qualification system that integrates expert knowledge and web crawling data to improve lead conversion rate. Specifically, the spherical clustering algorithm was applied to classify documents into a few groups based on a certain similarity measure and discover latent concepts in unstructured text documents. Prospects were clustered according to a spherical clustering. Since each document contains multiple concepts, assigning documents to a single cluster was problematic. Thus, a fuzzy clustering algorithm was also utilized to assist the clustering process [19]. For instance, Wei et al. [69] applied cluster analysis to identify the characteristics of loyal customers, which can be utilized with the RFM model for analyzing customers' values to determine potential customers with a higher profit.

2.4.2.3 Regression

Given the categorical/discrete outcomes of lead scoring (i.e., qualified or not qualified), applications of regression models have been scarce on this subject. Xu et al. [71] proposed a data-driven system by applying linear, exponential regression, and time series seasonal ARIMA model, as well as neural networks to forecast lead conversion rates and estimate sales revenue from opportunities. They claimed their proposed model is applicable to different sales patterns, products, and sales teams.

In summary of the current state of traditional and predictive lead scoring studies, we note that formal validation using statistical means is absent from the traditional lead scoring research stream [26, 44, 67]. Moreover, existing studies on predictive lead scoring only focus on conversion steps in the marketing-

sales funnel [17] from the perspectives of selling organizations [7, 17, 20, 48], hence neglecting insights in the buying decision-making process from the leads' perspectives.

2.4.3 Lead scoring models in grey literature

Regarding the reviewed grey literature, a report by the Aberdeen Group [42] identified the best practices in lead scoring and prioritization by analyzing the top-performing companies' processes, models, capabilities, and performances. Lead scoring and prioritization is the path to higher conversion, ultimately increasing companies' annual revenue and sales figures while reducing the cost spent per lead [42]. Additionally, Jaskaran [29] compared the rules-based model to predictive lead scoring and explained why the rules-based approach is not popular. The conclusion was that the impact exerted by traditional rules-based models on sales was not significant. Furthermore, Lattice guide, which considers the integration of predictive lead scoring, statistically showed that predictive lead scoring models enhance sales [39]. Also, Brown [13] proposed a traditional rules/points-based model to score and segment B2B sales leads and showed the benefits of lead scoring applications in the financial services industry. Finally, Boogar [11] discussed the three stages of lead scoring and concluded that lead scoring models evolve as marketing and sales departments grow. The positive impact imposed by traditional lead scoring on sales was not as significant as those exerted by predictive lead scoring.

2.4.4 The preferred model: supervised vs. unsupervised

Despite the broad choices offered by predictive lead scoring models, the literature is short of knowledge on the decision of which to use given different data sources (i.e., situations). This paper provides a way to classify these predictive lead scoring models and insights on when to use them. The classification and regression models are considered supervised learning models, while clustering is considered an unsupervised learning model. Supervised learning models can estimate the relationship between various prospects' attributes and the identified purchase behaviors (i.e., purchased, not purchased, and hesitation on purchase) [15]. Lead scoring models built using supervised learning models can score prospects based on the propensity to achieve the targeted purchase behaviors. On the other hand, unsupervised learning models can identify similar cases without target output; the pattern recognition is undirected [15]. Unsupervised learning models aim to uncover data patterns in a set of prospects' attributes.

Data identification is key in deciding which learning model to use for building a predictive lead scoring model. Data can be broadly categorized into two types: commercial data and internet textual data [19, 41, 63]. Commercial data includes profile data (i.e., demographic information), account profile data (i.e., firmographic attributes), prospect intent data, and activity data. In building predictive lead scoring models using supervised learning, integrating diverse commercial data sources is essential to identify patterns and features associated with lead conversion. Key data from CRM systems, such as lead profiles

(e.g., company name, industry location, job title, and lead source), sales activity history (e.g., records of interactions with salespeople), and lead stage and status (e.g., current stage in the sales funnel), form the core foundation for understanding the journey of each lead through the sales process. Additionally, website behavior data (e.g., page views, time spent on site, and number of downloads and form submissions) is crucial for gauging engagement and interest levels, which are strong indicators of conversion potential. Third-party data sources can further enrich this picture by providing additional insights that may not be available internally, contributing to a more comprehensive understanding of lead behavior and purchasing intent.

However, it can be challenging to establish commercial data integrity due to the normality of missing information [19]. To compensate for incomplete commercial data, internet textual data extracted from prospects' websites can be utilized. For instance, unsupervised learning methods can leverage data from website crawlers, advertising and social media responses, behavioral interactions, demographics, and product engagement. Behavioral data captured from website activity, content interactions, and browsing patterns can help cluster leads based on their engagement levels and interests. Social media engagement data, including followers, likes, shares, and comments, can help identify clusters of highly engaged users. Furthermore, analyzing email communications can reveal which leads are most responsive to email campaigns, enabling more effective prioritization. These data sources are invaluable for identifying distinct lead segments. By clustering leads based on these features, sellers could better understand lead behavior, prioritize leads, and improve lead conversion outcomes.

Given the unstructured nature of internet textual data, applying supervised classification models directly is impractical. Because unstructured text data usually contains numerous unique words or tokens, which are treated as features, resulting in high dimensionality, many of these features (words) may have a value of zero for a given document, leading to sparse data matrices that are difficult for supervised classification models to process. Techniques such as principal component analysis or s-SNE could be applied to reduce the high dimensionality of unstructured text data and improve the performance of classification modeling [19]. Since unstructured text consists of strings, sentences, or paragraphs, it first needs to be transformed into numeric input before supervised classification can be applied. This process requires feature extraction techniques such as Bag of Words, Term Frequency-Inverse Document Frequency, or word embeddings to convert text into meaningful numeric representations that capture both semantics and contextual information. Additionally, text data usually contains many stop words (e.g., "the", "are", "and"), irrelevant information (e.g., punctuation, unrelated phrases), misspellings, slang, and abbreviations that do not contribute significantly to understanding the text's meaning or classification. They need to be removed to reduce noise in the modeling process. Furthermore, unstructured text data needs an

extensive preprocessing (e.g., tokenization, stemming, lemmatization, etc.) to minimize noise before classification. Moreover, supervised classification requires large amounts of labeled data for training, and manually labeling unstructured text data can be time-consuming, costly, and error prone. Instead, unsupervised learning models (i.e., clustering algorithms) with textual data transformation techniques shall be applied to find clusters consisting of similar textual information [15, 19]. Then, latent semantic concepts extracted from each cluster along with expert knowledge (i.e., a set of binary variables about prospects) can be used as input for supervised learning models to estimate the likelihood of lead conversion and profitability of new potential customers [19, 41, 63].

In conclusion, if commercial data is available and mostly complete, then the direct application of supervised learning models is the recommended option to build predictive lead scoring models. However, when the quality of existing commercial data is low (i.e., too much missing data), then the recommended option is first to apply unsupervised learning models on internet textual data to identify key latent semantic concepts. The next step is to utilize supervised learning models with latent semantic concepts and incomplete commercial data as input to build predictive lead scoring models.

2.4.5 Impact of lead scoring models on sales performance

After investigating all lead scoring models found during the search, we examined the impact of lead scoring models on sales performance by studying each case further. We identified several metrics for assessing the impact of lead scoring models on sales performance. The most used performance metric is the lead conversion rate (see **Table 2.7**). It is calculated as the total number of conversions divided by the total number of leads [19].

In predictive lead scoring, D'Haen and Van den Poel [17] proposed a model consisting of three iterative phases, which produced a ranked list of high-quality prospects. After testing the proposed model on a telecom service company's dataset, the model had a lead conversion rate (from prospects to qualified leads) of 15.73%, which was higher than the average conversion rate of 10% [17]. A higher conversion rate indicated a specific increase in sales. Also, the proposed sales force automation tool was designed to be implemented in a web application; users only need to pay a small membership fee to access the application instead of paying for the entire database of prospects. Thus, costs can be reduced. Three years later, D'Haen et al. [19] proposed a lead scoring system to integrate text mining on web data. The lead conversion rate (i.e., from prospect to customer) of the experiment was 6.4%, better than the previous result (i.e., 3.5%) [17] without text mining. We noticed a scarcity of studies regarding text mining in lead scoring models during our search. As text mining techniques become more mature in data science applications, more textual data sources become freely available, and because lead conversion rates can be improved, we call on future research to consider various text mining techniques when building lead scoring models.

Furthermore, the probabilistic lead scoring models suggested by Duncan and Elkan [20] increased the lead conversion rate from 8% to 17% in a three-month experimental period. The experimental results showed that the models in question have additional benefits, including the reduced average time needed to qualify leads, reduced number of calls placed to schedule a product demo, increased number of successful sales, as well as increased total revenue [20]. The prospective lead scoring models discovered high-quality leads at an early stage in the sales process because they focused on features that measure the fit of leads with the products being sold, in addition to leads' behaviors.

The customer acquisition process using a data mashup approach (i.e., high scale mobile consumer data and customers' online food ordering transactional data) improves business performance from two aspects [58]. First, the new customers are twice as likely to be high-impact potential customers than before. Second, the correctly predicted high-impact potential customers have 21.41% higher average revenue per user than the overall value. This reveals the importance of targeting the right group of consumers for acquisition over randomly picking them.

In traditional lead scoring, Lindahl's study [37] showed that lead scoring could contribute to a more efficient marketing process by saving time for the sales and marketing departments, improving the lead conversion rate, reducing cost per lead, and enabling the automatization and personalization in digital markets. Monat [44] demonstrated that using a lead characterization model could significantly increase sales effectiveness and the accuracy of sales projections by increasing lead conversion rates and the total number of leads converted into sales. The research conducted by the Aberdeen Group revealed that companies that successfully implement effective lead scoring models deliver excellent performance in lead conversion rates (i.e., 26% average increase), annual revenue (i.e., 50% average increase), and cost per lead (i.e., 25% average decrease) [42]. These three metrics reflect the improvement in sales performance after deploying suitable lead scoring models. The improvements in sales performance indicate that lead scoring models can improve the effectiveness of lead management and sales and marketing efficiency. It is essential for companies to incorporate both implicit (i.e., behavioral information of leads) and explicit (i.e., manifest information of leads) attributes in the lead scoring model in order to influence sales performance significantly [42, 43]. Thus, we propose that:

Proposition 1 *The use of lead scoring models improves the lead conversion rate.*

Cost reductions/monetary savings stand for money that can be saved per lead conversion. Kazemi et al. [31] proposed a predictive lead scoring model considering effective identification factors to increase sales and customer satisfaction by using a decision tree and a basket purchase technique to analyze which potential customers were the "real" ones. The suggested model was tested on data from a furniture producer. In the post-performance evaluation, administrative costs were reduced by 8%, and the profit was increased

by 15% [31]. These results indicated that the proposed model positively affects sales in terms of cost reduction and profit growth. Additionally, Peng and Xu [53] integrated the rules generated from correlation analysis and decision trees to identify potential and sustainable customers for mobile communication companies. As results showed, marketing effectiveness has increased to 16.1%, compared to 2.1% without the model. Also, cost savings have been reached. Moreover, Meire et al. [41] showed the economic value of the advised customer acquisition decision support system by integrating social media data into the monetary savings and financial gains. They tested the system during a real-life field study at Coca-Cola Refreshments USA. The results showed that, on average, an increased lead response percentage of 4.75% can be achieved, which equaled 2376 extra leads that were likely to convert into customers without extra cost or an additional financial gain of 11 million dollars [41]. The above-mentioned studies showed that marketing and sales effectiveness could be enhanced if appropriate predictive lead scoring models were applied. Therefore, sales performance could be improved. For instance, Meire et al. [41] showed that using social media data in lead scoring within a B2B sales context can improve sales effectiveness. However, we did not find any literature in a Business-to-Customer (B2C) context in this regard. Thus, we call on further research to consider the application of predictive lead scoring models in a B2C context by using social media data and studying its impact on marketing and sales effectiveness.

Proposition 2 *The use of lead scoring models reduces costs spent on converting leads.*

Another metric to estimate sales performance is the customer value matrix. It was employed to analyze customer value by classifying potential customers and developing different marketing strategies accordingly [69]. This approach can enhance efficiency and effectiveness when prioritizing potential customers, and positively influencing sales. Furthermore, activity statistics (e.g., website visits, email click-throughs, form submits, etc.) were measured before and after implementing lead scoring models to measure leads' purchase likelihood [48]. Outputs such as median activity amount per purchase probability group enabled sales to further understand trends and insights on customer groups. Moreover, a hybrid customer prediction system that combined forward stepwise and multiple logistic regression was proposed by Soroush et al. [60] to identify the top 20% of potential customers. This system was tested on an insurance company's dataset. The results showed that the system selected 50% of the original features, reducing the computational cost and complexity [60]. Also, the number of insurance purchasers, the percentage of total purchases and predicted customers all increased, which proved that using the alleged system can improve prediction. They also demonstrated that as the prediction accuracy of the lead scoring models increases, sales performance improves. In other words, model prediction accuracy directly influences sales performance positively. This finding highlights the importance of checking the prediction accuracy of lead scoring models during the evaluation before model deployment.

In Luk et al.'s study [38], surveys were conducted to com- pare customer satisfaction before and after implementing an intelligent customer identification model (ICIM) for a company in the e-commerce logistics industry. Specifically, after adopting the ICIM, which integrated k-means clustering and decision tree classification, a 36.4% increase in overall satisfaction, a 50% increase in the number of customers who are willing to establish a close relationship, a 60% increase in the expected order frequency, and a 300% increase in the expected order spending amount have been observed [38]. With the potential customer classification rules produced by ICIM, the company can classify and prioritize potential customers with minimum time and resources. Because the model consisted of a historical view and analysis of all the existing customers, it can help companies prioritize leads based on the most valued attributes, maximizing profits and increasing sales. Thus, we propose that:

Proposition 3 *The use of lead scoring models increases profits and revenue.*

Thorleuchter et al. [63] proposed a content-based lead scoring model to support a mail-order company's customer acquisition process. They compared the success rate of the traditional customer acquisition process and the suggested strategy to measure the improvement in sales performance. The results showed refinements in both profitable customer acquisition success rate and sales while reducing the cost of paying for brokers' provided list of potential customers. Also, the density of profitable customers (18%) in the prioritized list of potential customers generated by this approach out- performed the density in brokers' lists (5%) [63]. For a customer acquisition process in a queue-based LMS, marketing and sales teams collaborate to increase the density of profit- able customers as one of the goals in lead scoring. Because an acquisition process is both time and cost-consuming and budgets are usually limited, identifying profitable customers in the top 20%-30% of a list is essential when assessing the excellence of a lead scoring model [60, 63]. Being able to identify more profitable customers in a more significant portion of a list can increase sales performance. However, as an inevitable result, more company and human resources will be spent, increasing costs. Apparently, there is a trade-off. Maintaining the balance between the two while steadily enhancing sales performance is an intriguing future research topic.

In addition to the metrics mentioned above, Kim et al. [32] used the hit rate on the actual number of customers who purchased recreational vehicle insurance to scale the impact on sales performance. The results showed that the advised Evolutionary Local Selection Algorithm & Artificial Neural Networks (ELSA/ANN) model has the highest hit rate in training and testing datasets. Additionally, the small number of features provided by the ELSA/ANN model implied that companies could reduce data collection and storage costs considerably. In a case study, Bohanec et al. [9] applied post-results analysis using various visualization tools on their proposed predictive lead scoring model. The hit rate on the "won" deals was around 45% [9]. However, a year later, they integrated a more advanced explanation method (i.e., IME and

EXPLAIN) with the same lead scoring model to better interpret and understand the results. After adjusting according to the what-if analysis and discovering more customers' insights, they increased the hit rate up to 60% [10]. Thus, we argue for a better way to improve sales performance by applying explanation methods to analyze lead scoring results and modify attributes' weights in the model accordingly. Thus, we propose that:

Proposition 4 *The use of lead scoring models increases the number of high-quality leads.*

The four abovementioned propositions summarize the major sales performance metrics that are directly and positively influenced by the application of lead scoring models, and answer RQ3 (i.e., *how do lead scoring models influence sales performance?*) Using lead scoring models when scoring, prioritizing, and managing leads is expected to enhance sales performance from various dimensionalities while reducing cost.

As another interesting sales performance measure in a real business case, the Gain curve/score [15] examined the distribution of won cases on the ranked prediction output list. Yan et al. [72] applied the Gain curve/score as the sales performance metric to compare the performance within a period. The results indicated that using data-driven predictive models is a promising way to drive better sales performance.

After analyzing sales performance in the selected studies, we can conclude that predictive lead scoring models impact sales performance positively in various ways. However, the impact posed by traditional lead scoring on sales performance may not be as significantly positive as predictive lead scoring. For example, the typical conversion rate from leads to customers is only 5% on average in traditional lead scoring systems, whereas the average conversion rate is 15% in predictive lead scoring systems [20]. Additionally, some results generated by traditional lead scoring can be inaccurate, which exerts a minimum positive impact on sales performance [39, 42]. For instance, in traditional lead scoring systems, salespeople spend too much time dealing with a large volume of low-quality MQLs that will not become sales qualified leads (SQLs) [20, 39]. Instead of hiring more salespeople, which is expensive, prediction lead scoring models can produce a much-refined list of MQLs for sales to contact so that efforts can be focused on leads that are most likely to convert. Predictive lead scoring models are especially beneficial to small-medium-sized businesses (SMB). By concentrating the limited inside sales resources on leads with the highest conversion probability in SMBs, marketing could forward fewer MQLs to sales but yield higher lead conversion rates.

Another factor is the limited capability of processing a large amount of data [39]. A successful prediction of lead conversion relies on vast amounts of data for analysis. Traditional lead scoring does not have the ability to analyze vast amounts of lead data due to the lack of computational power. However,

using predictive lead scoring is the right solution for forecasting the likelihood of leads converting to customers, given its high computational and analytical ability.

Meanwhile, traditional lead scoring models may fail to capture nonlinear effects between independent and dependent variables or complicated interactions between features [20]. These disadvantages mean spending resources on converting low-quality leads who are unlikely to convert at the end, which may degrade sales performance. However, predictive lead scoring can find various patterns and relationships between variables as well as identify trends and the most determinant features in the leads' data [15].

Finally, traditional lead scorecards are heavily reliant on behavioral data while negligent on demographic data, which may prevent the early discovery of high-quality leads [20]. A reliable predictive result of the likelihood that leads convert into customers should consider both the demographic and behavioral data of leads at different stages of the conversion process, not to mention that data on existing customers, old leads, and new leads should also be analyzed in calculating a lead score. In conclusion, predictive lead scoring is better than traditional lead scoring as it exerts more of a positive impact on sales performance.

2.4.6 Limitations

This study has limitations. The first one is the restricted year range of publications, as we only considered studies published between 2005 and 2022. In addition, the studies were extracted based on concepts and search keywords, as shown in **Table 2.1**. Hence, publications investigating lead scoring models without a keyword index could have been missed during the search phase. Indeed, there might be a threat to the completeness and adequacy of the selected studies. As a second limitation, the search for papers was limited to six online databases. However, there might be more articles related to lead scoring models in other academic journals or grey literature databases. Finally, this review only included studies that were published in English. We believe that studies regarding lead scoring models might have been discussed and published in other languages as well.

2.5 Discussion and conclusions

Lead scoring is critical to a successful inside sales process, as it helps sales teams to prioritize their efforts and identify which prospects are most likely to convert [7, 20, 48]. However, implementing an effective lead scoring model into LMS can be challenging.

The first issue is that lead scoring models can be too time-consuming [20, 39]. With so many leads to review and prioritize, it can be difficult for sales representatives to properly assess and score each lead in a timely manner. To address this issue, organizations should look for ways to automate the lead scoring process and consider artificial intelligence (AI)-based lead scoring models. By leveraging AI-based

technologies, organizations can reduce the time it takes to score leads and ensure that each lead receives the attention it deserves [11, 29, 39]. Our study reviews all existing lead scoring models, summarizes benefits, and recommends what AI-driven models can be implemented and when.

The second issue with lead scoring models is that they can be ineffective since they are too simplistic [17, 20, 29]. Many organizations rely on basic metrics such as firmographics and demographics to prioritize leads [29]. While these metrics can be essential in providing a broad overview of a lead's potential value, they can be limited and often fail to capture the nuances of an individual lead's situation and preferences [20]. As a result, leads may be incorrectly scored and not given the attention they deserve. Additionally, existing studies on predictive lead scoring only focus on conversion steps in the marketing-sales funnel from the perspective of selling organizations [7, 17, 20, 48], hence neglecting insights in the purchase decision-making process from the buyer's perspective. To address this problem, our study recommends employing more sophisticated lead-scoring models that consider a wider range of factors including both the seller's and buyer's perspectives that will help understand what phase of purchasing journey buyers are in and what their preferences are.

The third issue with lead scoring models is companies can be too resistant to change. Many organizations rely on predetermined criteria to assign scores to leads, which can lead to leads being incorrectly scored or overlooked. To address this issue, we recommend creating scoring models that are dynamic/flexible and allow adjusting models on the fly taking into account new coming data and sales representatives' inputs.

The last issue with lead scoring models is that they are mostly built based on low-quality data or insufficient, imbalanced datasets [22, 32, 74]. Organizations often build their scoring models based on historical data from previous sales cycles, which may not be reflective of the current market conditions. As a result, leads may be incorrectly scored or overlooked altogether. Our study recommends using industry and company-specific data up-to-date sources and provides recommendations when it is to deploying specific models considering data specifications.

There has been a growing body of literature on lead scoring models published in the past few years (see **Figure 2.3**). There are many reasons for this phenomenon, such as the advancement of computational capabilities, the introduction of various LMSs that implement lead scoring models, and the availability of large sales datasets. In addition, trends in applications and development of DM, ML, and AI-based approaches for business in general and sales in particular, are other factors. Also, the COVID-19 pandemic accelerated the adoption of remote selling through LMSs. As a result, the importance and urgency of implementing suitable lead scoring models skyrocketed.

More studies appear to focus on building data-driven predictive lead scoring models to predict the probability of lead conversion to prioritize leads for further steps in the sales process. Indeed, predictive lead scoring models have attracted the attention of academics and practitioners. An array of DM and ML techniques have been used to discover trends, find insights, detect relationships between variables within data, and predict lead conversion outcomes. Eventually, these help decision-makers optimize business processes and enhance sales performance. This study conducted a comprehensive review of lead scoring models and their impact on sales performance. During our review of existing lead scoring models, we found few studies that examine the performance of a small number of supervised learning algorithms on estimating the purchase probabilities of leads [6, 21, 48]. In addition to the small number of supervised learning algorithms examined, these studies did not address the impact of lead scoring models on sales performance.

This review paper has identified 44 published studies between 2005 and 2022 relevant to lead scoring models in LMSs. Our goal is to provide a research summary on the lead scoring models, their impact on sales performance, and their applications in the CRM domain. The qualified studies in this review test different lead scoring models on various experimental and real datasets (i.e., company-provided datasets). The qualified studies use multiple metrics to measure the impact of the proposed models on sales performance. The results show that lead scoring models positively impact companies' sales performance in various ways. Notably, predictive lead scoring is more effective and efficient than traditional lead scoring in many aspects, which results in a more positive impact on sales performance.

The results of this study carry the following significant implications:

- Based on increasing interest and past publication ratio in the area of lead scoring models, research in this area will increase significantly in the future, particularly, in the area of predictive lead scoring.
- Most of the reviewed papers are in the predictive lead scoring domain (i.e., 85%, 39 articles). This number indicates the rising importance of predictive lead scoring models as tools in LMSs. In addition, these studies provide insights to decision-makers on the common DM and ML practices adopted in the process of customer acquisition.
- Among all the DM models, classification is the most used model for predicting a lead's propensity to make a purchase.
- Among the 44 studies, it is surprising that neural network is not the most popular algorithm (i.e., seven studies). Neural networks can be applied to classification, regression, and clustering tasks given their flexibility and capability of studying complicated relationships [17]. Maybe this is due to the hardship of interpreting results generated by neural networks since they are "black box"

algorithms. However, the EXPLAIN and IME explanation methods can be applied to interpret complicated modeling processes in "black box" algorithms and their generated results [10]. Thus, more research could be conducted on lead scoring models by using neural networks coupled with explanation methods.

- Decision tree and logistic regression are tied as the most applied algorithms in lead scoring studies. The modeling processes of both algorithms are easier to understand than any "black box" algorithms such as neural networks. In addition, the results of these two techniques can be interpreted easily. Hence, these two algorithms might be preferable to non-DM experts in the business field.
- The most used metric to measure the impact of lead scoring models on sales performance is lead conversion rate. A growth in lead conversion rates shows that the application of lead scoring models indeed converts more leads into customers [56]. Furthermore, companies examined in this review (e.g., GE Capital and DocuSign Inc.) [1, 39] indicated that the productivity and effectiveness of sales and marketing functions have improved after the deployment of lead scoring models. This underlines the role of lead scoring models in improving the internal collaboration between the marketing and sales functions [56].
- There is not enough research discussing traditional lead scoring models. This implies that as DM models and ML algorithms become increasingly powerful in analyzing lead and customer data, the traditional way of scoring and prioritizing will probably be in decline.

From a theoretical perspective, the results of this review point to a few potential future work items:

1. Future research on lead scoring should be extended to study the methodology in each lead scoring model. An interesting research direction is to review methodologies in the lead scoring studies and their practical, real-life applications. Specifically, researchers should focus on analysis techniques in the lead scoring studies, research approaches and frameworks, and data collection steps. Discussion of the pros and cons of each methodology should be highlighted.
2. Additionally, various IT drivers can impact sales performance, such as automated lead generation, lead distribution, lead nurturing tools, and lead scoring models. Another possible research direction is exploring these IT drivers' impact on sales performance and their intra-relationships. To be specific, relevant survey questions could be sent out to participants. For data (i.e., survey question responses) analysis, statistical tools should be used to calculate and analyze correlation, T-Test, and regression.
3. Many factors influence the result of lead scoring, for example, customer characteristics, salesperson traits, organization types, and environmental attributes. A meta-analysis review could be carried

out to examine the relationship between these factors and lead scoring outcomes. This could help identify the key determinants of lead scoring success in a B2B inside sales context. Another future research direction is using meta-analysis to validate the formulated propositions with regards to lead scoring models' impact on sales performance proposed in this study.

4. As marketing has become increasingly customer-centric, another possible research avenue is to integrate insights from the buyers' perspective into building lead scoring models. This will likely bring more insights into the implementation of sophisticated lead scoring models.

From a practical perspective, the results of this review suggest the following future work items:

1. Deep learning methods have appeared and proved to be powerful enough when handling large data sets with a high degree of complexity. With the availability of larger data sets across industries, the application of deep learning models to improve sales performance seems feasible. Furthermore, with the help of explanation methods, the interpretation difficulty of predictions generated by "black box" models, such as deep learning, should be reduced.
2. Moreover, another interesting research question could be measuring and comparing various lead scoring models under the same environment/setting to find the most effective DM model and ML algorithm. The evaluation and comparison could be executed in two phases, pre-campaign, and post-campaign. In the pre-campaign phase, researchers should focus on comparing confusion matrices and evaluation graphs (i.e., Gains, Response, Lift charts, ROC curves) to estimate models' performances. After deploying models on marketing campaign datasets, the actual performances are evaluated and compared by using measures such as lead conversion rate and revenue gained in the post-campaign phase.
3. Finally, as lead scoring models have been widely adopted across industries, future research should investigate the application of various lead scoring models to different types/sizes of companies, and how to adjust these models to adapt to different industries. Investigating how to improve the suitability and practicability of lead scoring models in an industry-wide context is another research direction.

Acknowledgments This study was funded by grants from Mitacs Canada (No. 310636).

Table 2.1. Concepts and keywords

Concepts	Keywords
Lead Scoring	Lead scoring; lead ranking; lead prioritization; prospect scoring; prospect ranking; prospect prioritization; customer acquisition; customer identification
Modeling	Modeling; approach; concept; algorithm; data mining; machine learning; prediction
Sales	Sales; business; commerce; marketing; selling; b2b; b2c

Performance	Performance; outcome; output; result; lead conversion; conversion rate

Table 2.2. Number of retrieved articles from each database

Database names	Number of articles
Scopus	450
IEEE Xplore	234
ABI/INFORM Global	168
Web of Science	150
Business Source Complete	105
OpenGrey	43

Table 2.3. Distribution of studies according to the proposed classification framework

lead scoring approach	Lead scoring model	Algorithm	References
Predictive	Classification model	Decision tree (13)	[1, 6, 17, 18, 21–23, 31, 38, 45, 48, 53, 73]
		Logistic regression (13)	[6, 17–19, 22, 23, 45, 48, 57, 60, 63, 68, 72]
		Random forest (9)	[6, 9, 10, 21, 23, 35, 41, 45, 48]
		Gradient boosted trees (8)	[6, 20–23, 30, 45, 58]
		Neural network (7)	[17, 22, 23, 32, 48, 71, 74]
		Bayesian network (4)	[6, 7, 64, 73]
		K-nearest-neighbor (2)	[6, 17]
		Fuzzy logic (1)	[33]
		Support vector machine (1)	[21]
	Clustering model	K-means (3)	[35, 38, 69]
		Self-organizing maps (1)	[69]
		Expectation maximization (1)	[63]
		Spherical (1)	[19]
		Fuzzy (1)	[19]
	Regression model	Linear/exponential (2)	[35, 71]
		Correlation analysis (1)	[53]
		Seasonal ARIMA (1)	[71]
		Weighted average (1)	[40]
Traditional	Rules/points-based (7)	N/A	[11, 13, 24, 26, 29, 39, 44]
	Scorecard (3)	N/A	[37, 44, 67]
	Lamb/Spam (1)	N/A	[11]

Table 2.4. Types and approaches of the qualified lead scoring studies

Approach / Type of study	Predictive lead scoring	Traditional lead scoring
Peer-reviewed papers	34	5
Grey literature	5	5

Table 2.5. Predictive lead scoring models

Algorithms	Count	References
Decision tree classification	13	[1, 6, 17, 18, 21–23, 31, 38, 45, 48, 53, 73]
Logistic regression	13	[6, 17–19, 22, 23, 45, 48, 57, 60, 63, 68, 72]
Random forest classification	9	[6, 9, 10, 21, 23, 35, 41, 45, 48]
Gradient boosted trees classification	8	[6, 20–23, 30, 45, 58]
Neural network classification	7	[17, 22, 23, 32, 48, 71, 74]
Bayesian network classification	4	[6, 7, 64, 73]
K-means clustering	3	[35, 38, 69]
K-nearest-neighbor classification	2	[6, 17]
Linear/exponential regression	2	[35, 71]
Fuzzy logic classification	1	[33]
Support vector machine classification	1	[21]
Self-organizing maps clustering	1	[69]
Expectation maximization clustering	1	[63]
Spherical clustering	1	[19]
Fuzzy clustering	1	[19]
Correlation analysis	1	[53]
Seasonal ARIMA time series model	1	[71]
Weighted average formulas/functions	1	[40]

Table 2.6. Traditional lead scoring models

Traditional lead scoring models	Count	References
Rules/points-based	7	[11, 13, 24, 26, 29, 39, 44]
Scorecard	3	[37, 44, 67]
Lamb or Spam	1	[11]

Table 2.7. Metrics used to measure sales performance in all qualified studies

Sales performance metrics	Count	References
Lead conversion rate	12	[17–20, 22, 35, 37, 39, 40, 42, 44, 71]
Cost reduction/monetary savings	10	[11, 13, 22, 31, 35, 37, 42, 53, 57, 64, 68]
Number of qualified leads	9	[6, 23, 26, 29, 30, 41, 58, 60, 74]
Hit rate on number of customers who purchase	8	[9, 10, 21, 32, 44, 45, 68, 73]
Annual revenue	7	[1, 13, 20, 38, 42, 58, 71]
Profit/financial gains	7	[18, 24, 24, 31, 33, 38, 41]
Density of profitable customers in the list	2	[38, 63]
Response percentage	2	[41, 68]
Customer value matrix	1	[69]
Customer overall satisfaction	1	[38]
Average time needed to qualify a lead	1	[20]
Activity level (e.g., website visits, log-ins)	1	[48]
Equilibrium percentage	1	[35]
Gain curve/score	1	[72]

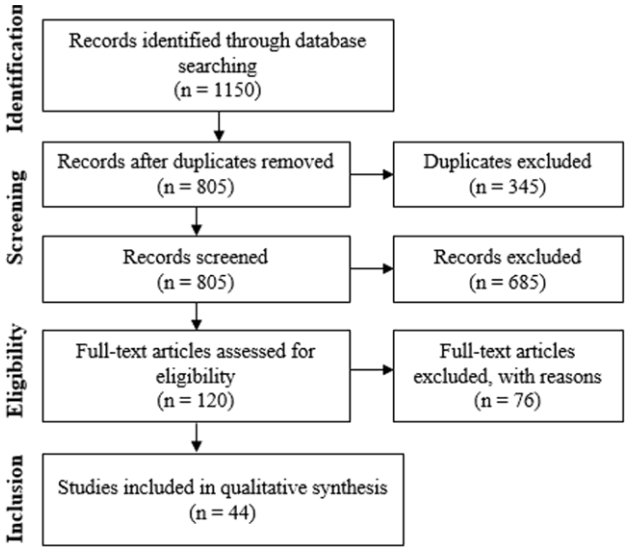

Figure 2.1. PRISMA

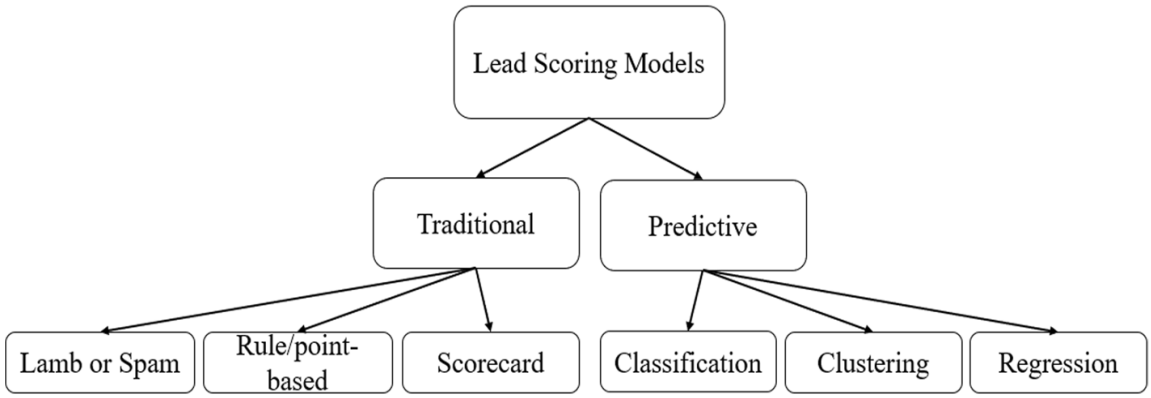

Figure 2.2. Classification framework



Figure 2.3. Trend of studies on lead scoring models

Appendix: table of extracted literature data

ID	Title	Authors	Methods/Models/Algorithms	Approach	Evaluation metrics	Sales performance metrics	Impact on sales
p1	Customer Targeting: A Neural Network Approach Guided by Genetic Algorithms	Kim et al. (2005)	Neural network & evolutionary local selection algorithm	Predictive	Hit rate (predictive accuracy), complexity (number of features) & lift curves	The hit rate on the actual number of customers who purchase recreational vehicle insurance	With the information on campaign costs & profit per additional actual customer, a marketer can use the proposed model to choose the best selection point where expected total revenue is maximized
p2	Sizing up prospects	Hornstein (2005)	Rules/points-based	Traditional	Accuracy of total lead score value	Number of qualified leads	A simple linear model combines points assigned to rate leads from four factors. Based on the net score, specific follow-up activities are offered
p3	What is a qualified sales lead	Grandy (2005)	Rules/points-based	Traditional	Number of qualified leads/calls	Profit/financial gains	The author examines leads generated in the equipment replacement industry and argues key determinants of lead quality, which affect profits in sales
p4	Grey Literature: Lead Prioritization and Scoring: The Path to Higher Conversion	Michiels (2008)	Scorecard & automated lead scoring model	Traditional & Predictive	N/A	Annual revenue, lead conversion rate & cost per lead	This study proves that lead scoring and prioritization are the paths to higher conversion, which ultimately increase companies'annual revenue and sales while reducing cost per lead
p5	Grey Literature: How real-time online sales lead scoring drives a competitive edge	Brown (2009)	Rules/points-based & automated lead scoring model	Traditional & Predictive	N/A	Cost reduction & annual revenue	The method scores and segments B2B sales leads. The author details the revolution of online sales lead scoring system and extols benefits of lead scoring in financial services industry
p6	Modeling of Potential Customers Identification Based on Correlation Analysis and Decision Tree	Peng & Xu (2011)	Correlation analysis & decision tree	Predictive	Support index, confidence index, accuracy & lift curve	Economic benefits gain (i.e., marketing effectiveness, cost savings)	The proposed model identifies potentially sustainable leads and facilitates the company achieving significant economic benefits while saving costs
p7	Industrial Sales Lead Conversion Modeling	Monat (2011)	Rules/points-based & scorecard	Traditional	Accuracy	Lead conversion rate & sales outcomes (i.e., won or lost status on leads)	A theory-based sales lead conversion model is proposed to predict the probability that a sales lead will convert to a customer based on observable traits. The model increases selling effectiveness as well as the accuracy of sales projections
p8	Towards an Optimal Classification Model against Imbalanced Data for Customer Relationship Management	Tu et al. (2011)	Decision tree, Bayesian network, support vector machine & neural network	Predictive	Accuracy, sensitivity, specificity & AUC	Cost reduction	The proposed classification framework by combining CFS method for processing data set with bagging ensemble cost-based Bayesian network classifier is applied to identify customers that are most valuable to the business. The high efficiency and accuracy of the model indicates cost reduction can be achieved
p9	A hybrid customer prediction system based on multiple forward stepwise logistic regression model	Soroush et al. (2012)	Multiple logistic regression & forward stepwise regression	Predictive	Root mean square error, mean square error & accuracy	The number of purchasers, percentage of total purchasers & percentage of predicted customers	The proposed hybrid system can reduce model computation complexity and improve prediction accuracy, which in turn significantly brings positive impact to sales

ID	Title	Authors	Methods/Models/Algorithms	Approach	Evaluation metrics	Sales performance metrics	Impact on sales
p10	Analyzing existing customers' websites to improve the customer acquisition process as well as the profitability prediction in B-to-B marketing	Thorleuchter et al. (2012)	Latent semantic indexing, expectation-maximization clustering & logistic regression	Predictive	Lift curve, precision, recall, AUC, sensitivity & specificity	Profitable customer acquisition success rate & density of the number of profitable customers in the list	The proposed model identifies more profitable customers at lower costs and distinguishes profitable from non-profitable customers, improving the acquisition process and sales of companies
p11	Model-supported business-to-business prospect prediction based on an iterative customer acquisition framework	D'Haen & Van den Poel (2013)	K-nearest neighbours, decision tree, logistic regression & neural networks	Predictive	AUC & efficiency of the iterative process in the proposed model	Conversion rate	The proposed model outputs a ranked list of prospects with high quality, making it easier for sales to qualify and convert them into customers. Also, costs can be reduced by embedding this tool into a web-based application
p12	Financing Lead Triggers: Empowering Sales Reps Through Knowledge Discovery and Fusion	Aggour & Hoogs (2013)	Decision tree	Predictive	Accuracy	Total revenue, net income & long-term debt	The system identifies actionable sales leads who are most likely in need of financing support. It increases the productivity of salespeople and effectiveness of sales force
p13	Predicting customer profitability during acquisition: Finding the optimal combination of data source and data mining technique	D'Haen et al. (2013)	Bagged decision trees, logistic regression & latent semantic indexing	Predictive	ROC curves & AUC	Conversion rate & profit/financial gains	The authors show which data (web, commercial, or combined) is the optimal source of predicting customer conversion and profitability during acquisition. The bagged decision trees on the combined data sources estimate the most accurate conversion rate and profitability as the results of a case study conducted in a German B2B mail order company
p14	Grey Literature: The Evolution from Traditional to Predictive Lead Scoring: A how-to guide for considering predictive scoring	McDonnell (2014)	Traditional rules/points-based & predictive data pattern-based	Traditional & Predictive	N/A	Number of additional buying signals, number of sales-qualified leads & conversion rate	This grey literature includes evidence to show the power of predictive lead scoring models with statistics proving the evolutionary improvement that predictive lead scoring brings to sales
p15	A data mining approach for turning potential customers into real ones in basket purchase analysis	Kazemi et al. (2015)	Decision tree & basket purchase analysis technique	Predictive	Prediction accuracy & most optimal variable	Cost reduction & profit/financial gains	The proposed approach combines decision tree and basket purchase techniques to identify and classify factors for turning leads into "real" customers. The approach positively influences sales by reducing cost and increasing profit gains
p16	Probabilistic Modeling of a Sales Funnel to Prioritize Leads	Duncan & Elkan (2015)	Gradient boosted trees & probabilistic classifier	Predictive	AUC & lift curve	Conversion rate, the average time needed to qualify a lead, number of successful sales & total revenue	The proposed models find high-quality leads earlier in the sales process, which can be applied in sales funnel for ranking initial leads faster and predicting successful sales. The proposed models bring positive impact to sales in multiple aspects as proved by using company-provided data
p17	Sales pipeline win propensity prediction: a regression approach	Yan et al. (2015)	Logistic regression	Predictive	AUC score & ROC curve	Gain curve and score & sales outcomes (i.e., won or lost status on leads)	A ML-based framework that estimates the win propensity of sales leads overtime. The model captures impact of sales activities and lead profiles on the outcome

ID	Title	Authors	Methods/Models/Algorithms	Approach	Evaluation metrics	Sales performance metrics	Impact on sales
p18	Applying Data Mining and RFM Model to Analyze Customers' Values of A Veterinary Hospital	Wei et al. (2016)	Self-organizing maps clustering, K-means clustering & RFM model	Predictive	RFM statistics (i.e., max, min, average & standard deviation)	RFM model descriptive statistics on clusters & customer value matrix	The cluster analysis identifies traits of loyal customers that can be utilized to determine potential customers with higher profit, which results in higher sales
p19	Identifying profitable clientele using the analytical hierarchy process	Verma et al. (2016)	Analytical hierarchy process-based quantitative framework & scorecard	Traditional	Quantitative Saaty scale to list the pair-wise comparison of weightage of each factor	The total score of each manufacturer	This AHP-based framework presents a scorecard model which helps companies rank and prioritize prospective partners based on the importance of factors, achieving profitable and sustainable growth
p20	Integrating expert knowledge and multilingual web crawling data in a lead qualification system	D'Haen et al. (2016)	Spherical clustering, fuzzy clustering & logistic regression	Predictive	AUC & F1 measure	Lead conversion rate	This lead qualification system integrates expert knowledge and web crawling data, improving the lead conversion rate as a result of the lead qualification process, which indeed brings positive impact on sales
p21	Integration of Machine Learning Insights into Organizational Learning: A Case of B2B Sales Forecasting	Bohanec et al. (2016)	Random forest & prediction explanation method EXPLAIN	Predictive	Confusion matrix, accuracy & AUC	Hit rate & sales outcomes (i.e., won or lost status on leads)	An explanation methodology is proposed for ML models in predicting B2B sales. The model enhances trust in forecasts and bridges the gap between users and technology
p22	An optimization approach to services sales forecasting in a multi-staged sales pipeline	Megahed et al. (2016)	Weighted average formulas/functions	Predictive	Relative error & absolute error	Lead conversion rate & win ratio	The authors develop a multi-stage sales pipeline maturity model for forecasting sales opportunity outcomes. The model helps increase the conversion rates in all stages of the sales pipeline
p23	A new transferred feature selection algorithm for customer identification	Zhu et al. (2017)	Neural networks & transfer learning on feature selection	Predictive	Accuracy, AUC, sensitivity & specificity	The actual number of customers identified in each class	The proposed algorithm selects key features that can be used to differentiate and rank various customers in the lead scoring phase, which in turn can facilitate sales growth
p24	A qualitative examination of lead scoring in B2B marketing automation, with a recommendation for its practice	Lindahl (2017)	Scorecard	Traditional	Total lead score value & interviews	Lead conversion rate & cost per lead	This qualitative study shows lead scoring maximizes the efficiency and effectiveness of the marketing process, which in turn brings positive impact to sales
p25	Customer Relationship Management and Small Data - Application of Bayesian Network Elicitation techniques for building a Lead scoring model	Benhaddou & Leray (2017)	Bayesian networks	Predictive	Precision, recall & accuracy	Calculate lead scores on a small dataset and request experts to evaluate them	In a case study of a small size company with a small amount of available customer data, Bayesian networks built based on expert knowledge to model a lead scoring process influences sales positively
p26	The added value of social media data in B2B customer acquisition systems: A real-life experiment	Meire et al. (2017)	Random forest	Predictive	AUC & lift curve	Monetary savings, financial gains, response percentage & the number of customers	This system proves that integrating social media data in the customer acquisition process can improve predictive performance, which, in turn, can increase sales while making the sales process more efficient
p27	A Binary Classification Approach to Lead Identification and Qualification	Gokhale & Joshi (2017)	Two-class logistic regression, gradient boosted trees, neural network & random forest	Predictive	Accuracy, precision, recall, F1 score & root mean square error	Number of qualified leads	The proposed framework solves the lead identification and qualification problems for small-medium-sized businesses, and accurately predicts number of qualified leads

ID	Title	Authors	Methods/Models/Algorithms	Approach	Evaluation metrics	Sales performance metrics	Impact on sales
p28	Explaining machine learning models in sales predictions	Bohanec et al. (2017)	Random forest, EXPLAIN & IME explanation methods	Predictive	Classification accuracy & AUC	Hit rate & sales outcomes (i.e., won or lost status on leads)	The authors propose a learning and explanation model to interpret the findings of Black box ML models in B2B sales forecasting. The transparency imposed by explanation models help domain experts better understand sales outcomes
p29	Hitting your number or not? A robust & intelligent sales forecast system	Xu et al. (2017)	Seasonal ARIMA, linear/exponential regression & neural network	Predictive	Accuracy, volatility & computational time	Lead conversion rate, win ratio & estimated revenue	The proposed data-driven system forecasts lead conversion rate and estimates revenue which can be applicable to different sales designs, products or service and teams
p30	Grey Literature: Predictive Lead Scoring: Why, How & Where	Jaskaran (2018)	Traditional rules/points-based	Traditional & Predictive	N/A	Number of qualified leads	Several reasons for rules/points-based lead scoring have made the concept not as popular as the predictive approach. Therefore, the impact on sales is not significant
p31	Guidelines for assessing the value of a predictive algorithm: a case study	Ylijoki (2018)	Naïve Bayes & decision tree	Predictive	Confusion matrix & accuracy	Hit rate on number of customers who purchase	The author presents a ML-based model to estimate the results of lead qualification in a case study of a large IT service provider. In addition, guidelines are provided for assessing business values of the predictive models and algorithms
p32	Design of an Intelligent Customer Identification Model in e-Commerce Logistics Industry	Luk et al. (2019)	K-means clustering & decision tree classification	Predictive	Kappa, absolute error, RMSE, confusion matrix & ROC	Annual revenue, profit/final gains, density of profitable customers in the list & customer satisfaction	This model helps companies classify and prioritize potential customers of e-commerce logistics with minimum time and human resources, which in turn can maximize the profit
p33	Grey Literature: The Three Stages of Lead Scoring: Lambs, Ducks & Kudus	Boogar (2019)	Traditional Lamb or Spam lead scoring, rules/points-based lead scoring & predictive lead scoring	Traditional & Predictive	N/A	Cost reduction	This grey literature discusses three stages of lead scoring and concludes lead scoring models evolving as marketing and sales organization grow. Impact on sales brought by traditional models is not as significant as predictive models
p34	Predicting and Defining B2B Sales Success with Machine Learning	Mortensen et al. (2019)	Random forest, multiple logistic regression, decision tree & XGBoost (gradient boosted decision trees)	Predictive	Accuracy, precision, recall, MSE & computational time	Hit rate & sales outcomes (i.e., won or lost status on leads)	The proposed model predicts win propensities for individual sales leads and opportunities. The improvement in forecasting accuracy increases overall B2B sales success
p35	Business analytics for sales pipeline management in the software industry: a machine learning perspective	Eitle & Buxmann (2019)	Gradient boosted trees, random forest, decision tree & support vector machine	Predictive	Confusion matrix, accuracy, sensitivity, specificity, precision, F1 score & AUC	Hit rate & sales outcomes (i.e., won or lost status on leads)	The authors propose a sales pipeline end-to-end process model for assisting salespeople to manage leads in the phased sales pipeline. Also, an explanation model is adapted to facilitate the understanding of the impact of different features on sales outcomes
p36	High Value Customer Acquisition & Retention Modelling - A Scalable Data Mashup Approach	Sangaralingam et al. (2019)	Gradient boosted trees & RFM model	Predictive	Precision, recall & accuracy	Number of qualified leads & annual revenue	The proposed model with a scalable data mashup approach is able to identify more profitable leads. In a case study of online food ordering, an increase in average revenue per user indicates positive impact exerted by the model on sales performance

ID	Title	Authors	Methods/Models/Algorithms	Approach	Evaluation metrics	Sales performance metrics	Impact on sales
p37	Automating Lead Scoring with Machine Learning: An Experimental Study	Nygård & Mezei (2020)	Logistic regression, decision tree, random forest & neural networks	Predictive	Confusion matrix, accuracy, precision, recall, sensitivity, specificity & AUC	Statistics of activity level	This empirical study evaluates the performance of four ML algorithms for automating lead scoring and shows how automated lead scoring can improve the sales process by revealing insights into sales
p38	Prediction of Potential Bank Customers: Application on Data Mining	Başarslan & Argun (2020)	Decision tree, Naïve Bayes, K-nearest-neighbour, logistic regression, random forest & gradient boosted trees	Predictive	Confusion matrix, accuracy, precision, sensitivity & F1-score	Number of qualified leads	Multiple models are created to estimate potential bank customers in a case study of bank customer acquisition. The best classifier in this study, random forest, is expected to influence sales performance by identifying qualified leads more accurately and effectively
p39	A screening method for lowering customer acquisition cost in small commercial building energy efficiency projects	Safari & Asadi (2020)	Logistic regression	Predictive	Goodness-of-fit: Hosmer-Lemeshow test	Cost reduction	The proposed methodology is applied to screen leads and prioritize targets for business recruitment in a building retrofit project. This data-driven approach to client acquisition can be used to improve the screening process and reduce costs of customer acquisition
p40	Potential Customer Classification in Customer Relationship Management Using Fuzzy Logic	Kulkarni et al. (2020)	Fuzzy logic	Predictive	Accuracy & AUC	Profit/financial gains	An application of fuzzy logic supervised learning method is proposed to classify and prioritize leads into three distinct categories for contacting. The proposed method identifies important leads who have the potential to increase future sales of business
p41	Lead management optimization using data mining: A case in the telecommunications sector	Espadinha-Cruz et al. (2021)	Gradient boosted trees, decision tree, logistic regression & neural network	Predictive	Accuracy, sensitivity, specificity, precision & ROC curves	Lead conversion rate & cost reduction	The methodology improves efficiency on each stage of leads management, particularly, lead qualification. In a telecommunications company case study, result of the proposed ensemble neural networks indicates a significant improvement in customer conversion effectiveness
p42	Target customer identification method of integrated energy service based on logistic regression	Wang et al. (2021)	Logistic regression	Predictive	Accuracy, precision, recall & Kappa coefficient	Hit rate on number of customers who purchase, response percentage & cost reduction	The prediction model with a target customer evaluation index framework provides support for integrated energy service companies to optimize the allocation of marketing resources more efficiently. This model helps companies target key customers, improves response rate at early stage of market, and reduces marketing costs

ID	Title	Authors	Methods/Models/Algorithms	Approach	Evaluation metrics	Sales performance metrics	Impact on sales
p43	Generic automated lead ranking in dynamics CRM	Kasturi et al. (2021)	Gradient boosted trees	Predictive	ROC curves, AUC & accuracy	Number of qualified leads	The Dynamics 365 generic lead ranking system with human inputs in feature engineering and selection has been used to increase the number of qualified leads by generating more accurate predictions
p44	A social CRM analytic framework for improving customer retention, acquisition, and conversion	Lamrhari et al. (2022)	Random forest, LDA topic modeling, K-means & linear optimization	Predictive	Accuracy, precision, recall, F1-score, specificity & sensitivity	Conversion rate, cost reduction & equilibrium percentage	The social CRM analytic framework with three modules aims at improving customer acquisition and conversion. The results suggest that the framework integrating social media data is able to drive effective marketing strategies and influence sales positively

References

1. Aggour KS, Hoogs B (2013) Financing lead triggers: empowering sales reps through knowledge discovery and fusion. Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining pp. 1141-1149
2. Ahmed SR (2004) Applications of data mining in retail business. Inform Technol: Coding and Comput 2:455–459
3. Badampudi D, Wohlin C, Petersen K (2015) Experiences from using snowballing and database searches in systematic literature studies. Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering pp. 1-10
4. Baecke P, Van Den Poel D (2012) Including spatial interdependence in customer acquisition models: A cross-category comparison. Expert Syst Appl 39(15):12105–12113
5. Baecke P, Van Den Poel D (2013) Improving customer acquisition models by incorporating spatial autocorrelation at different levels of granularity. J Intell Inform Syst 41(1):73–90
6. Başarslan MS, Argun İD (2020) Prediction of potential bank customers: application on Data Mining. Artificial Intelligence and Applied Mathematics in Engineering Problems pp. 96-106
7. Benhaddou Y, Leray P (2017) Customer relationship management and small data-application of Bayesian network elicitation techniques for building a lead scoring model. 2017 IEEE/ACS 14th International Conference on Computational Science and Its Applications pp. 251-255
8. Blei DM (2012) Probabilistic topic models. Commun ACM 55(4):77–84
9. Bohanec M, Borštnar MK, Robnik-Šikonja M (2016) Integration of machine learning insights into organizational learning: a case of B2B sales forecasting. Lecture Notes in Information Systems and Organisation pp. 71-85
10. Bohanec M, Bortnar MK, Robnik-ikonja M (2017) Explaining machine learning models in sales predictions. Expert Syst Appl 71:416–428

11. Boogar L (2019) The three stages of lead scoring: Lambs, Ducks & Kudus. The MadKudu Blog post. https://blog.madkudu.com/ three-stages-of-lead-scoring/ Accessed 28 June 2019
12. Brings J, Daun M, Kempe M, Weyer T (2018) On different search methods for systematic literature reviews and maps: experiences from a literature search on validation and verification of emergent behavior. Proceedings of the 22nd International Conference on Evaluation and Assessment in Software Engineering pp. 35-45
13. Brown G (2009) How real-time online sales lead scoring drives a competitive edge. Bloor Research White Paper No.1021. https://www.yumpu.com/en/document/read/18066356/ real-time-lead-scoring-by-analyst-firm-bloor-research-ebureau/. Accessed 05 June 2021
14. Carroll B (2006) What 's a Lead? Target Marketing, Philadelphia 29(11)
15. Chorianopoulos A (2016) Effective CRM using predictive analytics. John Wiley & Sons Ltd, United Kingdom
16. Coe JM (2004) The integration of direct marketing and field sales to form a new B2B sales coverage model. J Interact Mark 18(2):62–74
17. D'Haen J, Van den Poel D (2013) Model-supported business- to-business prospect prediction based on an iterative customer acquisition framework. Ind Mark Manage 42(4):544–551
18. D'Haen J, Van den Poel D, Thorleuchter D (2013) Predicting customer profitability during acquisition: Finding the optimal combination of data source and data mining technique. Expert Syst Appl 40(6):2007–2012
19. D'Haen J, Van den Poel D, Thorleuchter D, Benoit DF (2016) Integrating expert knowledge and multilingual web crawling data in a lead qualification system. Decis Support Syst 82:69–78
20. Duncan B, Elkan C (2015) Probabilistic modeling of a sales funnel to prioritize leads. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 1751-1758
21. Eitle V, Buxmann P (2019) Business analytics for sales pipeline management in the software industry: a machine learning perspective. Proceedings of the 52nd Hawaii International Conference on System Sciences pp. 1013-1022
22. Espadinha-Cruz P, Fernandes A, Grilo A (2021) Lead management optimization using data mining: a case in the telecommunications sector. Computers and Industrial Engineering p. 154
23. Gokhale P, Joshi P (2018) A binary classification approach to lead identification and qualification. Communications in Computer and Information Science pp. 279-291
24. Grandy T (2005) What is a qualified sales lead. Reeves J: Plumbing, Heating, Cool 85(11):22–23

25. Hasselwander A (2006) B2B Pipeline Management. B2B Marketing Confidential. http://b2bmarketingconfidential.blogspot.com. Accessed 11 June 2021
26. Hornstein S (2005) Sizing up prospects. Sales & Market Manag 157:22
27. Hou JL, Yang ST (2006) A critical customer identification model for technology and service providers. Int J Services and Standards 2(4):417–436
28. Järvinen J, Taiminen H (2016) Harnessing marketing automation for B2B content marketing. Ind Mark Manage 54:164–175
29. Jaskaran (2018) Predictive Lead Scoring: Why, How & Where. Inbound Mantra. https://inboundmantra.com/blog/predictive-lead-scoring-why-how-where. Accessed 26 Nov 2019
30. Kasturi G, Ezzour AJ, Berezin H, Bhanavase S, Preizler R, Hauon E, Nir O, Ronen R (2021) Generic automated lead ranking in dynamics CRM. RecSys 2021 - 15th ACM Conference on Recommender Systems 757-759
31. Kazemi A, Babaei ME, Javad MOM (2015) A data mining approach for turning potential customers into real ones in basket purchase analysis. Int J Bus Inform Syst 19(2):139–158
32. Kim Y, Street WN, Russell GJ, Menczer F (2005) Customer targeting: a neural network approach guided by genetic algorithms. Manage Sci 51(2):264–276
33. Kulkarni T, Mokadam P, Bhat J, Devadkar K (2020) Potential customer classification in customer relationship management using Fuzzy Logic. Innovative Data Communication Technologies and Application 67-75
34. Kitchenham B (2004) Procedures for performing systematic reviews. Technical Report, Keele University and NICTA, Staf- fordshire, UK 33:1–26
35. Lamrhari S, Ghazi HE, Oubrich M, Faker AE (2022) A social CRM analytic framework for improving customer retention, acquisition, and conversion. Technological Forecasting and Social Change p. 174
36. Liberati A (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med 151(4):65–94
37. Lindahl E (2017) A qualitative examination of lead scoring in B2B marketing automation, with a recommendation for its practice. Research Project, KTH Vetenskap Och Konst
38. Luk CC, Choy KL, Lam HY (2019) Design of an intelligent customer identification model in e-commerce logistics industry. MATEC Web of Conf 255:04003
39. McDonnell D (2014) The Evolution From Traditional To Predictive Lead Scoring. Demand Gen Report. https://www.demandgenreport.com/industry-resources/ebooks/2892-the-evolution-from traditional-to-predictive-lead-scoring. Accessed 23 Feb 2019

40. Megahed A, Yin P, Nezhad HR (2016) An optimization approach to services sales forecasting in a multi-staged sales pipeline. IEEE Int Conf Services Comput (SCC) 2016:713–719
41. Meire M, Ballings M, Van den Poel D (2017) The added value of social media data in B2B customer acquisition systems: a real-life experiment. Decis Support Syst 104:26–37
42. Michiels I (2008) Lead Prioritization and Scoring: The Path to Higher Conversion. Aberdeen Group. https://pdfslide.tips/documents/lead-prioritization-and-scoring-the-path-to-higher-conversion.html. Accessed 25 May 2019. Aberdeen Group, research report
43. Michiels I (2009) Lead lifecycle management: building a pipe- line that never leaks. Aberdeen Group. https://www.ontargetpartners.com/wpcontent/uploads/2010/02/Building_A_Pipeline_That_Never_Leaks_by_AberdeenGroup.pdf. Accessed 28 May 2019. Aberdeen Group, research report
44. Monat JP (2011) Industrial sales lead conversion modeling. Market Intell Plan 29(2):178–194
45. Mortensen S, Christison M, Li BC, Zhu AL, Venkatesan R (2019) Predicting and defining B2B sales success with machine learning. 2019 Systems and Information Engineering Design Symposium (SIEDS)
46. Ngai EWT, Xiu L, Chau DCK (2009) Application of data mining techniques in customer relationship management: a literature review and classification. Expert Syst Appl 36(2–2):2592–2602
47. Nguyen TT, Nguyen VLH, Nguyen PK (2012) Identifying customer characteristics by using rough set theory with a new algorithm and posterior probabilities. Fourth Int Conf Comput Inform Sci 2012:594–597
48. Nygård R, Mezei J (2020) Automating lead scoring with machine learning: an experiemntal study. In: Proceedings of the 53rd Hawaii International Conference on System Sciences pp. 1439-1448
49. Ohiomah A, Andreev P, Benyoucef M, Hood D (2019) The role of lead management systems in inside sales performance. J Bus Res 102:163–177
50. Ohiomah A, Benyoucef M, Andreev P (2016) Driving inside sales performance with lead management systems: a conceptual model. J Inform Syst Appl Res 9(1):4–15
51. Ohiomah A, Benyoucef M, Andreev P (2020) A multidimensional perspective of business-to-business sales success: a meta-analytic review. Ind Mark Manage 90:435–452
52. Papaioannou D, Sutton A, Carroll C, Booth A, Wong R (2010) Literature searching for social science systematic reviews: consideration of a range of search techniques: Literature searching for social science systematic reviews. Health Info Libr J 27(2):114–122
53. Patma TS, Wardana LW, Wibowo A, Narmaditya BS, Akbarina F (2021) The impact of social media marketing for Indonesian smes sustainability: Lesson from covid-19 pandemic. Cogent Business & Management 8(1)

54. Peng K, Xu D (2011) Modeling of potential customers identification based on correlation analysis and decision tree. Advances in Neural Networks - ISNN 2011 Lecture Notes in Computer Science 566-575
55. Pullins EB, Timonen H, Kaski T, Holopainen M (2017) An investigation of the theory practice gap in professional sales. J Market Theory Practice 25(1):17–38
56. Rutherford BN, Marshall GW, Park J (2014) The moderating effects of gender and inside versus outside sales role in multifaceted job satisfaction. J Bus Res 67(9):1850–1856
57. Sabnis G, Chatterjee SC, Grewal R, Lilien GL (2013) The sales lead black hole: on sales reps' follow-up of marketing leads. J Mark 77(1):52–67
58. Safari M, Asadi S (2020) A screening method for lowering customer acquisition cost in small commercial building energy efficiency projects. Energ Effi 13(8):1665–1676
59. Sangaralingam K, Verma N, Ravi A, Bae SW, Datta A (2019) High value customer acquisition retention modelling - A scalable data mashup approach. IEEE Int Conf Big Data 2019:1907–1916
60. Smith TM, Gopalakrishna S, Chatterjee R (2006) A three-stage model of integrated marketing communications at the marketing-sales interface. J Mark Res 43(4):564–579
61. Soroush A, Bahreininejad A, Van Den Berg J (2012) A hybrid customer prediction system based on multiple forward stepwise logistic regression mode. Intell Data Anal 16(2):265–278
62. Syam N, Sharma A (2018) Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. Ind Mark Manage 69:135–146
63. Thaichon P, Surachartkumtonkun J, Quach S, Weaven S, Palmatier RW (2018) Hybrid sales structures in the age of e-commerce. J Personal Sell Sales Manag 38(3):277–302
64. Thorleuchter D, Van Den Poel D, Prinzie A (2012) Analyzing existing customers' websites to improve the customer acquisition process as well as the profitability prediction in B-to-B marketing. Expert Syst Appl 39(3):2597–2605
65. Tu Y, Yang Z, Benslimane Y (2011) Towards an optimal classification model against imbalanced data for customer relationship management. 2011 7th International Conference on Natural Computation pp. 2401-2405
66. VanillaSoft (2014) 4 Ways queue-based lead management is shaping the inside sales industry. VanillaSoft.https://vanillasoft.com/resource-center/white-papers-ebooks/queue-based-lead-management. Accessed 17 July 2021
67. Velocity (2013) Best lead distribution methods for optimal sales performance. Velocity. https://www.slideshare.net/Velocify/vel-best-leaddistromthdbooklet_nal. Accessed 16 July 2021

68. Verma R, Koul S, Pai SS (2016) Identifying profitable clientele using the analytical hierarchy process. Int J Bus Syst Res 10(2–4):220–237
69. Wang M, Li Y, Li Q (2021) Target customer identification method of integrated energy service based on logistic regression. The 9th China International Conference on Electricity Distribution pp. 1025-1029
70. Wei JT, Lin SY, Yang YZ, Wu HH (2016) Applying data mining and rfm model to analyze Customers' Values of a veterinary hospital. Int Symp Comput Consum Control 2016:481–484
71. Xiaowen L, Lili T, Zuohao H (2014) Research on value assessment-based accurate identification of government and corporate customers of telecom operators. China Commun 11(11):168–173
72. Xu X, Tang L, Rangan V (2017) Hitting your number or not? A robust & intelligent sales forecast system. IEEE International Conference on Big Data (BIGDATA) 2017:3612–3622
73. Yan J, Gong M, Sun C, Huang J, Chu SM (2015) Sales pipeline win propensity prediction: A regression approach. 2015 IFIP/ IEEE International Symposium on Integrated Network Management (IM2015) 854-857
74. Ylijoki O (2018) Guidelines for assessing the value of a predictive algorithm: a case study. J Market Anal 6(1):19–26
75. Zhu B, Niu Y, Xiao J, Baesens B (2017) A new transferred feature selection algorithm for customer identification. Neural Comput Appl 28(9):2593–2603

Chapter 2 Key Points Summary:

- A systematic literature review was conducted to review and analyze the existing literature on lead scoring models and their impact on sales performance;
- Provided a comprehensive discussion on which machine learning method should be used to build predictive lead scoring models based on the availability of different types of data sources;
- Proposed a classification framework to classify lead scoring models and suggest ways of improving sales performance;
- Identified gaps in the current implementation of predictive lead scoring models.

Chapter 3. Unlocking B2B buyer intentions to purchase: Conceptualizing and validating inside sales purchases



Article Reference: Wu, M., Andreev, P., Benyoucef, M., & Hood, D. (2024). Unlocking B2B buyer intentions to purchase: Conceptualizing and validating inside sales purchases. *Decision Support Systems, 179*. 114165. https://doi.org/10.1016/j.dss.2023.114165

Article Abstract: This study focuses on understanding the purchase decision-making process of B2B buyers in the context of inside sales. While many studies have explored this topic in a B2C context, there has been little attention given to the unique features of B2B inside sales. To address this gap, we developed and empirically validated a buyer' s intention to purchase (BIP) model that integrates the B2B purchase decision-making process and buying behavior theories. Using PLS-SEM analysis on data collected from 126 B2B buyers, we found that trust in sellers and connection flexibility are crucial factors in determining buyers' intention to purchase. Specifically, these factors influence the purchase intention via appointment and contact willingness during different phases of the B2B purchase process. Additionally, our multigroup analyses suggest that buyers' demographic and firmographic attributes affect the purchase decision-making process in different ways. By identifying different groups of buyers and their distinctive needs and preferences, our findings inform more effective approaching and selling strategies. This study contributes to both academic and sales practices by providing critical insights into the factors that influence the B2B inside sales process, which ultimately improve sales performance.

Keywords: B2B inside sales; Purchase decision-making process; Buyer's intention to purchase; Sales performance; PLS-SEM

Note: This chapter is an author accepted manuscript submitted to the journal referred to above.

3.1 Introduction

In today's rapidly changing business landscape, understanding buyers' needs and preferences is vital for achieving sales success (Homburg et al., 2011; Rapp et al., 2012). However, while previous sales research has primarily focused on the organizational and salesperson's perspective (e.g., Ohiomah et al., 2019; Pöyry et al., 2017; Sabnis et al., 2013; Smith et al., 2006), a crucial aspect has been largely overlooked: the buyer's perspective. In particular, there is a lack of knowledge on modeling the buying behavior and purchase decision-making process in business-to-business (B2B) inside sales. As opposed to conventional/outside sales where sales representatives sell products/services based on in-person (i.e., face-to-face) interactions with potential buyers (Rapp et al., 2012), "inside sales" refers to selling activities conducted from a distance by utilizing one or more Information and Communication Technology (ICT) tools (Harmon & Funk, 2014). B2B inside sales reps remain indoors (e.g., call center, office) and try to sell products/services through interactions with potential buying organizations via phone, email, and/or the Internet. (Ohiomah et al., 2019, 2020; Seley & Holloway, 2008). With the COVID-19 pandemic accelerating the shift to online/remote sales (Patma et al., 2021), it is more critical than ever for B2B inside sales organizations to gain a more comprehensive understanding of their buyers' perspectives to improve their sales performance. In this paper, we address these limitations by introducing a conceptual model that integrates B2B inside sales purchase decision-making process and buying behavior theories, and empirically validating this model while considering distinct phases of the buying process and buyers' various characteristics. We are motivated by the aspiration to unravel deeper insights into the phased purchase decision-making process from the buyer's angle. Our findings reveal critical factors that influence a B2B inside sales buyer's purchase intention during the process, shedding light on how organizations improve selling strategies to better meet their buyers' needs and preferences.

This study makes three contributions to the literature: (1) the introduction of a conceptual model that integrates B2B inside sales purchase decision-making process and buying behavior theories, capturing drivers and enablers of a buyer's purchase intention; (2) the examination of buyers' purchase behavior among distinct groups and buyer profiles by considering distinct phases in the process; and (3) the uncovering of the influence of buyers' trust in sellers and connection flexibility on the intention to purchase via appointment and contact willingness in each phase.

Studies on purchase decision-making in consumer markets are well-established (e.g., Chaudhuri et al., 2021; Chiu et al., 2012; Lu et al., 2020; Zhang et al., 2014) and have developed fundamental theories on buying behavior (e.g., Howard & Sheth, 1969; Sheth et al., 2000; Sheth et al., 2009). However, research on purchase decision-making in B2B markets has not received the same attention (Monat, 2011). It is critical for B2B inside sales organizations to understand their customers' journey to gain competitive

advantage not just through product/service offerings but also through the entire purchasing experience (Steward et al., 2019). This is especially critical now, as the purchase decision-making process is increasingly moving towards online and remote formats. With B2B customers willing to complete over 50% of the purchasing process online (Huer, 2013; Maycock et al., 2012), understanding and optimizing the online purchasing experience is more important than ever. Particularly in inside sales, distant communication can make it challenging to understand the needs and preferences of buyers at each phase of the process. Such understanding is pivotal in enhancing the decision-making process of salespeople and helping managers apply appropriate strategies to their interactions with buyers by considering the current phase in the purchase decision-making process. Subsequently, this elevates sales performance. This study serves as a fundamental building block for subsequent studies on designing and implementing information technology artifacts (e.g., lead management systems) that can help sales practitioners make informed decisions and drive tangible improvements in inside sales.

By addressing these gaps in the literature and understanding drivers and enablers that influence a B2B inside sales buyer's purchase intention, sellers produce competitive advantages not only through product/service offerings but also through the entire purchasing journey. This study provides insights into how to improve the sales process by developing a more comprehensive understanding of buyers' needs and preferences and ultimately boosting sales performance.

This paper is organized as follows. Section 3.2 Background discusses the background and Section 3.3 B2B buyer's intention to purchase (BIP) decision-making model develops the conceptual model. The methodology is discussed in Section 3.4 Methodology and the empirical validation of the model is presented in Section 3.5 Results. Lastly, Section 3.6 Discussion and implications summarizes our findings and discusses contributions, implications, and future research directions.

3.2 Background

3.2.1 B2B purchase decision-making process

In B2B environments, purchasing decisions are made by a group of stakeholders, often referred to as the Decision Making Unit (Webster & Wind, 1972). Usually, the process is complex and involves multiple hierarchal approvals from different departments and stakeholders (Zolkiewski et al., 2017). The literature identifies two approaches to differentiate the steps in the B2B purchase decision-making process. One approach consists of three steps: pre-purchase, purchase, and post-purchase (Lemon & Verhoef, 2016; Neslin et al., 2006). The other approach models the process through a buying funnel consisting of five phases: problem recognition, information search, solution evaluation and selection, purchase, and post-purchase (Jansen & Schuster, 2011; Puccinelli et al., 2009; Sirakaya & Woodside, 2005).

This study adopts the latter approach, which provides a more granular view of the buyer's purchase decision-making process. As shown in **Figure 3.1**, in the need recognition phase (NRP), the buyer becomes aware of a new need for the organization. In the information search phase (ISP), the buyer confirms the need and begins searching for information on products/services. In the evaluation and selection phase (ESP), the buyer considers and compares specific offerings, brands, promotions, vendors, and conditions. After consideration and discussion, the buyer finalizes the buying decision regarding the products/services and makes the purchase. Finally, in the post-purchase phase, the buyer reflects on the purchase process and prepares for subsequent purchases.

Figure 3.1 appears here

3.2.2 B2B buying behavior

B2B buying behavior is complex and differs from B2C buying behavior. For example, Lu et al. (2020) examined how online forum browsing behavior affects consumers' purchase decisions. Chaudhuri et al. (2021) identified several customer engagement variables as significant contributors to a purchase decision. To understand B2B buyers' purchase intention, it is important to consider their preferences and attitudes toward a specific product/brand from multiple perspectives (Chang et al., 2020; Peña-García et al., 2020). While there are various theories and models on B2B buying behavior (Ohiomah et al., 2020; Steward et al., 2019), common determinants exist, such as customer, informational, environmental, group, seller, purchase, organizational characteristics, conflict/negotiation, and process (Robinson et al., 1967; Sheth, 1973; Webster & Wind, 1972). To model B2B inside sales buying behavior, this study proposes a conceptual model that incorporates customer and informational determinants (Johnston & Lewin, 1996) which are identified as the most significant among the common ones. Customer determinants, particularly customer trust, play a critical role in B2B sales success (Johnston & Lewin, 1996; Ohiomah et al., 2020). The quality of a salesperson-customer relationship, which is affected by customer trust, significantly influences sales outcomes (Ohiomah et al., 2020; Palmatier et al., 2006), while customer behavior and connection flexibility positively impact customers' purchase intention in a B2B environment (Chatterjee et al., 2022; Zheng et al., 2022).

Leveraging the relationship-based theory (Dwyer et al., 1987; Sengupta et al., 2000; Steward et al., 2019) which focuses on investigating strategic relationship variables between buyers and sellers, this study uses a "trust" construct in the proposed conceptual model. Trust has been emphasized as a major element that affects purchase intention and attitude towards sellers (Dwyer et al., 1987). Additionally, as suggested by Lu et al. (2020), customers' behaviors affect purchase decisions. Thus, as a behavioral attribute measuring the intensity of a buyer's participation and engagement in and connection with a selling organization's offerings or activities (Vivek et al., 2012), a "connection flexibility" construct is included in

this conceptual model. It measures the intensity and flexibility of a buyer's contacts with selling organizations, which is key in the purchase process. We use connection flexibility to represent the informational determinant (Johnston & Lewin, 1996) which measures "*the source and type of information each decision-maker is exposed to and his/her participation in the active search for this information*" (p.2). Sheth (1973) and Quigley et al. (1993) pointed out that an active search for information on product/service offerings is important for purchase intention. By including these constructs in the conceptual model, this study aims to examine the factors that influence a B2B inside sales buyer's intention to purchase.

3.3 B2B buyer's intention to purchase (BIP) decision-making model

We propose a conceptual model (**Figure 3.2**) for studying the B2B purchase decision-making process by integrating literature on customer trust (Chang et al., 2020; Dwyer et al., 1987; Sengupta et al., 2000) and informational determinants (Johnston & Lewin, 1996; Lemon & Verhoef, 2016; Quigley et al., 1993; Sheth, 1973). Focusing on the first three phases (NRP, ISP, & ESP), we aim to understand how B2B inside sales buying behavior influences purchase intention in the decision-making process. While in the purchase phase, a buyer has considered vendors' solutions and is ready to make a purchase (i.e., a purchase decision has already been made), hence neither independent nor dependent variables exist. The same reasoning applies in the post-purchase phase. Thus, both purchase and post-purchase phases are not considered in this study.

Intention to Purchase (IP): this is the dependent variable, defined as the likelihood of a buyer deciding to purchase at the end of each phase (Zeithaml, 1988). An increased IP can enhance the probability to purchase, improving sales success. This aligns with Monat (2011)'s lead conversion study representing purchase intention as the likelihood of making a purchase.

Figure 3.2 appears here

Seller Trust (ST): the literature identifies three types of trust: (1) trust in the seller/vendor/partner (Chiu et al., 2012; Dwyer et al., 1987; Sengupta et al., 2000); (2) trust in the product/service (Homburg et al., 2011; Rapp et al., 2006); and (3) trust in the buying process (Håkansson, 1982). In B2B inside sales, buyers' purchase activities are professional and formalized, which leads to trust in the process and product assessment. However, in most cases, trust in the seller/vendor becomes more crucial for B2B buyers. Therefore, this study focuses on trust in the seller/vendor. Trust is defined as "*a willingness to rely on an exchange partner in whom one has confidence"* (Moorman et al., 1993, p.82). It is the expected trustworthiness resulting from acting professionally with reliability (Moorman et al., 1993). Additionally, trust is perceived as "*confidence in the exchange partner's reliability and integrity"* (Morgan & Hunt, 1994, p.23), and it can be created when a buyer has confidence in the seller's reliability and integrity to deliver what has been promised while protecting the buyer's interests (Crosby et al., 1990; Dwyer et al., 1987). In this study, seller trust (ST) refers to a B2B buyer's reliance on the credibility and reliability of the

seller/vendor during the purchase decision-making process. Trust in sellers is critical for maintaining the healthy relationship between buyers and sellers (Chiu et al., 2012) and improving sales success (Dwyer et al., 1987; Sengupta et al., 2000). The customer's trust is the most significant factor influencing sales success (Ohiomah et al., 2020). Buyers receive information from sellers. The perceived credibility of the information source is a heuristic cue that affects purchase behavioral intention (Zhang et al., 2014). In fact, buyers' trust in sellers increases purchase intention while reducing perceived risk and decreasing consideration time as buyers select desired products/services (Ba & Pavlou, 2002; Chang et al., 2020). Based on this, we propose the following hypothesis:

H1. *A B2B buyer's purchase intention is positively impacted by the buyer's trust in sellers/vendors.*

Connection Flexibility (CF): we define CF as a B2B buyer's ability to connect using various information sources, channels, and time and frequency to obtain necessary information during the buying process. It is important for buyers to participate in information search during the decision-making process, which is the job responsibility of purchasing agents (Sheth, 1973). Various information sources and channels, such as personal recommendations, conferences, advertisements, phone, email, text messages, and social media define a buyer's connectivity. Moreover, a buyer's time flexibility in obtaining information is also critical. Buyers who are more flexible with regards to contact channels and time have more opportunities to learn and absorb information, and those who are open to more sources and contact day preferences rely on information from a wide variety of sources. An active search for information on product/service offerings is a positive indication of purchase intention (Quigley et al., 1993; Sheth, 1973). Therefore, we hypothesize that:

H2. *A B2B buyer's connection flexibility positively impacts intention to purchase.*

Willingness to be Contacted (WC): information exchange between B2B buyers and sellers is crucial to the buying process (Quigley et al., 1993). A buyer's willingness to be contacted (WC) by a salesperson during a particular phase is an important factor that affects purchase decisions (Quigley et al., 1993; Monat, 2011). Proactive inquiries made by buyers are indicators of purchase urgency and their contact willingness, which is essential for sales success (Jolson, 1988; Monat, 2011). We propose including WC in our conceptual model, as it may have a positive impact on purchase intention. Our hypothesis is as follows:

H3. *A B2B buyer's purchase intention is positively impacted by the buyer's willingness to be contacted by a salesperson.*

A buyer's contact willingness is positively related to the contact frequency and preferred contact days of the buyer (Ohiomah et al., 2018). Buyers may be more receptive to be contacted at certain phases in the buying process or during specific time windows. The choice of contact channels can also positively impact contact willingness (Ohiomah et al., 2019). Buyers who have many contact channel preferences

(e.g., email, phone, social media, video) may have stronger contact willingness. Therefore, we propose the following hypothesis:

H4. *A B2B buyer's willingness to be contacted is positively impacted by the buyer's connection flexibility.*

Research has shown that trust plays a vital role in contact willingness (Ohiomah et al., 2020; Schoorman et al., 2007). Buyers are more willing to interact with salespeople whom they trust. Salespeople who invest time in understanding the buyer's needs, providing relevant information, and building trustworthy relationships rather than solely pushing for a sale are more likely to positively impact the contact willingness of buyers during the buying process (Monat, 2011; Ohiomah et al., 2020). Thus, we propose the following hypothesis:

H5. *A B2B buyer's willingness to be contacted is positively impacted by the buyer's trust in sellers/vendors.*

Willingness to Make an Appointment (WA): refers to a B2B buyer's likelihood of scheduling an appointment with a salesperson to discuss products/services (Ohiomah et al., 2018). According to Smith et al. (2006), a B2B sales process involves three stages: generating sales opportunities, converting opportunities into appointments, and finally, turning appointments into actual sales. In this three-stage model, appointment willingness is crucial as it serves as a precursor to purchase commitment. For B2B inside sales, appointment willingness is particularly important because making an appointment is often the primary goal of the inside sales process (Ohiomah et al., 2018).

After conducting an information search, a buyer starts building rapport with the seller and becoming open to contact, but there is no commitment yet. A buyer's WA reflects the likelihood of making a commitment to the seller. Once an appointment has been made, the seller takes most control of the selling. Thus, it is essential to include appointment willingness in our conceptual model, given its importance in inside sales. We propose that:

H6. *A B2B buyer's purchase intention is positively impacted by the buyer's willingness to make an appointment with a salesperson.*

Some studies have shown that when a prospect requests an appointment during the buying process, it indicates their interest in the products/services offered by sellers (Ohiomah et al., 2018; Smith et al., 2006). For instance, when a prospect requests an appointment for a sales demo, it usually represents a pivotal point in the inside sales process. Bovée and Arens (1986) highlighted the role of flexibility and openness to being connected in establishing and maintaining moderate customer engagement including booking and appointment. A buyer's engagement in activities with salespeople is positively related to the contact time and frequency (i.e., connection flexibility) of the buyer (Chatterjee et al., 2022). Hence, we propose that:

H7. *A B2B buyer's willingness to make an appointment is positively impacted by the buyer's connection flexibility.*

Trust not only plays a significant role in building and maintaining long-term relationships, but also influences outcomes of relation building (Morgan & Hunt, 1994). Appointment booking is deemed an important outcome of relationship building between parties in inside sales (Ohiomah et al., 2018). Thus, we argue that the perceived trust in the credibility of sellers may affect buyers' willingness to make appointments. Hence, we propose that:

H8. *A B2B buyer's willingness to make an appointment is positively impacted by the buyer's trust in sellers/vendors*.

According to the segmentation approach (Baron & Kenny, 1986; Rungtusanatham et al., 2014), there are three steps to develop and demonstrate the role of a mediator. In the first step, we developed hypotheses for the effect of ST on WC (H5) and the effect of CF on WC (H4). Then, we developed a hypothesis for the impact of WC on IP (H3). Finally, we could conclude by hypothesizing the mediation effect of WC in the relationships between ST, CF, and IP using the segmentation approach. Thus, we propose that:

H9*. Willingness to be contacted mediates the relationship between seller trust and purchase intention;*

H10*. Willingness to be contacted mediates the relationship between connection flexibility and purchase intention.*

Appointment willingness is a broader concept that encompasses the buyer's interest and readiness to engage further in the remaining phases of the buying process. After applying the same segmentation approach on hypothesizing the mediation role of WA, we argue that:

H11. *Willingness to make an appointment mediates the relationship between seller trust and purchase intention;*

H12. *Willingness to make an appointment mediates the relationship between connection flexibility and purchase intention.*

We investigate the associations between seller trust, connection flexibility, and purchase intention, as well as the mediating roles of appointment and contact willingness.

3.4 Methodology

We used PLS-SEM with SmartPLS (v.3.3.8) software to validate our BIP model and test hypotheses. PLS-SEM is suitable for exploratory studies and has less restrictive sample size requirements than covariance-based PLS (Andreev et al., 2009). We followed recommended guidelines to validate the model in the NRP, ISP, and ESP phases (Andreev et al., 2009, Hair Jr et al., 2016). We considered survey participants' demographics and firmographics such as employment, seniority, education, sex, and company size as

control variables relevant to the endogenous variables (i.e., WA, WC, and IP). Because some control variables are partially significant, multigroup analyses (MGAs) were conducted to further assess differences in parameter estimates and path coefficients for additional insights. Two approaches were employed: i) grouping by employment, seniority, education, sex, and company size, and ii) a clustering modeling technique using Machine Learning (ML) to understand buyers' behavior in heterogeneous groups (Amoozad Mahdiraji et al., 2022; Chorianopoulos, 2016).

We designed an online survey with AA-ISP (American Association of Inside Sales Professionals) (https://www.aa-isp.org/) and conducted a preliminary pilot with three experts to validate the survey and measurement items. We refined the survey and used 11 measurement items associated with the target constructs of the BIP model that were employed in this study for each buying phase (a total of 33 measurement items for the three phases) (see **Table 3.2**).

Table 3.2 appears here

3.5 Results

3.5.1 Data collection

We randomly selected 1000 professional B2B buyers based in the USA using a stratified sampling technique, with the sample population provided by AA-ISP. The sample was stratified by industry type and included operations specialists, marketing and sales managers, and executives. As per the requirements of this study, the email list provided to us by AA-ISP contained the emails of professionals with B2B inside sales experience. An Amazon gift card ($25) was offered as an incentive for participation, and follow-up emails were sent after two weeks. The six-week online survey was conducted in May 2019. We received 126 complete responses after discarding 27 incomplete responses, resulting in an overall response rate of 12.6% which is considered acceptable (Visser et al., 1996). The sample size was sufficient for our model, being three times above the minimum requirement (Hair Jr et al., 2016; Ohiomah et al., 2019) and more than the minimum required sample size (113) with a significance level of 10% and a minimum path coefficient of 0.2 based on the more conservative inverse square root method (Hair Jr et al., 2022). Our online survey participants come from various industries with the majority from business and professional services.

Table 3.1 shows the demographic and firmographic profile of participants. Most respondents were middle-level employees (55.56%), followed by executive-level (31.74%) and operational-level (12.70%). Age was uniformly distributed. Company size was measured by the number of employees, with 45.24% of respondents from small-sized companies, and the rest from medium-sized or large-sized companies. Before participants started answering the online survey questions, they were introduced to the definition of the buyer purchase decision-making process with diagram demonstration to avoid a terminology bias. Then we

provided and explained specific scenarios for participants to consider while answering each question in the NRP, ISP, and ESP (see **Table 3.1**). Hence all participants being professional B2B buyers answered the sets of questions for different stages putting themselves in the situation of being at this stage.

Table 3.1 appears here

To mitigate and manage the potential bias introduced by posing modified questions for each stage, we documented the modifications made to questions and reasons for them when we introduced each stage to participants. Furthermore, using PLS-SEM to model the associations between latent variables and observed indicators allows us to account for measurement errors and modifications in questions during each stage. Finally, we consulted three B2B inside sales experts (from AA-ISP) to identify any potential bias introduced by modified questions in each stage before sending out the survey. The experts helped to refine questions and avoid potential bias.

3.5.2 BIP model evaluation

To validate the proposed BIP model for each phase (i.e., NRP, ISP, & ESP), we used a two-step approach. In the first step, we assessed the reliability and validity of constructs and indicators by focusing on the measurement models. In the second step, we evaluated the structural model to assess the relationships between constructs. To determine whether there were any significant differences in parameter estimates and path coefficients among pre-defined data groups, we conducted multigroup analyses (MGAs).

3.5.2.1 Measurement models evaluation

Reflective construct evaluation: we evaluated the reflective construct seller trust (ST) in our measurement model for reliability and validity. The internal consistency reliability was assessed using Cronbach's alpha (CA) and composite reliability (CR). **Table 3.3** shows that reflective constructs for all three phases meet internal consistency and reliability criteria as values are above 0.7 for CA and CR (Chin, 1998; Cronbach, 1951). Additionally, the Average Variance Extracted values are above minimum required level of 0.5 (Chin, 1998; Fornell & Larcker, 1981).

Table 3.3 appears here

We calculated the Fornell-Larcker criterion and the Heterotrait-monotrait (HTMT) ratio to assess the discriminant validity of reflective constructs. Because the shared variance with associated indicators of each construct (i.e., diagonal values in each phase) is higher than with other constructs, this verifies the uniqueness of each construct in the model (Hair Jr et al., 2016) (see **Table 3.4**). Additionally, **Table 3.5** indicates that all HTMT ratios between two constructs are less than 0.85 (Hair Jr et al., 2016; Henseler et al., 2015). The distinctiveness of constructs is confirmed.

Table 3.4 appears here

Table 3.5 appears here

Formative construct evaluation: we defined connection flexibility (CF) as a formative construct based on the rules (Andreev et al., 2009; Hair Jr et al., 2016). In our opinion, contact channel preference (CF1), contact frequency preference (CF2), information source influence (CF3), and contact day preference (CF4) (see **Table 3.1** for definitions) define/form CF. Implied by the conceptual definition of CF (Section 3.2 Background), the direction of causality is from these four indicators to CF. Also, considering direction of change, the change in indicators (e.g., contact data preference) impacts CF, not visa versa. Hence CF is formative. Additionally, all measures of a formative construct are not interchangeable. Furthermore, CF is the consequence of measurement items, thus CF is formative. It reflects a B2B buyer's flexibility/openness for connectivity in terms of CF1, CF2, CF3, and CF4. We assessed the validity of the formative construct CF (**Table 3.6**).

Table 3.6 appears here

We found that all formative indicators' weights were above 0.1 and statistically significant in the NRP, while in the ISP, only one indicator (CF1) was significant. In the ESP, all indicators except two (CF1 & CF2) were significant, indicating a difference in connection flexibility between the three phases. In the NRP and ISP, contact channel flexibility dominated in forming connection flexibility ($\gamma = 0.520$ and $\gamma = 0.735$), while in the ESP, the flexibility of information sources (CF3) was dominant ($\gamma = 0.630$). Contact frequency and day preferences also played a role in connection flexibility in some phases. We kept all formative indicators despite some being statistically insignificant because they were important in theory (Andreev et al., 2009). Also, we assessed the multicollinearity of indicators and found no evidence of it across the three phases (**Table 3.6**). Finally, for assessing discriminant validity, all correlations of formative indicators were found higher with their construct than with any other construct.

3.5.2.2 Structural model evaluation

We validated the structural model across the three phases for the coefficient of determination, the significance of path coefficients, mediation and moderation effects, effect size, and predictive relevance (Chin, 1998; Hair Jr et al., 2016; Henseler et al., 2009, 2015).

Coefficient of determination (R^2): in the NRP and ISP, the BIP model explains 14% and 11% of the variance in intention to purchase, respectively. However, in the ESP, the model's predictive power improves significantly, explaining 40% of the variance. Note that the model does not aim to comprehensively explain purchase intention, which is a multidimensional concept. On the contrary, the BIP model tries to understand the buyer's intention to purchase based on the limited information the selling organization can obtain and control. Although it was anticipated that the model could not explain the dependent variable with high R^2 in the initial phases of the buying process, the explained variance of the purchase intention is considered high in the ESP. Additionally, in the NRP, connection flexibility (CF) explains 11.2% of the variance in

the willingness to make an appointment (WA) and 21.9% of the variance in the willingness to be contacted (WC). CF explains 15.4% of the variance in WA and 13.9% of the variance in WC in the ISP. Finally, in the ESP, CF explains 18.6% of the variance in WA and 20.2% of the variance in WC.

Significance of path coefficients: in NRP, H4 and H6 - 7 were supported and H1 - 3, H5, and H8 were rejected. In ISP, H3 - 4 and H7 were supported and H1 - 2, H5 - 6, and H8 were rejected. In ESP, H3 - 4, H6 - 7 were supported and H1 - 2, H5, and H8 were rejected (**Table 3.7** and **Figure 3.3**).

Table 3.7 appears here

Mediator analysis: we examined the importance of mediators (WA and WC) across all three phases (**Table 3.8**). In the NRP, WA significantly mediates the relationship between CF and IP, while WC is not significant. This means that the total indirect effect is mediated through WA rather than WC. In the ISP, WC becomes a significant mediator, and WA is no longer significant. This implies that the total indirect effect between CF and IP is mediated through WC instead of WA. Finally, both WA and WC are significant mediators in the ESP. This also explains why connection flexibility only indirectly affects purchase intention (H2) in all three phases.

Table 3.8 appears here

Moderator analysis: additionally, we examined the moderating role of phase (i.e., coded as dummy category variable) on the hypothesized relationships. The results (see **Table 3.9**) indicate the only significant moderating effect ($b = 0.114$, $t = 1.914$) of phase is on the relationship between WC and IP (H3). With the inclusion of the interaction term, the R^2 increased from 21.1% to 32.5%. The relationship between WC and IP is strengthened by each standard deviation of the phase by 0.114. Phase is an amplifying moderator which enhances the effect of WC on IP.

Table 3.9 appears here

Effect size (f^2): we inspected effect size (f^2) of each hypothesized path relationship, with a small effect ($f^2 \geq 0.02$), a medium effect ($f^2 \geq 0.15$), and a large effect ($f^2 \geq 0.35$) indicating the level of impact (Cohen, 1988). **Table 3.10** shows that in NRP, CF has a small effect on WA and a medium effect on WC, while WA has a small effect on IP. In ISP, CF has a medium effect on WA and a small effect on WC, while WC has a small effect on IP. In ESP, CF has a medium effect on WA and WC, while WA and WC have a small effect on IP.

Table 3.10 appears here

Predictive relevance (Q^2): we conducted the Stone-Geisser's test to assess predictive relevance of the structural model. A structural model is assured to have predictive relevance if the Q^2 values of all endogenous constructs are greater than zero (Hair Jr et al., 2016), which has been illustrated in **Table 3.11**. Therefore, the predictive relevance of all path models has been verified.

Table 3.11 appears here

Common method bias evaluation: because we gathered similar data for the three phases, we checked the common method bias to ensure the measurement methods in the survey were unbiased using the full collinearity assessment approach (Kock, 2015; Kock & Lynn, 2012). As shown in **Table 3.12**, since all variance inflation factors of the inner model resulting from a full collinearity test are less than 3.3 threshold, our model is considered free of common method bias.

Table 3.12 appears here

3.5.3 Multigroup analyses

Unexpectedly, our study revealed that the hypothesized relationships between seller trust and other constructs were consistently disproven across all phases of the BIP model, standing in contrast to the pivotal role of trust in B2B buying process models. This finding raises curiosity and seems somewhat paradoxical, given the recognized significance of trust in these contexts. Plausible reasons for this outcome could be attributed to the potential non-homogeneity of data utilized in our analysis. This diversity may have obscured the trust relationship's prevalence when considering the entire sample. However, it's plausible that within specific homogeneous groups, trust might still exhibit importance and maintain statistically significant associations with other constructs.

With respect to the control variables, our results indicate that some of them have a significant effect (see **Table 3.13**). Further exploration is warranted to uncover nuanced insights within these potential subgroup dynamics. To further investigate seller trust and other constructs, we performed 15 MGAs and used a clustering modeling ML approach to differentiate our sample into three dominant buyer profiles, conducting 9 additional analyses. We conducted formal power analysis to support the use of sample sizes in these MGAs. Since the majority of the observed statistical powers are above 0.8 (see **Table 3.14**), we have enough statistical power to detect effects with confidence and our models are strong statistically (Lowry & Gaskin, 2014).

Table 3.13 appears here

Table 3.14 appears here

3.5.3.1 Comparison of homogeneous groups

The results of the BIP model explain the behavioral patterns of an average B2B buyer in the purchase intention when progressing through the buying phases. However, to understand the impact of individual differences between buyers we differentiated them according to employment level, age (seniority), education level, sex, and company size.

Employment level: first, we divided all participants into two groups, executive-level and non- executive level (including middle & operational-level). **Table 3.15** shows that seller trust positively affects WA and

WC for non-executive level employees in the NRP, and the effects are significant. However, the same effects are non-significant for executive-level employees.

Table 3.15 appears here

Seniority: second, we divided all respondents into two groups by seniority. The first group contains senior buyers, while the second group contains non-senior buyers (youth and junior buyers). **Table 3.16** illustrates that WA positively affects IP for senior buyers and non-senior buyers in the NRP, and the difference is significant. Furthermore, in the ESP, seller trust positively affects WC for senior buyers and the effect is significant. However, it is rejected for non-senior buyers.

Table 3.16 appears here

Education level: third, we divided all survey participants into binary groups based on education level. The first group contains buyers with master's degrees (graduate), while the second group (i.e., below graduate) contains buyers with bachelor's degrees and below bachelor's degrees. **Table 3.17** shows that seller trust positively affects IP for buyers with graduate degrees and the effect is significant in the ISP. However, it is rejected for buyers without graduate degrees.

Table 3.17 appears here

Sex: furthermore, we formed two groups based on the respondents' sex (i.e., male or female). In the ISP, WA positively affects IP for female buyers (**Table 3.18**). In the ESP, seller trust positively affects WC for B2B male buyers, and the effect is significant. However, this is not the case for B2B female buyers. Also, connection flexibility positively affects IP for male buyers; however, not for female buyers, and the difference between the two groups is significant.

Table 3.18 appears here

Company size: finally, we formed two groups of companies, small and medium-sized enterprises (SMEs) and large-sized ones. **Table 3.19** illustrates that the hypothesized path relationship from seller trust to purchase intention is positive and significant for buyers in large companies. However, this positive relationship is not significant for buyers who work in SMEs.

Table 3.19 appears here

3.5.3.2 Comparison of dominant buyers' profiles

To better understand the differences between buyers, we used a ML technique called Partitioning Around Medoids (PAM) clustering algorithm (Gower, 1971; Reynolds et al., 2006) in addition to traditional PLS MGA. We used the PAM clustering algorithm to group the 126 participants into clusters based on their demographic and firmographic attributes. We determined the optimal number of clusters to be three using the "elbow" method (Syakur et al., 2018). We assessed the quality of the clusters using the Silhouette coefficient, which was 0.42, indicating a satisfactory clustering solution (Chorianopoulos, 2016; Han et al.,

2012). We then created profiles for each cluster by identifying the most frequent category (i.e., mode) for each attribute and used these profiles for subsequent MGAs. We used profiling to group participants into three main profiles: John (#1), Lara (#2), and Samuel (#3), based on their demographic and firmographic attributes. **Table 3.20** displays the three main participants' profiles. John is an executive-level senior male buyer with a graduate degree who works in SMEs. Lara is a non-executive female employee who works for large-sized companies and does not have a graduate degree, while Samuel is a non-executive male employee who works in SMEs and is younger without a graduate degree.

Table 3.20 appears here

We compared these profiles across all three phases and found significant differences. In the NRP, the paths of WA to IP, CF to WA, and CF to WC are significantly positive for both John and Lara profiles, but the differences are significant. In the ISP, the effect of CF on IP is different for the two profiles and contact with sellers is important for Lara's profile. In the ESP, making appointments with salespeople is important for buyers with Lara's profile, while higher connection flexibility leads to IP for buyers with John's profile. **Table 3.21** shows a summary of the differences.

Table 3.21 appears here

We found significant differences between John and Samuel's profiles in all three phases (**Table 3.22**). In the NRP, the effect of WA on IP is significantly positive for buyers with John's profile, but not for Samuel's. In the ISP, seller trust is more important to buyers with Samuel's profile, while CF is more important for buyers with John's profile. In the ESP, connection flexibility has a greater impact on IP for buyers with John's profile, while seller trust remains important for buyers with Samuel's profile. Differences were statistically significant in all these relationships.

Table 3.22 appears here

Finally, we compared the profile of Lara (#2) and Samuel (#3) across the three phases of the buying process (**Table 3.23**). In the NRP, WA is critical for buyers with Lara's profile; however, not critical for buyers with Samuel's profile, and the difference is significant. In the ISP, seller trust is important for Samuel's profile; however, not important for Lara's profile. For Lara's profile, WC and CF are important, and the differences are significant. In the ESP, appointments are important for Lara's profile; however, not important for Samuel's profile. Moreover, WA directly influences IP in Lara's profile. The seller trust is still important in Samuel's profile.

Table 3.23 appears here

3.6 Discussion and implications

Our study highlights that in a B2B context, intention to purchase is impacted by the buyer's willingness to make an appointment and be contacted, which in turn is influenced by two key factors - seller trust and

connection flexibility. The extent to which these predictor variables explain the variance in intention to purchase varies based on the phases of the buying process. In the ESP, the BIP model's predictor variables explain 40% of the variance in intention to purchase, which is significant. However, during the initial two phases (NRP and ISP), the explanatory power is significantly lower, indicating that buyers are often unwilling to make a purchase at these stages.

This finding suggests that B2B sellers need to tailor their selling practices to match the buyers' current phase in the buying process and their known characteristics, rather than relying on the traditional "push" selling approach (Unni & Harmon, 2007). The lower explanatory power of the BIP model in the first two phases demonstrates that sellers need to use more "pull" strategies to understand the buyer instead of pushing for a deal. This aligns with the marketing strategies proposed in Zhang et al.'s (2022) study on multichannel firms and Molitor et al.'s (2018) study on mobile targeting firms, which suggest that "pull" strategies increase the effectiveness of marketing actions. This means that sellers should focus on "pull" strategies during the selling process, especially in the NRP and ISP. In contrast, during the ESP, sellers should use "push" strategies trying to engage more with buyers and increase conversion rates eventually.

3.6.1 Theoretical implications

This paper offers several theoretical implications that contribute to the existing body of knowledge in the field of B2B inside sales and buyer behavior. The development and validation of the BIP model provide a comprehensive conceptual framework for understanding the complex interplay of factors influencing a B2B buyer's purchase intention. The model integrates theories of the B2B purchase decision-making process and buying behavior, filling a significant gap in the literature. Furthermore, this study sheds light on the unique roles played by trust in sellers and connection flexibility in influencing the purchase intention across different phases. This enhances our understanding of how these constructs interact and contribute to the overall purchase intention.

Specifically, this study contributes to the theoretical and empirical understanding of the relationship between B2B buyers' connection flexibility and their appointment willingness (H7 is supported) during the three phases (Chatterjee et al., 2022). Our findings demonstrate that B2B buyers who are open to connectivity in terms of contact channels, frequency, and information sources are more willing to make appointments, indicating their interest in learning more about the products/services. Additionally, our results support H4 (Ohiomah et al., 2018, 2019), which suggests that B2B buyers who are flexible in connectivity when searching for information are more willing to be contacted by salespeople during the purchase process.

In consistency with prior research (Sabnis et al., 2013; Smith et al., 2006), our study supports the association between appointment willingness and intention to purchase (H6 is supported), except in the ISP.

During ISP, buyers are mainly engaged in information search about products/services and are not ready to make a purchase commitment. Furthermore, our study supports the relationship between contact with a salesperson and purchase intention (Jolson, 1988; Monat, 2011) (H3 is supported), except in the NRP. In NRP, buyers are aware of their needs and are not yet ready to make a purchase decision, despite being willing to be contacted by a salesperson. As buyers move on to the subsequent phases, more information on products/services and providers becomes available, leading to a higher likelihood of making a purchase decision.

While our study found no significant direct relationship between CF and IP (H2), we identified significant positive indirect relationships between them via mediators. These findings indicate that appointment and contact willingness act as enablers of the purchase intention.

3.6.1.1 Multigroup and profiles analyses

This study makes a valuable contribution by utilizing comprehensive MGAs as well as ML approaches to gain a deeper understanding of the differences among the most dominant buyers' profiles. The results of our MGAs reveal that the effects of seller trust and connection flexibility on intention to purchase, as well as the two mediators, differ among various groups of B2B buyers.

Non-executive level employees are more likely to make appointments and be contacted by salespeople in NRP when they have more trust in the credibility and reliability of the sellers, whereas the same relationships do not hold for executive-level employees. Trust still plays a role for non-executive level employees during their search for information in ISP, while executive-level employees prefer to be contacted by phone rather than by official appointments. Appointments are important for both groups in ESP.

Our study found that senior B2B buyers are more inclined to be contacted by sellers during ESP when they have greater trust in the sellers. However, we did not find a significant relationship between trust and contact willingness for non-senior B2B buyers. The effect of trust on contact willingness depends on demographic characteristics, such as age and education level (Chang et al., 2016; Menidjel & Bilgihan, 2022). For example, younger buyers favor appointments and are more likely to be contacted by salespeople than senior buyers during ESP, indicating that younger buyers may need more reassurance before making a purchase decision.

In terms of education level, B2B buyers with graduate degrees who trust a seller are more likely to purchase without evaluating or comparing with other sellers and are also eager to make appointments with salespeople. On the other hand, buyers without graduate degrees are more flexible regarding connectivity and are willing to be contacted by salespeople.

Our study also found that there are differences between male and female B2B buyers. In ISP, female B2B buyers prefer to make appointments with salespeople; however, this is not true for male B2B buyers. In ESP, male buyers are more willing to be contacted by sellers when they have greater trust in them, whereas female buyers' purchase intention is less impacted by seller trust.

Finally, our study also compared company sizes. In NRP, buyers from SMEs are more willing to make appointments with salespeople after they become aware of a need for a product/service. In ISP, buyers in large-sized companies who trust the sellers are more inclined to make a purchase quickly, but this is not the case for buyers in SMEs. In ESP, B2B buyers in large-sized companies who are more inclined to be contacted by salespeople exhibit a higher purchase intention.

Using clustering ML techniques yielded fascinating insights from cross-comparisons of identified profiles (i.e., John, Lara, and Samuel – see **Table 3.20**). During the NRP, John and Lara's profiles indicate that buyers are willing to make appointments with salespeople and are flexible in connectivity. Successful appointments can have a positive impact on purchase intention. There is a good chance that buyers like John and Lara are confident about which sellers to approach to make a purchase. This makes the "push" strategy successful for these types of buyers during the first phase (NRP). However, buyers like Samuel are not ready to commit in the NRP, regardless of salespeople's efforts. They are not willing to make appointments yet, and although they are open to being contacted, it does not necessarily mean they are close to making a purchase.

In the ISP, buyers with John and Lara's profiles remain flexible in connectivity, but their appointment willingness is not as strong as in the NRP. Samuel's profile is different as trust can have a positive impact on their purchase intention and appointment willingness, as Samuel is less experienced and works for an SME. However, our results show that making an appointment does not necessarily lead to a purchase. Therefore, the ISP is a preparation phase for Samuel to make a purchase decision in the subsequent phase. The more trust Samuel has in a seller, the more likely Samuel is to make an appointment and eventually make a purchase decision. In contrast, trust has no statistical direct or indirect impact on purchase intention for buyers with John and Lara's profiles.

In the ESP, buyers with John and Lara's profiles remain flexible in connectivity, and higher connection flexibility leads to a purchase. Buyers with Lara's profile become willing to make appointments with salespeople again after they finalize their information search and analyses of products/services. Samuel's profile shows a similar trend, where trust in sellers leads to more appointment-making and eventual purchase commitment.

In summary, trust is crucial for Samuel's profile, but not for John and Lara's profiles. This could be due to the specificity of the B2B environment and the buyer's role in the organization. For example, John

and Lara are seniors with experience in dealing with salespeople, so trust is not as essential to them. Connection flexibility is important for John and Lara's profiles during all three phases, but for Samuel's profile, it is only essential in the ESP. All three profiles indicate an increase in the purchase intention as they progress through the phases. These results demonstrate that, regardless of the profile, salespeople can increase their success by implementing "pull" strategies to identify buyers' current phase in the purchase decision process and provide them with the necessary information about needs (products/services) during the first two phases. They can then rely on "push" strategies in the ESP. These findings can be of significant value to salespeople and marketers, who can tailor their approaches based on the demographic characteristics of buyers.

3.6.2 Practical implications

The following major practical implications can be drawn for sales practitioners and organizations engaged in B2B inside sales. Salespeople should distinguish between a buyer's NRP and ISP. In the NRP, salespeople should avoid assuming that buyers who are more willing to be contacted will make a purchase. The ability to recognize the boundaries between the two phases is key in training programs for salespeople, as well as optimizing appointment and contact strategies, and aligning with buyers' preferences. In the ISP, salespeople should allow prospects the freedom to look for information on their own. Salespeople should not interfere too much with prospects in the ISP unless they express interest in being contacted. It has been found that the "average" B2B buyer performs half of the purchasing steps online before they engage with salespeople directly (Maycock et al., 2012). Salespeople need to be available to make contacts with buyers in the ESP. Higher connection flexibility increases the purchase intention, appointment preference, and contact willingness. Salespeople should realize that dealing with a flexible B2B buyer may not guarantee a purchase. Salespeople should understand buyers' expectations and goals at every phase and ensure the quality of conversations with buyers is acceptable in each phase.

With regards to connection flexibility, contact channel, frequency, and contact day are key components in NRP. Sales practitioners should provide as much relevant information as possible through various channels at a high frequency for prospects to learn and acquire information without directly contacting salespeople. In ISP, the dominating role of the contact channel becomes even more significant. Sellers should provide more relevant information via various channels for buyers to easily find the information without breaking a sweat. In ESP, information source and contact day are the key components of connection flexibility. Sellers should engage with buyers by providing them with information through various sources on weekdays. Additionally, sellers can leverage the insights on connection flexibility and appointment willingness to optimize communication strategies. This includes providing relevant

information via various channels, adapting appointment scheduling approaches, and enhancing the overall customer experience.

In a remote B2B sales environment, prioritizing connection flexibility (including contact channels, day, frequency, and information sources) outweighs the significance of seller trust. With the absence of face-to-face interactions and limited chances to establish trust, remote B2B sellers must emphasize efficient and effective selling methods. Allocating additional resources to enhance connection flexibility (e.g., increasing the variety of contact channels and information sources, expanding the sales team, and providing training to salespeople on optimal calling times) can be a strategic investment, contributing to improved overall sales performance. Nonetheless, trust plays an important role in some groups. B2B inside sales practitioners should pay attention to leveraging positive customer experiences and feedback from successful purchases to foster stronger relationships and boost confidence in sellers. Specific selling strategies should be targeted to buyers based on their profiles and preferences at different phases. For instance, a "push" selling strategy should not be used for buyers with Samuel's profile. Instead, a "pull" strategy may be better to sell in the first phase and nurture the lead and follow with lead through all the phases. Sellers should focus on communicating and building trust with buyers with Samuel's profile during ISP and ESP. Buyers with John and Lara's profiles prefer appointments in NRP, while buyers like John prefer discussing everything through phone calls in ESP. Sellers should focus on preparing well-organized appointments when meeting with buyers like John and Lara to increase the chance of selling. Buyers who work in large-sized companies like Lara's profile do not prefer appointments in ISP; however, they prefer appointments before finalizing a purchase in ESP. Sellers can use these insights to devise more targeted approaches that resonate with specific buyers' characteristics and needs. Also, the identified factors that influence purchase intention, can be integrated into existing lead management systems. This enhancement equips salespeople with more relevant information to guide interactions with buyers, which may eventually result in improved conversion rates.

Compared with offline B2B sales, more communication channels (e.g., email, online chat, webinars, etc.) are present in the IT-enabled B2B inside sales. This unique factor provides an opportunity for sellers in inside sales. By remotely offering various and flexible communication options to buyers, sellers who are more accessible through communication channels and actively responsive to inquires will more likely build trust and connections with buyers. In summary, findings of this study provide valuable insights for B2B inside sales practitioners. By recognizing the boundaries between different phases, providing product/service information through various channels, building trust, and applying appropriate selling strategies to different profiles, salespeople can increase their chances of success in the selling process.

3.6.3 Limitations and future work

While this study provides valuable insights into the factors affecting B2B buyers' intention to purchase in the context of inside sales, several limitations should be addressed in future research. First, the entire sample was collected from companies in the USA, which limits the generalizability of our findings to the global B2B inside sales industry. To improve the external validity of future studies, it would be beneficial to collect data from a more diverse range of companies across different regions of the world. Future studies could employ larger samples to increase the precision of the estimates. Additionally, the insignificance of seller trust in the collected data is not caused by the sample size as justified by results of the formal power analysis (see **Table 3.14**). There is sufficient statistical power to detect insignificant paths of seller trust with confidence. However, future studies should consider using larger sample sizes and explore this issue.

Second, there are single-item constructs in our model, which may reduce measurement reliability and validity. Although single-item measures are commonly used in survey research (Hair Jr et al., 2016), future studies could consider developing more comprehensive measures for constructs to enhance psychometric properties. Third, this study aims to build and empirically validate a conceptual BIP model that captures drivers and enablers of a buyer's purchase intention. Logically, we included purchase intention as the dependent variable and measured it through an online survey. However, we realize that purchase intention may not translate into actual purchase. Future studies could collect sales/transaction data to measure the actual purchase behavior and analyze it further. Furthermore, although our study provides an additional moderation analysis of phase and finds a significant moderating effect, we did not develop formal hypotheses regarding phase due to a lack of relevant literature support. Thus, we call on future studies to formally develop hypotheses regarding phase and explore its moderating role by using this study's results as evidence. Moreover, we did not consider repeated purchase situations when designing the survey's questionnaire. As repeated purchase may differ greatly from a new purchase, we recommend future studies to consider incorporating repeated purchase scenarios in studying purchase intention.

Finally, we suggest that future research should consider including additional independent variables in the model. Our study is exploratory by nature, and as suggested by Falk & Miller (1992) and Hair Jr et al. (2016), parsimonious models are sometimes more powerful than complex models in exploratory studies on developing theoretical or conceptual models using PLS-SEM.

3.6.4 Conclusion

The traditional sales literature has focused on understanding the sales process from salespeople's perspective, neglecting the importance of buyer's perspective (e.g., Ohiomah et al., 2019; Pöyry et al., 2017; Sabnis et al., 2013). Although some progress has been made in this area, improving sales performance, especially with respect to lead conversion rates, remains a challenge (Ohiomah et al., 2020). This is

particularly true in B2B inside sales, where there is a lack of knowledge about the factors that influence a buyer's purchase intention at each phase of the purchase decision-making process.

This study addresses this gap by developing and empirically validating a conceptual BIP model that integrates theories of the B2B purchase decision-making process and buying behavior to capture the drivers and enablers of a buyer's purchase intention. Our analysis of data collected from a survey using PLS-SEM revealed that buyers' trust in sellers and connection flexibility are critical factors that affect the purchase intention via appointment and contact willingness across the need recognition, information search, and evaluation and selection phases. Furthermore, our multigroup and profile analyses revealed various relationships between seller trust, connection flexibility, mediating variables, and purchase intention within distinct groups during each phase.

We provide a model that explains the relationships between variables in the phased purchase decision-making process and offer several strategic tips that sales practitioners can use to improve their inside sales success. By focusing on the buyer's perspective and understanding the factors that influence their purchase intention, sellers can tailor their selling strategies and improve their sales performance, especially in B2B inside sales.

Acknowledgment: this work was supported by grants from Mitacs Canada (No. 310636).

Table 3.1. Descriptions of constructs and indicators

Constructs	Indicators	Measures/Definitions	Scales	Sources
Connection flexibility (CF)	CF1	Contact channel preference (CF1) measures a buyer's preference on a channel to obtain information. Through which of the following channels do you prefer to be contacted by a salesperson? Options: phone, email, social media, SMS, and video.	Five-point scale: 1 = "do not prefer" & 5 = "prefer a great deal".	(Ohiomah et al., 2019)
	CF2	Contact frequency preference (CF2) refers to the number of times a buyer wants to be contacted. How many times do you prefer to be contacted by a salesperson if you do not answer the first contact attempt?	Numerical scale: 0 - 30	(Ohiomah et al., 2018)
	CF3	Information sources influence (CF3) refers to the influence of information sources on a buyer's connectivity. How would you assess the influence of the following information sources? Options: word of mouth, media appearances, social media postings, paid search advertising, organic search results, review sites, exhibitions & trade shows, contact by a salesperson, and potential users of	Five-point scale: 1 = "very weak" & 5 = "very strong".	(Sheth, 1973)

		the product/service in your organization.		
	CF4	Contact day preference (CF4) refers to a buyer's preferred weekdays to be contacted. Which of the following days do you prefer to be contacted? Options: Monday, Tuesday, Wednesday, Thursday, and Friday.	Five-point scale: 1 = "do not prefer" & 5 = "prefer a great deal".	(Ohiomah et al., 2018)
Seller Trust (ST)	ST1	Usually, how easy is it for you to make a buying decision for your organization considering the credibility and reliability of the sellers/vendors you are dealing with?	Five-point scale: 1 = "very hard" & 5 = "very easy".	(Halim & Karsen, 2020; Monat, 2011)
	ST2	Usually, how fast do you make a buying decision for your organization considering the credibility and reliability of the sellers/vendors you are dealing with?	Five-point scale: 1 = "very slow" & 5 = "very fast".	(Halim & Karsen, 2020; Monat, 2011)
	ST3	How would you assess your previous buying process experience with the sellers/vendors?	Five-point scale: 1 = "very hard" & 5 = "very easy".	(Monat, 2011; Zolkiewski et al., 2017)
	ST4	How would you assess your overall satisfaction with your past buying process of products/services offered by sellers/vendors?	Five-point scale: 1 = "extremely dissatisfied" & 5 = "extremely satisfied".	(Dickson et al., 2009; Homburg et al., 2011)
Willingness to make an appointment (WA)	WA	How likely would you make an appointment with a salesperson to learn more about the product/service?	Five-point scale: 1 = "extremely unlikely" & 5 = "extremely likely".	(Sabnis et al., 2013; Smith et al., 2006)
Willingness to be Contacted (WC)	WC	How likely would you want to be contacted by a salesperson?	Five-point scale: 1 = "extremely unlikely" & 5 = "extremely likely".	(Jolson, 1988; Monat, 2011)
Intention to Purchase (IP)	IP	How likely are you to make a purchase decision?	Five-point scale: 1 = "extremely unlikely" & 5 = "extremely likely".	(Monat, 2011)

Notes: indicators CF1 – 4, WA, WC, and IP in each stage were measured by asking questions with slightly different phrases to depict the presumed situation that a B2B buyer is currently in. Specifically, in the need recognition phase (NRP), a B2B buyer needs to assume that s/he is learning about a new need for the organization when answering questions (CF1-4). To answer questions related to WA, WC, and IP in the NRP, the respondent needs to assume the situation where the need has been recognized (end of NRP), before searching for information on products/services that can satisfy the need (i.e., before entering the information search phase).

In the information search phase (ISP), a B2B buyer needs to assume that s/he has already recognized a need and is now searching for information on products/services that can satisfy the need when answering questions related to CF1-4. For answering questions related to WA, WC, and IP in the ISP, the respondent needs to assume that s/he has obtained information about several products/services (end of ISP) but yet to consider any of them (i.e., before entering the evaluation and selection phase).

In the evaluation and selection phase (ESP), a B2B buyer needs to assume that s/he has already decided on the type of product/service that the organization needs and is now considering and comparing specific brands, vendors, promotions, and conditions when answering questions related to CF1-4. To answer questions related to WA, WC, and IP in the ESP, the respondent needs to assume that s/he has considered and compared several vendors' offers (end of ESP) and is ready to make a purchase decision (i.e., before entering the purchase phase).

Table 3.2. Socio-demographics and firmographics of survey participants

Demographics of participants	Count	Percentage
Employment Level		
Executive-level employment	40	31.74%
Middle-level employment	70	55.56%
Operational-level employment	16	12.70%
Sex		
Male	80	63.49%
Female	41	32.54%
Prefer not to say	5	3.97%
Seniority		
Youth (18-34 years old)	45	35.72%
Junior (35-44 years old)	40	31.74%
Senior (44+ years old)	41	32.54%
Education		
Below Bachelor's Degree	26	20.63%
Bachelor's Degree	74	58.74%
Master's Degree	26	20.63%
Company size		
Small-sized (1-99 employees)	57	45.24%
Medium-sized (100-499 employees)	22	17.46%
Large-sized (500+ employees)	47	37.30%
Industry		
Business & Professional Services	57	45.24%
IT Services	33	26.19%
Healthcare	13	10.32%
Manufacturing	5	3.97%
Financial Services	4	3.17%
Others (including Education Telecom, Construction, etc.)	14	11.10%
Total	**126**	**100%**

Table 3.3. Reliability and validity of seller trust

Construct	Convergent validity	Internal consistency reliability		Indicator reliability			
	AVE	CR	CA	Indicator	Loadings	t-Stat	p-value
ST (NRP)	0.641	0.877	0.816	ST1	0.849	2.732	0.006
				ST2	0.822	2.394	0.017
				ST3	0.800	2.648	0.008
				ST4	0.727	2.270	0.023
ST (ISP)	0.640	0.876	0.814	ST1	0.808	3.542	0.000
				ST2	0.758	2.892	0.004
				ST3	0.830	3.776	0.000
				ST4	0.809	3.426	0.001
ST (ESP)	0.637	0.875	0.814	ST1	0.839	2.991	0.003
				ST2	0.779	2.657	0.008
				ST3	0.779	2.971	0.003
				ST4	0.803	2.775	0.006

ST: Seller Trust; NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase; AVE: Average variance extracted; CR: Composite reliability; CA: Cronbach's alpha

Table 3.4. Fornell-Larcker criterion for the three models/phases

Construct	NRP				ISP				ESP			
	IP	ST	WA	WC	IP	ST	WA	WC	IP	ST	WA	WC
IP	1.000				1.000				1.000			
ST	0.056	**0.801**			0.107	**0.800**			0.080	**0.798**		
WA	0.373	0.108	1.000		0.223	0.110	1.000		0.602	0.047	1.000	
WC	0.270	0.110	0.663	1.000	0.290	0.087	0.682	1.000	0.567	0.135	0.683	1.00

IP: Intention to purchase; ST: Seller trust; WA: Willingness to make an appointment; WC: Willingness to be contacted;
NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase

Table 3.5. Heterotrait-monotrait ratio test for the three models/phases

Construct	NRP			ISP			ESP		
	IP	ST	WA	IP	ST	WA	IP	ST	WA
ST	0.112			0.123			0.083		
WA	0.373	0.115		0.223	0.115		0.602	0.058	
WC	0.270	0.119	0.663	0.290	0.095	0.682	0.567	0.140	0.683

IP: Intention to purchase; ST: Seller trust; WA: Willingness to make an appointment; WC: Willingness to be contacted;
NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase

Table 3.6. Multicollinearity and indicator validity of formative construct: connection flexibility

Construct	Indicator	Weight	STD	t- Stat	p-value	VIF
CF (NRP)	CF1(NRP)	0.520***	0.063	8.215	0.000	1.511
	CF2(NRP)	0.328***	0.088	3.732	0.000	1.074
	CF3(NRP)	0.174*	0.100	1.745	0.081	1.220
	CF4(NRP)	0.383***	0.066	5.765	0.000	1.343
CF (ISP)	CF1(ISP)	0.735***	0.205	3.591	0.000	1.511
	CF2(ISP)	–	–	–	–	1.074
	CF3(ISP)	–	–	–	–	1.220
	CF4(ISP)	–	–	–	–	1.343
CF (ESP)	CF1(ESP)	–	–	–	–	1.379
	CF2(ESP)	–	–	–	–	1.105
	CF3(ESP)	0.630***	0.147	4.285	0.000	1.236
	CF4(ESP)	0.421***	0.156	2.697	0.007	1.290

***: 1% significance, **: 5% significance, *: 10% significance, and -: Non-significant;
STD: Standard deviation; VIF: Variance inflation factor;
CF: Connection flexibility; NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase

Table 3.7. Significance of hypothesized path relationships in the three models/phases

Hypothesis	NRP						ISP						ESP					
	Direct effect		Indirect effect		Total effect		Direct effect		Indirect effect		Total effect		Direct effect		Indirect effect		Total effect	
	Path	t-Stat	Path	t-Stat	Path	t-Stat	Path	t-Stat	Path	t-Stat	Path	t-Stat	Path	t-Stat	Path	t-Stat	Path	t-Stat
H1. ST → IP	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H2. CF → IP	–	–	0.121**	2.377	–	–	–	–	0.116**	2.071	–	–	–	–	0.292***	4.700	0.308***	3.273
H3. WC → IP	–	–	–	–	–	–	0.271**	2.140	–	–	0.271**	2.140	0.282**	2.285	–	–	0.282**	2.285
H4. CF → WC	0.457***	7.004	–	–	0.457***	7.004	0.364***	4.369	–	–	0.364***	4.369	0.428***	5.811	–	–	0.428***	5.811

H5. ST → WC	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H6. WA → IP	0.345***	3.318	–	–	0.345***	3.318	–	–	–	–	–	–	0.399***	3.087	–	–	0.399***	3.087
H7. CF → WA	0.317***	4.407	–	–	0.317***	4.407	0.379***	4.872	–	–	0.379***	4.872	0.431***	5.904	–	–	0.431***	5.904
H8. ST → WA	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

***: 1% significance, **: 5% significance, *: 10% significance, and -: Non-significant;
IP: Intention to purchase; CF: Connection flexibility; ST: Seller trust; WA: Willingness to make an appointment; WC: Willingness to be contacted

Table 3.8. Mediation analysis

Indirect path	Original (O)	t-Stat	CI lower (5%)	CI upper (95%)
H12 (NRP): CF → WA x WA → IP	0.110**	2.313	0.037	0.224
H10 (NRP): CF → WC x WC → IP	–	–	–	–
H12 (ISP): CF → WA x WA → IP	–	–	–	–
H10 (ISP): CF → WC x WC → IP	0.099*	1.865	0.010	0.216
H12 (ESP): CF → WA x WA → IP	0.172***	2.822	0.060	0.298
H10 (ESP): CF → WC x WC → IP	0.121**	1.990	0.013	0.247

***: 1% significance, **: 5% significance, *: 10% significance, and -: Non-significant; CI: Confidence interval;
NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase;
IP: Intention to purchase; WA: Willingness to make an appointment; WC: Willingness to be contacted

Table 3.9. Moderation analysis

Relationships	Original (O)	Standard Deviation	T-stat
Moderating Effect (Phase x WC) -> IP	0.114*	0.060	1.914
Phase -> IP	0.323***	0.044	7.320
WC -> IP	0.205***	0.066	3.126
Moderating Effect (Phase x CF) -> WA	-0.009	0.046	0.200
Moderating Effect (Phase x CF) -> WC	-0.058	0.046	1.256
Moderating Effect (Phase x WA) -> IP	-0.001	0.062	0.023
Moderating Effect (Phase x CF) -> IP	0.010	0.048	0.206
Moderating Effect (Phase x ST) -> IP	-0.004	0.056	0.077
Moderating Effect (Phase x ST) -> WA	-0.025	0.053	0.464
Moderating Effect (Phase x ST) -> WC	0.014	0.049	0.294

***: 1% significance, **: 5% significance, *: 10% significance, and -: Non-significant
IP: Intention to purchase; CF: Connection flexibility; ST: Seller trust; WA: Willingness to make an appointment; WC: Willingness to be contacted.

Table 3.10. Effect size (f^2) statistics

Hypothesis	NRP		ISP		ESP	
	f^2	Effect	f^2	Effect	f^2	Effect
H1. ST → IP	0.000	–	0.007	–	0.001	–
H2. CF → IP	0.001	–	0.000	–	0.001	–
H3. WC → IP	0.000	–	0.038	Small	0.067	Small
H4. CF → WC	0.265	Medium	0.138	Small	0.220	Medium
H5. ST → WC	0.006	–	0.001	–	0.010	–
H6. WA → IP	0.077	Small	0.001	–	0.137	Small
H7. CF → WA	0.112	Small	0.161	Medium	0.215	Medium
H8. ST → WA	0.007	–	0.003	–	0.000	–

IP: Intention to purchase; CF: Connection flexibility; ST: Seller trust; WA: Willingness to make an appointment; WC: Willingness to be contacted;
NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase

Table 3.11. Predictive relevance (Q^2) for the three path models/phases

Construct	NRP			ISP			ESP		
	SSO	SSE	Q^2	SSO	SSE	Q^2	SSO	SSE	Q^2
IP	126.00	118.37	0.06	126.00	124.23	0.01	126.00	85.12	0.32
WA	126.00	120.71	0.04	126.00	113.63	0.10	126.00	108.64	0.14
WC	126.00	103.49	0.18	126.00	114.09	0.09	126.00	109.15	0.13

IP: Intention to purchase; WA: Willingness to make an appointment; WC: Willingness to be contacted;
SSO: Sum of squared observations; SSE: Sum of squared errors; Q^2: (=1-SSE/SSO);
NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase

Table 3.12. Common method bias evaluation: full collinearity variance inflation factors

	Variance Inflation Factors		
Constructs	**IP**	**WTA**	**WTC**
CF	1.511	1.023	1.023
ST	1.031	1.023	1.023
WTA	2.193	–	–
WTC	2.075	–	–

IP: Intention to purchase; CF: Connection flexibility; ST: Seller trust; WA: Willingness to make an appointment; WC: Willingness to be contacted

Table 3.13. Results of control variables

EVs ← CVs		NRP		ISP		ESP	
		Original (O)	T-Stat	Original (O)	T-Stat	Original (O)	T-Stat
IP	Employment level	–	–	0.344**	1.991	–	–
	Seniority	–	–	0.387**	2.014	0.633*	1.825
	Sex	–	–	–	–	–	–
	Company size	1.094*	1.944	0.819*	1.940	-0.726*	1.870
	Education	0.496*	1.953	0.371**	1.966	0.201*	1.927
WA	Employment level	–	–	–	–	–	–
	Seniority	–	–	–	–	–	–
	Sex	–	–	0.339*	1.933	–	–
	Company size	-0.750*	1.780	-0.995***	2.502	–	–
	Education	–	–	–	–	–	–
WC	Employment level	0.250**	1.961	–	–	–	–
	Seniority	–	–	–	–	0.602*	1.768
	Sex	–	–	–	–	0.428**	1.961
	Company size	–	–	-0.892***	2.543	–	–
	Education	–	–	–	–	–	–

***: 1% significance, **: 5% significance, *: 10% significance, and -: Non-significant; EV: Endogenous variables; CV: Control variables
NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase;
IP: Intention to purchase; WA: Willingness to make an appointment; WC: Willingness to be contacted

Table 3.14. Power analysis on sample sizes across phases

Phases	Groups	Observed R^2	Probability Level	Sample Size	Observed Statistical Power
NRP	Original (no group split)	0.141	0.05	126	0.963
	Executive Group	0.280	0.05	40	0.857
	Non-Executive Group	0.186	0.05	86	0.950
	Senior Group	0.372	0.05	41	0.974
	Non-Senior Group	0.132	0.05	85	0.813
	Male Group	0.147	0.05	80	0.838
	Female Group	0.295	0.05	41	0.893
	SME-sized Group	0.154	0.05	79	0.854
	Large-sized Group	0.217	0.05	47	0.817
	Graduate Group	0.259	0.05	26	0.563
	Below-Graduate Group	0.170	0.05	100	0.960
ISP	Original (no group split)	0.105	0.05	126	0.874
	Executive Group	0.257	0.05	40	0.810
	Non-Executive Group	0.138	0.05	86	0.840
	Senior Group	0.314	0.05	41	0.921
	Non-Senior Group	0.136	0.05	85	0.830
	Male Group	0.141	0.05	80	0.820
	Female Group	0.354	0.05	41	0.962
	SME-sized Group	0.149	0.05	79	0.840
	Large-sized Group	0.288	0.05	47	0.930
	Graduate Group	0.364	0.05	26	0.801
	Below-Graduate Group	0.131	0.05	100	0.877
ESP	Original (no group split)	0.404	0.05	126	0.999
	Executive Group	0.586	0.05	40	0.999
	Non-Executive Group	0.370	0.05	86	0.999
	Senior Group	0.257	0.05	41	0.821
	Non-Senior Group	0.527	0.05	85	0.999
	Male Group	0.460	0.05	80	0.999
	Female Group	0.423	0.05	41	0.992
	SME-sized Group	0.390	0.05	79	0.999
	Large-sized Group	0.579	0.05	47	0.999
	Graduate Group	0.542	0.05	26	0.988
	Below-Graduate Group	0.429	0.05	100	0.999

Table 3.15. Employment level: executive-level vs. non-executive level

Predicted ← Predictor		NRP			ISP			ESP		
		1 Exec	2 Non-Exec	Diff. (1-2)	1 Exec	2 Non-Exec	Diff. (1-2)	1 Exec	2 Non-Exec	Diff. (1-2)
IP	CF	–	–	–	–	–	–	–	–	–
	ST	–	–	–	–	0.205*	–	–	–	–
	WA	0.464**	0.247*	–	–	–	–	0.433**	0.428**	–
	WC	–	–	–	0.448*	–	–	–	–	–
WA	CF	0.562***	0.310***	–	0.618***	0.325***	0.292*	0.545***	0.388***	–
	ST	–	0.237**	-0.237**	–	–	–	–	–	–
WC	CF	0.649***	0.365***	0.283*	0.552***	0.345***	–	0.576***	0.438***	–
	ST	–	0.197**	-0.197**	–	–	–	–	–	–

***: 1% significance, **: 5% significance, *: 10% significance, and -: Non-significant; Diff: Difference; NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase

IP: Intention to purchase; ST: Seller trust; WA: Willingness to make an appointment; WC: Willingness to be contacted.

Table 3.16. Seniority: senior vs. non-senior

Predicted ← Predictor		NRP			ISP			ESP		
		1 Senior	2 Non-Senior	Diff. (1-2)	1 Senior	2 Non-Senior	Diff. (1-2)	1 Senior	2 Non-Senior	Diff. (1-2)
IP	CF	–	–	–	–	–	–	–	–	–
	ST	–	–	–	–	–	–	–	–	–
	WA	0.667***	0.215*	0.452**	0.372*	–	–	–	0.502***	–
	WC	–	–	–	–	0.339*	–	–	0.297**	–
WA	CF	0.397**	0.316***	–	–	0.413***	–	0.389*	0.507***	–
	ST	–	–	–	–	–	–	–	–	–
WC	CF	0.541***	0.448***	–	–	0.420***	–	0.378**	0.471***	–
	ST	–	–	–	–	–	–	0.373***	–	–

***: 1% significance, **: 5% significance, *: 10% significance, and -: Non-significant; Diff: Difference;
NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase
IP: Intention to purchase; ST: Seller trust; WA: Willingness to make an appointment; WC: Willingness to be contacted.

Table 3.17. Education: graduate vs. below graduate

Predicted ← Predictor		NRP			ISP			ESP		
		1 Graduate	2 Below Graduate	Diff. (1-2)	1 Graduate	2 Below Graduate	Diff. (1-2)	1 Graduate	2 Below Graduate	Diff. (1-2)
IP	CF	–	–	–	–	–	–	0.562**	–	0.562**
	ST	–	–	–	0.588**	–	0.588**	–	–	–
	WA	–	0.329***	–	0.513*	–	0.513*	–	0.396**	–
	WC	–	–	–	–	0.487***	-0.487***	–	0.34**	–
WA	CF	–	0.388***	–	0.649***	0.344***	–	0.629***	0.395***	–
	ST	–	–	–	-0.382**	–	-0.382**	-0.345*	–	–
WC	CF	0.538**	0.481***	–	–	0.472***	–	0.655***	0.358***	–
	ST	–	–	–	–	–	–	–	–	–

***: 1% significance, **: 5% significance, *: 10% significance, and -: Non-significant; Diff: Difference;
NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase
IP: Intention to purchase; ST: Seller trust; WA: Willingness to make an appointment; WC: Willingness to be contacted.

Table 3.18. Sex: male vs. female

Predicted ← Predictor		NRP			ISP			ESP		
		1 Male	2 Female	Diff. (1-2)	1 Male	2 Female	Diff. (1-2)	1 Male	2 Female	Diff. (1-2)
IP	CF	–	–	–	–	–	–	0.187*	–	0.187*
	ST	–	–	–	–	–	–	–	–	–
	WA	0.225*	0.591***	–	–	0.498**	-0.498**	–	0.482***	–
	WC	–	–	–	–	–	–	0.366*	0.391**	–
WA	CF	0.285***	0.494***	–	0.377***	–	–	0.422***	0.479***	–
	ST	–	–	–	–	–	–	–	–	–
WC	CF	0.485***	0.551***	–	0.360***	–	–	0.418***	0.408***	–
	ST	–	–	–	–	–	–	0.257***	–	0.257***

***: 1% significance, **: 5% significance, *: 10% significance, and -: Non-significant; Diff: Difference;
NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase
IP: Intention to purchase; ST: Seller trust; WA: Willingness to make an appointment; WC: Willingness to be contacted.

Table 3.19. Company size: SME vs. large.

Predicted ← Predictor		NRP			ISP			ESP		
		1 SME	2 Large	Diff. (1-2)	1 SME	2 Large	Diff. (1-2)	1 SME	2 Large	Diff. (1-2)
IP	CF	–	–	–	–	–	–	–	–	–
	ST	–	–	–	–	0.377**	-0.377**	–	–	–
	WA	0.385***	–	–	–	–	–	0.417**	0.298*	–
	WC	–	–	–	0.429***	–	–	–	0.495***	–
WA	CF	0.365***	–	–	0.459***	–	–	0.399***	0.447***	–
	ST	–	–	–	–	–	–	–	–	–
WC	CF	0.490***	0.417*	–	0.416***	–	–	0.378***	0.531***	–
	ST	–	–	–	–	–	–	–	–	–

***: 1% significance, **: 5% significance, *: 10% significance, and -: Non-significant; Diff: Difference;
NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase
IP: Intention to purchase; ST: Seller trust; WA: Willingness to make an appointment; WC: Willingness to be contacted.

Table 3.20. Profiles of online survey participants

Profile #	Profile Name	Size	Employment Level	Seniority	Education Level	Sex	Company Size
1	John	40	Executive	Senior	Graduate	Male	SME
2	Lara	40	Non-Executive	Non-Senior	Below Graduate	Female	Large
3	Samuel	46	Non-Executive	Non-Senior	Below Graduate	Male	SME

Table 3.21. Profiles' comparison: profile #1 John vs. profile #2 Lara

Predicted ← Predictor		NRP			ISP			ESP		
		1 John	2 Lara	Diff. (1-2)	1 John	2 Lara	Diff. (1-2)	1 John	2 Lara	Diff. (1-2)
IP	CF	–	–	–	–	-0.433*	0.433***	0.495**	–	0.495***
	ST	–	–	–	–	–	–	–	–	–
	WA	0.385**	0.736***	-0.351***	–	–	–	–	0.605***	-0.605***
	WC	–	–	–	–	0.372**	-0.372***	–	–	–
WA	CF	0.387***	0.471***	-0.084***	0.489**	0.521***	–	0.447**	0.549***	-0.102***
	ST	–	–	–	–	–	–	–	–	–
WC	CF	0.485***	0.592***	-0.107***	0.402**	0.520***	-0.118***	0.528***	0.483***	–
	ST	–	–	–	–	–	–	–	–	–

***: 1% significance, **: 5% significance, *: 10% significance, and -: Non-significant; Diff: Difference;
NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase
IP: Intention to purchase; ST: Seller trust; WA: Willingness to make an appointment; WC: Willingness to be contacted.

Table 3.22. Profiles' comparison: profile #1 John vs. profile #3 Samuel

Predicted ← Predictor		NRP			ISP			ESP		
		1 John	2 Samuel	Diff. (1-2)	1 John	2 Samuel	Diff. (1-2)	1 John	2 Samuel	Diff. (1-2)
IP	CF	–	–	–	–	–	–	0.495**	–	0.495***
	ST	–	–	–	–	0.345**	-0.345***	–	–	–
	WA	0.385**	–	0.385***	–	–	–	–	–	–

	WC	–	–	–	–	–	–	–	–	–
WA	CF	0.387***	–	0.387***	0.489**	–	0.489***	0.447**	0.412***	–
	ST	–	–	–	–	0.369**	-0.369***	–	0.219**	-0.219***
WC	CF	0.485***	0.477**	–	0.402**	0.263*	0.139***	0.528***	0.391***	0.137***
	ST	–	–	–	–	0.251*	-0.251***	–	0.307***	-0.307***

***: 1% significance, **: 5% significance, *: 10% significance, and -: Non-significant; Diff: Difference;
NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase
IP: Intention to purchase; ST: Seller trust; WA: Willingness to make an appointment; WC: Willingness to be contacted.

Table 3.23. Profiles' comparison: profile #2 Lara vs. profile #3 Samuel

Predicted ← Predictor		NRP			ISP			ESP		
		1 Lara	2 Samuel	Diff. (1-2)	1 Lara	2 Samuel	Diff. (1-2)	1 Lara	2 Samuel	Diff. (1-2)
IP	CF	–	–	–	-0.433*	–	-0.433***	–	–	–
	ST	–	–	–	–	0.345**	-0.345***	–	–	–
	WA	0.736***	–	0.736***	–	–	–	0.605***	–	0.605***
	WC	–	–	–	0.372**	–	0.372***	–	–	–
WA	CF	0.471***	–	0.471***	0.521***	–	0.521***	0.549***	0.412***	0.137***
	ST	–	–	–	–	0.369**	-0.369***	–	0.219**	-0.219***
WC	CF	0.592***	0.477**	0.115***	0.520***	0.263*	0.257***	0.483***	0.391***	0.092***
	ST	–	–	–	–	0.251*	-0.251***	–	0.307***	-0.307***

***: 1% significance, **: 5% significance, *: 10% significance, and -: Non-significant; Diff: Difference;
NRP: Need Recognition Phase; ISP: Information Search Phase; ESP: Evaluation and Selection Phase
IP: Intention to purchase; ST: Seller trust; WA: Willingness to make an appointment; WC: Willingness to be contacted.



Figure 3.1. Buyer's purchase decision-making process

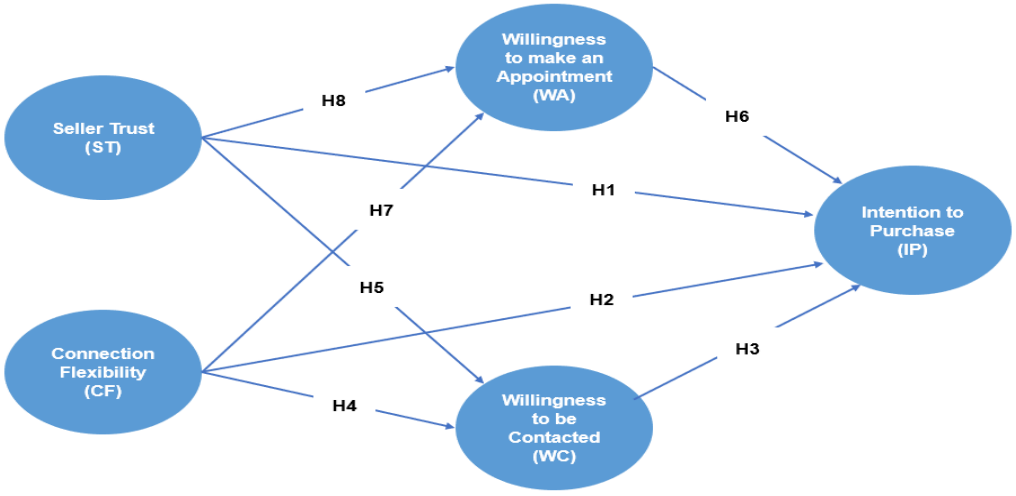

Figure 3.2. Buyer's Intention to Purchase (BIP) conceptual model

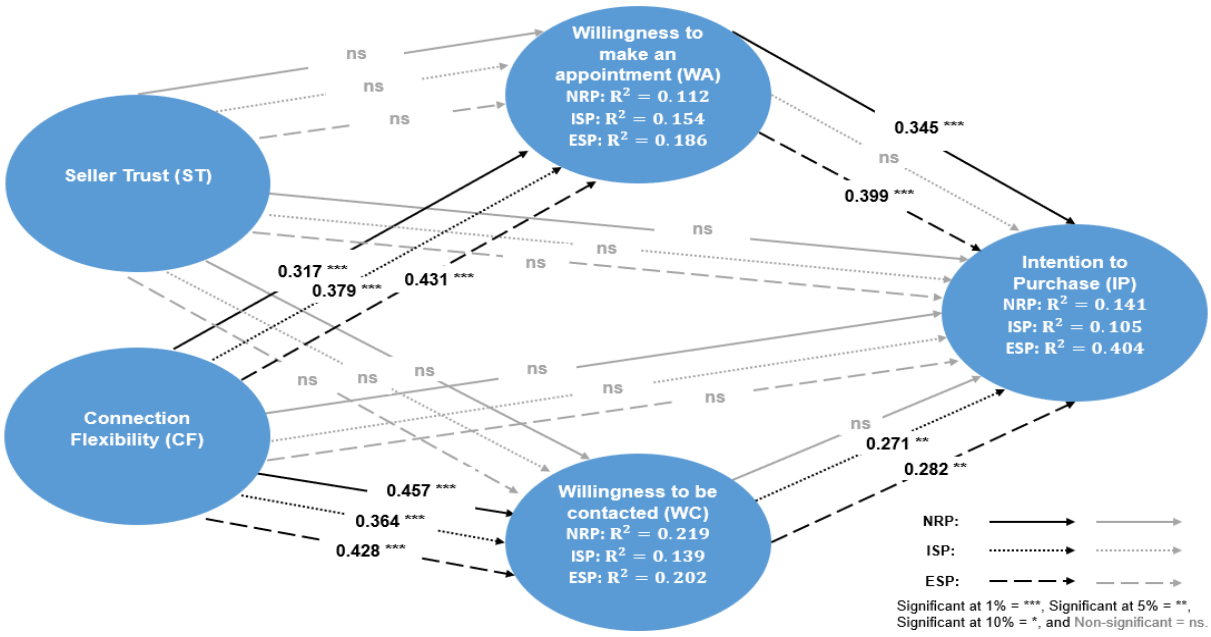

Figure 3.3. Structural model assessments for the three phases

References

Amoozad Mahdiraji, H., Tavana, M., Mahdiani, P., & Abbasi Kamardi, A. A. (2021). A multi-attribute data mining model for rule extraction and service operations benchmarking. *Benchmarking: An International Journal*, *29*(2), 456–495.

Andreev, P., Heart, T., Maoz, H., & Pliskin, N. (2009). Validating formative partial least squares (PLS) models: Methodological review and empirical illustration. *ICIS 2009 proceedings* (pp. 193).

Ba, S., & Pavlou, P. A. (2002). Evidence of the effect of Trust Building Technology in electronic markets: Price premiums and buyer behavior. *MIS Quarterly*, *26*(3), 243.

Baron, R.M., Kenny, D.A., 1986. The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182.

Bovée, & Arens, W. F. (1986). *Contemporary advertising (2nd ed.).* Irwin.

Chang, Y.-Y., Lin, S.-C., Yen, D. C., & Hung, J.-W. (2020). The trust model of enterprise purchasing for B2B e-marketplaces. *Computer Standards & Interfaces*, *70*, 103422.

Chang, J., O'Neill, G., & Travaglione, A. (2016). Demographic influences on employee trust towards managers. *International Journal of Organizational Analysis*, *24*(2), 246–260.

Chatterjee, S., Chaudhuri, R., & Vrontis, D. (2022). Examining the role of cross-cultural factors in the international market on Customer Engagement and purchase intention. *Journal of International Management*, *28*(3), 100966.

Chaudhuri, N., Gupta, G., Vamsi, V., & Bose, I. (2021). On the platform but will they buy? Predicting customers' purchase behavior using deep learning. *Decision Support Systems, 149*, 113622.

Chin, W. W. (1998). Commentary issues and opinion on structural equation modeling. *MIS Quarterly, 22*(1), Vii–Xvi.

Chiu, C.-M., Hsu, M.-H., Lai, H., & Chang, C.-M. (2012). Re-examining the influence of trust on online repeat purchase intention: The moderating role of habit and its antecedents. *Decision Support Systems, 53*(4), 835–845.

Chorianopoulos, A. (2016). *Effective CRM using predictive analytics*. John Wiley & Sons Inc.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). *Vol.20*, Hillsdale, NJ.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334.

Crosby, L. A., Evans, K. R., & Cowles, D. (1990). Relationship Quality in Services Selling: An Interpersonal Influence Perspective. *Journal of Marketing*, *54*(3), 68-81.

Dwyer, F. R., Schurr, P. H., & Oh, S. (1987). Developing buyer-seller relationships. *Journal of Marketing, 51*(2), 11–27.

Falk, R. F., & Miller, N. B. (1992). A primer for soft modeling. Akron, OH: University of Akron Press.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39.

Gower, J. (1971). A General Coefficient of Similarity and Some of Its Properties. *Biometrics*, *27*(4), 857–871.

Hair, J. F., Jr., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM).* Sage Publications.

Hair, J. F., Jr., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2022). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R*. Springer.

Harmon, L. L., & Funk, D. (2014). *42 Rules for Building a High-velocity Inside Sales Team: Actionable Guide to Creating Inside Sales Teams that Deliver Quantum Results*: Super Star Press.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *New challenges to international marketing* (pp.277–319). Emerald Group Publishing Limited.

Homburg, C., Maller, M., & Klarmann, M. (2011). When Should the Customer Really Be King? On the Optimum Level of Salesperson Customer Orientation in Sales Encounters. *Journal of Marketing, 75*, 55-74.

Howard, J. A., & Sheth, J. N. (1969). *The theory of buyer behavior*. Wiley.

Huer, M. (2013). Three Myths of the "67 Percent" Statistic. Forrester. https://go.forrester.com/blogs/three-myths-of-the-67-percent-statt-Statistic [Accessed Feb. 22]

Jansen, B. J., & Schuster, S. (2011). Bidding on The Buying Funnel for Sponsored Search and Keyword Advertising. *Journal of Electronic Commerce Research, 12*(1), 1-18.

Johnston, W. J., & Lewin, J. E. (1996). Organizational buying behavior toward an integrative framework. *Journal of Business Research, 35*(1), 1–15.

Jolson, M. (1988). Qualifying sales leads: The tight and loose approaches. *Industrial Marketing Management*, *17*(3), 189–196.

Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of e-Collaboration, 11*(4), 1-10.

Kock, N., & Lynn, G.S. (2012). Lateral collinearity and misleading results in variance-based SEM: An illustration and recommendations. *Journal of the Association for Information Systems*, 13(7), 546-580.

Lemon, & Verhoef, P. C. (2016). Understanding Customer Experience Throughout the Customer Journey. *Journal of Marketing, 80*(6), 69–96.

Lowry, P. B., & Gaskin, J. (2014). Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it. *IEEE Transactions on Professional Communication, 57*(2), 123–146.

Lu, X., He, S., Lian, S., Ba, S., & Wu, J. (2020). Is user-generated content always helpful? The effects of online forum browsing on consumers' travel purchase decisions. *Decision Support Systems, 137*, 113368.

Maycock, M., Spenner, P., Hamshar, R., West, S., Brown, T., Daniels, N., & Choudhury, N. V. (2012). The Digital Evolution in B2B Marketing. *CEB Marketing Leadership Council.* 46.

Menidjel, C., & Bilgihan, A. (2022). The determinants of retail customers' purchase intent. *International Journal of Consumer Studies (2022,* in press*)*.

Molitor, D., Ghose, A., Spann, M., & Reichhart, P. (2018). Measuring the effectiveness of location-based mobile push vs. pull targeting. Paper presented at the *International Conference on Information Systems 2018, ICIS 2018*

Monat, J. P. (2011). Industrial sales lead conversion modeling. *Marketing Intelligence & Planning, 29*(2), 178–194.

Moorman, C., Deshpande, R., & Zaltman, G. (1993). Factors Affecting Trust in Market Research Relationships. *Journal of Marketing*, *57*(1), 81-101.

Morgan, R. M., & Hunt, S. D. (1994). The Commitment-Trust Theory of Relationship Marketing. *Journal of Marketing*, *58*(3), 20-38.

Neslin, Grewal, D., Leghorn, R., Shankar, V., Teerling, M. L., Thomas, J. S., & Verhoef, P. C. (2006). Challenges and Opportunities in Multichannel Customer Management. *Journal of Service Research, 9(*2), 95–112.

Ohiomah, A., Benyoucef, M., & Andreev, P. (2020). A multidimensional perspective of business-to-business sales success: A meta-analytic review. *Industrial Marketing Management*, *90*, 435–452.

Ohiomah, A., Benyoucef, M., Andreev, P., Kuziemsky, C., Hood, D., & Le Bon, J. (2018). Best Practices for Inside Sales Professionals: An Historical Analysis. *Developments in Marketing Science: Proceedings of the Academy of Marketing Science*, 449–461.

Ohiomah, A., Andreev, P., Benyoucef, M., & Hood, D. (2019). The role of lead management systems in inside sales performance. *Journal of Business Research, 102*, 163–177.

Palmatier, R. W., Dant, R. P., Grewal, D., & Evans, K. R. (2006). Factors influencing the effectiveness of relationship marketing: A meta-analysis. *Journal of Marketing, 70*(4), 136–153.

Patma, T. S., Wardana, L. W., Wibowo, A., Narmaditya, B. S., Akbarina, F., & Awan, U. (2021). The impact of social media marketing for Indonesian SMEs sustainability: Lesson from Covid-19 pandemic. Cogent Business & Management, 8(1), 1953679.

Peña-García, N., Gil-Saura, I., Rodríguez-Orejuela, A., & Siqueira-Junior, J. R. (2020). Purchase intention and purchase behavior online: A cross-cultural approach. *Heliyon*, 6(6), e04284.

Pöyry, E., Parvinen, P., & McFarland, R. G. (2017). Generating leads with sequential persuasion: Should sales influence tactics be consistent or complementary? *Journal of Personal Selling & Sales Management*, *37*(2), 89–99.

Puccinelli, Goodstein, R. C., Grewal, D., Price, R., Raghubir, P., & Stewart, D. (2009). Customer Experience Management in Retailing: Understanding the Buying Process. *Journal of Retailing, 85*(1), 15–30.

Quigley, C., Bingham, F., & Patterson, M. (1993). The Information Flow for a Business to Business Buying Decision Process: A Modeling Approach. *Journal of Marketing Theory and Practice*, *2*(1), 103–121.

Rapp, A., Ahearne, M., Mathieu, J., & Schillewaert, N. (2006). The impact of knowledge and empowerment on working smart and working hard: The moderating role of experience. *International Journal of Research in Marketing, 23*, 279-293.

Rapp, A., Beitelspacher, L. S., Schillewaert, N., & Baker, T. L. (2012). The differing effects of technology on inside vs. Outside sales forces to facilitate enhanced customer orientation and interfunctional coordination. *Journal of Business Research*, *65*(7), 929–936.

Reynolds, A., Richards, G., de la Iglesia, B., & Rayward-Smith, V. (2006). Clustering Rules: A Comparison of Partitioning and Hierarchical Clustering Algorithms. *Journal of Mathematical Modelling and Algorithms*, *5*(4), 475–504.

Robinson, P.J., Faris, C.W., & Wind, Y. (1967). *Industrial Buying and Creative Marketing*. Boston, MA: Allyn and Bacon.

Rungtusanatham, M., Miller, J. W., & Boyer, K. K. (2014). Theorizing, testing, and concluding for mediation in SCM Research: Tutorial and procedural recommendations. *Journal of Operations Management, 32*(3), 99–113.

Sabnis, G., Chatterjee, S. C., Grewal, R., & Lilien, G. L. (2013). The Sales Lead Black Hole: On Sales Reps' Follow-Up of Marketing Leads. *Journal of Marketing, 77*(1), 52–67.

Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An integrative model of Organizational Trust: Past, Present, and future. *Academy of Management Review, 32*(2), 344–354.

Seley, A., & Holloway, B. (2008). Sales 2.0: *Improve business results using innovative sales practices and technology.* John Wiley & Sons.

Sengupta, S., Krapfel, R. E., & Pusateri, M. A. (2000). An empirical investigation of key account salesperson effectiveness. *The Journal of Personal Selling and Sales, 20*, 253-261.

Sheth, J. (1973). A Model of Industrial Buyer Behaviour. *Journal of Marketing*, *37*(4), 50–56.

Sheth, J., Sharma, A., & Iyer, G. (2009). Why integrating purchasing with marketing is both inevitable and beneficial. *Industrial Marketing Management*, *38*(8), 865–871.

Sheth, J., Sisodia, R., & Sharma, A. (2000). The antecedents and consequences of customer-centric marketing. *Journal of the Academy of Marketing Science*, *28*(1), 55–66.

Sirakaya, E., & Woodside, A. (2005). Building and testing theories of decision making by travellers. *Tourism Management (1982)*, *26*(6), 815–832.

Smith, T. M., Gopalakrishna, S., & Chatterjee, R. (2006). A Three-Stage Model of Integrated Marketing Communications at the Marketing-Sales Interface. *Journal of Marketing Research, 43*(4), 564–579.

Steward, Narus, J. A., Roehm, M. L., & Ritz, W. (2019). From transactions to journeys and beyond: The evolution of B2B buying process modeling. *Industrial Marketing Management, 83*, 288–300.

Unni, R., & Harmon, R. (2007). Perceived effectiveness of push vs. pull mobile location based advertising. *Journal of Interactive Advertising*, *7*(2), 28–40.

Visser, P. S., Krosnick, J. A., Marquette, J., & Curtin, M. (1996). Mail surveys for election forecasting? An evaluation of the Columbus Dispatch poll. *Public Opinion Quarterly, 60*(2), 181–227.

Webster, F. E., & Wind, Y. (1972). A General Model for Understanding Organizational Buying Behaviour. *Journal of Marketing*, *36*(2), 12-19.

Zeithaml, V. A. (1988). Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence. *Journal of Marketing*, *52*(3), 2-22.

Zhang, P., Kim, S. & Chakravarty, A. (2022). Influence of pull marketing actions on marketing action effectiveness of multichannel firms: A meta-analysis. *Journal of the Academy of Marketing Science.* (2022, in press).

Zhang, K. Z. K., Zhao, S. J., Cheung, C. M. K., & Lee, M. K. O. (2014). Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model. *Decision Support Systems, 67*, 78–89.

Zheng, R., Li, Z., & Na, S. (2022). How customer engagement in the live-streaming affects purchase and customer acquisition. *Journal of Retailing and Consumer Services*, *68*, 103015.

Chapter 3 Key Points Summary

- B2B inside sales model explores factors influencing purchase intention in the phased purchase decision-making process;
- PLS-SEM validates relationships between seller trust, connection flexibility, and purchase intention;
- Demographic and firmographic attributes affect the purchase intention differently in each phase;
- Seller trust and connection flexibility drive appointment and contact willingness in the B2B inside sales purchase process.

Chapter 4. Smart Sales: Amplifying the Power of Predictive Lead Scoring in B2B Sales

Article Reference: Wu, M., Andreev, P., & Benyoucef, M. (2024). Smart Sales: Amplifying the Power of Predictive Lead Scoring in B2B Sales. *The Pacific Asia Conference on Information Systems (PACIS) 2024 Proceedings. 3.* https://aisel.aisnet.org/pacis2024/track09_digittrans/track09_digittrans/3

Article Abstract: In B2B sales, assessing and selecting leads to pursue have traditionally relied on sales practitioners' expertise. However, due to inherited limitations and subjectivity, this approach has been widely acknowledged as ineffective and inefficient for lead scoring/qualification. Therefore, scholars have shifted their focus to more objective data-driven approaches that could increase lead conversion rate and, eventually, sales performance. However, there is a lack of studies on lead scoring models that consider the leads' willingness and purchase intentions. Therefore, this study develops a "smart sales" predictive lead scoring model using a combination of both unsupervised and supervised machine learning methods and incorporates leads' purchase intentions from their perspective in the phased sales process. The empirical findings of this study claim that data-driven models can improve decision-making in the B2B sales process. Implementation of the proposed solution will help enhance the effectiveness and efficiency of the B2B sales process and its performance.

Keywords: Smart Sales; Lead Scoring; Machine Learning; Leads' Perspective; Data-driven Decision-making

Note: This chapter is an author accepted manuscript submitted to the *Pacific-Asia Conference on Information Systems, Ho Chi Minh City, Vietnam, 2024* referred to above.

4.1 Introduction

The evolution of information technology (IT) has fundamentally reshaped Business-to-Business (B2B) sales, establishing inside (remote) selling as the predominant method for connecting with leads (Rutherford et al., 2014; Thaichon et al., 2018). In the rapidly evolving B2B sales world, the "competition" bar has been set higher than ever due to the growing sophistication of technology and buyers' purchasing behavior (Karhunsalo, 2021). Companies need to find better ways to increase sales performance to be competitive, especially considering technological advancements (Paschen et al., 2020; Syam & Sharma, 2018). The critical stage of a B2B sales process is prospecting (Paschen et al., 2020), which involves the marketing and sales functions and consists of lead generation (i.e., finding leads who have the potential to be converted to customers) and lead scoring/qualification (i.e., evaluating leads' likelihood to be converted to customers). Lead scoring relies heavily on data collected and managed through various information systems (e.g., customer relationship management, lead management system, marketing automation platform). These systems enable the aggregation and analysis of data attributes such as customer behavior, demographics,

and purchase history, which are crucial for estimating a lead's score (Nygård & Mezei, 2020). Selecting and contacting the most promising leads from a vast pool has posed critical challenges in the lead scoring process (D'Haen et al., 2013; Ohiomah et al., 2019; Sabnis et al., 2013), primarily because identifying the attributes that distinguish high-quality leads is inherently difficult. Typically, the successful identification of these critical characteristics can occur only after a salesperson has initiated contact with a lead, thus limiting the proactive potential of the lead scoring process (Paschen et al., 2020).

Traditionally, the lead scoring process relied mainly on marketing and salespeople's judgment and intuition. This approach inherited personal biases, subjectivity, and misjudgment into the scoring process (D'Haen et al., 2016; Eitle & Buxmann, 2019) since it was primarily based on an individual's professional expertise, cognitive abilities, and accumulated personal experiences with past leads (Bohanec et al., 2017; D'Haen & Van den Poel, 2013; Gokhale & Joshi, 2018). Salespeople often argue that many leads generated by marketing lack the potential to result in sales, while marketing complains about salespeople's poor lead follow-up (Järvinen & Taiminen, 2016; Smith et al., 2006). Consequently, despite the strenuous efforts, salespeople often misallocate company resources by dedicating them to "cold" leads, resulting in fewer sales opportunities, reduced sales productivity, decreased organizational revenue, and increased customer acquisition costs (D'Haen & Van den Poel, 2013; Smith et al., 2006). The subjective experience/intuition-based approach for scoring and choosing leads has been criticized for its ineffectiveness and inefficiency (D'Haen & Van den Poel, 2013; Duncan & Elkan, 2015; Nygård & Mezei, 2020), resulting in poor connection between marketing and sales and low conversion rates, leading to decreased sales performance (Xu et al., 2017). On average, the conversion rates of prospects into qualified leads stand at around 10%, with only 1% to 5% of these leads eventually being converted into sales and resulting in customer acquisition (Coe, 2004; D'Haen & Van den Poel, 2013; Eitle & Buxmann, 2019).

Effective and efficient lead scoring is critical for sellers to improve lead conversion rates, expand their market share, explore new customer markets, and improve sales performance (Bradford et al., 2016; Eitle & Buxmann, 2019). In the last two decades, the B2B sales process has substantially transformed, primarily driven by advancements in digitalization, information and communications technology (ICT), and social shifts (Paschen et al., 2020; Syam & Sharma, 2018). An important factor behind this transformation is the increased usage of artificial intelligence (AI), robotics, and machine learning (ML)(Paschen et al., 2020). Advanced information systems that integrate ML and AI can enhance lead scoring by analyzing historical data and continuously improving the effectiveness of the scoring criteria. As a result, the traditional personal selling and sales functions gradually shifted towards AI-driven supported sales (Syam & Sharma, 2018). The key challenge is finding qualified leads smartly (i.e., technology-driven) and ultimately converting them into customers. Predictive lead scoring models have

emerged as a solution to this challenge by assessing the leads' likelihood of making a purchase through scoring, ranking, and prioritizing them for sales (Benhaddou & Leray, 2017; Duncan & Elkan, 2015; Nygård & Mezei, 2020). These models have the potential to streamline sales activities and enhance lead conversion rates while reducing the use of arbitrary rules by sales practitioners.

Existing studies on lead scoring models are mainly built on the foundation of the marketing-sales funnel (also known as the sales process) (e.g., D'Haen & Van den Poel, 2013; Duncan & Elkan, 2015; Eitle & Buxmann, 2019), which is seller-centric (**Figure 4.1**). The marketing-sales funnel is utilized to control the progression of leads throughout different phases of the sales process (Ang & Buttle, 2006; Patterson, 2007). The literature suggests that from a seller's perspective, a typical suspect goes through a series of phases, which include lead generation (i.e., from suspect to prospect), lead scoring (i.e., from prospect to market-qualified lead (MQL)), approach/presentation (i.e., from MQL to sales-qualified lead (SQL)), and exclusive management/close (i.e., from SQL to customer) (D'Haen & Van den Poel, 2013; Xu et al., 2017).

As shown in **Figure 4.1**, the marketing function deploys lead scoring models to generate lists of MQLs which are continuously sent to salespeople for action (i.e., contact). When salespeople are overwhelmed by the number of MQLs they need to contact, some may not be contacted at the best possible time. Research has shown that the best time to contact leads is a key factor in lead conversion (Ohiomah et al., 2018). If a salesperson misses the best time to contact a lead, the company may lose a genuine sales opportunity. In addition, constantly putting pressure on salespeople to contact leads forces them to get hold of leads continually via all possible means.

Figure 4.1 appears here

Wu et al. (2024) suggest that leads may not in favor of being constantly contacted during the process. Recognizing leads as the starting point of marketing activities and driver of the sales process is strong evidence that the marketing-sales funnel has become increasingly lead-centric (Shannahan et al., 2013; Sheth et al., 2000, 2009). Understanding leads' perspective, especially leads' purchase intention for sellers is paramount since it can customize communications and offers to align with the particular needs and preferences of leads (Peña-García et al., 2020). This personalized approach increases the likelihood of conversion and enhances customer experience. Additionally, incorporating purchase intention into lead scoring enables sellers to allocate resources more efficiently. They can focus efforts on leads who are most likely to be converted, thus optimizing the use of time and resources. Furthermore, considering leads' purchase intentions in lead scoring facilitates lead segmentation. This segmentation based on leads' characteristics allows for more targeted marketing campaigns and sales initiatives (e.g., early sales opportunities) to adapt to different groups (Jadli et al., 2022), therefore improving response rates and sales success. However, to the best of our knowledge, the literature lacks studies that consider the leads'

perspective (i.e., the extent of a lead's willingness to purchase from his/her perspective at a specific phase in the sales process) incorporated and implemented in a lead scoring model. Syam and Sharma (2018, p.141) called for research on how ML and AI can "*help develop systems that recommend customer engagement strategies based on analysis of past engagements and their success/failure and depending on salesperson, industry and customer characteristics*".

To fill this gap in the existing literature and respond to this research call, we propose and develop a "*smart sales*" predictive lead scoring model. This model demonstrates the suitability of AI for addressing the lead scoring issue during the prospecting stage of the sales process[1]. It considers leads' purchase intentions based on their perspective in the phased sales process. We define "smart sales" as a forward-thinking approach to sales management and lead prioritization that leverages the capabilities of advanced technologies and data-driven strategies (e.g., AI and ML) to optimize sales strategies and performance outcomes. Also, we verify that sales performance can be improved using data on leads' characteristics and their interactions with salespeople within a B2B legal service consulting industry.

This paper documents that AI-driven models can enhance decision-making during the sales process and contributes by developing a "smart sales" predictive lead scoring model demonstrating ML algorithms and data mining (DM) models' roles in transitioning from traditional subjective decision-making sales to AI-driven objective decision-making sales. From a practical perspective, our study offers a solution that helps sales practitioners generate a concise and high-quality list of leads to improve lead scoring, enhance the effectiveness of the sales process, reduce costs, and improve sales performance.

The structure of this paper is as follows. The background section briefly discusses the importance of lead scoring models and the transition from conventional to automated lead scoring in the B2B sales process. The methodology section outlines our data collection process and analysis method. We discuss the performance of our "smart sales" lead scoring model in the results section. In the conclusion section, we discuss our findings' theoretical and practical implications. Finally, we discuss the limitations of our study and suggest potential areas for future research.

4.2 Background

4.2.1 From traditional to predictive lead scoring models

During the prospecting stage in the B2B sales process, the objective is to identify potential customers (leads) and surface high-quality leads who are more likely to be converted into customers. Historically, marketers and salespeople have used various factors to decide which leads to pursue from a large pool (D'Haen et al., 2013). Typically, they have depended heavily on prior experience with leads, personal judgment, and professional competence for such decisions. Such practices could lead to the inefficient use of company

[1] "sales process" and "B2B sales process" are used interchangeably in this paper.

resources by focusing on less promising leads (D'Haen & Van den Poel, 2013; McDonnell, 2014; Monat, 2011). A superior method is lead scoring, which qualifies the quality of leads in queue-based Lead Management Systems (LMSs) (D'Haen & Van den Poel, 2013; Duncan & Elkan, 2015). Lead scoring prioritizes marketing and sales efforts toward leads that are more likely to convert into customers (Benhaddou & Leray, 2017; Duncan & Elkan, 2015; Nygård & Mezei, 2020), which is an essential task at the initial stage of a successful sales process. A lead scoring model is an IT tool that assesses the likelihood of leads making a purchase by ranking them on a scale to differentiate and prioritize them for sales efforts (Boogar, 2019; Duncan & Elkan, 2015; McDonnell, 2014). An effective lead scoring model requires close collaboration between two functions to define what constitutes a good lead. Companies that implement lead scoring models in their LMSs could potentially see up to a 70% increase in return on investment from lead generation compared to those that do not use them (Duncan & Elkan, 2015; McDonnell, 2014).

There are two types of lead scoring models: traditional and predictive (Duncan & Elkan, 2015; McDonnell, 2014). The goal of traditional models is to prioritize leads for sales and to develop scalable approaches for leads that meet minimum qualification requirements. Since the beginning of the 21st century, the Lamb or Spam and rules/points-based models (i.e., scorecard model) have been frequently utilized to score and quantify the quality of leads (Wu et al., 2023). The former filters out low-quality leads and surfaces relatively good ones by assessing leads' information (e.g., email domain, company size, and location) (Boogar, 2019). In the rules/points-based model, sales practitioners assign points to leads' demographic and behavioral traits according to specific rules (McDonnell, 2014). However, these rules are set based on the sellers' subjective experiences. The lead score is the outcome of a weighted function of these traits. Monat (2011) introduced a practical qualitative modeling tool designed to predict the likelihood of an industrial sales lead converting into a customer according to observable attributes of the lead. The author asserted that this is the first theoretical lead characterization model which includes a rules/points-based evaluation scorecard for assessing leads. It considers eight determinants consisting of sixteen manifest attributes. The traditional scorecard model has been used to calculate lead scores. An Analytical Hierarchy Process-based framework assists companies in ranking and prioritizing leads according to the importance levels of factors (Verma et al., 2016). The lead scorecard model allows companies to evaluate the critical factors for acquiring potential customers. Traditional lead scoring models in the sales process can be error prone because of the manual selection of rules and human intervention (Boogar, 2019; Duncan & Elkan, 2015). In this regard, some results produced by traditional lead scoring models may be biased and inaccurate. Furthermore, traditional lead scoring models may not be able to capture nonlinear relationships between leads' attributes. Finally, they rely on leads' behavioral information, which may not always be appropriate (Duncan & Elkan, 2015).

In traditional lead scoring, most decisions on which leads to pursue were made by humans based on their experience-driven predictions and intuitions. Such an approach has been criticized for its ineffectiveness and inefficiency (Wu et al., 2023). With the introduction of ICT, ML, and AI, decision-making is shifting from humans to machines (Syam & Sharma, 2018), and the lead scoring task is no exception. Predictive lead scoring models use data-driven predictive analytics to discover insights within "cold" leads data, discover hidden relationships between predictors and outcomes, and estimate a score for each new lead (Chorianopoulos, 2016; Nygård & Mezei, 2020). Decisions on which leads to pursue can now be made by data-driven predictive ML models, which constitute an essential element in the transition from human-based to machine-based decision-making. In the existing literature, predictive lead scoring models have been categorized into classification, regression, and clustering models (Ahmed, 2004; Chorianopoulos, 2016; Ngai et al., 2009). A few common ML algorithms that leverage the classification modeling technique have been implemented in industrial practice and academic research. For instance, a financing company, GE Capital, used a two-class decision tree model to identify patterns that are common across positive leads but not present in negative leads (Aggour & Hoogs, 2013). Furthermore, as an ensemble learning technique built on the concept of decision trees, random forest has been used to classify leads with social media data as input (Meire et al., 2017). Random forest has also been used with explanation models to qualify and classify leads based on features and historical customer data (Bohanec et al., 2017). Yan et al. (2015) introduced a model to predict the likelihood of winning sales leads over a defined timeframe. They employed logistic regression to capture and quantify the influence of a salesperson's daily activities and lead-related information on sales outcomes (won or lost). Additionally, Xu et al. (2017) developed a data-driven system by utilizing linear and exponential regression, and seasonal time series models to predict lead conversion rates and estimate sales revenue from opportunities. Furthermore, Wei et al. (2016) used clustering analysis to group potential customers into clusters for scoring purposes and developed distinct marketing strategies. One common characteristic of these studies is that they only output a single prioritized list of leads ranked based on conversion probabilities generated by one classic ML model. However, we argue that multiple short and high-quality lists of leads considering the leads' perspective could outperform the original long list of leads.

4.3 Methodology

In this study, we employed a "smart sales" approach to construct predictive lead scoring models. Because leads have different purchase intentions at each phase in the B2B sales process (Wu et al., 2024), it is essential to identify leads with the highest purchase intention in each phase in order to optimize the sales process and sell smartly and efficiently. Research has shown that the purchase intentions of potential customers can be represented by their demographics (Bhat et al., 2021; Mahajan & Mahajan, 2018). Thus,

incorporating the leads' perspective (i.e., leads' purchase intentions) and applying the unsupervised clustering modeling learning method, we first clustered leads into groups[2] using demographic attributes. Each cluster contains similar leads with the highest purchase intention at a specific phase in the sales process. Then, we integrated insights discovered from this unsupervised clustering modeling into the supervised classification modeling process to predict leads' conversion probabilities and rank them accordingly. For example, the modeling process on the left-hand side in **Figure 4.2** shows the "classification-only" approach implemented in the current literature. The result is a single long ranked list of leads. On the right-hand side of the "smart sales" approach, we initially apply clustering to divide leads into groups based on demographic attributes. For example, cluster 0 contains leads with similar demographics who usually make a purchase decision at an early stage in the sales process (Wu et al, 2024). Then, for each cluster, we apply classification to train the model and calculate leads' conversion probabilities. Finally, we generate a short and high-quality list of leads out of each cluster and rank leads in each cluster according to their conversion probabilities/scores.

Instead of directly incorporating lead demographics into a single classification model, we chose the proposed modeling process for the following reasons. First, given that each cluster represents a subgroup with potentially unique characteristics of leads, by developing separate classification models for each cluster, we could create specialized models that are more suitable to specific features and patterns of these subgroups. This customized approach may enhance both accuracy and predictive performance within each cluster (Rodriguez et al., 2019). Second, it is common in many real-world datasets for some classes to be underrepresented. Clustering helps recognize these minority classes as distinct groups. Separate models can then address the class imbalance directly within each cluster, which may improve the sensitivity and specificity for minority classes (Chorianopoulos, 2016). Third, the process of clustering may discover patterns/relationships within the data that might remain hidden under a single-model approach. Such insights could drive further analysis and inform decision-making (Han et al., 2012). Finally, salespeople could benefit from the proposed approach by specializing in working with specific groups of leads.

The results of these models provide recommendations for sales practitioners when choosing which sales lead to pursue next. From an ML standpoint, exploring lead scoring is particularly intriguing due to the significant class imbalance present in leads data. Typically, in leads data, only about 5% of all cases result in a positive outcome (i.e., converted), while the remaining either have a negative outcome (i.e., not converted) or unknown (D'Haen & Van den Poel, 2013; Duncan & Elkan, 2015; Eitle & Buxmann, 2019). The small number of cases with positive outcomes is a challenge for ML to perform classification

[2] "groups" and "clusters" are used interchangeably in this paper.

effectively, which requires thorough preprocessing and careful model development. Next, we describe the data employed in this study, emphasizing the key variables and terms pertinent to the research design.

Figure 4.2 appears here

4.3.1 Data collection

We employed two datasets (lead's field and call history) provided by a top North American LMS software provider. The datasets encompass 15 fields, including leads' characteristics and contact history information. The data used in this study covers the call history that captures call interactions between salespeople and leads of a selling company that provides legal consulting services from March to May 2023. The data was generated through the activities of salespeople who utilized the LMS software during their interactions (including lead qualification, lead nurturing, and presentation of sales proposals) with leads.

We merged the call history with the lead's field datasets and pre-processed the data to remove redundant and irrelevant variables and outliers, impute missing values, and reduce the excessive number of classes within certain categorical variables. We ended up with a sample of 7,623 call interactions with 2,489 leads. We derived new variables through self-calculations on the existing variables. The data transformation process resulted in a more concise and enhanced dataset, which could potentially reduce the complexity of our data analysis and improve the generalizability of our conclusions (Yu et al., 2006). We used Python Spyder IDE (version 5.4.0) for data cleaning, preparation, model development, and analysis.

Table 4.1 describes the variables used in this study. There are six variables on the demographics of leads (i.e., revenue, company size, title, management level, primary lead source, and secondary lead source), and three variables that conceptualize salesperson effort (i.e., total duration, number of calls, and result group). The revenue and company size are of numerical type. The rest of the demographic attributes of leads are categorical. The categories of leads' titles are admin, assistant, attorney, associate attorney, legal secretary, marketer, office manager, owner, and partner. The management levels of leads are C-suite, director, non-manager, or manager. In the data collection process, leads were obtained from one of the primary lead sources (i.e., Zoominfo, others) and one of the following secondary lead sources (i.e., Cognism, Zoominfo, or Zoominfo-SalesIntel). The total duration of phone calls and the number of calls made between a lead and a salesperson are also of numerical types.

Table 4.1 appears here

Next, we selected the result group variable as the outcome of interactions with leads. Initially, the data had 5 different interaction outcomes due to abbreviations and synonyms. Consequently, after consulting with one of our industrial partner's data analysts, we categorized the outcomes of interactions into two possible outcomes: *positive* or *negative contact.* An outcome is classified as *negative contact* if a salesperson was unable to qualify or convert a lead during the sales process. Conversely, a *positive contact*

outcome includes qualified sales leads who demonstrated interest, made requests, booked appointments/demos, etc. Additionally, a *positive contact* includes cases where a sales goal was met, such as the successful sale of a legal service.

4.3.2 Machine learning (ML) algorithms

As mentioned above, we first need to identify distinct groups of leads. This is an unsupervised clustering ML problem, which partitions a heterogeneous sample into several homogeneous groups (Chorianopoulos, 2016). We applied clustering to partition the leads into groups based on demographics. We relied on a partitioning clustering algorithm using K-Prototypes (Gupta et al., 2022), a hierarchical clustering algorithm using the agglomerative (AGNES) approach (Reynolds et al., 2006), and a density-based clustering algorithm using the density-based spatial clustering of applications with noise (DBSCAN) (Dalmaijer et al., 2022; Han et al., 2012). These clustering algorithms are discussed below.

K-Prototypes aims to partition a dataset into K distinct (non-overlapping) clusters by assigning each instance to the nearest cluster while keeping clusters as distinct as possible (Gupta et al., 2022). Unlike traditional partitioning algorithms (e.g., K-means), K-Prototypes is capable of handling mixed data types, making it ideal for datasets that contain both numerical and categorical variables (which is in our case). The algorithm calculates the centroid of each cluster, which represents the average profile of the cluster, using the mean for numerical data and mode for categorical data. The process iteratively assigns each data point to the nearest cluster based on a combination of Euclidean distance for numerical data and a dissimilarity measure for categorical data. The clusters become stabilized when no more changes occur on the centroids. We choose K-Prototypes because of its ability to handle mixed data types effectively and its computational efficiency. The AGNES hierarchical clustering algorithm initially treats each data instance as its own cluster. Then it merges these clusters based on a criterion (distance or dissimilarity) between clusters progressively until all instances are merged into one single cluster or until a stopping criterion is reached (Chorianopoulos, 2016). As the result of visualization, a tree-like diagram called a dendrogram shows the clusters' merging process. The visualization assists in understanding the structure and deciding where to cut the tree to determine clusters. We chose AGNES algorithm for its superior interpretability through visualization and its ability to help identify outliers or anomalies within the dataset (Reynolds et al., 2006). DBSCAN generates clusters based on high-density areas of data instances that are separated by low-density areas. The clusters formed by DBSCAN typically have arbitrary shapes and sizes, which is different from the spherical clusters generated by K-Prototypes (Chorianopoulos, 2016). Given a radius of a core point, the process of DBSCAN starts from the core point and expands the cluster by recursively including all density-reachable data points (Han et al., 2012). The process stops when no more data points can be included in the cluster.

We chose DBSCAN in this study because of its flexibility in cluster shapes and its ability to identify and separate noise points that do not belong to any cluster (Dalmaijer et al., 2022).

Next, we trained separate classifiers on each identified group to predict the odds of conversion (i.e., scoring the leads) and produce ranked lists of leads. Specifically, we developed lead scoring models using independent variables to predict/classify the outcome of interactions (i.e., *positive* or *negative contact*) with leads in the sales process. This constitutes a supervised classification ML problem, wherein a model is trained using labeled instances and subsequently applied to map and classify new instances (Chorianopoulos, 2016; Kotu & Deshpande, 2015). We employed supervised ML algorithms based on their distinct learning processes and methodologies. These include a decision-based classifier using Decision Trees (Aggour & Hoogs, 2013; Başarslan & Argun, 2020), an ensemble-based classifier using Gradient Boosting Trees (Duncan & Elkan, 2015; Eitle & Buxmann, 2019), a regression-based classifier using Logistic Regression (D'Haen & Van den Poel, 2013), and a probabilistic-based classifier using Naïve Bayes (Benhaddou & Leray, 2017). The selection of these algorithms is rooted in their suitability for addressing the research problem and dataset, as recommended by the existing literature (Wu et al, 2023). We discuss the detailed rationale for selecting these algorithms below.

Decision Tree (DT) scans the dataset and constructs the tree by dividing the attributes into groups and subgroups based on the similarity of their observed values (Crawford, 1989). This results in dissimilar datasets being divided into subgroups, where each subgroup consists of similar instances. We choose DT because it is non-parametric and does not assume a specific distribution for the input or outcome variables, which enhances its applicability across various types of data (Abbott, 2014). Gradient Boosting Trees (GBT) involves aggregating multiple classifiers, each with varying performance levels, to form a single ensemble classifier aimed at achieving more accurate and consistent predictions than any of the individual classifiers alone (Gokhale & Joshi, 2018). The GBT algorithm begins with a simple tree that makes initial predictions on the data. Next, it calculates the errors of the model. Then, it trains a new decision tree specifically to predict these errors and update predictions to improve accuracy. In the process, it gradually builds a more accurate model by adding more trees (i.e., repeating these steps) until a set number of trees is reached or improvement becomes minimal. We chose the GBT algorithm for its high classification accuracy, effectiveness in handling various types of data, and ability to model complex non-linear relationships in practical applications (Duncan & Elkan, 2015). Logistic Regression (LR) is a widely utilized ML and statistical method that models the relationship between multiple predictor variables and discrete outcome variables with minimal constraints (Kotu & Deshpande, 2015; Murphy, 2012). LR evaluates the probability of an event's occurrence based on the values of the predictor variables (Bewick et al., 2005). Furthermore, LR employs a non-linear transformation known as the logit or sigmoid function to effectively map the odds

ratio of the outcome probabilities (Murphy, 2012). We chose LR in this study because it effectively explains variations in the outcome variable and assesses the significance of each independent variable. Naïve Bayes (NB) is a probabilistic classifier that estimates probabilities by counting the frequency and combinations of values within a dataset (McCallum & Nigam, 1998) based on Bayes' theorem. The assumption of this algorithm is the impact of a predictor variable's value on a particular class is independent of other predictor values, even if those predictors are interdependent or reliant on other predictor variables (Murphy, 2012). Using the training data, the algorithm determines the probability of each class and the conditional probability for each class given each predictor value (Mohammed et al, 2017). We chose NB for its speed, ease of implementation, and suitability for handling large datasets.

4.3.3 Model evaluation and imbalanced dataset

In this study, a ten-fold cross-validation method is employed to assess and compare the performance of the chosen algorithms. The utilization of a ten-fold cross-validation is supported by existing literature which sees it as a robust approach for evaluating models (Başarslan & Argun, 2020; Chorianopoulos, 2016; Kohavi, 1995). We used standard metrics to evaluate model performance. The gathered data is imbalanced (i.e., unequal distribution of instances across outcome variables). Specifically, *positive contact* outcomes form 7% of all instances while the remaining instances fall under the *negative contact* category. The small number of data instances within the minority class poses a challenge for algorithms in achieving effective classification results. This class imbalance may also result in a trade-off between sensitivity and specificity for some models (Akosa, 2017; Napierala & Stefanowski, 2012). Thus, this research also considers balanced accuracy and F1-score to compare model performance. Data imbalance can result in higher misclassification rates for the minority class during the evaluation, primarily because accuracy calculations tend to be skewed toward the majority class and may not adequately reflect performance for the minority class (Fernández et al., 2018). In addition, distinguishing between a minority class and a noisy instance may become difficult in the case of an imbalanced dataset.

To address the data imbalance issue, we resampled the data using three techniques. First, we used the oversampling technique which randomly adds instances to the minority class (i.e., *positive contact*) to yield a totally balanced dataset (Chorianopoulos, 2016). Second, we applied the undersampling technique which randomly deletes instances from the majority class (i.e., *negative contact*) to yield a completely balanced dataset (Jo & Japkowicz, 2004). Last, we resampled using the Synthetic Minority Oversampling Technique (SMOTE) which uses the k-nearest neighbor method to generate random synthetic instances for the minority class, while simultaneously reducing samples from the majority class until both classes achieve a comparable data sample proportion (Chawla et al., 2002). The literature suggests that the oversampling technique could result in model overfitting, and undersampling might potentially eliminate instances that

are critical for the classification objective (Chawla, 2005; Napierala & Stefanowski, 2012). The SMOTE technique aids in creating an alternative balanced dataset.

4.4 Results

4.4.1 Modeling procedure

First, we applied clustering modeling algorithms (i.e., K-Prototypes, AGNES, DBSCAN) to separate leads into groups that reflect similarities among them. These groups were formed based on the demographic features of leads because this is typically the only information a seller can have about leads before contacting them. We applied the Elbow method to determine the optimal number of clusters (Syakur et al., 2018). Three, four, and five are candidates for the optimal number of clusters. Second, we applied classification modeling using four distinct classifiers: DT, GBT, LR, and NB. Our dataset was partitioned by randomly selecting 70% of the instances as a training set, with the remaining 30% reserved as a testing set for assessing the performance of each model. Furthermore, as mentioned above, to address the class imbalance issue, we experimented with oversampling, undersampling, and SMOTE techniques. We evaluated the performance of all classifiers using the metrics outlined above.

4.4.2 Modeling results

We assessed the quality of clusters using the Silhouette coefficient (Han et al., 2012) and the Davies-Bouldin index (Davies & Bouldin, 1979). These metrics evaluate how well-separated and compact the clusters are. A higher Silhouette coefficient (from -1 to 1) and a lower Davies-Bouldin index (from 0 to 1) indicate better clustering quality. We pruned the dendrograms generated by AGNES in a manner that results in the same number of clusters as K-Prototypes. Finding the best DBSCAN clustering result is not trivial. It involves tuning hyperparameters which are “eps” (the maximum distance between two samples for one to be considered as in the neighborhood of the other) and “min sample” (the number of samples in a neighborhood for a data point to be a core point) (Han et al., 2012). We looped through different parameter combinations and selected the best ones (the best eps is 1.600 and the minimum samples are 3) based on the Silhouette and Devis-Bouldin scores. Finally, we applied DBSCAN with the best hyperparameters to cluster the instances (k=5). **Table 4.2** shows the performance of the three clustering algorithms with different numbers of clusters. K-Prototypes (k = 3) has the best Silhouette and Davies-Bouldin scores. We selected three as the optimal number of clusters and segmented leads accordingly using K-Prototypes.

Table 4.2 appears here

After splitting the dataset into training and testing sets, we first calculated the lead conversion rate in the training dataset. The lead conversion rate is a commonly used metric to examine the impact of lead scoring models on sales performance (Wu et al., 2023). The number of conversions (i.e., the outcome is a *positive contact*) in the training dataset is 119, the number of training instances is 1742, and the conversion

rate of the training dataset is 6.83%. Ideally, salespeople strive to contact every lead in the test dataset and attempt to convert all of them into sales. However, in practice, a salesperson is unable to contact the entire list (Ohiomah et al., 2018). Additionally, some leads have higher conversion odds than others, which can be determined simply by looking at their basic information (Monat, 2011). Thus, to simulate a real-life sales situation, only half of the testing dataset is considered in subsequent evaluations. We randomly took 50% of the testing set as the random policy. The conversion rate under the random policy is 7.24%. Without employing lead scoring, these conversion rates are in line with the average conversion rate (6%) in the literature (Coe, 2004; D'Haen & Van den Poel, 2013; Duncan & Elkan, 2015; Eitle & Buxmann, 2019).

Table 4.3 shows the performance of the data resampling techniques. Overall, the GBT classifier with SMOTE sampling had the best performance considering the highest balanced accuracy and F1-score with a conversion rate of 13.67%. The NB classifier had the worst performance. Furthermore, the oversampling and undersampling techniques of all classifiers produced similar results, while SMOTE had the best performance. Also, we noticed that the lift of the predictive model's conversion rate (13.67%) over the random model's (7.24%) is almost doubled.

Table 4.3 appears here

Next, we compared the performance of the "classification-only" lead scoring model in the literature to our "smart sales" approach (i.e., clustering-then-classification). Specifically, we applied the GBT classifier (the best-performing classifier in **Table 4.3**) to each resulting cluster (cluster 0 contains 380 leads, cluster 1 contains 273 leads, and cluster 2 contains 1,836 leads). Since the optimal number of clusters is three, we had three clusters/lists of leads (see **Table 4.4**). In addition to the metrics used above, we considered the cost of false positives when choosing our "champion" (i.e., best) model in the following analysis. False positives (FPs) are cases in which we predicted "yes" but the actual outcome was "no". In our case, FPs are leads that our model predicted would convert, however, they ended up being not converted. In practice, there is a cost associated with these FPs, which is $8 USD per FP (Lima, 2018). In general, a lower cost of FPs is better.

In **Table 4.4**, the GBT cluster 0 model has a lower cost of FP than the GBT "classification-only" model, and the former had a higher conversion rate. Moreover, the GBT cluster 2 model had the best numbers in model performance evaluation metrics such as balanced accuracy, sensitivity, and F1-score. To underline the superiority and effectiveness of the proposed lead scoring approach, research shows that the lead conversion rate increases by 5% on average as a result of applying the original predictive lead scoring (i.e., classification-only) (D'Haen & Van den Poel, 2013; Duncan & Elkan, 2015). The best model (GBT cluster 0) generated by our "smart sales" approach was able to reach a conversion rate of 19.30% which is

more than the average 5% increment compared to 13.67% (the best conversion rate of the classification-only approach).

Table 4.4 appears here

Moreover, we inspected the average performance of all classifiers on the derived clusters (**Table 4.5**). Despite a slightly lower specificity, GBT clusters outperformed others in conversion rate, cost of FP, balanced accuracy, sensitivity, and F1-score. In summary, based on our "smart sales" approach, the combination of K-Prototypes and GBT classifier is the best-performing predictive lead scoring model.

Table 4.5 appears here

4.5 Conclusion

Before the fourth industrial revolution, decision-making in B2B sales has been predominantly guided by methods that relied on individual judgment, professional competence, and prior experience. Particularly in lead scoring, such a subjective approach could result in low conversion rates as well as company resources being wasted by dedicating them to wrong leads (D'Haen & Van den Poel, 2013; Monat, 2011). To address this issue, AI-driven and ML-based lead scoring models have emerged as an alternative solution that shifts the majority of decision-making to predictive analytical models. To answer the research call proposed by Syam & Sharma (2018) and further improve lead conversion rates, we developed and implemented a "smart sales" predictive lead scoring model by integrating insights from the leads' perspective. We first identified groups of leads generated by clustering. Each group contains similar leads with higher purchase intentions at a specific phase in the sales process. Then we built separate classification models to predict leads' conversion probabilities in each group and provide salespeople with ranked lists of leads. To make effective contacts with leads and sell smartly, salespeople need to involve with leads who have higher purchase intentions at each phase. Different combinations of clustering algorithms (K-Prototypes, hierarchical clustering, DBSCAN) and classification algorithms (DT, GBT, LR, NB) were tested. Our assessments of the models indicate that the combination of partitioning-based K-Prototypes clustering with ensemble-based Gradient Boosting Trees classifier generated the best-performing "smart sales" lead scoring model. Next, we discuss the derived implications, interesting findings, contributions, limitations, and future work.

4.5.1 Implications

Wu et al. (2024) suggest that leads with different demographics have distinct purchase intentions at each phase in the B2B sales process. Adopting the same perspective, we applied clustering to identify such groups with different demographics. Each group has a different purchase intention at each phase. For example, a particular group may have the highest purchase intention at an early phase and another group may have the highest purchase intention at the final phase. In fact, identifying such groups can be used to indicate the most representative leads of the highest purchase intention at a specific phase in the sales

process, which can improve sales performance significantly. Without considering interactions between salespeople and leads, we first applied clustering only to the leads' demographic attributes. The assessment of the "smart sales" predictive lead scoring model along with the enhanced lead conversion rates and reduced false positive costs verified the effectiveness and feasibility of applying clustering before classification in the lead scoring process to surface high-quality leads.

We clustered leads into representative groups. For example, 380 leads in cluster 0 who are titled partners at the director management level working in small-sized[3] companies have the highest conversion rate (19.30%). Typically, these leads are most likely to make a purchase decision at an early phase of the sales process (Wu et al., 2024). Thus, this group of leads should be contacted by salespeople first as a priority. Furthermore, 1,836 leads in cluster 2 with the title of attorney/legal secretary at the manager level working in small-sized companies are most likely to make a purchase decision during the interim phase. Salespeople should approach this list of leads next. The final list of leads that salespeople should be involved with contains 273 leads in cluster 1 who are predominantly titled partners at the non-manager management level working in large-sized companies. These leads tend to make a purchase decision when they reach the final phase in the sales process. Thus, they should not be treated by salespeople as prioritized contacts.

Indeed, the demographic traits (e.g., large-sized companies) of leads in cluster 1 indicate that they are not ready to make an early purchase decision. This is partially because in large-sized companies decision-making is usually complex and involves multiple approvals from various stakeholders (Webster & Wind, 1972; Zolkiewski et al., 2017). Leads in cluster 1 should be approached last. As Wu et al. (2024) suggest, regardless of salespeople's efforts, leads in cluster 1 are not ready to make a purchase commitment early in the process. Instead, salespeople should focus on leads in clusters 0 and 2 who are typically authorized with a certain degree of buying power (given their roles and management levels) to make a purchase decision in the early and interim phases. Especially, more attention should be devoted to leads in cluster 0, given that they are partners/owners who work in small-sized companies. Salespeople would have a better chance of converting these leads if they are engaged at the right time. From a salesperson's perspective, short lists of high-quality leads are better than long lists in order to sell smartly. Given the enhanced lead conversion rate (19.30%) and performance of the "smart sales" predictive lead scoring model, we conclude that the effectiveness of lead scoring can be significantly improved by first clustering leads into groups, followed by predicting conversion outcomes separately and ranking leads in each group.

In the marketing-sales funnel (**Figure 4.1**), the sizes of leads' pools shrink as we go down from prospects to customers, which is the nature of the funnel (i.e., unqualified leads are removed from the pool

[3] The company sizes were categorized into small-sized, medium-sized, and large-sized (Nooteboom, 1994) in this study.

at each phase). To improve the bonding between marketing and sales functions and enhance the sales process, it is imperative to optimize the quality of pools. In this study, we increase the quality of pools by reducing embedded noise. Noise in the pool is low-quality leads who are predicted to convert, however, they end up not converting (i.e., false positives). Thus, the lower the cost of false positives, the better quality the pools would be. Comparing the cost of false positives ($56) of the best list of leads (i.e., GBT cluster 0) to the classification-only method ($72) (**Table 4.4**), we can conclude that the proposed "smart sales" approach outperformed the original classification-only method in the aspect of optimizing the pools' quality.

4.5.2 Research contributions

The findings of this study contribute to the theory and practice of lead scoring in the field of B2B sales. The rapid advancements in digital transformation technologies, such as artificial intelligence technologies, machine learning algorithms, and data mining models have fundamentally changed how decisions are made by stakeholders in organizations. We demonstrated that the lead scoring task in the prospecting stage of the B2B sales process has been positively and radically changed since the emergence of predictive ML models under the wave of digital transformation. By suggesting a "smart sales" predictive lead scoring approach to increase the efficiency and effectiveness of the B2B sales process, the empirical findings of this study demonstrated the important roles that data-driven predictive ML models play in the transition from traditional human decision-making to model-based decision-making.

Second, we provided sales professionals with reasonable and actionable insights. The derived short and high-quality lists of leads serve as guidelines that could help sales professionals choose, contact, and pursue leads in the B2B sales process. Moreover, from a short-term perspective, a high-quality lead scoring model with superior performance could convince salespeople to contact more MQLs and convert those "ready-to-buy" leads into customers. From a long-term perspective, having a high-quality lead scoring model can also improve the internal collaboration between marketing and sales, potentially mitigating the negative effect of the "sales lead black hole" (Sabnis et al., 2013) and ultimately improving sales performance. Lead scoring models allow industries to segment their customer base more effectively. By scoring leads based on their conversion likelihood and purchase intention, sellers can design customized marketing and service strategies for various segments, enhancing personalization and increasing leads' satisfaction. Also, the predictive lead scoring model is capable of adapting to variations in the market by continually learning from new data sources. Additionally, lead scoring helps prioritize resources by focusing on leads with the highest potential for returns. This targeted approach ensures that resources are not wasted on leads with little conversion likelihood. However, to implement the proposed "smart sales" model in practice, we discuss some potential practical challenges and propose tentative solutions in the following section.

4.5.3 Limitations and future research

The model's performance highly depends on the availability and quality of the data. Real data issues such as inaccuracy, incompleteness, and obsolescence could significantly reduce the model's effectiveness, which potentially lead to dissatisfaction among stakeholders. Future studies could improve the "smart sales" model by gathering commercial and social media textual data for training, and by regularly updating the model based on feedback and changing conditions. Moreover, integrating a predictive lead scoring model into the existing sales processes, such as CRM systems, could be challenging. Any integration issues may hinder the effective use of the model's outputs in the sales process. Implementing the model to provide real-time scoring in a dynamic sales environment could also pose technical difficulties. Future research could focus on developing an application programming interface for the "smart sales" lead scoring model that facilitates real-time scoring of leads and modifying the existing CRM systems to accommodate new data fields and ensure smooth data flow. Furthermore, resistance from sales practitioners in adopting the new model (system) may occur. It will be crucial to provide proper trainings and clear communication for sales practitioners, not only on how to use the new model (system) but also on understanding its benefits.

While this study provides insights into the effective and efficient lead scoring practice, we acknowledge its limitations and suggest additional opportunities for future research. First, the size of the dataset in this study is considered small. However, this study's purpose is to use this small dataset to demonstrate the feasibility of a new approach in predicting lead conversion. This small dataset can serve as proof of concept before scaling up. Also, Lindstrom (2016) presented a collection of case studies to show that small datasets can offer substantial insights into marketing and customer behavior studies. The lead scoring model generated by this small dataset can be treated as a baseline model. Future research can compare the performance of larger datasets against this baseline to show improvements and justify findings. Second, the best-performing resampling technique in this study is SMOTE which generates synthetic samples by interpolating between minority class (positive contact) instances. However, this resampling approach could lead to model overfitting as the classifier may memorize these artificial data points. This issue is particularly concerning with smaller datasets, where the synthetic data points might not capture the full variability of real-world data, causing the model to excel on the training datasets but underperform on unseen datasets. Also, since SMOTE creates synthetic data along the lines between minority class (positive contact) instances, it could potentially place synthetic points in areas where the majority class (negative contact) is dominant. This could blur class boundaries and lead to misclassifications of the majority class, especially when the classes overlap. Future research should explore this concern by investigating various overfitting detection techniques, such as out-of-sample testing, model regularization, and model simplification. Third, the models and the insights derived from this study are limited to the legal service

industry. Although this limited scope helped mitigate potential confounding factors, it constrains the generalizability of results to service industries. Future research can consider applying the same approach to other industries to test the applicability of this study's findings. Finally, the empirical testing results and insightful implications were produced by applying the "smart sales" predictive lead scoring model to historical data used in this study. Future studies could test the proposed "smart sales" model in a field study to enhance the validity of this model. For example, researchers could adopt the proposed "smart sales" method in this study to generate lists of ranked leads to salespeople for contact. The field results can be compared to the empirical results in this study.

Table 4.1. Variable description

Name	Type	Description
Contact ID	Numeric	A unique identifier key for a lead.
Call history ID	Numeric	A unique identifier key for a call made.
Revenue	Numeric	The revenue of the company that a lead belongs to.
Company size	Numeric	The number of employees of the company that a lead belongs to.
Title	Categorical	The titles of leads/contacts (i.e., admin, assistant, attorney, associate attorney, legal secretary, marketer, office manager, owner, and partner).
Management level	Categorical	The management levels of leads/contacts (i.e., C-suite, director, non-manager, and manager).
Primary lead source	Categorical	The primary source where a lead was obtained from.
Secondary lead source	Categorical	The secondary source where a lead was obtained from.
Total duration	Numerical	The amount of time spent on calls between a lead and a salesperson.
Number of calls	Numerical	The number of calls made between a lead and a salesperson.
Result group	Categorical	The outcome of conversion (i.e., positive or negative).

Table 4.2. Evaluation of Clustering Algorithms

Algorithms & Number of Clusters	Silhouette Coefficient	Davies-Bouldin Index
K-Prototypes (k = 3)	0.698	0.599
K-Prototypes (k = 4)	0.602	0.658
K-Prototypes (k = 5)	0.587	0.692
AGNES (k = 3)	0.569	0.772
AGNES (k = 4)	0.579	0.648
AGNES (k = 5)	0.559	0.651
DBSCAN (k = 5)	0.578	0.600

Table 4.3. Evaluation of Resampling Techniques ("classification-only" lead scoring model)

Resampling Techniques	Conversion Rate	Balanced Accuracy	Specificity	Sensitivity	F1-score
DT Baseline	11.53%	0.86	0.94	0.79	0.86

DT Oversampling	12.87%	0.85	0.93	0.77	0.84
DT Undersampling	13.14%	0.85	0.73	0.96	0.83
DT SMOTE	12.60%	0.89	0.94	0.85	0.89
GBT Baseline	13.67%	0.87	0.98	0.76	0.86
GBT Oversampling	13.67%	0.88	0.87	0.88	0.88
GBT Undersampling	13.67%	0.86	0.76	0.96	0.85
GBT SMOTE	13.67%	0.90	0.92	0.88	0.90
LR Baseline	13.67%	0.88	0.98	0.78	0.87
LR Oversampling	13.67%	0.87	0.79	0.96	0.86
LR Undersampling	13.67%	0.83	0.78	0.96	0.86
LR SMOTE	13.67%	0.89	0.84	0.96	0.90
NB Baseline	11.14%	0.68	0.35	0.90	0.50
NB Oversampling	11.14%	0.59	0.18	0.95	0.30
NB Undersampling	11.14%	0.73	0.97	0.49	0.65
NB SMOTE	12.87%	0.50	0.40	0.95	0.56

Table 4.4. Evaluation Comparison of Lead Scoring Models

Lead scoring Models	Conversion Rate	Cost of FP	Balanced Accuracy	Specificity	Sensitivity	F1-score
GBT classification-only	13.67%	$72	0.88	0.98	0.78	0.87
GBT Cluster 0 (380 leads)	19.30%	$56	0.83	0.85	0.82	0.83
GBT Cluster 1 (273 leads)	9.67%	$56	0.91	0.81	0.94	0.87
GBT Cluster 2 (1,836 leads)	13.45%	$100	0.92	0.92	0.95	0.93

Table 4.5. Average Performance Evaluation of Classifiers on Clusters

Algorithms	Conversion Rate	Cost of FP	Balanced Accuracy	Specificity	Sensitivity	F1-score
DT Clusters	10.44%	$48	0.87	0.92	0.77	0.83
GBT Clusters	14.20%	$43	0.90	0.86	0.92	0.89
LR Clusters	13.25%	$70	0.83	0.94	0.72	0.80
NB Clusters	11.30%	$253	0.58	0.55	0.72	0.55

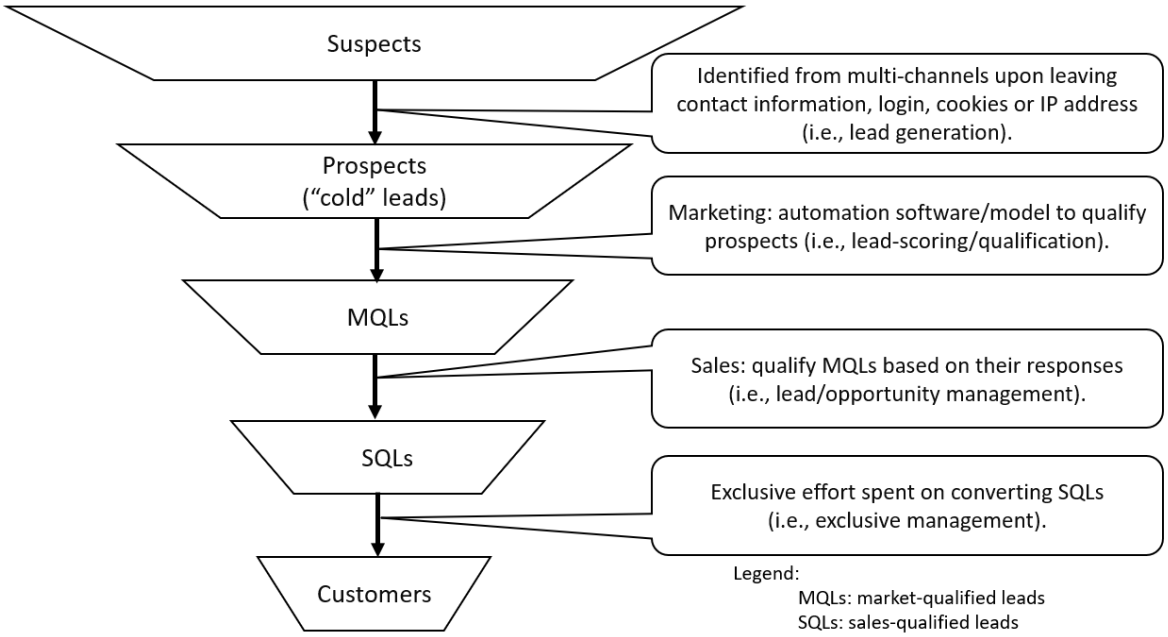

Figure 4.1 Marketing-sales funnel

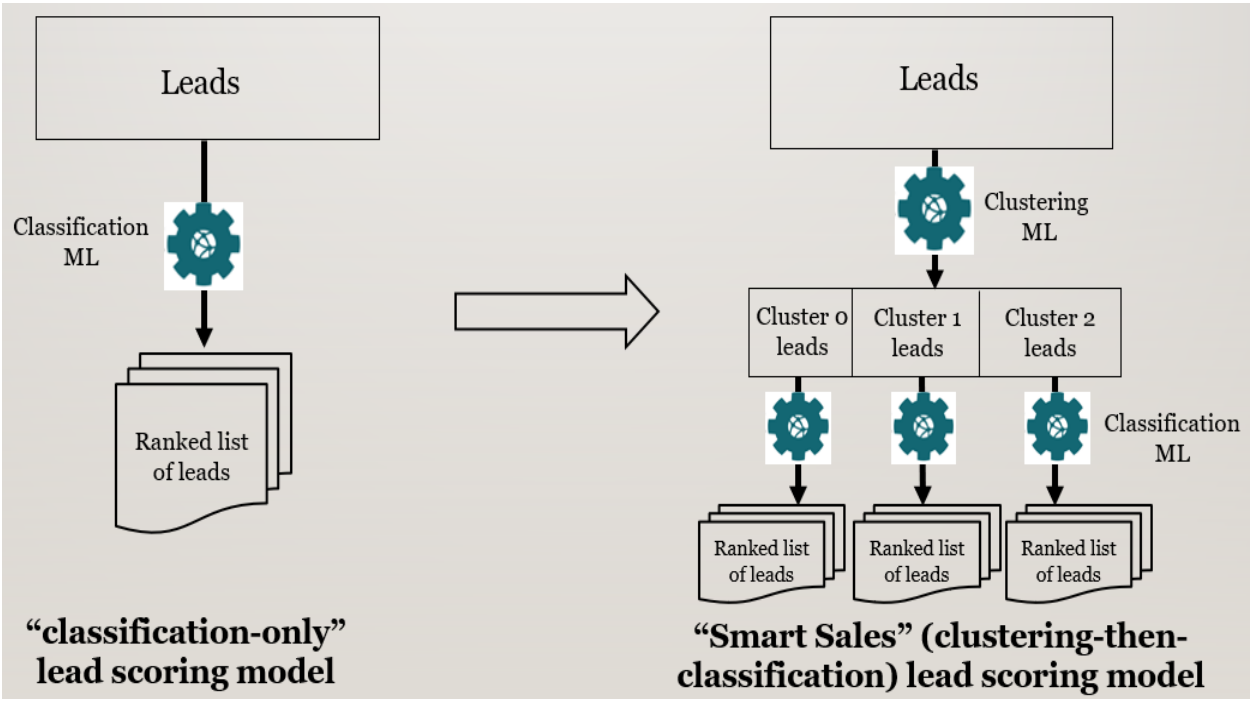

Figure 4.2. Lead Scoring Modeling Processes

References

Abbott, D. (2014). *Applied predictive analytics: principles and techniques for the professional data analyst*. John Wiley & Sons.

Ahmed, S. R. (2004). Applications of data mining in retail business. *International Conference on Information Technology: Coding Computing, ITCC*, *2*, 455-459 Vol.2. **https://doi.org/10.1109/ITCC.2004.1286695**

Aggour, K. S., & Hoogs, B. (2013). Financing lead triggers: Empowering sales reps through knowledge discovery and fusion. *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1141–1149. **https://doi.org/10.1145/2487575.2488190**

Akosa, J. (2017). *Predictive Accuracy: A Misleading Performance Measure for Highly Imbalanced Data*. Paper presented at the *Proceedings of the SAS Global Forum.* **https://www.semanticscholar.org/paper/Predictive-Accuracy-%3A-A-Misleading-Performance-for-Akosa/8eff162ba887b6ed3091d5b6aa1a89e23342cb5c**

Ang, L., & Buttle, F. (2006). Managing For Successful Customer Acquisition: An Exploration. *Journal of Marketing Management*, *22*(3–4), 295–317. **https://doi.org/10.1362/026725706776861217**

Başarslan, M. S., & Argun, İ. D. (2020). Prediction of potential bank customers: Application on Data Mining. *Artificial Intelligence and Applied Mathematics in Engineering Problems*, 96–106. **https://doi.org/10.1007/978-3-030-36178-5_9**

Benhaddou, Y., & Leray, P. (2017). Customer Relationship Management and Small Data—Application of Bayesian Network Elicitation Techniques for Building a Lead Scoring Model. *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)*, 251–255. **https://doi.org/10.1109/AICCSA.2017.51**

Bewick, V., Cheek, L., & Ball, J. (2005). Statistics review 14: Logistic regression. *Critical Care (London, England)*, *9*(1), 112–118. **https://doi.org/10.1186/cc3045**

Bhat, S. A., Islam, S. B., & Sheikh, A. H. (2021). Evaluating the Influence of Consumer Demographics on Online Purchase Intention: An E-Tail Perspective. *Paradigm (Ghāziabād, India)*, *25*(2), 141–160. **https://doi.org/10.1177/09718907211045185**

Bohanec, M., Kljajić Borštnar, M., & Robnik-Šikonja, M. (2017). Explaining machine learning models in sales predictions. *Expert Systems with Applications*, *71*, 416–428. **https://doi.org/10.1016/j.eswa.2016.11.010**

Boogar, L. (2019). The three stages of lead scoring: Lambs, Ducks & Kudus. The MadKudu Blog post. **https://blog.madkudu.com/three-stages-of-lead scoring/**. Accessed 04 March 2024

Bradford, W. R., Johnston, W. J., & Bellenger, D. (2016). The impact of sales effort on lead conversion cycle time in a business-to-business opportunity pipeline. *SSRN Electronic Journal*. **https://doi.org/10.2139/ssrn.2866954**

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, *16*, 321–357. https://doi.org/10.1613/jair.953

Chawla, N. V. (2005). Data mining for Imbalanced Datasets: An overview. In *Data Mining and Knowledge Discovery Handbook* (pp. 853–867). Springer. https://doi.org/10.1007/0-387-25465-x_40

Chorianopoulos, A. (2016). *Effective CRM using Predictive Analytics*. John Wiley & Sons.

Coe, J. M. (2004). The integration of direct marketing and field sales to form a new B2B sales coverage model. *Journal of Interactive Marketing*, *18*(2), 62–74. https://doi.org/10.1002/dir.20005

Crawford, S. L. (1989). Extensions to the CART algorithm. *International Journal of Man-Machine Studies*, *31*(2), 197–217. https://doi.org/10.1016/0020-7373(89)90027-8

Dalmaijer, E. S., Nord, C. L., & Astle, D. E. (2022). Statistical power for cluster analysis. *BMC Bioinformatics*, *23*(1), 205. https://doi.org/10.1186/s12859-022-04675-1

Davies, D. L., & Bouldin, D. W. (1979). A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-1*(2), 224–227. https://doi.org/10.1109/TPAMI.1979.4766909

D'Haen, J., & Van den Poel, D. (2013). Model-supported business-to-business prospect prediction based on an iterative customer acquisition framework. *Industrial Marketing Management*, *42*(4), 544–551. https://doi.org/10.1016/j.indmarman.2013.03.006

D'Haen, J., Van den Poel, D., & Thorleuchter, D. (2013). Predicting customer profitability during acquisition: Finding the optimal combination of data source and data mining technique. *Expert Systems with Applications*, *40*(6), 2007–2012. https://doi.org/10.1016/j.eswa.2012.10.023

D'Haen, J., Van den Poel, D., Thorleuchter, D., & Benoit, D. F. (2016). Integrating expert knowledge and multilingual web crawling data in a lead qualification system. *Decision Support Systems*, *82*, 69–78. https://doi.org/10.1016/j.dss.2015.12.002

Duncan, B., & Elkan, C. (2015). Probabilistic Modeling of a Sales Funnel to Prioritize Leads. *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1751–1758. https://doi.org/10.1145/2783258.2788578

Eitle, V., & Buxmann, P. (2019). Business analytics for sales pipeline management in the software industry: A machine learning perspective. *Proceedings of the Annual Hawaii International Conference on System Sciences*. https://doi.org/10.24251/hicss.2019.125

Fernández, A., García, S., Galar, M., Prati, R. C., Krawczyk, B., & Herrera, F. (2018). Imbalanced Classification with Multiple Classes. In A. Fernández, S. García, M. Galar, R. C. Prati, B. Krawczyk,

& F. Herrera (Eds.), *Learning from Imbalanced Data Sets* (pp. 197–226). Springer International Publishing. **https://doi.org/10.1007/978-3-319-98074-4_8**

Gokhale, P., & Joshi, P. (2018). A binary classification approach to lead identification and qualification. *Communications in Computer and Information Science*, 279–291. **https://doi.org/10.1007/978-981-13-1423-0_30**

Gupta, R., Kumar, H., Jain, T., Shrotriya, A., & Sinha, A. (2022). Demographic Customer Segmentation of banking users Based on k-prototype methodology. *2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 578–584. **https://doi.org/10.1109/Confluence52989.2022.9734169**

Han, J.W., Kamber, M., & Pei, J. (2012). 10 - Cluster Analysis: Basic Concepts and Methods. In *Data Mining: Concepts and Techniques* (Third Edition, pp. 443–495). Elsevier Inc. **https://doi.org/10.1016/B978-0-12-381479-1.00010-1**

Jadli, A., Hamim, M., Hain, M., & Hasbaoui, A. (2022). Toward a smart lead scoring system using machine learning. *Indian Journal of Computer Science and Engineering*, *13*(2), 433–443. **https://doi.org/10.21817/indjcse/2022/v13i2/221302098**

Lima C. D. (2018). *End-to-end-lead scoring.* GitHub. **https://github.com/calvdee/end-to-end-lead scoring/commits/main/docs**

Järvinen, J., & Taiminen, H. (2016). Harnessing marketing automation for B2B content marketing. *Industrial Marketing Management*, *54*, 164–175. **https://doi.org/10.1016/j.indmarman.2015.07.002**

Jo, T., & Japkowicz, N. (2004). Class imbalances versus small disjuncts. *ACM SIGKDD Explorations Newsletter*, *6*(1), 40–49. **https://doi.org/10.1145/1007730.1007737**

Karhunsalo, M. (2021). Transforming B2B sales essentials: how to reshape traditional sales procedures to engage modern organizational buyers (thesis). **https://urn.fi/URN:NBN:fi:amk-2021111320188**

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, 1137–1143.

Kotu, V., & Deshpande, B. (2015). Chapter 4—Classification. In V. Kotu & B. Deshpande (Eds.), *Predictive Analytics and Data Mining* (pp. 63–163). Morgan Kaufmann. **https://doi.org/10.1016/B978-0-12-801460-8.00004-5**

Lindstrom, M. (2016). *Small Data: The Tiny Clues That Uncover Huge Trends*. St. Martin's Press.

Mahajan, R., & Mahajan, D. (2018). Impact of Customers Demographic Characteristics during buying homes. *TIJ's Research Journal of Social Science & Management - RJSSM, 7*.

McCallum, A. and Nigam, K. (1998) A Comparison of Event Models for Naive Bayes Text Classification. *Proceedings in Workshop on Learning for Text Categorization*, AAAI'98, 41-48.

McDonnell, D. (2014). The Evolution From Traditional To Predictive Lead scoring. Demand Gen Report. **https://www.demandgenreport.com/industry-resources/ebooks/2892-theevolution-from-traditional-to-predictive-lead scoring**. Accessed 04 March 2024

Meire, M., Ballings, M., & Van den Poel, D. (2017). The added value of social media data in B2B customer acquisition systems: A real-life experiment. *Decision Support Systems*, *104*, 26–37. **https://doi.org/10.1016/j.dss.2017.09.010**

Mohammed, M., Bashier, E. B. Mohammed., & Khan, M. Badruddin. (2017). *Machine learning: algorithms and applications*. CRC Press. **https://doi.org/10.1201/9781315371658**

Monat, J. P. (2011). Industrial sales lead conversion modeling. *Marketing Intelligence & Planning*, *29*(2), 178–194. **https://doi.org/10.1108/02634501111117610**

Nooteboom, B. (1994). Innovation and Diffusion in Small Firms: Theory and Evidence. *Small Business Economics*, *6*(5), 327–347. **https://doi.org/10.1007/BF01065137**

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. The MIT Press.

Napierala, K., & Stefanowski, J. (2012). BRACID: A comprehensive approach to learning rules from imbalanced data. *Journal of Intelligent Information Systems*, *39*(2), 335–373. **https://doi.org/10.1007/s10844-011-0193-0**

Ngai, E. W. T., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications*, *36*(2), 2592–2602. **https://doi.org/10.1016/j.eswa.2008.02.021**

Nygård, R., & Mezei, J. (2020). Automating lead scoring with Machine Learning: An experimental study. *Proceedings of the Annual Hawaii International Conference on System Sciences*. **https://doi.org/10.24251/hicss.2020.177**

Ohiomah, A., Andreev, P., Benyoucef, M., & Hood, D. (2019). The role of lead management systems in inside sales performance. *Journal of Business Research*, *102*, 163–177. **https://doi.org/10.1016/j.jbusres.2019.05.018**

Ohiomah, A., Benyoucef, M., Andreev, P., Kuziemsky, C., Hood, D., & Le Bon, J. (2018). Best practices for inside sales professionals: An historical analysis. *Developments in Marketing Science: Proceedings of the Academy of Marketing Science*, 449–461. **https://doi.org/10.1007/978-3-319-99181-8_149**

Paschen, J., Wilson, M., & Ferreira, J. J. (2020). Collaborative intelligence: How human and artificial intelligence create value along the B2B sales funnel. *Business Horizons*, *63*(3), 403–414. **https://doi.org/10.1016/j.bushor.2020.01.003**

Patterson, L. (2007). Marketing and sales alignment for improved effectiveness. *Journal of Digital Asset Management*, *3*(4), 185–189. **https://doi.org/10.1057/palgrave.dam.3650089**

Peña-García, N., Gil-Saura, I., Rodríguez-Orejuela, A., & Siqueira-Junior, J. R. (2020). Purchase intention and purchase behavior online: A cross-cultural approach. *Heliyon*, *6*(6), e04284–e04284. **https://doi.org/10.1016/j.heliyon.2020.e04284**

Reynolds, A. P., Richards, G., de la Iglesia, B., & Rayward-Smith, V. J. (2006). Clustering Rules: A Comparison of Partitioning and Hierarchical Clustering Algorithms. *Journal of Mathematical Modelling and Algorithms*, *5*(4), 475–504. **https://doi.org/10.1007/s10852-005-9022-1**

Rodriguez, M. Z., Comin, C. H., Casanova, D., Bruno, O. M., Amancio, D. R., Costa, L. da F., & Rodrigues, F. A. (2019). Clustering algorithms: A comparative approach. *PloS One*, *14*(1), e0210236–e0210236. **https://doi.org/10.1371/journal.pone.0210236**

Rutherford, B. N., Marshall, G. W., & Park, J. (2014). The moderating effects of gender and inside versus outside sales role in multifaceted job satisfaction. *Journal of Business Research*, *67*(9), 1850–1856. **https://doi.org/10.1016/j.jbusres.2013.12.004**

Sabnis, G., Chatterjee, S. C., Grewal, R., & Lilien, G. L. (2013). The Sales Lead Black Hole: On Sales Reps' Follow-Up of Marketing Leads. *Journal of Marketing*, *77*(1), 52–67. **https://doi.org/10.1509/jm.10.0047**

Shannahan, R. J., Bush, A. J., Moncrief, W. C., & Shannahan, K. L. J. (2013). Making Sense of the Customer's Role in the Personal Selling Process: A Theory of Organizing and Sensemaking Perspective. *Journal of Personal Selling & Sales Management*, *33*(3), 261–275. **https://doi.org/10.2753/PSS0885-3134330302**

Sheth, Sharma, A., & Iyer, G. R. (2009). Why integrating purchasing with marketing is both inevitable and beneficial. *Industrial Marketing Management*, *38*(8), 865–871. **https://doi.org/10.1016/j.indmarman.2008.12.021**

Sheth, Sisodia, R. S., & Sharma, A. (2000). The Antecedents and Consequences of Customer-Centric Marketing. *Journal of the Academy of Marketing Science*, *28*(1), 55–66. **https://doi.org/10.1177/0092070300281006**

Smith, T. M., Gopalakrishna, S., & Chatterjee, R. (2006). A Three-Stage Model of Integrated Marketing Communications at the Marketing-Sales Interface. *Journal of Marketing Research*, *43*(4), 564–579. **https://doi.org/10.1509/jmkr.43.4.564**

Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018). Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster. *IOP*

Conference Series: Materials Science and Engineering, *336*(1), 012017. https://doi.org/10.1088/1757-899X/336/1/012017

Syam, N., & Sharma, A. (2018). Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. *Industrial Marketing Management*, *69*, 135–146. https://doi.org/10.1016/j.indmarman.2017.12.019

Thaichon, P., Surachartkumtonkun, J., Quach, S., Weaven, S., & Palmatier, R. W. (2018). Hybrid sales structures in the age of e-commerce. *The Journal of Personal Selling & Sales Management*, *38*(3), 277–302. https://doi.org/10.1080/08853134.2018.1441718

Verma, R., Koul, S., & Pai, S. S. (2016). Identifying profitable clientele using the analytical hierarchy process. *International Journal of Business and Systems Research*, *10*(2/3/4), 220. https://doi.org/10.1504/ijbsr.2016.075758

Webster, F. E., & Wind, Y. (1972). A general model for understanding organizational buying behavior. *Journal of Marketing*, *36*(2), 12. https://doi.org/10.2307/1250972

Wei, J. T., Lin, S.-Y., Yang, Y.-Z., & Wu, H.-H. (2016). Applying data mining and RFM model to analyze customers' values of a veterinary hospital. *2016 International Symposium on Computer, Consumer and Control (IS3C)*. https://doi.org/10.1109/is3c.2016.126

Wu, M., Andreev, P., & Benyoucef, M. (2023). The state of lead scoring models and their impact on sales performance. *Information Technology and Management*, *25*(1), 69–98. https://doi.org/10.1007/s10799-023-00388-w

Wu, M., Andreev, P., Benyoucef, M., & Hood, D. (2024). Unlocking B2B buyer intentions to purchase: Conceptualizing and validating inside sales purchases. *Decision Support Systems*, *179*. 114165. https://doi.org/10.1016/j.dss.2023.114165

Xu, X., Tang, L., & Rangan, V. (2017). Hitting your number or not? A robust & intelligent sales forecast system. *2017 IEEE International Conference on Big Data (Big Data)*, 3613–3622. https://doi.org/10.1109/BigData.2017.8258355

Yan, J., Zhang, C., Zha, H., Gong, M., Sun, C., Huang, J., Chu, S., & Yang, X. (2015). On machine learning towards Predictive Sales Pipeline Analytics. *Proceedings of the AAAI Conference on Artificial Intelligence*, *29*(1). https://doi.org/10.1609/aaai.v29i1.9455

Yu, L., Wang, S., & Lai, K. K. (2006). An integrated data preparation scheme for neural network data analysis. *IEEE Transactions on Knowledge and Data Engineering*, *18*(2), 217–230. https://doi.org/10.1109/TKDE.2006.22

Zolkiewski, J., Story, V., Burton, J., Chan, P., Gomes, A., Hunter-Jones, P., O'Malley, L., Peters, L. D., Raddats, C., & Robinson, W. (2017). Strategic B2B customer experience management: the importance

of outcomes-based measures. *The Journal of Services Marketing*, *31*(2), 172–184. https://doi.org/10.1108/JSM-10-2016-0350

Chapter 4 Key Points Summary:

- Implemented and validated a "Smart Sales" predictive lead scoring model incorporating purchase intention of leads from leads' perspective on a small dataset from a B2B service industry;
- A two-step approach (i.e., the combination of K-Prototypes clustering with Gradient Boosting Trees classification) generated the best-performing model;
- Each group characterized by different attributes has different levels of purchase intentions in each purchase phase;
- The proposed model generated multiple concise and high-quality lists of sales leads comparing to the original singular long list.

Chapter 5. Data-Driven Predictive Lead Scoring Model With Lead Profiling

Article Reference: Wu, M., Andreev, P., & Benyoucef, M. (2024). Data-driven predictive lead scoring model with lead profiling. *Decision Support Systems.* (working paper)

Article Abstract: In inside sales, processing leads and selecting the ones to pursue have traditionally relied on salespeople's intuition and experience-driven predictions. However, this rather subjective method has been widely acknowledged as ineffective and inefficient for lead scoring and conversion. Moreover, this method could result into the notorious "sales lead black hole" (aka "gap") between the marketing and sales functions. Therefore, scholars have shifted their focus to the more objective data-driven approaches to guide the decision-making of salespeople as leads progress through the sales pipeline. Nevertheless, the existing literature lacks comprehensive lead scoring models with solid theoretical foundations. Consequently, this study leverages real datasets in three service industries to develop a cascading predictive lead scoring model using a combination of clustering and classification methods. Moreover, this study relies on two theoretical foundations, namely the buyer purchase decision-making process from the lead's perspective, and the marketing-sales funnel from the seller's perspective. The empirical results of this study suggest the importance of integrating lead profiling into the lead scoring process, highlighting that data-driven models can enhance decision-making in B2B sales. The practical findings support sales practitioners with concise and high-quality lists of leads.

Keywords: Lead Scoring; Lead Profiling; Machine Learning; Marketing-Sales Funnel; Purchase Decision-Making Process

Note: This chapter is a working paper and will be submitted to the journal referred to above.

5.1 Introduction

In the rapidly evolving realm of the sales industry, competition has intensified, driven by technological advancements and changes in buyer behavior (Paschen et al., 2020; Syam & Sharma, 2018). Companies are challenged to enhance their performance to remain competitive, particularly in prospecting which is a critical initial stage involving lead generation (i.e., seeking leads that have the potential to become customers) and lead scoring (i.e., assessing leads' likelihood to become customers) (Paschen et al., 2020). However, identifying high-quality leads from a big pool is a significant challenge, as distinguishing the attributes of exceptional leads is inherently difficult (D'Haen et al., 2013; Ohiomah et al., 2019; Sabnis et al., 2013). The successful identification of these attributes typically happens after salespeople have made initial contact with leads, thus reducing the proactive effectiveness of the lead scoring process (Paschen et al., 2020). In addition, over the past decade, the traditional (i.e., "in-person", "face-to-face", "outside") sales function has undergone a substantial transformation, primarily influenced by advancements in digital technology and social shifts (Eitle & Buxmann, 2019; Paschen et al., 2020; Syam & Sharma, 2018). An

important factor behind this transformation is the increased significance and expansion of inside (remote) sales (Rutherford et al., 2014; Thaichon et al., 2018). The practice of inside sales involves remote engagement between salespeople and leads through information communication technology (ICT) (e.g., phone, Internet) (Kawas et al., 2013; Ohiomah et al. 2019, 2020). For some industries, inside sales (in the rest of the paper, sales means inside sales) has become the only way to sell products due to its advantages with regards to cost reduction and opportunity maximization (Wu et al., 2023).

The sales pipeline divides the sales process into multiple stages, namely suspects, prospects (aka "cold" leads), market-qualified leads (MQLs), sales-qualified leads (SQLs), and customers (Ang & Buttle, 2006; D'Haen et al., 2013; Patterson, 2007) (**Figure 5.1**). These stages represent a progressive process from the initial contact with a suspect to the final sales deal. The effective and efficient management of the sales pipeline is critical for sales organizations to enhance lead conversion rates, expand their market share, explore new customer markets, and improve overall sales performance (Bradford et al., 2016; Eitle & Buxmann, 2019). Traditionally, scoring and managing leads across the sales pipeline have been mostly guided by marketers' and salespeople's prior experience with leads, personal intuition, professional competence, and cognitive capability (Bohanec et al., 2017; D'Haen & Van den Poel, 2013), which might introduce subjectivity, personal biases, and misjudgment into the process (D'Haen et al., 2016; Eitle & Buxmann, 2019). Meanwhile, marketers and salespeople often disagree on the factors to consider to determine what lead to pursue out of a large pool of leads (D'Haen et al., 2013). Salespeople frequently claim that numerous leads generated by marketing lack the potential to convert into sales, whereas marketers criticize salespeople for their purportedly inadequate follow-up on leads (Järvinen & Taiminen, 2016; Smith et al., 2006). Consequently, despite efforts, salespeople often misallocate company resources by dedicating them to the wrong (aka. "cold") leads, resulting in fewer sales opportunities, reduced sales productivity, decreased revenue, and increased prospecting costs (D'Haen & Van den Poel, 2013; Smith et al., 2006).

Figure 5.1 appears here

The conventional intuition-based approach for qualifying and choosing leads in the sales pipeline has been criticized for its ineffectiveness and inefficiency (D'Haen & Van den Poel, 2013; Duncan & Elkan, 2015; Nygård & Mezei, 2020), leading to low conversion rates[4] and poor connection between marketing and sales (Sabnis et al., 2013; Smith et al., 2006), which eventually results in decreased sales performance. Relying on the subjective judgment, intuition, and experience of sales practitioners to predict which leads might convert is not an effective strategy for achieving sales success (Xu et al., 2017). The key challenge

[4] The average conversion rate of prospects to qualified leads is approximately 10% and merely 1-5% of leads will convert to sales and ultimately become customers (Coe, 2004; D'Haen & Van den Poel, 2013).

for sales practitioners is how to systematically and scientifically differentiate between high-quality and low-quality leads (Ohiomah et al., 2019). As a qualification approach, predictive lead scoring has emerged as an effective and efficient way of assessing the quality of a large number of leads (Duncan & Elkan, 2015).

Data-driven predictive approaches leverage data mining (DM) models and machine learning (ML) algorithms to uncover non-linear relationships between attributes and outcome, identify hidden patterns, and bring the most influencing attributes to the surface (Chorianopoulos, 2016; Duncan & Elkan, 2015; Nygård & Mezei, 2020). Indeed, predictive lead scoring models are an advanced tool for assessing the leads' likelihood of making a purchase by scoring and ranking them to differentiate and prioritize for sales (Benhaddou & Leray, 2017). Predictive lead scoring models can significantly enhance lead conversion rates and sales performance while minimizing the reliance on arbitrary rules by sales practitioners in the sales pipeline (Eitle & Buxmann, 2019). Although a few studies have applied ML-based lead scoring models to improve different stages in the sales pipeline (Bohanec et al., 2017; Duncan & Elkan, 2015), the literature remains limited in scope and lacks an adequate number of relevant studies. Specifically, we identified two predominant limitations in the existing literature.

The first limitation is *the absence of theoretical evidence on how predictive lead scoring models should be implemented.* Existing studies on lead scoring models are built on the foundation of the marketing-sales funnel (e.g., D'Haen & Van den Poel, 2013; Duncan & Elkan, 2015), which is based on the well-known seven steps of selling (Dubinsky, 1981). However, the marketing-sales funnel is seller-centric. The recognition of leads (or buyers) as the starting point of sales activities and the driver of the exchange process is strong evidence that the sales process has become increasingly lead-centric (or buyer-centric) (e.g., Shannahan et al., 2013; Sheth et al., 2009). Sellers have begun to consider how to better respond to leads' demands and purchasing behaviors, instead of aiming to influence them during purchase decision-making process. There is a lack of research on how to implement predictive lead scoring models that are both seller and lead-centric. This represents a unique theoretical research opportunity. By considering the potential buyer's perspective, we contend that sales performance (e.g., lead conversion rates) could be improved by integrating insights discovered from the phased purchase decision-making process (**Figure 5.2**) into the lead scoring modeling process.

Figure 5.2 appears here

The second limitation is *the effectiveness and efficiency of the current marketing-sales funnel are not optimized.* Traditionally, based on the marketing-sales funnel, the marketing function deploys lead scoring models to generate lists of MQLs which are continuously sent to salespeople for action (i.e., contact the leads). When salespeople are overwhelmed by the number of MQLs they need to contact, some MQLs may not be contacted at the best possible time. Research has shown that the best time to contact a lead is a

key factor in lead conversion (Ohiomah et al., 2018). If a salesperson misses the best time to contact a lead, then the company may lose a genuine sales opportunity. In addition, constantly putting pressure on salespeople to contact leads forces them to get hold of leads continually via all possible means. Based on the results of Wu et al. (2024), potential customers may not always in favor of being constantly contacted (i.e., pushed) during their purchase decision-making process. For example, leads without a management title do not like being constantly contacted during the early phases of the purchase decision-making process (Wu et al., 2024). Instead, they prefer to seek relevant product information first and being contacted by salespeople in the later phase when finalizing their purchase decisions. Thus, to improve the efficiency and effectiveness of the funnel, it is essential to correctly identify which phase a lead is currently at in order to engage with him/her in the most appropriate way, without delay. Current literature lacks studies that develop models in this regard.

This paper aims to address the abovementioned limitations while answering the following two calls for research. Syam and Sharma (2018, p.141) advocated for research into how ML and artificial intelligence (AI) can "*help develop systems that recommend customer engagement strategies based on analysis of past engagements and their success/failure and depending on salesperson, industry and customer characteristics?*" Ohiomah (2020, p.201) called for studies that explore leads' characteristics to build profiles of leads and investigate these profiles' potential applications in lead management. To that end, we apply lead profiling to develop a "smart sales" predictive lead scoring model that manages and processes leads more effectively. We define "smart sales" as an innovative way to lead prioritization that utilizes data-driven methods (e.g., ML and DM), to optimize sales performance. We demonstrate that sales performance can be enhanced by adopting lead profiling within the data-driven lead scoring process to support decision-making of sales practitioners. This paper highlights the role of ML and DM in shifting from traditional subjective to data-driven objective decision-making in sales.

Practically, this study contributes by integrating the lead profiling step into the predictive modeling process. We provide salespeople with concise and high-quality lists of sales leads to contact. Thus, by improving the lead scoring process, the effectiveness and efficiency of the marketing-sales funnel can be enhanced, which ultimately boosts sales performance. Theoretically, to the best of our knowledge, this is the first study that proposes lead scoring models using the marketing-sales funnel from a seller's perspective (D'Haen & Van den Poel, 2013), and the phased buyer purchase decision-making process from a lead's perspective (Puccinelli et al., 2009; Sirakaya & Woodside, 2005).

This paper is organized as follows. In Section 5.2 Background, we provide a brief overview of lead processing in the marketing-sales funnel and buyer purchase decision-making process, and review the existing literature on the application of DM models and ML algorithms in lead scoring. Section 5.3

Methodology details the data collection process and methodology. We present our analytical results in Section 5.4 Results. In Section 5.5 Discussion and conclusion, we discuss the theoretical and practical implications of our findings. We also acknowledge the limitations and suggest directions for future research.

5.2 Background

5.2.1 Marketing-sales funnel

The marketing-sales funnel approach (see **Figure 5.1**) is utilized for controlling the progression of potential customers throughout different phases of the sales process (Ang & Buttle, 2006; Patterson, 2007). In a typical marketing-sales funnel, a suspect goes through a series of steps which include lead generation (i.e., from suspect to prospect), lead scoring/qualification (i.e., from prospect to MQL), lead/opportunity management (i.e., from MQL to SQL), and exclusive management (i.e., from SQL to customer) (D'Haen & Van den Poel, 2013; Smith et al., 2006; Xu et al., 2017). Suspects who are identified from multiple channels (e.g., email, social media, search engines) upon leaving contact information, login, cookies, or IP address are deemed as prospects (i.e., "cold" leads). Then, the marketing team applies lead scoring models to evaluate and qualify these prospects based on predefined characteristics and predetermined thresholds of conversion probability, producing a ranked list of leads (i.e., MQLs) as output. MQLs contain a refined list of prospects who meet specific predefined characteristics. Next, salespeople contact MQLs in the refined list from top to bottom, utilizing various approaching and selling skills to pursue a deal. Some of these MQLs become SQLs after being assessed on their budget, authority, and requirements, while some will be nurtured further, and some will be discarded. Finally, customers are SQLs who end up making a purchase.

The literature suggests that, from a seller's perspective, in the lead generation stage, sellers initiate the sales process and encompass all marketing activities aiming at identifying leads from a pool of suspects through website visits, completion of information forms, attending webinars, and other similar approaches (Xu et al., 2017). During the lead scoring stage, sellers evaluate prospects' likelihood of purchasing by rating them on a scale, allowing for differentiation and prioritization in the sales process (Duncan & Elkan, 2015; Kawas et al., 2013). During the opportunity management stage, salespeople actively interact with and nurture leads through various communication channels such as email, phone, and SMS. Also, salespeople may arrange in-person or virtual meetings to present product demos to leads until they are successfully converted into sales (Kawas et al., 2013; Xu et al., 2017). On average, the entire sales cycle requires between three to nine months, contingent upon the complexity of the product, the industry, and the characteristics of leads. However, certain leads may take over a year to nurture before they can be successfully converted into sales (Syam & Sharma, 2018; Xu et al., 2017). Additionally, some leads are discarded from the funnel at a particular stage due to a lack of interest in the product. Some leads

successfully close and convert into sales, while the majority stay in the funnel until the seller eventually gives up and discards them (Bradford et al., 2016).

5.2.2 Buyer purchase decision-making process

The literature suggests that a buyer's purchase decision-making process can be modeled through a buying funnel, which is the process that a lead takes to purchase a product (Jansen & Schuster, 2011). As shown in **Figure 5.2**, the process consists of four phases, namely need recognition, information search, solution evaluation and selection, and purchase (Jansen & Schuster, 2011; Puccinelli et al., 2009; Sirakaya & Woodside, 2005). In the need recognition phase, the lead becomes aware of a need for products. In the information search phase, the lead confirms the need and begins searching for information on potential products that might address the need. In the evaluation and selection phase, the lead considers and compares specific offerings, brands, promotions, vendors, and conditions. After careful consideration and discussion, the lead finalizes the buying decision and makes the purchase. In this study, the purchase phase is not considered because a purchase decision has already been made.

As a foundation for this study, we align the buyer's purchase decision-making process with the marketing-sales funnel. As a result, we see potential buyers who become aware of a need in the need recognition phase as MQLs. Buyers who seek information on products that address the need (i.e., information search phase) are transitioning from MQLs to SQLs. After being contacted by salespeople, buyers who are in the process of considering and comparing different products, conditions, brands, vendors, and promotions (i.e., evaluation selection phase) are SQLs. Eventually, buyers who become customers are in the purchase phase.

5.2.3 Application of machine learning (ML) and data mining (DM) in lead scoring

Before the era of rapid advancements in digital technologies, big data, and AI, different views of the sales and marketing functions regarding lead scoring and processing have been studied in the literature (Sabnis et al., 2013; Smith et al., 2006). The different views particularly involved the qualification and pursuit of leads. Traditionally, lead scoring process in the sales pipeline mostly relied on salespeople's experience-driven predictions and intuitions. However, this traditional method has been criticized for its inefficiency and ineffectiveness (Wu et al., 2023). The evolution of AI, particularly ML, provides opportunities for strengthening the interface between the marketing and sales in the sales process (Järvinen & Taiminen, 2016), and data-driven predictive lead scoring models are one example (Ngai et al., 2009; Syam & Sharma, 2018). These models employ data-driven predictive analytics to uncover insights within "cold" leads data, reveal hidden relationships between predictors and outcomes, and calculate a score for each lead (Chorianopoulos, 2016; Nygård & Mezei, 2020).

Current studies have explored and categorized predictive lead scoring models into classification, regression, and clustering models (Ngai et al., 2009; Nygård & Mezei, 2020). A few classification ML algorithms have been used for building lead scoring models. Peng and Xu (2011) propose a model to identify potential leads by relying on the rules identified in a decision tree model. Moreover, as an ensemble learning method based on decision trees, random forest has been employed to categorize leads with social media data as input (Meire et al., 2017). It has also been utilized with explanation models to qualify and classify leads based on historical customer data (Bohanec et al., 2017). Yan et al. (2015) present a lead scoring model designed to forecast the likelihood of securing sales leads within a specified timeframe. They employ logistic regression to identify and quantify the impact of a salesperson's activities and lead information on sales outcomes (won or lost). Furthermore, Xu et al. (2017) implement a data-driven model utilizing linear and exponential regression, along with seasonal time series models, to estimate lead conversion rates and forecast sales revenue. Additionally, clustering algorithms such as Expectation-Maximization have been used to group potential prospects' websites according to terms that mainly appear on the websites of profitable prospects. The results of this clustering analysis help companies identify profitable leads (Thorleuchter et al., 2012).

We found that existing predictive lead scoring studies only focus on the conversion steps in the marketing-sales funnels from sellers' perspective, neglecting insights from the leads' purchase decision-making process. Additionally, the literature lacks studies on the segmentation of leads using clustering DM models. Hence, considering both the sellers and leads' points of view, this study integrates lead profiling into the implementation of a "smart sales" predictive lead scoring approach by using a combination of clustering and classification models.

5.3 Methodology

In this study, we apply a "smart sales" approach (i.e., clustering then classification) to build predictive lead scoring models. Given that leads exhibit varying levels of purchase intention at each phase of the sales process (Wu et al., 2024), it is crucial to pinpoint those with the highest likelihood of purchase in each phase to enable smarter and more efficient sales practices. Accordingly, considering the perspective of leads and utilizing an unsupervised clustering method, we initially segment leads into groups. Each group consists of similar leads who share higher levels of purchase intention specific to a phase. Subsequently, we use the insights from this unsupervised clustering to inform a supervised classification approach, aiming to predict the conversion probabilities of leads and rank them. For instance, the "classification only" approach in the current literature generates a singular and lengthy list of ranked leads. Conversely, our approach starts by categorizing leads into demographically similar groups. We then apply classification within each cluster ("clusters", "groups", and "profiles" are used interchangeably in this paper) for model training and predict

the conversion probabilities of the leads. Ultimately, this process allows us to produce multiple shorter and high-quality lists of leads from these clusters, ranked by their conversion likelihoods. The results aim to provide suggestions for salespeople when choosing which sales leads to approach next. Next, we delve into the data used in this study, describing the crucial variables and terms related to the research design.

Incorporating leads' demographic attributes directly into a single classification model is also a valid approach, however, we propose the "clustering then classification" approach because this tailored approach can result in higher accuracy and better predictive performance within each cluster (Thomassey & Fiordaliso, 2006; van Steenbergen & Mes, 2020; Wang, 2018). Additionally, the proposed approach can better handle the class imbalance issue that exists in datasets (Chorianopoulos, 2016). Furthermore, it could discover hidden patterns and relationships between variables within each cluster, which can be useful for further analysis (Han et al., 2012). Lastly, salespeople can benefit from this approach by specializing when engaging with particular groups of leads.

5.3.1 Data collection and preparation

We were provided with nearly 8 million data records by our industrial partner, a top North American Lead Management System (LMS) software provider and a marketing automation company. The dataset includes data from 40 inside sales companies covering three service Business-to-Business (B2B) industries, namely legal consulting service, commercial office building cleaning service, and custom payment solution service. We concentrate on these three service industries as the insights derived from them could potentially be generalizable to the broader service sector. Each industry is represented by two datasets (leads' characteristics and contact history). Together, the datasets encompass 19 fields. The contact history datasets capture call activities between salespeople and leads from February 2011 to March 2024. The salespeople used LMSs for their activities (including lead qualification, nurturing, the sales proposal presentation) with leads.

We merged the leads information and the call history datasets and pre-processed the data to eliminate redundant variables and irrelevant anomalies, impute missing values, and decrease the excessive number of categories within certain categorical fields. Furthermore, we derived new variables through calculations that involved existing variables. The process of data transformation produced a more concise and enhanced dataset, which could decrease the complexity of data analysis and enrich the generalizability of our conclusions (Yu et al., 2006). We utilized Python Spyder IDE (version 5.4.0) for data preparation, model development, analysis, and visualization. **Table 5.1** describes the variables employed in this study. There are six variables on leads' demographics (i.e., revenue, title, company size, management level, primary and secondary lead source), and three variables on salesperson effort (i.e., number of calls, total duration, and result group). Specifically, in the B2B legal consulting service industry, titles include admin,

assistant, attorney, associate attorney, legal secretary, marketer, office manager, owner, and partner. In the B2B cleaning service industry, titles contain chairman, controller, member, owner, partner, and president. In the B2B payment solution service industry, controller, owner, point of contact, president, and treasurer. Management levels include C-suite, director, non-manager, and manager across all three industries. Company size is categorized into small, medium, and large-sized. Primary lead sources contain Sage, Quickbooks, Intuit Campaign, Skyline Campaign, Hoovers, Hunter, and Zoominfo. Secondary lead sources include Cognism, Zoominfo, and Zoominfo-Salesintel. These lead sources indicate the lead platforms through which the leads are acquired.

Table 5.1 appears here

Next, we selected the Result Group variable from the data to consider the outcome (positive or negative) of activities with leads. Originally, the data had 6 different outcomes. After consulting with a data analyst from our industrial partner, we categorized outcomes into three possible values: *no contact*, *negative contact*, and *positive contact.* These outcomes align with Yan et al.'s (2015) categorization of sales outcomes. An activity outcome is categorized as *no contact* if nobody answered the phone call; the phone line was busy or disconnected; the contact was unreachable; the contact information was incomplete or inaccurate; another contact details were provided; or the lead requested a call back or send information. Leads with this outcome may still be undergoing qualification, with ongoing follow-up calls and contact attempts by the salesperson. An outcome is labeled as *negative contact* if a salesperson was unbale to qualify or convert a lead during the sales process. In contrast, a *positive contact* outcome includes qualified sales leads who have demonstrated interest, made requests, booked appointments/demos, etc. Additionally, *positive contact* includes cases that a sales goal was achieved, such as the successful sale of a service/product.

Moreover, we filtered the data to only include relevant instances where outcomes ended in either a *positive* or *negative contact*, excluding those with a *no contact* outcome. This enables us to identify the optimal lead approaching strategies that may lead to a *positive contact* outcome at the end of the sales process. Hence, we prioritize leads that eventually resulted in either a *positive contact* or *negative contact*, as leads may transition through various stages of the sales process with different outcomes. For example, a salesperson may initially contact a lead but only reach their voicemail (i.e., outcome: *no contact*). Subsequently, upon successfully connecting with the lead in the following day, the salesperson secures an appointment for a product demo, resulting in a *positive contact* outcome. However, after the demo, the lead decides not to proceed with purchasing the product, which results in a *negative contact* outcome. In such case, the initial outcome (*no contact*) differs from the final outcome (*negative contact*) for various reasons (e.g., budget limits, product mismatch). Activities with a last outcome of *no contact* were excluded from

this study as they fall outside the scope of the research. To ensure the selection of proper data for our analysis, we implemented the following procedures.

First, we mapped out the last activity instance with each lead in our dataset. Next, we filtered out leads whose final outcome was *no contact* and who had been contacted within 365 days of the data collection date (2024-03-28). This step eliminates data associated with leads who have not completed sales cycle yet. This approach aligns with the literature suggesting that the sales cycle for certain leads may extend up to one year (Syam & Sharma, 2018; Xu et al., 2017). Then, we selected leads that ended in either a *positive* or *negative contact* outcome, excluding those contacts where the last outcome was *no contact*. Finally, we matched all the leads with a final *positive* or *negative contact* to the original dataset. This allowed us to compile a dataset focused on leads that ended in either a *positive* or *negative contact* outcome, even if they had *no contact* outcome instances as they went through the sales process.

5.3.2 Machine learning algorithms

To assign leads to their corresponding profiles, we first need to establish distinct groups. This becomes an unsupervised clustering problem. Clustering segments a heterogeneous sample into several homogeneous groups (Chorianopoulos, 2016; Ngai et al., 2009). Initially, we applied clustering to split the leads into groups based on demographics. This step includes a partitioning clustering algorithm using K-Prototypes (Gupta et al., 2022; Ho et al., 2023; Meena et al., 2023), a hierarchical clustering algorithm using an agglomerative (AGNES) approach (Othayoth & Muthalagu, 2022; Reynolds et al., 2006), and a density-based clustering algorithm using Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Dalmaijer et al., 2022; Han et al., 2012).

K-Prototypes clusters data into K distinct groups by assigning each instance to the nearest cluster while keeping them as separate as possible (Gupta et al., 2022). Unlike traditional partition-based algorithms such as K-means, K-Prototypes handles both numerical and categorical data, which suits our dataset. It calculates cluster centroids utilizing the mean for numerical data and the mode for categorical data. The algorithm iteratively assigns data points to the nearest cluster according to a combination of Euclidean distance and categorical dissimilarity, stabilizing when centroids no longer change. We chose K-Prototypes for its effectiveness with mixed data types and computational efficiency. AGNES hierarchical clustering starts with each data instance as its own cluster, then progressively merges clusters based on distance or dissimilarity until one cluster remains or a stopping criterion is met. The merging process is visualized with a dendrogram which helps understand the structure and decide where to cut the tree to form clusters. We chose AGNES for its clear visualization and ability to identify outliers (Reynolds et al., 2006). Unlike the spherical clusters of K-Prototypes, DBSCAN forms clusters based on dense areas of data points and can create clusters of various shapes and sizes. Starting from a core point, DBSCAN expands the cluster

by including all nearby dense points and stops when no more points can be added (Han et al., 2012). We chose DBSCAN for its flexibility in cluster shapes and ability to detect and isolate noise points (Dalmaijer et al., 2022).

Furthermore, we trained individual classifiers on each identified group to predict the conversion probabilities (i.e., scoring leads) and generate prioritized lists. Specifically, we implemented lead scoring models by using several independent variables to predict and classify the outcome of activities (i.e., *positive* or *negative contact*) with leads in the sales pipeline. In essence, this forms a supervised classification ML problem, wherein a model is trained using labeled observations and subsequently applied to map and categorize new instances (Chorianopoulos, 2016; Kotu & Deshpande, 2015). We applied various supervised ML algorithms including a decision-based classifier using Decision Trees (Başarslan & Argun, 2020; Peng & Xu, 2011), and an ensemble-based classifier using the Gradient Boosting Trees (Duncan & Elkan, 2015; Eitle & Buxmann, 2019), a regression-based classifier using the Logistic Regression (D'Haen & Van den Poel, 2013), and a probabilistic-based classifier using the Naïve Bayes (Benhaddou & Leray, 2017). The decision to use these algorithms is rooted in their suitability for addressing the specific research problem and dataset, as recommended by the existing literature (Wu et al, 2023).

Decision Trees (DT) algorithm scans the dataset and creates a tree by grouping similar attribute values. This process divides the dataset into subgroups of similar instances (Xia et al., 2020). We chose DT because it is non-parametric, which means it doesn't assume a specific data distribution, making it versatile for various data types (Bastı et al., 2015; Olson & Chae, 2012). Gradient Boosting Trees (GBT) combine multiple classifiers to create a more accurate and consistent ensemble classifier. The algorithm starts with a simple tree to make initial predictions, estimates errors, and then trains new trees to predict these errors, improving accuracy at each step. We chose GBT for its high classification accuracy and ability to handle different data types and model complex relationships (Duncan & Elkan, 2015). Logistic Regression (LR) is a popular method for modeling relationships between predictor variables and discrete outcomes (Kotu & Deshpande, 2015). It calculates the probability of an event based on predictor values using the logit or sigmoid function. We chose LR because it explains variations in outcomes and assesses the significance of each predictor variable (Banerjee et al., 2017). Naïve Bayes (NB) is a probabilistic classifier that uses Bayes' theorem to estimate probabilities based on frequency and value combinations within the dataset (McCallum & Nigam, 1998). It assumes that the effect of a predictor's value on a class is independent of other predictors. We chose NB for its speed, ease of implementation, and suitability for large datasets (Wu et al., 2023).

5.3.3 Model evaluation and imbalanced dataset

In this study, we implemented a ten-fold cross-validation method to evaluate and compare the performance of the chosen algorithms. The application of a ten-fold cross-validation is supported by existing literature

as a robust approach for evaluating models (Başarslan & Argun, 2020; Kohavi, 1995). Besides employing a ten-fold cross-validation, we employed standard metrics to assess model performance. Because our datasets suffer from class imbalance, there exists the potential for models to effectively classify instances belonging to the majority class, which may lead to deceptive high classification accuracy. This class imbalance may also result in a trade-off between specificity and sensitivity for some models (Napierala & Stefanowski, 2012). Thus, this research also considers balanced accuracy and F1-score when comparing model performance. Furthermore, from the data perspective, *positive* and *negative contact* instances are rare but typically display more unique features compared to *no contact* instances. However, *no contact* instances are more frequently seen and generally possess fewer distinctive features that distinguish them from *positive contact* and *negative contact* instances. Also, data preparation method outlined above might alter the data distribution.

In response to the data imbalance issue, we utilized three resampling techniques to resample the data. First, we applied the oversampling technique which randomly inserts instances to the minority class to produce a balanced dataset. Second, we used the undersampling technique that randomly removes instance from the majority class to generate a balanced dataset (Chorianopoulos, 2016; Jo & Japkowicz, 2004). Finally, we conducted the Synthetic Minority Oversampling Technique (SMOTE) which is a popular resampling technique using the k-nearest neighbor method to create random synthetic instances for the minority class, while simultaneously deleting samples from the majority class until both classes reach a comparable data sample proportion (Chawla et al., 2002). The literature advocates that the oversampling technique could result in model overfitting and undersampling might remove instances that are crucial for the classification objective (Napierala & Stefanowski, 2012). The SMOTE technique facilitates in generating an alternative balanced dataset.

Because the labels in the datasets are binary outcomes (i.e., positive or negative contact), when evaluating the balanced datasets (after resampling) we selected a threshold of 0.5 to round probabilities in the balanced datasets, assuming the model is equally likely to predict either class (positive or negative), thus adopting a balanced approach. However, for imbalanced datasets (before resampling) using a 0.5 threshold could lead to poor performance, particularly for the minority class (positive contact). To improve the balance between precision and recall for both classes, it is necessary to adjust the threshold. We employed the precision-recall curve/trade-off to determine an optimal threshold by selecting the point that maximizes the F1-score, which is the harmonic mean of precision and recall (Chorianopoulos, 2016). This approach ensures equal weight is given to both precision and recall.

Additionally, we evaluated the model's sensitivity using feature importance analysis (tree-based method) and one-at-a-time (OAT) sensitivity analysis (Chorianopoulos, 2016; Han et al., 2012). The goal

of running sensitivity analysis was to examine how changes in input variables affect the output of a predictive model. This helps identify which features or parameters have the most significant influence on the model's predictions. Since our predictive lead scoring model is tree-based (i.e., gradient boosting trees), we used the model's built-in feature importance measures, which assessed the reduction in impurity caused by each feature, indicating how important each feature was for making predictions. We ranked the features based on their importance to the model's predictions. In the OAT sensitivity analysis, we adjusted the value of one feature at a time, while keeping other features constant, to observe its impact on the output (i.e., lead conversion probability). We repeated this process for other features to identify the ones with the greatest impact. Our analysis revealed that the number of calls, title, and management level of leads were the top three influential input features.

Furthermore, we assessed the robustness of our model using two tests: out-of-sample testing and cross-validation with shuffling (Chorianopoulos, 2016; Mohammed, 2017). The goal was to evaluate how well our predictive lead scoring model performs when exposed to uncertainties, noise, or variations in input data. In the out-of-sampling testing, we compared the model's performance across the training sets (**Table 5.4 - Table 5.6)** and the new unseen test sets (**Table 5.4 - Table 5.6)** to detect any overfitting. Because the performance gaps were small (0-5%) (Chorianopoulos, 2016) across all three industries, the models did not show a clear sign of overfitting. This was further verified in the cross-validation test. We tracked how performance metrics varied across different shuffled data sets. A model that demonstrates stable performance across all shuffles is considered more robust. Our models showed consistent and reliable results under various data conditions, which is critical for lead scoring applications where real-world data is often noisy, incomplete, or subject to change over time.

5.4 Results

After data pre-processing, the final data sample comprised data from 40 companies across 3 service industries. Initially, our datasets contained approximately 8 million records from these 40 companies. Following the procedure outlined in Section 5.3 Methodology, we ended up with a sample of 3,448,111 call interactions with 784,493 leads. We consider each interaction with a lead as an individual instance in our datasets. The lead distribution in industries is summarized in **Table 5.2**.

Table 5.2 appears here

5.4.1 Modeling procedure

First, we applied clustering modeling algorithms to split leads into clusters that reflect similarities among leads in each group. We only consider leads' demographic features when forming these clusters as sellers typically have access to only this information prior to initiating contact. We used the "Elbow" method to

find the candidates of the optimal number of clusters (Syakur et al., 2018) and experimented with various numbers of clusters when building clustering models.

Second, we built classification models using four classifiers: DT, GBT, LR, and NB. We divided the dataset by randomly choosing 70% of the records as the training dataset, with the remaining 30% held as the unseen testing dataset for assessing the performance of each model. Then, we applied oversampling, undersampling, and SMOTE resampling techniques on the training dataset to address the class imbalance issue. We built each model using 10-fold cross-validation and repeated the process ten times to enhance the stability and reliability of model performance. We applied resampling techniques to the training folds during each iteration of cross-validation, leaving the validation fold untouched. This ensures the model's evaluations remains valid and avoids data leakage.

5.4.2 Modeling results

To pick the optimal number of clusters for each industry, we evaluated the clusters' quality using the Silhouette coefficient (Chorianopoulos, 2016; Han et al., 2012) and Davies-Bouldin index (Davies & Bouldin, 1979). These metrics help examine how well-separated and compact the clusters are. Typically, a higher Silhouette score (-1 to 1) and a lower Davies-Bouldin score (0 to 1) indicate better qualify of clustering results. We trimmed the dendrogram produced by AGNES to achieve the same number of clusters as K-Prototypes. Finding the optimal DBSCAN clustering result is challenging. It requires fine-tuning hyperparameters, specifically "EPS" (the maximum distance between two points for them to be considered neighbors) and "min sample" (the minimum number of points in a neighborhood for a point to be considered as a core point) (Han et al., 2012). We experimented with different parameter combinations and picked the best ones based on the Silhouette and Devis-Bouldin scores. We employed DBSCAN with the best hyperparameters to group the instances.

Because the evaluation scores are close, we performed the Kruskal-Wallis statistical tests (Dalmaijer et al., 2022) to assess whether the Silhouette coefficient scores were statistically significantly different from one another. The same test was applied to the Davies-Bouldin scores to examine the statistical power for clustering analysis. We extracted individual scores for each cluster generated by various clustering algorithms and performed the statistical tests. The results indicated insufficient statistical power to conclude that any of the scores were significantly different. So, we tried other methods to determine the optimal number of clusters. We generated cluster graphs to visually assess which clustering setting produced the best results (i.e., where data points were closely grouped within each cluster, while the clusters themselves were well-separated). Based on the "Elbow" plot method (Syakur et al., 2018), cluster graph, and the evaluation scores, we determined the optimal number of clusters for each industry. **Table 5.3** shows the performance of the three distinct clustering algorithms with different numbers of clusters in

all three industries. For the B2B legal consulting service industry, K-Prototypes (k = 3) has the best evaluations scores, so three was chosen to be the optimal number of clusters and leads were partitioned accordingly. For the B2B cleaning service, despite a slightly higher evaluation score when k = 6 of K-Prototypes, we chose k = 5 because of a smaller number and an overall better quality of clusters. Finally, for the B2B custom payment solution service, K-Prototypes (k = 3) was chosen to be the optimal number of clusters.

Table 5.3 appears here

We first computed the lead Conversion Rate (CR) in the training dataset. The lead CR is a widely utilized metric to assess the effect of lead scoring models on sales performance (Wu et al, 2023). It is computed by dividing the total number of conversions by the total number of leads (D'Haen et al., 2016). Ideally, the goal of salespeople is to contact every lead and try to convert him/her into a sale. However, in reality, a salesperson is typically unable to go over the entire list when contacting leads due to restricted resources that can be allotted (Ohiomah et al., 2018). Additionally, some leads have higher conversion probabilities than others, which can be easily determined by examining their basic information (Monat, 2011). Thus, to replicate a real sales scenario, we only considered half of the testing dataset in the following evaluation comparisons. We randomly selected 50% of the testing dataset as the random policy (to mimic which leads that a salesperson would choose to contact). The resulting CR under this random policy was 7.24%. Note that this is the CR without employing lead scoring in the legal consulting service. We applied the same calculations to find the CRs under random policy (without employing lead scoring) for the other industries (i.e., cleaning service and payment solution service), and they are 68.26% and 2.70%, respectively.

The conversion rates of the three service industries examined in this study differ due to a few reasons inherent to each industry's nature (complexity and risk), expectations of leads, and decision-making process. The legal consulting services usually involve a complicated process, specialized knowledge, and require service providers to possess adequate expertise and experience. Leads need to trust these service providers. This trust-building process can result in longer decision-making cycles and lower conversion rates (cite). Also, the significant impact of legal issues on businesses makes leads more cautious in their selection process. This further contributes to lower conversion rates as leads may spend more time comparing and selecting service providers. Similar to the legal service industry, in the custom payment solution industry, payment solutions often need to be integrated with existing systems, which can be a complex and technical process. Businesses need to ensure compatibility and reliability of the payment solution, which may need a longer decision-making period and result in lower conversion rates. Additionally, custom payment solutions need to comply with various financial regulations and security

standards, which adds another layer of examination to the selection process. This can result in lower conversation rates. Moreover, the nature of payment solution industry involves ongoing fees such as transaction-based charges. Companies need to consider the cost implications carefully, which can slow down the decision-making process and influence conversion rates. In contrast, the commercial office building cleaning services are typically more standardized compared to the abovementioned industries. This standardization makes it easier for leads to compare providers based on price, availability, basic service offerings, and promotions. This can result in higher conversion rates. In addition, the impact of selecting a cleaning service provider is generally less critical than choosing a legal advisor or payment solution provider, which leads to quicker decision-making and higher conversion rates. Furthermore, the recurring needs of cleaning services can encourage faster decision-making and higher conversion rates. Understanding these differences among distinct industries helps explain the variations in conversion rates.

Table 5.4 - Table 5.6 demonstrate the performance of the data resampling techniques using the "classification only" lead scoring approach for each classifier on the testing datasets. For the legal consulting service, the GBT classifier with SMOTE resampling demonstrated the best performance, achieving the highest balanced accuracy and F1-score with a CR of 13.67%. The oversampling and undersampling models generated similar results, while SMOTE models demonstrated the best performance. For the cleaning service industry, from an all-around perspective, the GBT classifier with SMOTE sampling was the best performing model. It demonstrated the second-best CR of 74.55% with decent model performance scores. For the payment solution service, the GBT classifier with undersampling technique was the best performing model. It shares the best CR (5.39%) with other resampling methods, yet it has the best model metrics. In general, the NB classifier demonstrated the worst overall performance. Note that the conversion rates in these tables are based on the top 50% of the testing datasets (leads who are most likely to convert) to make reasonable and meaningful comparisons.

Table 5.4 appears here

Table 5.5 appears here

Table 5.6 appears here

Next, we compared the performance of the "classification only" lead scoring model from the current literature to our proposed approach (i.e., clustering then classification). Specifically, we applied the best performing classifier on the segmented datasets (i.e., resulting clusters). For each industry, the number of clusters/lists of leads is the same as the optimal number of clusters. In addition to the metrics employed above, we took into account the cost of false positives when choosing the best performing model in the subsequent analysis. False positives (FPs) occur when our model predicted a "yes" outcome, however, the actual result was "no". In our context, FPs refer to leads that our model identified as likely to convert (i.e.,

positive contact), but they ultimately did not convert (i.e., negative contact). These FPs were eventually contacted by salespeople, resulting in a negative outcome. Usually, a cost is associated with these negative contacts and varies depending on industries. For simplicity, we considered the cost to be $8 USD per FP across all three industries in this study (Lima, 2018). Generally, a lower cost of FPs indicates a more cost-friendly model. Since GBT was the best performing classifier across all three industries, we focus on lead scoring models generated by GBT.

Table 5.7 - **Table 5.9** show that the lead scoring approach proposed in this paper outperforms the original "classification only" approach in the literature on the testing datasets across all three industries. We considered all measuring metrics in the following analysis. In **Table 5.7**, GBT cluster 0 model demonstrated a lower cost of FP than the "classification only" model, and the former demonstrated a higher CR. Moreover, the GBT cluster 2 model exhibited the best values in performance metrics including balanced accuracy, F1-score, and sensitivity. In **Table 5.8**, GBT cluster 2 demonstrated the best values in all measurements. **Table 5.9** shows that GBT cluster 1 demonstrated the highest CR and lowest cost of FP with better model performance than the original "classification only" model.

Table 5.7 appears here

Table 5.8 appears here

Table 5.9 appears here

Studies in the existing literature indicate that the CR boosts by 5% on average as the result of employing predictive lead scoring (i.e., "classification only" approach in literature) (D'Haen & Van den Poel, 2013; Duncan & Elkan, 2015). The following underlines the superiority of the proposed lead scoring modeling approach. In the legal consulting service industry, the top-performing model (GBT cluster 0) reached a CR of 19.30%, which represents an increase of over 5% compared to 7.24% (CR under random policy) and 13.67% (CR of the "classification only" approach). For the cleaning service industry, 85.65% (CR of GBT cluster 2) is more than the average 5% increase compared to 68.26% (CR under random policy) and 74.62% (best CR under "classification only" approach). Finally, for the payment solution service industry, 14.55% (CR of GBT cluster 1) is higher than the average 5% increase compared to 2.70% (CR under random policy) and 5.39% (best CR under "classification only" approach). In summary, the results of our proposed lead scoring approach indicate that the combination of K-Prototypes and the GBT classifier is the top-performing predictive lead scoring model.

5.4.3 Profiling the clusters

After generating the clusters, we used them on profiling leads according to their demographics, which is one of the contributions in this study. To build profiles from the leads' perspective, we delved into each cluster to study the patterns. First, we inspected the group distribution of clustering. **Figure 5.3** shows the

number of leads in each cluster across all three industries. Next, we checked the detailed distribution of leads' titles, management levels, and corresponding company sizes. After forming these clusters and examining their demographics, we profiled each cluster and concluded the most indicative demographics in each cluster. We created profiles for each cluster in each industry by identifying the most frequent category (mode) for each demographic and used these profiles for subsequent analysis.

In **Figure 5.4 - Figure 5.6** for the legal consulting service industry, cluster 0 mainly contains leads with partner title, c-suite (i.e., CEO, CFO, CIO, etc.) management level, and working in small-sized companies. Cluster 1 contains leads with partner title, non-manager level, and working in large-sized companies. The majority in cluster 2 are leads with titles such as attorney and legal secretary who are managers working in small-sized companies. In conclusion, the three leads' profiles deduced as a result of clustering are cluster 0 (partner; c-suite; small-sized company), cluster 1 (partner; non-manager; large-sized company), and cluster 2 (attorney/legal secretary; manager; small-sized company).

Figure 5.3 appears here

For the cleaning service industry, the five profiles are cluster 0 (member; non-manager; small-sized company), cluster 1 (partner; non-manager; small-sized company), cluster 2 (controller; c-suite; small-sized company), cluster 3 (president; non-manager; large-sized company), and cluster 4 (controller; manager; small-sized company). Finally, for the payment solution service industry, the three profiles are cluster 0 (controller/treasurer; manager; small-sized company), cluster 1 (owner; c-suite; small-sized company), and cluster 2 (controller; manager; small-sized company).

Figure 5.4 appears here

Figure 5.5 appears here

Figure 5.6 appears here

5.5 Discussion and conclusion

Over the last two decades, decision-making in the sales pipeline has predominantly relied on methods that emphasize individual judgement, professional competence, and prior experience. Especially in lead scoring, such a subjective strategy could lead to low conversion rates and result in wasted company resources by allocating them to the wrong leads (D'Haen & Van den Poel, 2013; Monat, 2011). To resolve this issue, data-driven and ML-based lead scoring models have emerged as a solution to transfer most decision-making to predictive models. However, the existing literature lacks studies that offer theoretical evidence on how such models should be implemented to produce effective lists of leads to pursue.

In this study, utilizing one small dataset and two large datasets from three service industries, we developed a "smart sales" predictive lead scoring models using a cascading approach (i.e., clustering then classification) based on two theoretical foundations (see **Figure 5.1** and **Figure 5.2**). We first applied

clustering to determine groups of leads. Each group (i.e., profile) comprises similar leads of higher purchase intention at a particular phase in the sales process. Next, we identified patterns in each profile generated by the clustering modeling. Then we built separate prediction/classification models to predict the conversion likelihood of leads in each profile, therefore providing salespeople with prioritized lists of leads to pursue. To contact leads effectively, salespeople need to proactively approach leads who have a higher purchase intention at each phase of the sales process. We experimented with different combinations of clustering and classification algorithms.

Our model evaluation indicates that the combination of the partitioning-based K-Prototypes clustering and the ensemble-based classifier using Gradient Boosting Trees generated the best performing "smart sales" predictive lead scoring model. Conversely, the probabilistic-based classifier (i.e., NB) did not generate the most ideal model. Next, we discuss the interesting findings, research implications, contributions, limitations, and future research directions.

5.5.1 Implications

The problem with the lead processing as described in Section 5.2.1 **Marketing-sales funnel** and **Figure 5.1** is that it incites salespeople to contact leads with the aim of converting them, which may not be appropriate. According to Wu et al. (2024), not all potential customers are in favor of being constantly contacted during their purchase decision-making process. To improve the effectiveness of lead scoring in the sales process, we first built profiles of leads, recognizing that different profiles of leads exhibit distinct purchase intentions at each phase of the sales process (Wu et al., 2024). For example, a category of leads is inclined to make a purchase commitment at an early phase while another category may exhibit the highest purchase intention at the final phase. These profiles, formed by demographic characteristics, can effectively reflect the most representative leads with the highest purchase intention at each phase of the sales process. Without taking interaction data into consideration, we first applied clustering to the leads' demographic attributes to identify representative clusters and build profiles of leads. Each profile contains leads who are prone to make a buying decision at a particular phase in the sales process.

In the legal service industry, cluster 0 (3800 leads) titled as partners with c-suite management level who work for small-sized companies have the highest CR (19.30%). In the cleaning service industry, cluster 2 (6659 leads) titled as presidents with c-suite management level who work for small-sized companies have the highest CR (85.65%). In the payment solution industry, cluster 1 (12764 leads) titled as owners with c-suite management level who work for small-sized companies have the highest CR (14.55%). The three most prominent clusters display notable similarities. Individuals categorized as partners, presidents, or owners exhibit a certain degree of authority in purchase decision-making, and they all hold c-suite positions withing small-sized companies. Indeed, these leads are most likely to make a purchase commitment at an early

phase in the sales process as suggested by Wu et al. (2024). Thus, these lists of leads should be treated as a top priority by salespeople.

Furthermore, in the legal service, cluster 2 (18360 leads) who are titled as attorney/legal secretary with manager level working in small-sized companies have the second highest CR (13.45%). For the cleaning service, cluster 4 (26626 leads) who are titled as controller with manager level working in small-sized companies have the second highest CR (84.25%). For the payment solution service industry, cluster 2 (110195 leads) who are titled as controller with manager level working in small-sized companies have the second highest CR (5.89%). According to Wu et al. (2024), leads characterised as having a management level and a moderate level of authority in purchase decision-making, particularly those employed in small-sized companies, are more inclined to make a purchase commitment in the interim phase. Thus, salespeople should engage with these lists of leads next.

Finally, when leads are reaching the final stage in the funnel, for the legal service, the final list of leads that salespeople should engage with is cluster 1 (2730 leads) in which the majority are titled as partners with non-manager level working in large-sized companies. For the cleaning service industry, cluster 0 (128091 leads) with non-manager level working in small-sized companies should be engaged last by salespeople. In the payment solution service industry, cluster 0 (333276 leads) with treasurer title and non-manager level working in small-sized companies should be engaged last by salespeople. Leads with the non-manager level attribute tend to make a decision when they attain the final phase of the sales process (Wu et al., 2024). Hence, salespeople should not treat them as priorities.

Indeed, the demographic attribute “non-manager” indicates that leads do not usually make purchase decisions. Additionally, the “large-sized company” attribute in cluster 1 within the legal service suggests that these leads are not prepared to make an early purchase decision. This is partly due to the complexity of purchase decision-making in large organizations, which typically requires multiple approvals from different stakeholders (Zolkiewski et al., 2017). Therefore, these leads should be contacted last. As Wu et al. (2024) advocates, regardless of salespeople’s efforts, these leads are not set to make a buying decision early in the sales process. Instead, salespeople should prioritize their efforts on trying to convert leads in clusters 0, 2, and 1, which correspond to the legal consulting, cleaning, and payment solution industries, respectively. This prioritization is warranted for several reasons. Firstly, leads in these lists exhibit the highest CR. Secondly, leads within these lists are usually authorized with a certain level of purchasing power, as indicated by their titles, roles, management levels, and affiliation with small-sized companies, enabling them to make purchase decisions during the early and interim phases. Salespeople stand a higher chance of converting these leads by engaging with them at the right moment. Thirdly, from a salesperson’s

perspective, working with concise lists of high-quality leads entails less stress compared to working with long lists.

As illustrated in the marketing-sales funnel (see **Figure 5.1**), the size of pools shrinks progressively from prospects to customers, mirroring the funnel's inherent nature of filtering out unqualified leads at each stage. To enhance the sales process and deliver high-quality lists of leads to salespeople, it is important to optimize the quality of these pools. A key objective is to enhance pool efficiency and effectiveness by mitigating embedded noise. The noise in the pool consists of low-quality leads initially predicted to convert (i.e., *positive contact*), however, failing to do so (i.e., *negative contact*), eventually resulting in false positives (FPs). Consequently, the lower the cost associated with FPs, the higher the pool quality. In the legal consulting service industry, the cost of FPs ($560) associated with the best list (i.e., cluster 0) is lower compared to the "classification only" method ($720), Similarly, in the cleaning service industry, the top-performing list (i.e., cluster 2) boasts the lowest cost of FPs ($688). Furthermore, in the payment solution service industry, the best performing list (i.e., cluster 1) exhibits the lowest cost of FPs ($4,296). Hence, it can be inferred that our proposed "smart sales" approach outperforms the original "classification only" method in reducing the embedded noise associated with these pools and optimizing pool quality.

The assessment of the proposed "smart sales" predictive lead scoring model, coupled with the observed enhancements in lead conversion rates and reductions in costs associated with false positives, confirms the efficacy and viability of incorporating clustering prior to classification modeling in the lead scoring task to identify high-quality leads. The notable improvements in the performance metrics of the proposed cascading models (i.e., clustering followed by classification) across all sampled industries underscore the potential for significantly enhancing lead scoring effectiveness. The insights discovered from leads profiles built based on the buyers' purchase decision-making process can be integrated into lead scoring to improve sales performance.

We conclude that conducting lead profiling first, followed by predicting conversion independently within each list of leads would offer a promising approach for optimizing lead scoring outcomes. Therefore, this paper fills the literature gap of theoretical evidence on how predictive lead scoring models should be implemented by suggesting a "smart sales" approach (clustering then classification) based on two theoretical foundations from the perspectives of both sellers and buyers.

5.5.2 Research contributions

The results of this study contribute to both the theory and practice of lead scoring within the domain of inside sales. Rapid progress in digital technologies, including artificial intelligence, machine learning, and data mining has dramatically altered the decision-making process within organizations. Our analysis reveals

that predictive analytical models significantly and positively transform the lead scoring activity during the prospecting phase of the sales process.

First, to the best of our knowledge, this is the first study that uses two theoretical foundations for modeling lead scoring, the marketing-sales funnel from the seller's perspective and the purchase decision-making process from the lead's perspective. The existing literature on lead scoring models only focuses on the conversion steps in the marketing-sales funnel and only considers the seller's perspective. The empirical evidence from this study proves that sales performance could be further improved by integrating insights from lead profiling into the scoring process. Thus, we answer the research call made by Ohiomah (2020, p.201) on exploring the profiles of leads and investigating these profiles' potential application in the lead scoring task.

Second, by investigating leads' demographics and interactions with salespeople, we answer the research call on how ML and AI can assist in developing models that suggest customer engagement strategies by analyzing past activities and their outcomes, taking into account salesperson, industry, and potential customer characteristics (Syam & Sharma, 2018, p.141). We reveal that potential customers' characteristics and salespeople's performance attributes are significant predictors of lead conversion outcomes. The empirical validation from this study supports the adoption of a "smart sales" predictive lead scoring method to enhance the efficiency and effectiveness of the sales process. This research highlights the pivotal role that data-driven predictive machine learning models play in shifting from traditional human decision-making to a more model-based approach.

Finally, our research offers actionable insights for sales professionals. The refined and exceptional quality lists of leads can act as practical guidelines, assisting sales practitioners in selecting, contacting, and pursuing leads throughout the sales process. Furthermore, in the short term, an exceptional lead scoring model has the potential to encourage salespeople to engage with more market-qualified leads, thereby converting those "ready-to-buy" leads to customers. In the long haul, the implementation of a lead scoring model with exceptional quality could enhance the collaboration between marketing and sales departments, potentially alleviating the detrimental impact of the "sales lead black hole" (Sabnis et al., 2013) and eventually improving overall sales efficacy.

5.5.3 Limitations and future research

Our research has several limitations that should be addressed in future research. First, this research is restricted to a few variables included in the datasets. We recommend that future studies investigate data sources with additional attributes of leads and variables of interactions between salespeople and leads. In addition, variables that can depict seller characteristics such as salesperson effort and experience could be incorporated in the model. For example, a salesperson's effort could be measured by taking the average

number of calls made by this salesperson in a day. A salesperson's experience could be measured by counting the number of working days in the sales department. Second, our analysis is confined to the service industry. Future research would benefit from a broader industry and cross industry comparison. Furthermore, in the model's sales performance evaluation, we focused on the lead conversion rate and cost of false positives, overlooking another important metric that is false negatives. In a binary classification, a false negative happens when a model incorrectly predicts a negative outcome for an instance that should have been a positive outcome. A false negative represents a missed opportunity, as the model fails to act on a positive case. For example, if a lead likely to make a purchase is incorrectly classified as uninterested, the resulting opportunity cost is the lost revenue from not engaging that lead with marketing and sales efforts. This opportunity cost of false negative represents the potential loss due to the incorrect prediction. Thus, future research should consider incorporating the opportunity cost of false negatives into the model's sales performance assessment. Moreover, although SMOTE resampling performs well for the legal consulting and cleaning service industries, it may present potential drawbacks. First, in situations where the minority class (positive contact) is scattered across feature space (imbalanced distributions), SMOTE may generate synthetic data points that do not represent meaningful or realistic data, thus reducing the model's effectiveness. Second, there might be a risk of multicollinearity. By creating synthetic points through interpolation, SMOTE could increase the correlation between features, which may compromise the model's stability and interpretability. Future research should address these potential issues. Finally, to implement the "smart sales" model in practice and integrate it into the current system could be challenging. Any potential integration issues may hinder the effectiveness of the lead scoring during the sales process. Furthermore, technical issues could occur when integrating the model to provide real-time scoring in a dynamic sales environment. Future research could explore on how to develop a smooth integration process to enable real-time lead scoring.

Table 5.1. Variable description

Name	Type	Description
Contact ID	Numeric	A unique identifier key for a lead.
Call History ID	Numeric	A unique identifier key for a call made.
Industry	Categorical	The industry that the lead resides in.
Revenue	Numeric	The revenue of the company that a lead belongs to.
Company Size	Numeric	The number of employees of the company that a lead belongs to.
Title	Categorical	The titles of a lead (i.e., admin, controller, member owner, partner, president, etc.).
Management Level	Categorical	The management level of leads (i.e., C-suite, director, president, non-manager, and manager).
Primary Lead Source	Categorical	The primary source where a lead was obtained from.
Secondary Lead Source	Categorical	The secondary source where a lead was obtained from.

Total Duration	Numeric	The amount of time spent on call (in minutes) between a lead and a salesperson.
Number of calls	Numeric	The number of calls made between a lead and a salesperson.
Result Group	Categorical	The outcome of conversion (i.e., positive or negative).

Table 5.2. Lead distribution by industry

Industry	**Count**	**Percentage**
Legal Consulting Service	24,890	3.2%
Commercial Office Building Cleaning Service	303,368	38.6%
Custom Payment Solution Service	456,235	58.2%
Total	**784,493**	

Table 5.3. Evaluation of clustering algorithms

Clustering Algorithms & Number of Clusters	**Silhouette Coefficient**	**Davies-Bouldin Index**	**Industry**
K-Prototypes (k = 3)	0.698	0.599	Legal Consulting Service
K-Prototypes (k = 4)	0.602	0.658	
K-Prototypes (k = 5)	0.587	0.692	
AGNES (k = 3)	0.569	0.772	
AGNES (k = 4)	0.579	0.648	
AGNES (k = 5)	0.559	0.651	
DBSCAN (k = 5)	0.578	0.600	
K-Prototypes (k = 4)	0.529	0.678	Commercial Office Building Cleaning Service
K-Prototypes (k = 5)	0.562	0.650	
K-Prototypes (k = 6)	0.564	0.658	
K-Prototypes (k = 7)	0.552	0.660	
AGNES (k = 4)	0.522	0.687	
AGNES (k = 5)	0.511	0.698	
AGNES (k = 6)	0.509	0.695	
AGNES (k = 7)	0.498	0.690	
DBSCAN (k = 6)	0.501	0.666	
K-Prototypes (k = 3)	0.603	0.512	Custom Payment Solution Service
K-Prototypes (k = 4)	0.591	0.525	
K-Prototypes (k = 5)	0.563	0.537	
K-Prototypes (k = 6)	0.552	0.633	
AGNES (k = 4)	0.532	0.643	
AGNES (k = 5)	0.543	0.655	
AGNES (k = 6)	0.521	0.660	
AGNES (k = 7)	0.501	0.658	
DBSCAN (k = 6)	0.510	0.665	

Table 5.4. Evaluation of "classification only" lead scoring model (legal consulting service)

Resampling Techniques	Conversion Rate	Balanced Accuracy	Specificity	Sensitivity	F1-score
DT Baseline	11.53%	0.86	0.94	0.79	0.86
DT Oversampling	12.87%	0.85	0.93	0.77	0.84
DT Undersampling	13.14%	0.85	0.73	0.96	0.83
DT SMOTE	12.60%	0.89	0.94	0.85	0.89
GBT Baseline	13.67%	0.87	0.98	0.76	0.86
GBT Oversampling	13.67%	0.88	0.87	0.88	0.88
GBT Undersampling	13.67%	0.86	0.76	0.96	0.85
GBT SMOTE	13.67%	0.90	0.92	0.88	0.90
LR Baseline	13.67%	0.88	0.98	0.78	0.87
LR Oversampling	13.67%	0.87	0.79	0.96	0.86
LR Undersampling	13.67%	0.83	0.78	0.96	0.86
LR SMOTE	13.67%	0.89	0.84	0.96	0.90
NB Baseline	11.14%	0.68	0.35	0.90	0.50
NB Oversampling	11.14%	0.59	0.18	0.95	0.30
NB Undersampling	11.14%	0.73	0.97	0.49	0.65
NB SMOTE	12.87%	0.50	0.40	0.95	0.56

Table 5.5. Evaluation of "classification only" lead scoring model (cleaning service)

Resampling Techniques	Conversion Rate	Balanced Accuracy	Specificity	Sensitivity	F1-score
DT Baseline	68.53%	0.86	0.64	0.79	0.71
DT Oversampling	70.42%	0.83	0.88	0.76	0.82
DT Undersampling	69.81%	0.68	0.68	0.66	0.67
DT SMOTE	70.17%	0.75	0.75	0.73	0.74
GBT Baseline	70.20%	0.71	0.69	0.76	0.72
GBT Oversampling	74.54%	0.74	0.70	0.82	0.75
GBT Undersampling	74.62%	0.73	0.70	0.82	0.75
GBT SMOTE	74.55%	0.75	0.71	0.83	0.76
LR Baseline	68.26%	0.64	0.73	0.52	0.61
LR Oversampling	68.26%	0.65	0.56	0.52	0.54
LR Undersampling	68.26%	0.65	0.55	0.52	0.54
LR SMOTE	68.26%	0.66	0.56	0.53	0.55
NB Baseline	68.88%	0.58	0.65	0.11	0.19
NB Oversampling	68.85%	0.52	0.64	0.11	0.19
NB Undersampling	68.87%	0.52	0.63	0.11	0.18
NB SMOTE	68.84%	0.52	0.65	0.09	0.16

Table 5.6. Evaluation of "classification only" lead scoring model (payment solution service)

Resampling Techniques	Conversion Rate	Balanced Accuracy	Specificity	Sensitivity	F1-score
DT Baseline	4.47%	0.64	0.65	0.79	0.71
DT Oversampling	4.43%	0.81	0.78	0.84	0.81
DT Undersampling	5.15%	0.81	0.75	0.87	0.81

DT SMOTE	4.49%	0.74	0.86	0.63	0.73
GBT Baseline	5.39%	0.61	0.63	0.73	0.68
GBT Oversampling	5.39%	0.82	0.73	0.90	0.81
GBT Undersampling	5.39%	0.82	0.74	0.91	0.81
GBT SMOTE	5.39%	0.82	0.73	0.91	0.81
LR Baseline	5.31%	0.82	0.81	0.83	0.82
LR Oversampling	5.37%	0.87	0.81	0.83	0.82
LR Undersampling	5.36%	0.87	0.81	0.83	0.82
LR SMOTE	5.37%	0.88	0.81	0.83	0.82
NB Baseline	5.06%	0.72	0.65	0.11	0.19
NB Oversampling	5.06%	0.76	0.89	0.60	0.72
NB Undersampling	5.05%	0.68	0.84	0.40	0.55
NB SMOTE	5.01%	0.77	0.76	0.90	0.82

Table 5.7. Evaluation comparison of lead scoring models (legal consulting service)

Lead Scoring Models	Conversion Rate	Cost of FP	Balanced Accuracy	Specificity	Sensitivity	F1-score
GBT "classification only"	13.67%	$720	0.88	0.98	0.78	0.87
GBT Cluster 0 (3800 leads)	19.30%	$560	0.83	0.85	0.82	0.83
GBT Cluster 1 (2730 leads)	9.76%	$560	0.91	0.81	0.94	0.87
GBT Cluster 2 (18360 leads)	13.45%	$1000	0.92	0.92	0.95	0.93

Table 5.8. Evaluation comparison of lead scoring models (cleaning service)

Lead Scoring Models	Conversion Rate	Cost of FP	Balanced Accuracy	Specificity	Sensitivity	F1-score
GBT "classification only"	70.20%	$70,136	0.71	0.69	0.76	0.72
GBT Cluster 0 (128091 leads)	74.56%	$34,424	0.75	0.72	0.82	0.77
GBT Cluster 1 (140495 leads)	71.18%	$39,536	0.75	0.73	0.82	0.77
GBT Cluster 2 (6732 leads)	85.65%	$688	0.84	0.84	0.83	0.83
GBT Cluster 3 (1424 leads)	71.17%	$788	0.77	0.75	0.82	0.78
GBT Cluster 4 (26626 leads)	84.25%	$5432	0.70	0.73	0.65	0.68

Table 5.9. Evaluation comparison of lead scoring models (payment solution service)

Lead Scoring Models	Conversion Rate	Cost of FP	Balanced Accuracy	Specificity	Sensitivity	F1-score

GBT "classification only"	5.39%	$20,408	0.61	0.63	0.73	0.68
GBT Cluster 0 (333276 leads)	5.81%	$22,344	0.81	0.70	0.93	0.79
GBT Cluster 1 (12764 leads)	14.55%	$4,296	0.87	0.67	0.91	0.77
GBT Cluster 2 (110195 leads)	5.89%	$14,712	0.87	0.83	0.92	0.87

Table 5.10. Evaluation of lead scoring models on training sets (legal consulting service)

Lead Scoring Models	Accuracy	Specificity/Recall	Sensitivity/Precision	F1-score
GBT Baseline	0.91	0.89	0.76	0.82
GBT Oversampling	0.94	0.94	0.92	0.93
GBT Undersampling	0.92	0.90	0.92	0.91
GBT SMOTE	0.94	0.98	0.92	0.95
GBT Cluster 0	0.90	0.91	0.89	0.90
GBT Cluster 1	0.96	0.89	0.95	0.89
GBT Cluster 2	0.97	0.97	0.97	0.97

Table 5.11. Evaluation of lead scoring models on training sets (cleaning service)

Lead Scoring Models	Accuracy	Specificity/Recall	Sensitivity/Precision	F1-score
GBT Baseline	0.78	0.79	0.79	0.79
GBT Oversampling	0.74	0.76	0.82	0.79
GBT Undersampling	0.74	0.76	0.82	0.79
GBT SMOTE	0.76	0.77	0.83	0.80
GBT Cluster 0	0.75	0.77	0.82	0.79
GBT Cluster 1	0.75	0.74	0.82	0.78
GBT Cluster 2	0.85	0.84	0.84	0.84
GBT Cluster 3	0.78	0.79	0.83	0.81
GBT Cluster 4	0.77	0.78	0.75	0.77

Table 5.12. Evaluation of lead scoring models on training sets (payment solution service)

Lead Scoring Models	Accuracy	Specificity/Recall	Sensitivity/Precision	F1-score
GBT Baseline	0.70	0.69	0.78	0.73
GBT Oversampling	0.87	0.78	0.91	0.84
GBT Undersampling	0.89	0.79	0.93	0.86
GBT SMOTE	0.88	0.79	0.92	0.85
GBT Cluster 0	0.86	0.75	0.93	0.83
GBT Cluster 1	0.92	0.72	0.91	0.80
GBT Cluster 2	0.93	0.88	0.92	0.89

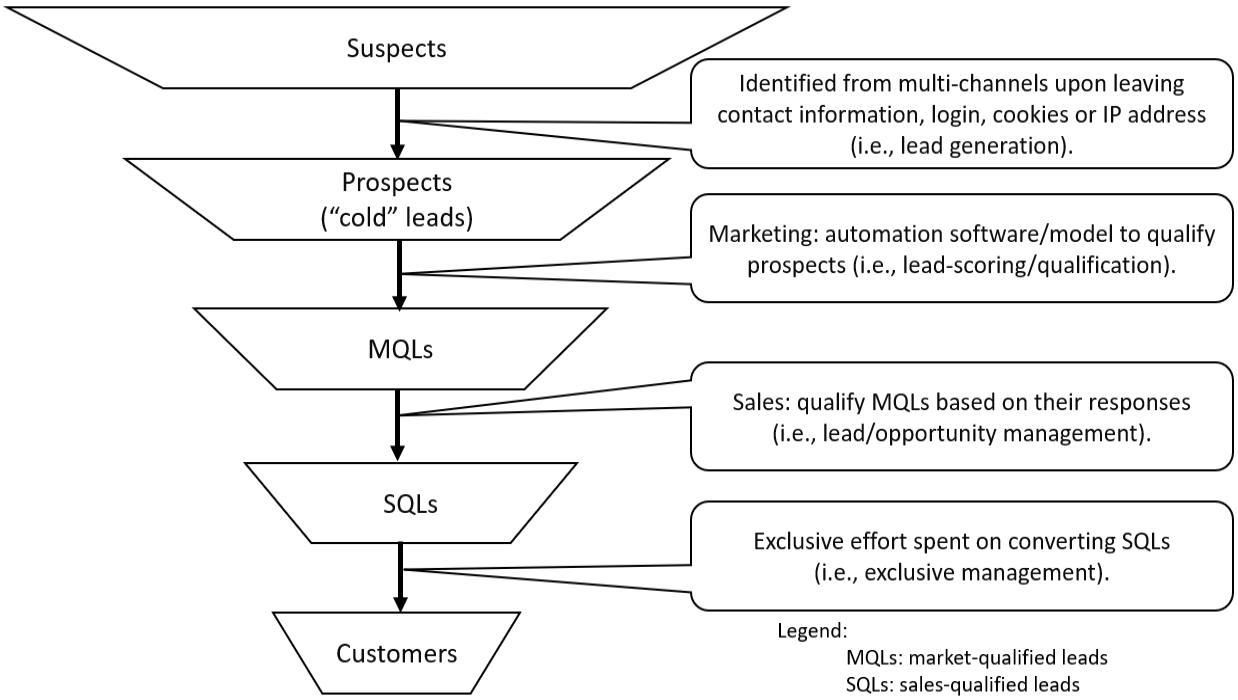

Figure 5.1. Marketing-sales funnel

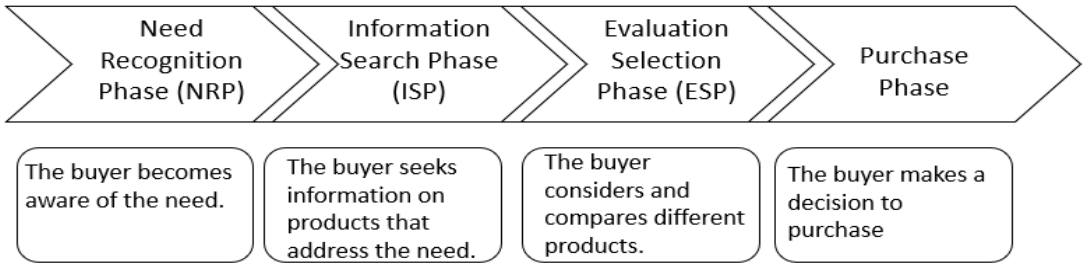

Figure 5.2. Buyer purchase decision-making process

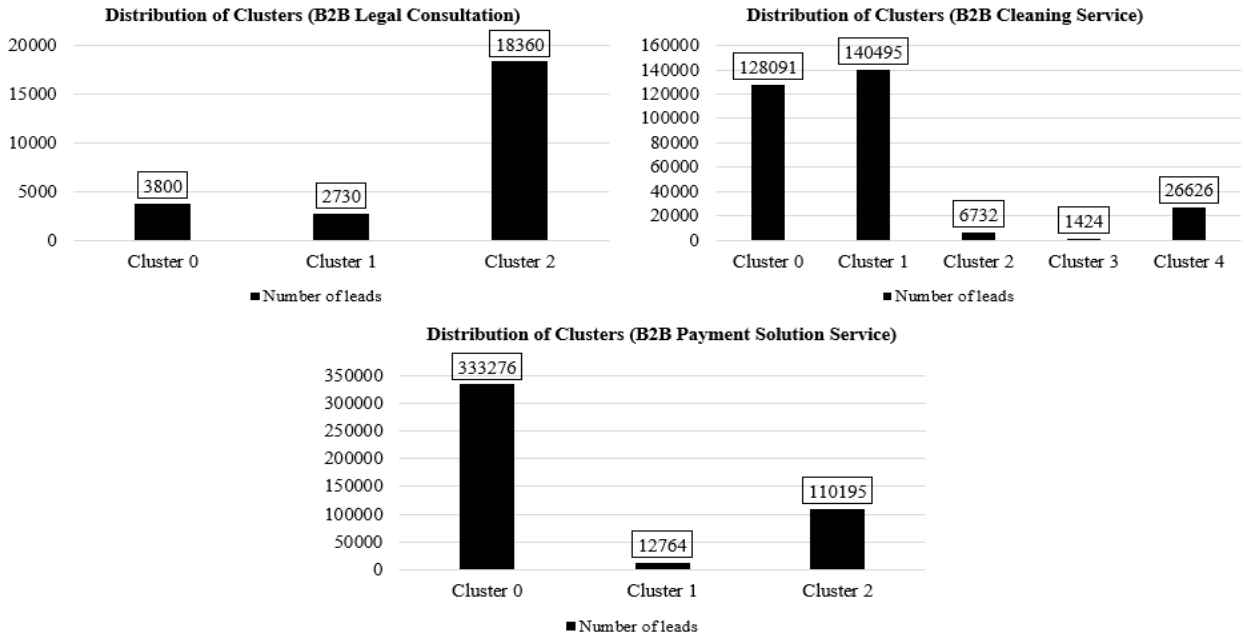

Figure 5.3. Distribution of clusters

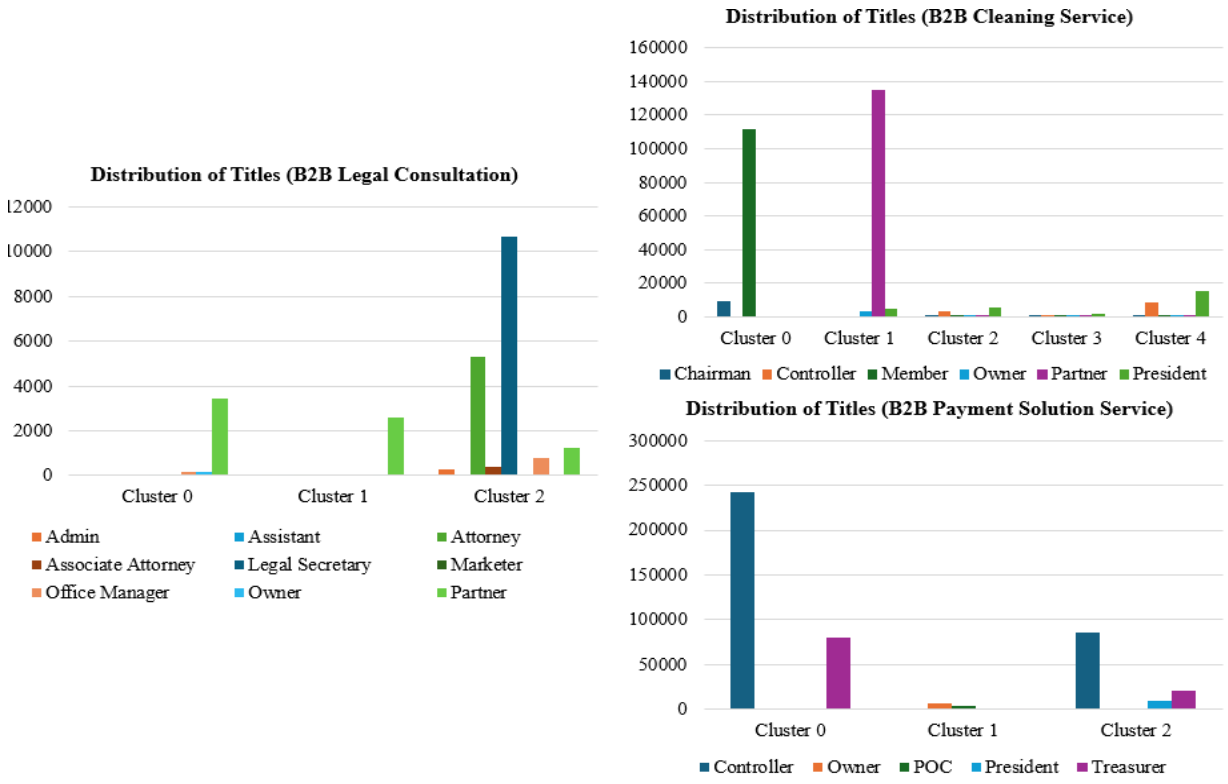

Figure 5.4. Distribution of titles

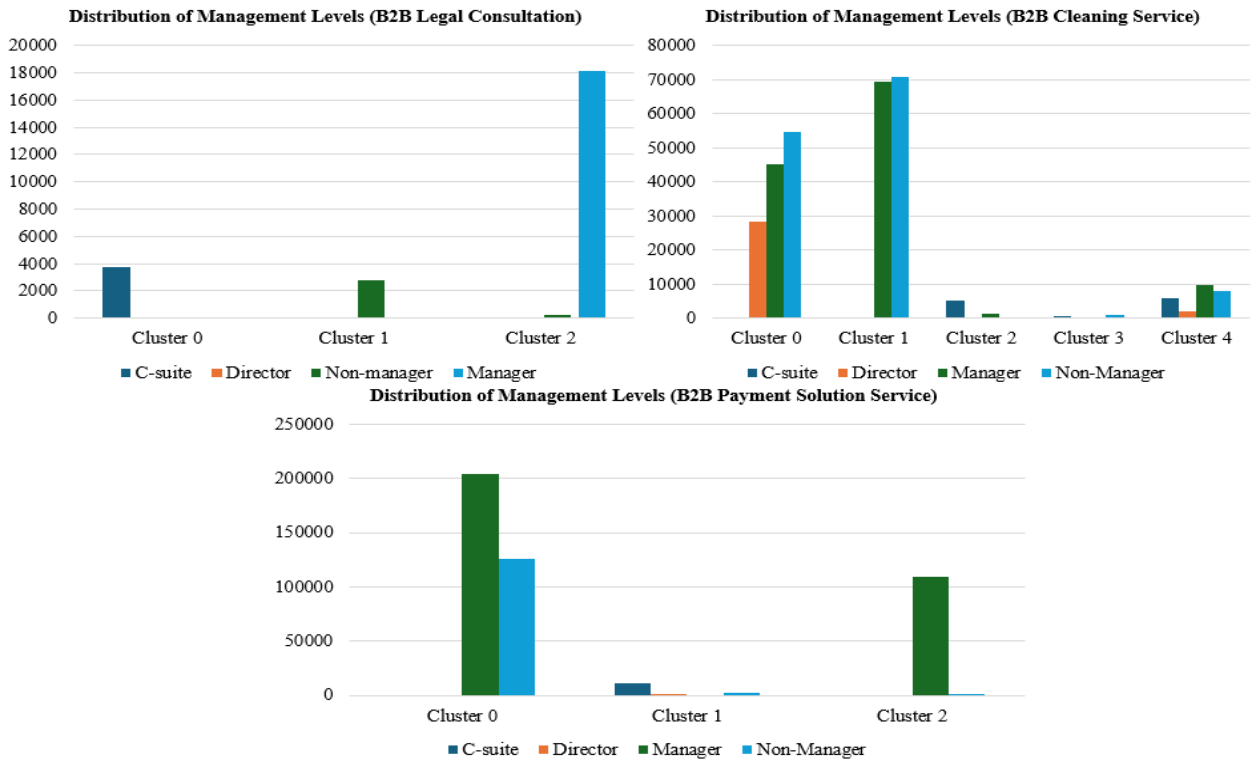

Figure 5.5. Distribution of management levels

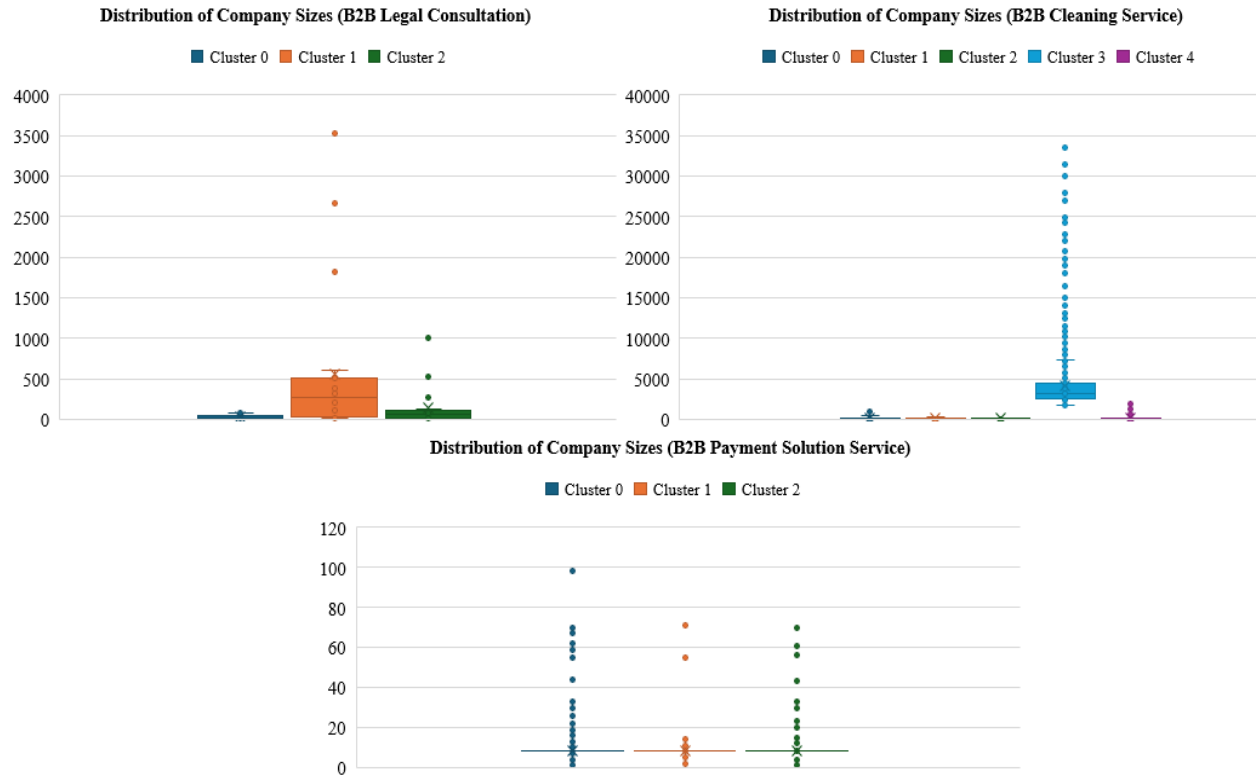

Figure 5.6. Distribution of company sizes

References

Ang, L., & Buttle, F. (2006). Managing For Successful Customer Acquisition: An Exploration. *Journal of Marketing Management, 22*(3–4), 295–317. https://doi.org/10.1362/026725706776861217

Banerjee, S., Bhattacharyya, S., & Bose, I. (2017). Whose online reviews to trust? Understanding reviewer trustworthiness and its impact on business. *Decision Support Systems, 96*, 17–26. https://doi.org/10.1016/j.dss.2017.01.006

Başarslan, M. S., & Argun, I. D. (Eds.). (2020). *Prediction of potential bank customers (Vol. 43)*. Springer International Publishing. https://doi.org/10.1007/978-3-030-36178-5

Bastı, E., Kuzey, C., & Delen, D. (2015). Analyzing initial public offerings' short-term performance using decision trees and SVMs. *Decision Support Systems, 73,* 15–27. https://doi.org/10.1016/j.dss.2015.02.011

Benhaddou, Y., & Leray, P. (2017). Customer Relationship Management and Small Data—Application of Bayesian Network Elicitation Techniques for Building a Lead Scoring Model. *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA),* 251–255. https://doi.org/10.1109/AICCSA.2017.51

Bohanec, M., Kljajić Borštnar, M., & Robnik-Šikonja, M. (2017). Explaining machine learning models in sales predictions. *Expert Systems with Applications, 71,* 416–428. https://doi.org/10.1016/j.eswa.2016.11.010

Bradford, W., Johnston, W. J., & Bellenger, D. (2016). The Impact of Sales Effort on Lead Conversion Cycle Time in a Business-to-Business Opportunity Pipeline (SSRN Scholarly Paper ID 2866954). *Social Science Research Network.* https://papers.ssrn.com/abstract=2866954

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research, 16,* 321–357. https://doi.org/10.1613/jair.953

Chorianopoulos, A. (2016). *Effective CRM using Predictive Analytics.* John Wiley & Sons.

Coe, J. M. (2004). The integration of direct marketing and field sales to form a new B2B sales coverage model. *Journal of Interactive Marketing, 18*(2), 62–74. https://doi.org/10.1002/dir.20005

Dalmaijer, E. S., Nord, C. L., & Astle, D. E. (2022). Statistical power for cluster analysis. *BMC Bioinformatics, 23*(1), 205. https://doi.org/10.1186/s12859-022-04675-1

Davies, D. L., & Bouldin, D. W. (1979). A Cluster Separation Measure. IEEE Transactions on Pattern Analysis and Machine Intelligence, *PAMI-1(2),* 224–227. https://doi.org/10.1109/TPAMI.1979.4766909

D'Haen, J., & Van den Poel, D. (2013). Model-supported business-to-business prospect prediction based on an iterative customer acquisition framework. *Industrial Marketing Management, 42*(4), 544–551. https://doi.org/10.1016/j.indmarman.2013.03.006

D'Haen, J., Van den Poel, D., & Thorleuchter, D. (2013). Predicting customer profitability during acquisition: Finding the optimal combination of data source and data mining technique. *Expert Systems with Applications, 40*(6), 2007–2012. https://doi.org/10.1016/j.eswa.2012.10.023

D'Haen, J., Van den Poel, D., Thorleuchter, D., & Benoit, D. F. (2016). Integrating expert knowledge and multilingual web crawling data in a lead qualification system. *Decision Support Systems, 82,* 69–78. https://doi.org/10.1016/j.dss.2015.12.002

Dubinsky, A. (1981). A Factor Analytic Study of the Personal Selling Process. *The Journal of Personal Selling & Sales Management*, *1*(1), 26–33. https://doi.org/10.1080/08853134.1981.10754192

Duncan, B., & Elkan, C. (2015). Probabilistic Modeling of a Sales Funnel to Prioritize Leads. *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1751–1758. https://doi.org/10.1145/2783258.2788578

Eitle, V., & Buxmann, P. (2019). Business analytics for sales pipeline management in the software industry: A machine learning perspective. *Proceedings of the Annual Hawaii International Conference on System Sciences*. https://doi.org/10.24251/hicss.2019.125

Gupta, R., Kumar, H., Jain, T., Shrotriya, A., & Sinha, A. (2022). Demographic Customer Segmentation of banking users Based on k-prototype methodology. *2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence),* 578–584. https://doi.org/10.1109/Confluence52989.2022.9734169

Han, J., Kamber, M., & Pei, J. (2012). 11—Advanced Cluster Analysis. In J. Han, M. Kamber, & J. Pei (Eds.), *Data Mining* (Third Edition) (pp. 497–541). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-381479-1.00011-3

Ho, T., Nguyen, S., Nguyen, H., Nguyen, N., Man, D.-S., & Le, T.-G. (2023). An Extended RFM Model for Customer Behaviour and Demographic Analysis in Retail Industry. *Business Systems Research Journal, 14*(1), 26–53. https://doi.org/10.2478/bsrj-2023-0002

Jansen, B. J., & Schuster, S. (2011). Bidding on the buying funnel for sponsored search and keyword advertising. *Journal of Electronic Commerce Research, 12*(1), 19.

Järvinen, J., & Taiminen, H. (2016). Harnessing marketing automation for B2B content marketing. *Industrial Marketing Management, 54*, 164–175. https://doi.org/10.1016/j.indmarman.2015.07.002

Jo, T., & Japkowicz, N. (2004). Class imbalances versus small disjuncts. *ACM SIGKDD Explorations Newsletter, 6*(1), 40–49. https://doi.org/10.1145/1007730.1007737

Kawas, B., Squillante, M. S., Subramanian, D., & Varshney, K. R. (2013). Prescriptive Analytics for Allocating Sales Teams to Opportunities. *2013 IEEE 13th International Conference on Data Mining Workshops*, 211–218. https://doi.org/10.1109/ICDMW.2013.156

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, 1137–1143.

Kotu, V., & Deshpande, B. (2015). Chapter 4—Classification. In V. Kotu & B. Deshpande (Eds.), *Predictive Analytics and Data Mining* (pp. 63–163). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-801460-8.00004-5

Lima C. D. (2018). End-to-end-lead scoring. GitHub. https://github.com/calvdee/end-to-end-lead-scoring/commits/main/docs

McCallum, A., & Nigam, K. (1998). A comparison of event models for naive bayes text classification. *AAAI Conference on Artificial Intelligence.*

Meena, P., Kumar, C., & Puri, S. (2023). Customer segmentation and behavioral systems through influential effective elements: *An E-satisfaction analysis using machine learning*. 020047. https://doi.org/10.1063/5.0154287

Meire, M., Ballings, M., & Van den Poel, D. (2017). The added value of social media data in B2B customer acquisition systems: A real-life experiment. *DECISION SUPPORT SYSTEMS, 104*, 26–37. https://doi.org/10.1016/j.dss.2017.09.010

Monat, J. P. (2011). Industrial sales lead conversion modeling. *Marketing Intelligence & Planning*. https://doi.org/10.1108/02634501111117610

Napierala, K., & Stefanowski, J. (2012). BRACID: A comprehensive approach to learning rules from imbalanced data. *Journal of Intelligent Information Systems, 39*(2), 335–373. https://doi.org/10.1007/s10844-011-0193-0

Ngai, E. W. T., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications, 36*(2, Part 2), 2592–2602. https://doi.org/10.1016/j.eswa.2008.02.021

Nygård, R., & Mezei, J. (2020). Automating lead scoring with Machine Learning: An experimental study. *Proceedings of the Annual Hawaii International Conference on System Sciences*. https://doi.org/10.24251/hicss.2020.177

Ohiomah, A. (2020). *Lead to Win: Recipes for Inside Sales Success* [Doctoral thesis, University of Ottawa]. https://ruor.uottawa.ca/items/49992d5d-c90e-49e5-8bce-3b15b7b2b4fa

Ohiomah, A., Andreev, P., Benyoucef, M., & Hood, D. (2019). The role of lead management systems in inside sales performance. *Journal of Business Research, 102*, 163–177. https://doi.org/10.1016/j.jbusres.2019.05.018

Ohiomah, A., Benyoucef, M., Andreev, P., Kuziemsky, C., Hood, D., & Le Bon, J. (2018). Best practices for inside sales professionals: An historical analysis. *Developments in Marketing Science: Proceedings of the Academy of Marketing Science*, 449–461. https://doi.org/10.1007/978-3-319-99181-8_149

Olson, D. L., & Chae, B. (2012). Direct marketing decision support through predictive customer response modeling. Decision Support Systems, 54(1), 443–451. https://doi.org/10.1016/j.dss.2012.06.005

Othayoth, S. P., & Muthalagu, R. (2022). Customer segmentation using various machine learning techniques. *International Journal of Business Intelligence and Data Mining*. https://www.inderscienceonline.com/doi/10.1504/IJBIDM.2022.123218

Paschen, J., Wilson, M., & Ferreira, J. J. (2020). Collaborative intelligence: How human and artificial intelligence create value along the B2B sales funnel. *Business Horizons, 63*(3), 403–414. https://doi.org/10.1016/j.bushor.2020.01.003

Patterson, L. (2007). Marketing and sales alignment for improved effectiveness. *Journal of Digital Asset Management, 3*(4), 185–189. https://doi.org/10.1057/palgrave.dam.3650089

Peng, K., & Xu, D. (2011). Modeling of potential customers identification based on correlation analysis and decision tree. *Advances in Neural Networks – ISNN 2011*, 566–575. https://doi.org/10.1007/978-3-642-21111-9_64

Puccinelli, N. M., Goodstein, R. C., Grewal, D., Price, R., Raghubir, P., & Stewart, D. (2009). Customer Experience Management in Retailing: Understanding the Buying Process. *Journal of Retailing, 85*(1), 15–30. https://doi.org/10.1016/j.jretai.2008.11.003

Reynolds, A. P., Richards, G., de la Iglesia, B., & Rayward-Smith, V. J. (2006). Clustering Rules: A Comparison of Partitioning and Hierarchical Clustering Algorithms. *Journal of Mathematical Modelling and Algorithms, 5*(4), 475–504. https://doi.org/10.1007/s10852-005-9022-1

Rutherford, B. N., Marshall, G. W., & Park, J. (2014). The moderating effects of gender and inside versus outside sales role in multifaceted job satisfaction. *Journal of Business Research, 67*(9), 1850–1856. https://doi.org/10.1016/j.jbusres.2013.12.004

Sabnis, G., Chatterjee, S. C., Grewal, R., & Lilien, G. L. (2013). The Sales Lead Black Hole: On Sales Reps' Follow-Up of Marketing Leads. *Journal of Marketing, 77*(1), 52–67. https://doi.org/10.1509/jm.10.0047

Shannahan, R. J., Bush, A. J., Moncrief, W. C., & Shannahan, K. L. J. (2013). Making Sense of the Customer's Role in the Personal Selling Process: A Theory of Organizing and Sensemaking Perspective. *Journal of Personal Selling & Sales Management, 33*(3), 261–275. https://doi.org/10.2753/PSS0885-3134330302

Sheth, Sharma, A., & Iyer, G. R. (2009). Why integrating purchasing with marketing is both inevitable and beneficial. *Industrial Marketing Management, 38*(8), 865–871. https://doi.org/10.1016/j.indmarman.2008.12.021

Sirakaya, E., & Woodside, A. G. (2005). Building and testing theories of decision making by travellers. *Tourism Management, 26*(6), 815–832. https://doi.org/10.1016/j.tourman.2004.05.004

Smith, T. M., Gopalakrishna, S., & Chatterjee, R. (2006). A Three-Stage Model of Integrated Marketing Communications at the Marketing-Sales Interface. *Journal of Marketing Research, 43*(4), 564–579. https://doi.org/10.1509/jmkr.43.4.564

Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018). Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster. *IOP Conference Series: Materials Science and Engineering, 336*, 012017. https://doi.org/10.1088/1757-899X/336/1/012017

Syam, N., & Sharma, A. (2018). Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. *Industrial Marketing Management*, *69*, 135–146. https://doi.org/10.1016/j.indmarman.2017.12.019

Thaichon, P., Surachartkumtonkun, J., Quach, S., Weaven, S., & Palmatier, R. W. (2018). Hybrid sales structures in the age of e-commerce. *Journal of Personal Selling & Sales Management, 38*(3), 277–302. https://doi.org/10.1080/08853134.2018.1441718

Thomassey, S., & Fiordaliso, A. (2006). A hybrid sales forecasting system based on clustering and decision trees. *Decision Support Systems, 42*(1), 408–421. https://doi.org/10.1016/j.dss.2005.01.008

Thorleuchter, D., Van Den Poel, D., & Prinzie, A. (2012). Analyzing existing customers' websites to improve the customer acquisition process as well as the profitability prediction in B-to-B marketing. *Expert Systems with Applications, 39*(3), 2597–2605. https://doi.org/10.1016/j.eswa.2011.08.115

van Steenbergen, R. M., & Mes, M. R. K. (2020). Forecasting demand profiles of new products. *Decision Support Systems, 139*, 113401. https://doi.org/10.1016/j.dss.2020.113401

Wang, S.-T. (2018). Integrating KPSO and C5.0 to analyze the omnichannel solutions for optimizing telecommunication retail. *Decision Support Systems, 109,* 39–49. https://doi.org/10.1016/j.dss.2017.12.009

Wu, M., Andreev, P., & Benyoucef, M. (2023). The state of lead scoring models and their impact on sales performance. *Information Technology and Management*, *25*(1), 69–98. https://doi.org/10.1007/s10799-023-00388-w

Wu, M., Andreev, P., Benyoucef, M., & Hood, D. (2024). Unlocking B2B buyer intentions to purchase: Conceptualizing and validating inside sales purchases. *Decision Support Systems, 179,* 114165. https://doi.org/10.1016/j.dss.2023.114165

Wu, Z., Chen, X., & Gao, Z. (2023). Bayesian non-parametric method for decision support: Forecasting online product sales. *Decision Support Systems, 174,* 114019. https://doi.org/10.1016/j.dss.2023.114019

Xia, H., Pan, X., Zhou, Y., & Zhang, Z. (Justin). (2020). Creating the best first impression: Designing online product photos to increase sales. Decision Support Systems, 131, 113235. https://doi.org/10.1016/j.dss.2019.113235

Xu, X., Tang, L., & Rangan, V. (2017). Hitting your number or not? A robust & intelligent sales forecast system. 2017 IEEE International Conference on Big Data (Big Data), 3613–3622. https://doi.org/10.1109/BigData.2017.8258355

Yan, J., Zhang, C., Zha, H., Gong, M., Sun, C., Huang, J., Chu, S., & Yang, X. (2015). On machine learning towards predictive sales pipeline analytics. 3, 1945–1951. Scopus.

Yu, L., Wang, S., & Lai, K. K. (2006). An integrated data preparation scheme for neural network data analysis. IEEE Transactions on Knowledge and Data Engineering, 18(2), 217–230. https://doi.org/10.1109/TKDE.2006.22

Zolkiewski, J., Story, V., Burton, J., Chan, P., Gomes, A., Hunter-Jones, P., O'Malley, L., Peters, L. D., Raddats, C., & Robinson, W. (2017). Strategic B2B customer experience management: The importance of outcomes-based measures. Journal of Services Marketing, 31(2), 172–184. https://doi.org/10.1108/JSM-10-2016-0350

Chapter 5 Key Points Summary:

- Developed and implemented predictive lead scoring models using a two-step approach (i.e., clustering then classification) that considers leads' profiles (demographics of leads) and the interactions between salespeople and leads;
- Empirically validated the proposed model using data from three different B2B service industries;

- Lead profiling is a critical step to understand different leads' purchase intentions in each purchase decision-making phase;
- Demonstrated ML algorithms and DM models' roles in transitioning from traditional subjective decision-making sales to model-based objective decision-making sales.

Chapter 6. Discussion

6.1 Conclusion and implications

The applications of digital transformation technologies are rapidly growing in the inside sales industry. Among these applications, predictive lead scoring has demonstrated significant potential and a crucial role in transforming the sales process, enhancing sales team efficiency, and increasing sales performance. Despite the apparent potential of predictive lead scoring, there remains a lack of knowledge about the factors and practices that improve its efficacy and ultimately boost sales performance. This thesis, comprising four articles, contributes to the lead scoring literature by developing and implementing a "smart sales" predictive lead scoring model that draws on two theoretical frameworks from the perspectives of both the seller and the lead. The resulting targeted, concise, and high-quality lead lists verify the success of the proposed method and its potential to improve sales performance. The implications and contributions of the four articles are summarized in **Figure 6.*1***.

Figure 6.1 appears here

Chapter 2 of this thesis conducts a systematic literature review and analysis to identify key lead scoring models and assess their impact on sales performance. This response to research calls about how these models can address challenges in the inside sales process reveals 44 studies and 14 metrics used to evaluate the effects of lead scoring models on sales performance. This study is likely the first comprehensive review to cover lead scoring models explored in both academic research and industrial applications. It introduces a classification framework that categorizes traditional and predictive lead scoring models, the latter of which employ various machine learning algorithms and data mining models. Moreover, this study recommends appropriate learning methods (i.e., supervised and/or unsupervised) based on the available data sources for constructing predictive models. Additionally, by comparing these models and analyzing their impact, this study suggests ways of improving sales performance through different lead scoring models. This contributes to achieving the first research objective and addresses the first research question (i.e., *"What are the advantages and disadvantages of the existing lead scoring models and how do they influence sales performance?"*).

While the lead conversion rate is a useful and popular metric for evaluating sales performance, it has several limitations when used in isolation, especially in industries with long sales cycles, complex lead qualification processes, or significant variations in lead quality and value. For instance, in sectors with longer sales cycles, a low conversion rate may not necessarily reflect poor overall performance. Complex sales processes usually require multiple touchpoints, meetings, and hierarchical approvals, which naturally extend the time before a conversion occurs. Additionally, the lead conversion rate does not account for the quality of leads. Salespeople dealing with higher-quality ("warmer") leads will likely achieve higher

conversion rates compared to those working with less qualified ("colder") leads. Therefore, focusing solely on conversion rates overlooks the nuances of different lead types. Moreover, different lead sources (e.g., inbound marketing, outbound calls, referrals) may yield significantly different conversion rates. Comparing conversion rates without considering lead sources or the quality of leads could lead to misleading interpretations of the overall sales performance. Furthermore, the conversion rate fails to consider customer lifetime value. Sales teams that prioritize quick lead conversions may close deals that are less profitable or prone to churn, while teams that focus on building long-term relationships may generate higher revenue and customer lifetime value, despite having a lower conversion rate. In some cases, sales teams may focus on converting a high volume of low value leads to boost their conversion rate, which can skew performance metrics. This is particularly problematic in industries where high-value deals are critical, such as luxury goods, industrial equipment, or enterprise software. In such cases, the conversion rate can be misleading if it does not account for revenue or deal size. To gain a comprehensive understanding of sales performance, companies should complement conversion rate metrics with other indicators, such as revenue, profit gains, customer lifetime value, customer acquisition cost, and lead quality. This ensures that they measure not just the number of conversions, but also the value and potential of each sale.

To address the knowledge gap concerning the purchase decision-making process of potential buyers in an inside sales context and to improve sales performance, Chapter 3 in this thesis proposes and empirically validates a potential buyer's intention to purchase (BIP) model based on the purchase decision-making process and buying behavior theories. The BIP model captures critical determinants influencing potential buyers' intention to purchase. Furthermore, using data from an online survey of 126 inside sales professionals, this study suggests that potential buyers' trust in sellers and connection flexibility affect their purchase intention via appointment and contact willingness during various phases in the purchase process. The findings of this study also uncover how varying demographic and behavioral attributes of potential buyers (i.e., multiple leads' profiles) distinctly affect the purchase decision-making at each phase of the process. This study answers the second research question and fulfills the second research objective by identifying key characteristics of leads for higher purchase intention and validating critical drivers and enablers for inside sales success in each purchase decision-making phase (i.e., "*How do drivers and enablers influence a buyer's purchase decision-making process in inside sales context?"*).

Finally, Chapters 4 and 5 in this thesis develop and implement a "smart sales" predictive lead scoring model that incorporates insights from leads' profiling into the lead scoring and qualification process. Utilizing a two-step learning methodology that begins with an unsupervised clustering followed by a supervised classification method on leads' characteristics and salespeople's performance data, our proposed "smart sales" model delivers explainable outcomes, revealing patterns in the data across stages of the sales

process. The empirical findings based on the small dataset (in Chapter 4) and the large datasets (in Chapter 5) collected from the three services-based industries underscore the critical role of data-driven predictive machine learning models in transitioning from traditional human decision-making to a more systematic ML model-based approach. In addition, the proposed data-driven "smart sales" model offers actionable insights for sales practitioners by generating concise and high-quality lists of leads for different stages of the sales process, which can help optimize resource allocation and boost sales performance. These chapters achieve the third objective and address the third research question (i.e., *"How can predictive analytics optimize lead scoring outcomes for inside sales?"*).

6.2 Limitation and future work

6.2.1 Potential biases in data

In Wu et al. (2024a), there may be potential biases and limitations in the self-reported survey-based data used for analysis. The first one is sampling bias in data collection. Although the online survey participants represent a range of industries, certain sectors may have been missed, leading to the underrepresentation of specific groups. Additionally, non-response bias could be present, as individuals who choose to respond may have different opinions or behaviors compared to those who do not, resulting in skewed outcomes. Response bias is another concern. The way questions are phrased or worded in an online survey can influence how participants answer. Moreover, surveys often capture data at a single point in time, which may not reflect changing sentiments or behaviors over time. Also, offering incentives to complete the survey could introduce bias, with some participants rushing through or providing insincere responses simply to claim the reward. To mitigate these biases, we consulted three B2B inside sales experts from the AA-ISP (American Association of Inside Sales Professionals) to identify any potential bias introduced by questions in each stage before sending out the survey. We also removed outliers in the data before conducting the analysis.

The second potential bias arises from using industry-specific data (i.e., service industry) in Wu et al. (2024b). Focusing on data from the service industry limits the generalizability of the model's findings to other industries. For example, industries such as manufacturing or technology may exhibit different customer behaviors, operational models, or market dynamics, making insights derived from the service sector potentially inapplicable to them. Furthermore, using data from one industry could lead to overfitting in predictive models, where the model becomes too specialized in the nuances of that specific industry and fails to perform well when applied to other contexts or industries. To mitigate these biases, it would be beneficial to incorporate additional data sources, conduct cross-industry comparisons, and apply rigorous validation techniques to enhance the model's robustness and applicability across industries.

6.2.2 Model generalization

To generalize the proposed "smart sales" model to other settings (e.g., outside sales, B2C contexts), several key differences in sales processes, lead-seller interactions, and data availability need to be considered. Regarding the sales process, outside sales often involve longer sales cycles due to the emphasis on face-to-face meetings and relationship-building activities. Additionally, outside sales usually engage multiple stakeholders and require more complex decision-making processes, which are harder to quantify (Zolkiewski et al., 2017). The longer sales cycles and multi-stage negotiations in outside sales could potentially make it harder to estimate a lead's score early in the sales process. This could decrease the predictive model's ability to make quick and accurate assessments. As a result, the predictive lead scoring model may need to consider the influence of various decision-makers, their internal dynamics, and external market factors. Also, the model may need to be adjusted to reflect the extended timeline and complexity. For instance, tracking sales progress over extended periods and incorporating variables such as relationship depth and multi-stakeholder engagement could improve the model's accuracy in these settings. Another aspect is data availability and sources. Outside sales may require incorporating qualitative data such as meeting notes, interactions with leads during field visits, and personal evaluations from salespeople (Rapp et al., 2012). We may need to integrate these new data points into the model. Additionally, the importance of relationships and personal interactions in outside sales means the model would benefit from including unstructured data, such as customer relationship insights or feedback from salespeople. However, predictive models rely heavily on quantitative data (e.g., travel frequency, time spent on each visit), integrating subjective factors (qualitative data) into the model can be challenging.

6.2.3 Real-world case study

Despite the extensive validations and performance evaluations of the model, this research lacks a case study demonstrating the real-world implementation and efficacy of the proposed "smart sales" predictive lead scoring model. Therefore, one direction for future work is to validate the proposed model in a practical field/case study, ensuring that the model performs well and generalizes effectively to real-world unseen data. For example, we can conduct a pilot test by implementing the model in a real sales scenario, where Group A (control group) follows the original lead scoring method, while Group B (test group) employs the proposed model (Chorianopoulos, 2016). By comparing the conversion rates and sales outcomes between the two groups, we could determine whether the proposed model yields better results. This approach would increase the model's immediate applicability and practical relevance by verifying its effectiveness, usability, and alignment with business goals, such as prioritizing and converting leads.

6.2.4 Cutting-edge alternative model and learning method

Deep learning models are a subset of machine learning models that are based on artificial neural networks with many hidden layers. These models are designed to learn hierarchical representations of data through multiple layers of abstraction. This enables them to handle complicated tasks such as natural language processing, speech recognition, and image recognition. Despite their powerful capabilities, they are not as widely used for analyzing market and sales data for the following reasons. First, deep learning models typically require large amounts of labeled data to perform effectively (Han et al., 2012). However, market and sales data, especially from small and medium-sized companies, may lack the volume needed for training effective deep learning models. While market and sales data are continuous, the datasets generally contain fewer instances than what is needed for a deep learning model to extract meaningful patterns without overfitting the data. In contrast, traditional machine learning models (e.g., decision trees, logistic regression, ensemble models, etc.) are usually better suited for smaller and structured sales datasets, making them more practical for marketing and sales predictions.

Second, market and sales data are highly volatile and noisy. Deep learning models may overfit historical data by capturing irrelevant fluctuations instead of learning generalizable patterns, resulting in poor performance on unseen data. Traditional machine learning models, such as random forests or gradient boosting could handle small and noisy datasets more effectively by controlling overfitting through mechanisms such as regularization or early stopping, which lead to more reliable and robust predictions.

Third, marketing and sales tasks such as lead scoring typically involve relatively simple features such as demographics, firmographics, customer engagement, or website behavior (Kotu & Deshpande, 2015). These features are often structured and do not require the extensive deep hierarchical feature extraction offered in deep learning. In contrast, traditional machine learning models could already model relationships between these structured features and the outcomes effectively without the added complexity of deep learning. While deep learning excels at feature extraction from unstructured data (e.g., images or text), this capability is less relevant for the structured data commonly found in market and sales contexts.

Finally, deep learning models are often referred to as “black boxes” due to their lack of interpretability (Chorianopoulos, 2016). In market and sales analyses, interpretability is a significant requirement, as investors and decision-makers need to understand why a model makes certain predictions or decisions. These stakeholders prefer models where the relationships between input features and the outcomes are more explainable and transparent. Models built using traditional machine learning techniques such as logistic regression, decision trees, or XGBoost are easy to interpret. Ensemble models, for example, are effective at capturing relationships in market data and are more interpretable and easier to tune than deep learning models. They provide decent performance without the complicated deep learning

architectures. Since these traditional models already perform well and are easier to implement, the added complexity of deep learning models may not provide enough advantages in the marketing and sales contexts.

However, when unstructured text data (e.g., email, social media content) becomes available, an important future work direction is to incorporate deep learning models (e.g., recurrent neural networks, convolutional neural networks). These models can effectively process complex unstructured data and extract valuable features from leads' emails, website logs, and social media interactions. Additionally, the application of reinforcement learning in predictive lead scoring for marketing and sales offers an innovative approach to optimizing lead prioritization and management (Chorianopoulos, 2016). In this context, reinforcement learning could enhance actions over time to help improve the long-term value of customer interactions and maximize conversions by learning from a sequence of decisions about how leads are processed over time. A hybrid approach combining deep learning models and reinforcement learning could potentially create a powerful system that improves decision-making in lead management. The output of deep learning models could be a conversion probability score which represents the likelihood that a lead will convert based on historical data and behavior patterns. This lead score and the extracted features could then be used as part of the state in the reinforcement learning model which then explores various sequences of actions and learns the best strategies to convert leads. Finally, reinforcement learning agents could interact with real leads or simulated environments, determining the best action at each stage based on the lead's current state and historical data. By integrating deep learning models and reinforcement learning methods, predictive lead scoring can move from static models to dynamic and adaptive systems. This hybrid approach enhances the model's ability to process complex data and optimize decisions over time, which could ultimately result in improved lead prioritizing, personalized actions, and higher conversion rates.

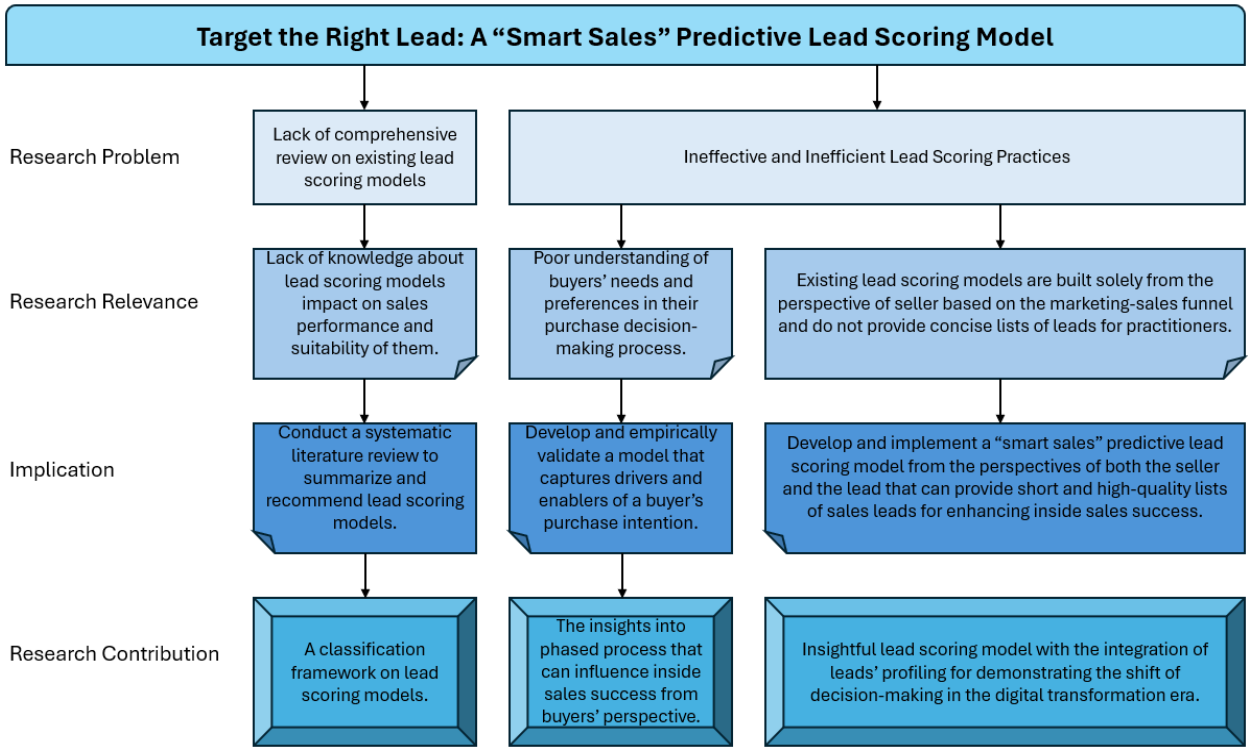

Figure 6.1. Overview of the Thesis by Articles

References

Başarslan, M. S., & Argun, İ. D. (2020). Prediction of potential bank customers: Application on Data Mining. *Artificial Intelligence and Applied Mathematics in Engineering Problems*, 96–106. https://doi.org/10.1007/978-3-030-36178-5_9

Boogar, L. (2019). The Three Stages of Lead Scoring: Lambs, Ducks & Kudus. *The MadKudu Blog post,* (accessed September 23, 2019), [Retrieved from https://blog.madkudu.com/three-stages-of-lead-scoring/].

Benhaddou, Y., & Leray, P. (2017). Customer Relationship Management and Small Data—Application of Bayesian Network Elicitation Techniques for Building a Lead Scoring Model. *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)*, 251–255. https://doi.org/10.1109/AICCSA.2017.51

Buttle, F. (2009). Managing the Customer Lifecycle: Customer Acquisition. In *Customer Relationship Management* (pp. 225–254). Routledge. https://doi.org/10.4324/9780080949611-15

Carroll, B. (2006). What 's a Lead? Target Marketing, Philadelphia 29(11)

Chorianopoulos, A. (2016). *Effective CRM using Predictive Analytics*. John Wiley & Sons.

Churchill, G. A., Ford, N. M., Hartley, S. W., & Walker, O. C. (1985). The Determinants of Salesperson Performance: A Meta-Analysis. *Journal of Marketing Research, 22*(2), 103–118. https://doi.org/10.1177/002224378502200201

Coe, J. M. (2004). The integration of direct marketing and field sales to form a new B2B sales coverage model. *Journal of Interactive Marketing*, *18*(2), 62–74. https://doi.org/10.1002/dir.20005

D'Haen, J., & Van den Poel, D. (2013). Model-supported business-to-business prospect prediction based on an iterative customer acquisition framework. *Industrial Marketing Management*, *42*(4), 544–551. https://doi.org/10.1016/j.indmarman.2013.03.006

D'Haen, J., Van den Poel, D., & Thorleuchter, D. (2013). Predicting customer profitability during acquisition: Finding the optimal combination of data source and data mining technique. *Expert Systems with Applications*, *40*(6), 2007–2012. https://doi.org/10.1016/j.eswa.2012.10.023

Dickson, P., Lassar, W., Hunter, G., & Chakravorti, S. (2009). The Pursuit of Excellence in Process Thinking and Customer Relationship Management. *The Journal of Personal Selling & Sales Management*, *29*(2), 111–124. https://doi.org/10.2753/PSS0885-3134290201

Dubinsky, A. (1981). A Factor Analytic Study of the Personal Selling Process. *The Journal of Personal Selling & Sales Management*, *1*(1), 26–33. https://doi.org/10.1080/08853134.1981.10754192

Duncan, B., & Elkan, C. (2015). Probabilistic Modeling of a Sales Funnel to Prioritize Leads. *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1751–1758. https://doi.org/10.1145/2783258.2788578

Eitle, V., & Buxmann, P. (2019). Business analytics for sales pipeline management in the software industry: A machine learning perspective. *Proceedings of the Annual Hawaii International Conference on System Sciences*. https://doi.org/10.24251/hicss.2019.125

Harmon, L. L., & Funk, D. (2014). *42 Rules for Building a High-velocity Inside Sales Team: Actionable Guide to Creating Inside Sales Teams that Deliver Quantum Results*: Super Star Press.

Homburg, C., Müller, M., & Klarmann, M. (2011). When Should the Customer Really Be King? On the Optimum Level of Salesperson Customer Orientation in Sales Encounters. *Journal of Marketing*, *75*(2), 55–74. https://doi.org/10.1509/jm.75.2.55

Järvinen, J., & Taiminen, H. (2016). Harnessing marketing automation for B2B content marketing. *Industrial Marketing Management*, *54*, 164–175.https://doi.org/10.1016/j.indmarman.2015.07.002

Jaskaran (2018) Predictive Lead Scoring: Why, How & Where. *Inbound Mantra.* (Accessed 26 Nov 2019) [Retrieved from https://inboundmantra.com/blog/predictive-lead-scoring-why-how-where]

Jansen, B. J., & Schuster, S. (2011). Bidding on The Buying Funnel for Sponsored Search and Keyword Advertising, *Journal of Electronic Commerce Research, 12*(1), 1-18.

Johnston, W. J., & Lewin, J. E. (1996). Organizational buying behavior: Toward an integrative framework. *Journal of Business Research*, *35*(1), 1–15. https://doi.org/10.1016/0148-2963(94)00077-8

Lindahl, E. (2017). A Qualitative Examination of Lead Scoring in B2B Marketing Automation, with a Recommendation for Its Practice, research project, KTH Vetenskap Och Konst. http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-213432

McDonnell, D. (2014). The Evolution From Traditional To Predictive Lead scoring. *Demand Gen Report.* (Accessed 04 March 2024), [Retrieved from https://www.demandgenreport.com/industry-resources/ebooks/2892-theevolution-from-traditional-to-predictive-lead scoring].

Michiels, I. (2008). *Lead Prioritization and Scoring: The Path to Higher Conversion*, research report, Aberdeen Group. https://pdfslide.tips/documents/lead-prioritization-and-scoring-the-path-to-higher-conversion.html

Michiels, I. (2009). *Lead Lifecycle Management: Building a Pipeline That Never Leaks*, research report, Aberdeen Group (July). https://www.ontargetpartners.com/wpcontent/uploads/2010/02/Building_A_Pipeline_That_Never_Leaks_by_AberdeenGroup.pdf

Monat, J. P. (2011). Industrial sales lead conversion modeling. *Marketing Intelligence & Planning*. https://doi.org/10.1108/02634501111117610

Ngai, E. W. T., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications*, *36*(2, Part 2), 2592–2602. https://doi.org/10.1016/j.eswa.2008.02.021

Nygård, R., & Mezei, J. (2020). Automating lead scoring with Machine Learning: An experimental study. *Proceedings of the Annual Hawaii International Conference on System Sciences*. https://doi.org/10.24251/hicss.2020.177

Ohiomah, A., Benyoucef, M., & Andreev, P. (2016). Driving inside sales performance with lead management systems: A conceptual model. *Journal of Information Systems Applied Research, 9*(1), 4–15. http://jisar.org/2016-9/

Ohiomah, A., Andreev, P., Benyoucef, M., & Hood, D. (2019). The role of lead management systems in inside sales performance. *Journal of Business Research*, *102*, 163–177. https://doi.org/10.1016/j.jbusres.2019.05.018

Ohiomah, A., Benyoucef, M., & Andreev, P. (2020). A multidimensional perspective of business-to business sales success: A meta-analytic review. *Industrial Marketing Management*, *90*, 435–452. https://doi.org/10.1016/j.indmarman.2020.08.011

Paschen, J., Kietzmann, J., & Kietzmann, T. C. (2019). Artificial intelligence (AI) and its implications for market knowledge in B2B marketing. *The Journal of Business & Industrial Marketing*, *34*(7), 1410–1419. https://doi.org/10.1108/JBIM-10-2018-0295

Paschen, J., Wilson, M., & Ferreira, J. J. (2020). Collaborative intelligence: How human and artificial intelligence create value along the B2B sales funnel. *Business Horizons*, *63*(3), 403–414. https://doi.org/10.1016/j.bushor.2020.01.003

Patma, T. S., Wardana, L. W., Wibowo, A., Narmaditya, B. S., & Akbarina, F. (2021). The impact of social media marketing for Indonesian smes sustainability: Lesson from covid-19 pandemic. *Cogent Business & Management*, *8*(1). https://doi.org/10.1080/23311975.2021.1953679

Pullins, E. B., Timonen, H., Kaski, T., & Holopainen, M. (2017). An investigation of the theory practice gap in professional sales. *Journal of Marketing Theory and Practice*, *25*(1), 17–38. https://doi.org/10.1080/10696679.2016.1236665

Puccinelli, N. M., Goodstein, R. C., Grewal, D., Price, R., Raghubir, P., & Stewart, D. (2009). Customer Experience Management in Retailing: Understanding the Buying Process. *Journal of Retailing, 85*(1), 15–30. https://doi.org/10.1016/j.jretai.2008.11.003

Sirakaya, E., & Woodside, A. G. (2005). Building and testing theories of decision making by travellers. *Tourism Management, 26*(6), 815–832. https://doi.org/10.1016/j.tourman.2004.05.004

Pöyry, E., Parvinen, P., & McFarland, R. G. (2017). Generating leads with sequential persuasion: Should sales influence tactics be consistent or complementary? *Journal of Personal Selling & Sales Management*, *37*(2), 89–99. https://doi.org/10.1080/08853134.2016.1275973

Rapp, A., Beitelspacher, L. S., Schillewaert, N., & Baker, T. L. (2012). The differing effects of technology on inside vs. outside sales forces to facilitate enhanced customer orientation and interfunctional coordination. *Journal of Business Research*, *65*(7), 929–936. https://doi.org/10.1016/j.jbusres.2011.05.005

Rodrigues Chagas, B. N., Nogueira Viana, J. A., Reinhold, O., Lobato, F., Jacob, A. F. L., & Alt, R. (2018). Current Applications of Machine Learning Techniques in CRM: A Literature Review and Practical Implications. *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI),* 452–458. https://doi.org/10.1109/WI.2018.00-53

Rutherford, B. N., Marshall, G. W., & Park, J. (2014). The moderating effects of gender and inside versus outside sales role in multifaceted job satisfaction. *Journal of Business Research*, *67*(9), 1850–1856. https://doi.org/10.1016/j.jbusres.2013.12.004

Sabnis, G., Chatterjee, S. C., Grewal, R., & Lilien, G. L. (2013). The Sales Lead Black Hole: On Sales Reps' Follow-Up of Marketing Leads. *Journal of Marketing*, *77*(1), 52–67. https://doi.org/10.1509/jm.10.0047

Shannahan, R., Bush, A., Moncrief, W., & Shannahan, K. (2013). Making Sense of the Customer's Role in the Personal Selling Process: A Theory of Organizing and Sensemaking Perspective. *The Journal of Personal Selling & Sales Management, 33*(3), 261–275. https://doi.org/10.2753/PSS0885-3134330302

Sharma, A. (1997). Who Prefers Key Account Management Programs? An Investigation of Business Buying Behavior and Buying Firm Characteristics. *The Journal of Personal Selling & Sales Management, 17*(4), 27–39. https://doi.org/10.1080/08853134.1997.10754108

Sheth, J. (1973). A Model of Industrial Buyer Behavior. *Journal of Marketing*, *37*(4), 50–56. https://doi.org/10.1177/002224297303700408

Sheth, J., Sharma, A., & Iyer, G. (2009). Why integrating purchasing with marketing is both inevitable and beneficial. *Industrial Marketing Management, 38*(8), 865–871. https://doi.org/10.1016/j.indmarman.2008.12.021

Sheth, J., Sisodia, R., & Sharma, A. (2000). The antecedents and consequences of customer-centric marketing. *Journal of the Academy of Marketing Science, 28*(1), 55–66. https://doi.org/10.1177/0092070300281006

Smith, T. M., Gopalakrishna, S., & Chatterjee, R. (2006). A Three-Stage Model of Integrated Marketing Communications at the Marketing-Sales Interface. *Journal of Marketing Research*, *43*(4), 564–579. https://doi.org/10.1509/jmkr.43.4.564

Syam, N., & Sharma, A. (2018). Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. *Industrial Marketing Management*, *69*, 135–146. https://doi.org/10.1016/j.indmarman.2017.12.019

Thaichon, P., Surachartkumtonkun, J., Quach, S., Weaven, S., & Palmatier, R. W. (2018). Hybrid sales structures in the age of e-commerce. *Journal of Personal Selling & Sales Management*, *38*(3), 277–302. https://doi.org/10.1080/08853134.2018.1441718

VanillaSoft. (2014). 4 Ways queue-based lead management is shaping the inside sales industry. *VanillaSoft* (Accessed August 20, 2019), [Retrieved from https://www.vanillasoft.com/resources/white-papers/queue-based-lead-management].

Velocity. (2013). Best lead distribution methods for optimal sales performance. *Velocity.* (Accessed November 24, 2020), [Retrieved from https://www.slideshare.net/Velocify/vel-best-leaddistromthdbooklet_nal. Accessed 16 July 2021].

Wu, M., Andreev, P., Benyoucef, M., & Hood, D. (2024a). Unlocking B2B buyer intentions to purchase: Conceptualizing and validating inside sales purchases. *Decision Support Systems, 179*. 114165. https://doi.org/10.1016/j.dss.2023.114165

Wu, M., Andreev, P., & Benyoucef, M. (2024b). Smart Sales: Amplifying the Power of Predictive Lead Scoring in B2B Sales. *The Pacific Asia Conference on Information Systems (PACIS) 2024 Proceedings. 3.* https://aisel.aisnet.org/pacis2024/track09_digittrans/track09_digittrans/3

Zoltners, A. A., Sinha, P. K., & Lorimer, S. E. (2013). The Growing Power of Inside Sales. *Harvard Business Review*.