

Decision Trees: Modeling with fast intuition and slow, deliberate analysis

Peter Darveau, P. Eng. - Research IT – Scientific Computing, University of Ottawa, Canada

Abstract - The Dual Nature of Decision Trees

Decision trees demonstrate a fascinating duality between human intuition and mathematical optimization. Psychologists like Kahneman and Tversky [1] revealed how people rely on mental shortcuts and biased, heuristic-based thinking. This mirrors how decision trees use simple, hierarchical branching based on key features - just like our minds categorize objects using decisive traits. Yet decision trees are also rigorously constructed by calculating metrics like information gain that maximize analytical power. This parallels the structured analysis of rational thinking, optimizing the tree mathematically. Supported by various works by D. Kahneman, Busemeyer et al. [2], and researchers at the university of Ottawa, this duality gives decision trees their interpretability and versatility. The visual tree structure appeals to intuitive pattern recognition, while optimized construction exploits powerful analytical techniques. Understanding this fusion between intuitive shortcuts and calculated reasoning is key to advancing decision tree capabilities and addressing their ethical and regulated use in AI (Artificial Intelligence) applications.

Keywords: ethics, psychology, machine learning, AI, decision trees

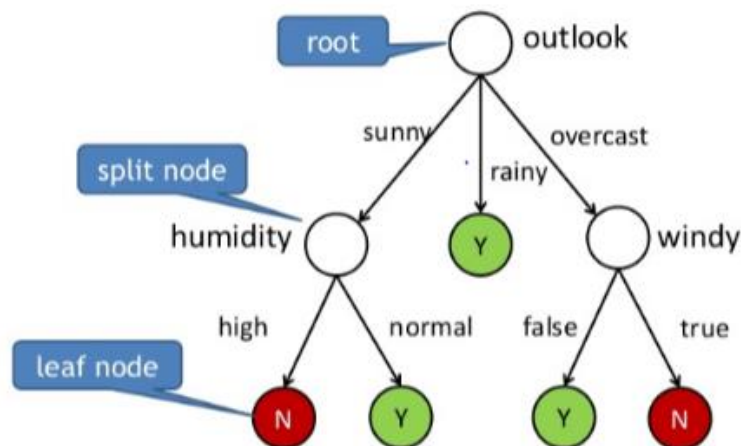
Introduction - The Psychology Within the Machine

Psychology professor Daniel Kahneman's work on fast and slow thinking indicates that the interplay between intuition and analysis is central to human cognition. His research with Amos Tversky on biases and heuristics provides insight into the intuitive side of the duality. The analytical side goes far back with Claude Shannon's information theory where notions of entropy, information quantifying and communication capacity provide framework on information dynamics that model cognitive processes core to many machine learning techniques like decision tree optimization. The cross-pollination of ideas between these fields continues to drive progress in designing artificial agents [3] that more closely mirror sophisticated cognitive abilities. By analysing these latent connections, this paper elucidates theoretical foundations that underpin both Shannon's and Kahneman's visionary contributions as well as ongoing development of machine learning techniques that bring us closer to truly cognitive systems [4].

How Decision Trees Work

Decision trees represent an intriguing fusion of mathematical optimization and nuances of human psychology. Tracing the path from root to leaf reveals both the logical and intuitive aspects of how these models operate.

Decision Tree



Following the Flow from Root to Leaf

The tree is constructed top-down, starting from the root node which contains the complete dataset. Features are evaluated to split the data into binary branches descending to child nodes. At each node, the feature that best separates the data is used to split the dataset for the left and right branches. This recursive partitioning continues until leaf nodes are reached that contain the most homogeneous partitions of the target variable.

The algorithm looks at all the features and calculates which one provides the most information gain about the target variable. This feature with the optimal split is used to divide the data into two groups based on whether they meet the split criteria or not. For example, a root node in a tree predicting credit risk may consider features like income, age, and credit history. If income provides the most information gain, the data is split into two child nodes using an income threshold, like $< \$50k$ income and $\geq \$50k$ income. These child nodes now contain a subset of the original data that is more homogeneous with respect to credit risk than the complete dataset. The algorithm repeats the process of evaluating features and splitting the data on each child node. This recursive partitioning continues until leaf nodes are reached that provide no improvement in information gain. The leaf nodes contain the most homogeneous partitions of credit risk targets.

Information Gain - Finding the Best Split

To determine the most decisive split at each node, decision trees employ information gain. This metric measures how much information a feature provides about the target variable. The feature that reduces entropy the most in the node's dataset has the highest information gain and is used for the split. Entropy represents the impurity of targets within the node. Higher entropy means more uncertainty. The decision tree aims to maximize information gain at each step so that child nodes are purer than parent nodes. To calculate information gain for a potential split on a feature, the original node entropy is compared to the weighted average entropy of the child nodes after splitting. The larger this reduction in entropy, the higher the

information gain. For example, income may have high information gain for the credit risk root node because splitting on income dramatically decreases the entropy in the child nodes. The tree is optimized by greedily selecting the feature with highest information gain at each step. This provides mathematical rigor to find the most informative splits. By maximizing information gain at each step, the tree optimizes its analytical power.

As described, the branching criteria arises from an optimization process during training. Information gain quantifies the reduction in entropy achieved by a given attribute split. Attributes that maximize information gain are higher in the tree. This mathematical optimization mirrors the cognitive biases that Kahneman and Tversky uncovered [5]. By maximizing information gain, decision trees minimize ambiguity and impose order, just as humans are intrinsically motivated to reduce uncertainty and perceive structure. The information gain metric bakes our tendency toward clarity and definitiveness into the model. Consider this example for a selected iris classification dataset using information gain to select sepal width and petal width as the top splits:

```
Sepal length:
- Entropy if sepal length ≤ 5.5 = 0.909 (mostly versicolor)
- Entropy if sepal length > 5.5 = 0.963 (mixed classes)
- Weighted entropy = (112/150 * 0.909) + (38/150 * 0.963) = 0.921
- Information gain = 1.577 - 0.921 = 0.656
Sepal width:
- Entropy if sepal width ≤ 3.0 = 0 (only setosa)
- Entropy if sepal width > 3.0 = 1.0 (mixed versicolor and virginica)
- Weighted entropy = (50/150 * 0) + (100/150 * 1.0) = 0.667
- Information gain = 1.577 - 0.667 = 0.91
Petal length:
- Entropy if petal length ≤ 4.95 = 0.918 (mostly versicolor)
- Entropy if petal length > 4.95 = 0.971 (mixed classes)
- Weighted entropy = (137/150 * 0.918) + (13/150 * 0.971) = 0.921
- Information gain = 1.577 - 0.921 = 0.656
Petal width:
- Entropy if petal width ≤ 1.75 = 0.592 (mostly versicolor)
- Entropy if petal width > 1.75 = 0.964 (mixed classes)
- Weighted entropy = (112/150 * 0.592) + (38/150 * 0.964) = 0.701
- Information gain = 1.577 - 0.701 = 0.876
```

Sepal width has the highest information gain, followed by petal width. Therefore, sepal width and petal width would be selected as the top two splits. However, modifying the width or length criteria can have a non-negligible effect on the selection.

As shown above, the flow from root to leaf uses human-like categorical thinking, while information gain provides mathematical rigor showing that decision trees exemplify two facets of cognition - intuition and analytics. This dual nature gives decision trees their intuitive interpretability combined with analytical optimization. Understanding this constructive collaboration is key to using them in application using these algorithms.

The Rational Side – The calculated approach to decision Trees

Decision trees follow a logical, analytical process of splitting data based on features that provide the most information gain. This data-driven approach is systematic and rational: The structure of the decision tree from root to leaf nodes lays out a clear sequence of rules/conditions to categorize data. This rule-based thinking aligns with rational analysis showing that Decision trees are deterministic because they will always produce the same output given a set of inputs. The very structure of decision trees reflects the measured, reasonable thought process of System 2. Their systematic analysis encapsulates how our brain methodically weighs decisions when we have time for reflective thinking.

Logical, Step-by-Step Analysis

Decision trees epitomize the systematic, analytical thinking of Kahneman's System 2. They break down a decision into a highly structured visual flowchart, weighing potential outcomes at each branch point. This step-by-step analysis reflects the deliberate and intentional approach of our slow, rational thinking. Constructing a decision tree requires methodically identifying key decision points, viable options at each point, probabilities, and payoffs for each outcome path, and calculating expected values. This level of rigorous logical analysis aligns closely with System 2's calculative and analytical nature. For example, a business might use a decision tree to systematically assess whether to launch a new product. The tree would map out critical go/no-go gates, like estimated manufacturing costs, projected customer demand based on market research, competitive landscape, marketing budget required, and potential revenue projections. Statistical models and historical data would feed the probabilities and financial outcomes at each branch. The optimal path is rationally derived by cascading through the tree. Decision trees are touted for removing emotions, biases, and gut feelings from the decision process. As Kahneman describes, System 2 thinking is our objective, statistical mode of thought. Decision trees quantify options and potential consequences in a dispassionate, empirical manner. The goal is to deliberately shape choices based on logic and facts vs intuitions.

Data-Driven Decisions

Proponents argue decision trees can overcome cognitive biases by formally structuring the decision analysis based on available data. For example, placing too much weight on recent events versus long-term averages is avoided by statistically modelling outcomes. Formal decision trees provide a rational framework to counter our irrational instincts. However, some argue that decision trees cannot entirely escape human biases or emotions. There is subjectivity in how the tree is constructed, which branches are included, the probabilities and outcomes estimated, and how options are framed. So, while decision trees aim for systematic analysis, elements of System 1 thinking inevitably influence the structure. In summary, the rigorous analytical approach of decision trees strongly reflects the measured, reasonable thought process of Kahneman's System 2. But our fast-thinking System 1 may still shape these

purportedly "rational" models in subtle ways. The debate continues how much unemotional logic can be achieved.

Here is an example of proponents arguing that decision trees can help overcome cognitive biases: In their book "Smart Choices", John Hammond, Ralph Keeney, and Howard Raiffa, leading decision analysis experts, argue that decision trees and other structured analytical approaches can counteract many ingrained human cognitive biases. For example, they cite anchoring bias, where people put too much emphasis on initial starting points or suggestions. Structured analysis using decision trees forces a broad exploration of options rather than anchoring on one. Another example is confirmation bias, where people favour information that confirms their existing beliefs. Decision trees require systematically weighing all evidence, counterarguments, and probabilities rather than just confirming one's initial hunch. As the authors state, "By formalizing the steps of good decision making, decision analysis disciplines thinking, counters the decision maker's tendency to irrationality, challenges premises, and frames issues in a way that leads to sound and innovative strategies. They argue decision analysis is "a defense against irrationality" by providing a structured approach that reduces cognitive biases and avoids relying solely on intuition. Using decision trees helps shift from fast, instinctive System 1 thinking towards slower, more analytical System 2 reasoning.

The Emotional Side - Unconscious Biases Built into Decision Trees

Selecting the right features or questions to include in the decision tree can involve intuition and heuristic-based thinking vs. pure data analysis. This introduces an element of human judgement. There are creative aspects in visualizing and building the tree structure that could be seen as subjective or instinctual vs. completely formulaic. Some argue the branching structure of decision trees loosely mimics how the brain thinks and wires connections, introducing a human/emotional element. The final predictions or classifications made by the decision tree require interpretation, which allows human emotions and biases to shape how we utilize the output.

Intuition in Heuristic Selection

Classifying an observation by starting at the root node and following branching rules appears orderly and analytical. However, Kahneman and Tversky' work on prospect theory [6] and decision behaviour [7] shows that human reasoning involves more than just conscious analysis. Their research revealed the predominance of instinctive mental shortcuts and gut reactions, especially under uncertainty [8]. Once trained, decision trees can quickly categorize observations by pattern matching down the learned branches. This rapid, reflexive route to a decision closely mirrors the biased heuristics and mental shortcuts that Kahneman and Tversky identified as a core part of human judgment [9]. Decision trees exhibit both the logic of analytical training and the intuition of experiential pattern recognition.

While decision trees aim to provide an objective, analytical approach to decisions, there are inherent challenges in entirely removing bias and emotion. One issue is framing effects. The way the decision question and options are framed can influence the optimal choice, and there are often multiple valid ways to frame a given decision. The structure of the decision tree steers the framing and introduces subjectivity. For example, a tree evaluating a new product launch could frame the decision as "launch product" or "kill project". Or branch outcomes could be framed as profits vs losses rather than raw dollar values. The framing shapes the perceived attractiveness of options. Another bias is overconfidence in projections. Estimating probabilities and outcomes often relies on subjective predictions and gut feelings. People tend to be overconfident in their own projections and assumptions. These overconfident assessments get built into the decision tree branches. Anchoring bias also creeps in, as existing data points or rules of thumb anchor estimates. For example, an optimistic sales projection for a new product may get anchored based on success of a previous product rather than objective data. Finally, people often construct trees to favour options they intuitively prefer. Branches may be added to justify rather than truly analyse a choice. This makes the tree more of a rationalization of a predetermined decision vs an unbiased analysis. While decision trees bring more structure and analytics to choices, Kahneman would argue System 1 thinking still permeates the subjective inputs and framing of the analysis. Our fast, instinctive biases shape elements of the rational, unemotional model. So, in practice, decision trees incorporate a blend of System 1's biases and heuristics with System 2's objective calculation. Pure rationality is limited by the inherent subjectivity in how these decision tools get constructed.

Human Judgement in Model Building

The framing of options and outcomes in a decision tree can impact the analysis. Different framings steer thinking and implied value judgements. For example, a new product decision could be framed as either "launch product" or "launch product if market research supports it". The latter frames the decision as more tentative and conditional, focusing on evidence needed to support launching. Or branch outcomes could be presented as profits vs losses, even if the raw dollar values are equivalent. Losses loom larger psychologically, so the frame sways the decision. Positive and negative framings elicit different risk preferences. Anchoring Bias-estimating probabilities and outcomes often anchor on available precedents or rules of thumb, even if newer data suggests those anchors are outdated or invalid. People cling to existing anchors due to familiarity. For instance, sales projections for a new product may anchor too closely to the results of an older product, vs gathering unbiased data on market demand for the new offering. Anchors exert an unconscious pull. Overconfidence occurs when subjective predictions and assumptions get embedded into decision trees. But research shows people are often overconfident in forecasts or probabilistic assessments when there is ambiguity or uncertainty. So, a project team may have unreasonable faith in the accuracy of their projections. Without realizing it, people build overconfidence into subjective branch estimates, skewing the analysis. People constructing decision trees are prone to confirmation bias, causing them to over-weight branches confirming their initial intuition and under-weight

contradictory evidence. Rather than conducting an open-minded analysis, tree builders subtly distort the structure to favour their preconceived notion. The final tree ends up rationalizing rather than truly analysing a decision. Decision tree builders cannot fully escape the influence of System 1's inherent biases, even as they strive for System 2's objectivity. Subjectivity permeates framing, estimation, and structuring of the analysis. Pure rationality is limited by human psychology.

When Decision Trees Conflict with Moral Intuition

There can be cases where decision tree analysis yields a solution that conflicts with moral intuition: Sometimes a statistically optimal decision tree prescribes an action that feels morally questionable or repugnant. For example, a tree may suggest segmenting customers by race to maximize profit. Or firing loyal employees to cut costs. The tree output can seem to violate ethics, values, and deeper human considerations beyond the data itself. Relying solely on decision tree guidance risks "the dehumanization of human life", as Phil Tetlock describes it. It frames everything as statisticized inputs to maximize utility functions. But people are more than just data points. There are social, emotional, and moral dimensions beyond the quantifiable branches. Consider healthcare rationing decisions during COVID-19 based on maximizing lives saved. These evaluation protocols sparked outrage as they assigned higher priority to the developed countries versus the ones under development. Sometimes going against the data to follow one's moral compass can pay off long-term. For example, Patagonia, a sportswear company, privately held to maintain its social values, though the decision tree based solely on business metrics would have said to sell out. This built tremendous customer loyalty and employee engagement over decades, driving growth. Tetlock argues System 1 moral intuition serves as a moderating force against the excesses of System 2 detached rationality. So, while decision trees bring order, they may miss the deeper human context. Wisdom means integrating the calculus of the mind with the compassion of the heart.

Here are some additional examples and thoughts:

- A marketing decision tree may suggest a viral advertising campaign that relies on controversial shock value or stirring up unhealthy peer pressure. This can boost sales but feels unethical.
- A hiring decision tree may indicate rejecting candidates based on protected class characteristics like race, gender or age due to statistical correlations, even if unlawful. The algorithm aims to maximize productivity, ignoring fairness.
- An insurance risk model may prescribe denying coverage or raising rates based on genetic test results, even if discrimination based on genetics is prohibited. The model exploits loopholes in the law.
- A political campaign decision tree may recommend spreading disinformation or preying on voters' fears if polling shows it will influence votes. Ignores integrity principles.

In these cases, blindly following the decision tree output without considering broader ethical implications can lead organizations and leaders astray. However, total reliance on moral intuition and instinct has its own perils. Gut feelings are prone to in-group favouritism, confirmation bias, and prejudice. So, ethics require a balancing of statistics and social values. Many argue ethical AI and algorithms should be designed to align with moral values from the start. For example, models could be optimized to maximize fairness or opportunity instead of just profit or productivity. In the end, human oversight and governance is crucial. Leaders must cultivate the wisdom to critically evaluate where decision tree guidance holds merit, and where it violates core human values. Statistics inform but cannot replace moral reasoning.

Blending logic and intuition in decision making

Decision trees exemplify the fusion of human intuition with mathematical optimization that makes machine learning so powerful. Tracing the path from root to leaf reveals the interplay between fast heuristic thinking and slower analytical judgment.

As said by Blaise Pascal "The heart has its reasons which reason knows not.". Decision making relies on a fusion of logic and intuition. Pure logic suffers from emotional blindness, while pure intuition lacks analytical rigor. Optimal choices emerge when we blend conceptual wisdom and empirical data (Klein, 2003). Like two overlapping circles in a Venn diagram, logic and intuition provide complementary strengths with a region of intersection containing our deepest insights. Intuition encompasses the conceptual leaps and emotional hunches that often precede conscious reasoning (Gladwell, 2005). It is the "gut feel" that guides our rapid judgments, harnessing experience into heuristic mental models (Gigerenzer, 2008). But without logical analysis, intuition risks irrational biases and false assumptions. For example, first impressions of people are intuitively assembled within seconds, but require measured reasoning to overcome our cognitive biases (Ambady, 2001). Formal logic provides an anchoring counterweight, demanding evidence, and statistical norms (Stanovich, 2011). Yet purely logical decisions struggle in novel situations with limited data. They also ignore the rich contextual and emotional factors that shape human choice. Strictly rational decisions, like choosing utilitarian outcomes in moral dilemmas, often clash with our moral intuitions (Greene, 2001). Optimal decision making embraces the symbiosis between logic and intuition (Dijksterhuis et al., 2006). As Nobel laureate Daniel Kahneman (2011) described in *Thinking Fast and Slow*, human cognition thrives on this dual processing. Our quick intuitive judgments benefit from deliberation, just as formal analysis succeeds when guided by meaning and emotional value. We see this fusion in practices like design thinking, which cycles divergent intuitive brainstorming with convergent analytical selectivity (Brown, 2008). We also see it in wise leadership, combining compassionate rapport with rigorous strategy (Nonaka & Toyama, 2007). Blending logic and intuition leads to our deepest insights and soundest choices. It enables innovation guided by empathy, and plans anchored in human values. By integrating the signals of our intuition with the focus of logical analysis, we gain the wisdom that arises when rationality meets emotional meaning. We access a deeper intelligence informed by both data and our innate humanity.

Balancing Fast and Slow Thinking in Decision Trees – Shaping a vision

The pioneering research of Daniel Kahneman and Amos Tversky on cognitive biases provides insight into how decision trees blend intuition and analysis. Their work on prospect theory identified two modes of thinking: Fast, instinctive System 1 which operates automatically and quickly, relying on heuristics and shortcuts; Slow, deliberate System 2 which allocates attention to mental activities requiring effort and logic. Decision trees utilize both systems. The hierarchical structure and branching decisions mirror System 1 heuristics to categorize observations. But the information gain metric to optimize splits relies on System 2 mathematical analysis. Pruning branches to avoid overfitting also balances fast thinking biases with slow analytical judgment. Ensemble methods like random forests further leverage the strengths of quick intuitive trees with rigorous statistical learning. The most effective decisions synthesize the constructive interaction between System 1 and 2. Consider these scenarios:

- A sales manager needs to prioritize leads. She first builds a lead scoring model analysing size, likelihood to buy, customer profiles, etc. This systematic analysis brings logic. But she also relies on her field reps' gut sense of hot prospects not fully captured by the data. They see enthusiasm and relationship strength that boosted unlikely deals before.
- A startup CEO uses decision tree analysis to quantify the odds of success across product, pricing, and partnership options. But he also taps his vision to imagine possibilities beyond the defined branches. This sparks an inventive pivot when the math said to stay the course.
- A chief strategist at a bank maps out acquisition targets based on growth potential, cost, culture fit, etc. But she also pays attention to subtle concerns raised in management meetings that aren't quantified, like low morale at one target. She adjusts valuation accordingly.
- A policymaker stress tests proposed pandemic lockdown models based on infection curve projections versus economic impact. But he also goes out in the community to get beyond the statistics, sensing people's patience wearing thin. This nudges him toward a more moderate plan.

Here are some ways to balance fast and slow thinking when building and applying decision tree models, drawing on Kahneman's theory:

- Slow thinking for model building: Take time to carefully select input features, tuning parameters, and methodology. Avoid rushing this process or making snap judgements. Consider interactions and nonlinearities.
- Fast thinking for prototyping: Use decision trees as a fast-prototyping tool when exploring new data or ideas. They allow quick iterations to gain insights.
- Slow thinking for interpretation: Carefully inspecting the decision tree structure and evaluate feature importance metrics. Don't jump to quick conclusions.
- Fast thinking for basic predictions: Use a simple pre-trained decision tree for fast predictions on new standard data. Leverage fast thinking.

- Slow thinking for anomaly detection: Have processes to detect anomalous new data points before predictions. Use slow thinking to investigate.
- Fast thinking with safeguards: Add safeguards like warning systems for unlikely predictions. This allows fast thinking while reducing risk.
- Slow thinking for regular reviews: Occasionally reviewing and retune the decision tree as new data comes in. Don't just always apply the same tree.

The key is to leverage the strengths of both modes of thinking: using fast thinking for efficient application and prototyping while relying on slow thinking for careful analysis and review. A clear decision thought process is part of a model's verification and validation framework to minimize Fast thinking (System 1) inherent biases.

Combining Logic and Instinct in Decision Making – Verify and Validate

Formal logic and rigorous analysis bring empirical discipline to decisions. They prevent unexamined assumptions and force us to confront hard truths. However, purely logical choices often lack meaning, given their emotional detachment. Decisions based solely on technical optimization rather than human values tend to backfire in practice (Flyvbjerg, 2006). Equally, intuitive instinct alone is prone to subjectivity and prejudice masquerading as wisdom. While conceptual leaps synthesize experience quickly, they require testing and verifying to avoid reinforcing false beliefs. Relying purely on emotion or charismatic rapport is as dangerous as detached rationality (Nonaka & Toyama, 2007). Therefore, sound judgment blends analytical rigor with compassionate intuition (Snowden & Boone, 2007). Logic structures and focuses our thinking, without drowning out our humanity. Intuition connects us to what we care about through insight and emotion. But it benefits from the discipline of impartial analysis. By braiding together System 1 fluidity and System 2 focus, we reach decisions both technically sound and emotionally aligned. This integration is verified through logic and validated with data to ensure robustness. However, statistical validation [10] is only one side of the coin. Humans must also apply critical thinking to question the model's assumptions and real-world applicability, drawing on instinct developed from domain expertise. For high-stakes decisions, model-predicted probabilities should inform but not determine outcomes. Ultimately, by synergistically integrating the logic of validation data and the human instinct of scrutiny, machine learning can be leveraged reliably. Ongoing monitoring and incremental improvement further allow the symbiotic fusion of logic and instinct to adapt models to new data, developing trusted and valuable AI solutions [11].

Below are some technical tips on tuning decision trees:

- Feature Selection:
 - Recursively eliminate features and compare model performance to find optimal subset. Removes irrelevant or redundant features.
 - Use a p-value threshold from statistical tests to identify noteworthy features. Removes

noise variables.

- Evaluate feature importance scores and remove low importance features. Keeps only predictive features.
- Hyperparameter Tuning:
 - Tune max depth by incrementally increasing and checking for overfitting. Find sweet spot before overfitting.
 - Tune min samples for splits by trying different values and checking impact on tree size and accuracy.
 - Tune leaf node size by incrementally lowering and evaluate overfitting. Avoids overly specialized leaves.
 - Tune pruning severity parameter if using cost complexity pruning. Simplifies over-complex trees.
- Ensemble Building:
 - Build multiple de-correlated trees by bagging or random forests and average for more robust estimates.
 - Incrementally increasing number of estimators and evaluate performance. Find sweet spot before overfitting ensemble.
 - Tune max features considered per split in random forest to control correlation and overfitting.
- Optimization:
 - Use grid search to systematically evaluate different hyperparameter values and find optimal settings.
 - Employ random search by sampling hyperspace to cover more ground and avoid local optima.
 - Use Bayesian hyperparameter optimization to focus search in promising areas based on previous results.

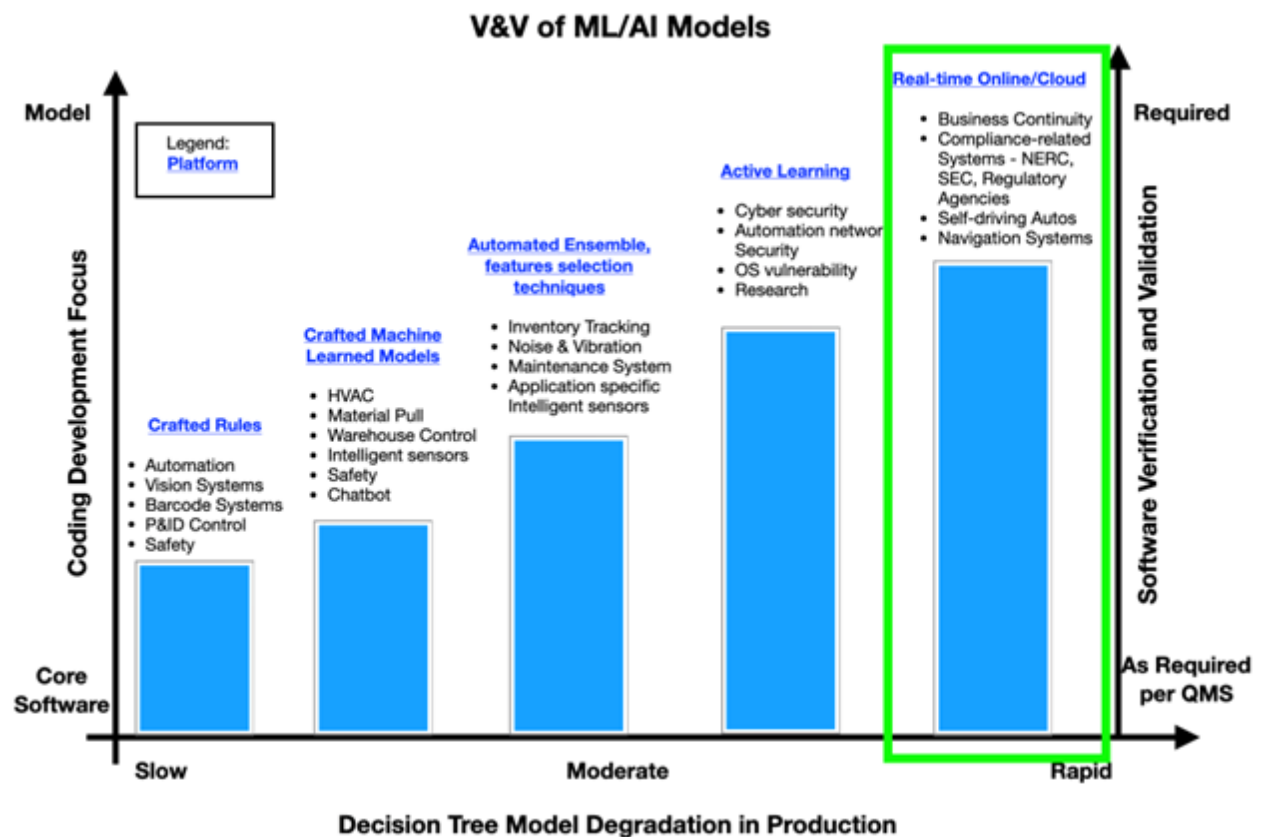
The goal is to leverage a mix of heuristics and systematic optimization to thoroughly search the hyperspace and find the ideal tree configuration without overfitting.

In combination with the above, here are some tips for using intuitive domain knowledge to judge if a decision tree model aligns with expectations [12]:

- Review the split variables and values at each node - do they make intuitive sense based on your understanding of the data relationships? If not, that's a flag.
- Look at the terminal leaf nodes - do the final predictions match what you would expect for those groups? Surprising predictions are a clue something may be off.
- Examine the most influential features - are these in line with your domain expertise? Counterintuitive prominent features are a sign of potential issues.
- Visualize the tree structure - does the flow match your mental model of the data? Overly complex or fragmented trees may be suspect.
- Look at individual predictions - spot check some samples against your intuition. Disagreements indicate a problem.

- Consider symmetry - does the tree look balanced based on your knowledge? Lopsided trees or overrepresented groups can be a clue.
- Check for confirmation bias - are there intangible reasons results may confirm your existing assumptions vs genuinely fitting the data?
- Evaluate segmentation - do any groups receive similar predictions that you would expect to be distinguished?
- Compare model variations - does adding/removing expert knowledge shift predictions as anticipated?
- Consider edge cases - are common exceptions handled properly according to your expertise?

Carefully trust your judgement as you thoroughly examine the model from different perspectives. Ensure it aligns with and enhances the model requirements of the domain subject matter expert but consider how dynamically the data changes. Models refreshed with higher ratios and volatility of new data are more susceptible to skewing and biases that require more thorough verification and validation to flush out as shown in the chart below [13].



Conclusion - Decision Trees Mimic Our Own Decision Processes

The step-by-step flow of decision trees appeals to the fundamental structures of human cognition. Their branching shape resonates with how our minds categorize information and make sequenced judgments. We naturally break down complex problems into hierarchical decisions, continuously partitioning information to reach conclusions (Klein, 1998). In many ways, decision trees automate this process of human reasoning. Their greedy, recursive splits emulate how we separate experiences into ever-purer groupings. For example, a doctor diagnosing a patient's illness [14] will often narrow down symptom clusters to home in on a specific disease. Their information gain metrics quantify the decisiveness of each partitioning, much like we intuitively gauge which factors matter most in making judgments. From disambiguating medical diagnoses to navigating social interactions, humans have always parsed complexity through trees of sequential choices (Peng & Reggia, 1990). Therefore, beyond their analytical power, decision trees also provide a window into our own minds. Their aesthetic design and optimization process mirror the structured flow of human reasoning. In training decision trees, we instil machines with a simplified version of our own mental processes for breaking down ambiguity (Quinlan, 1986).

The successful and responsible application of artificial intelligence relies heavily on collaboration with relevant subject matter experts. Their specialized knowledge is essential for defining the right problems for AI to solve, providing insights into key variables and complex aspects of the domain, and contextualizing raw data needed for training robust models. During development, they can provide validation by gauging whether an AI's behaviour aligns with real-world expectations. Experts play a crucial role in identifying potential risks or shortcomings that may arise from narrow AI capabilities. Their nuanced human judgment acts as a counterbalance to AI systems' data-driven recommendations. Given inherent biases in data or design oversights, subject matter expertise helps ensure AI is deployed only where appropriate and with the necessary governance. By partnering with those who deeply understand the field of application, AI builders can produce solutions that solve problems correctly, efficiently and reflect an authentic representation of reality. The practical wisdom of experts guides AI to augment human capabilities rather than simply mimic them. The resulting models can then augment human judgment with the benefits of mathematical rigor and scale. Just as our brains parse signals from noise, decision trees extract patterns and structure from vast data. Understanding decision trees grants insight into both mathematical modelling and the mechanisms of human choice.

Finally, as artificial intelligence technologies become more advanced and integrated into our lives, thought leadership [15][16] is crucial to steer these tools to benefit society. With AI's potential to automate jobs and impact many industries, we need diverse perspectives to shape policies and practices that are ethical and inclusive. Thought leaders who understand both technology's capabilities and limitations as well as wider social contexts are best equipped to advocate for responsible AI development. They can identify current gaps in representation, safety, accountability and provide expertise or recommendations to address them. Whether in academic research, corporate policies, government regulation or public discourse, forward-thinking leaders who proactively engage with emerging technologies can help direct innovation down paths aligned with human values and priorities. As AI grows more powerful, strong thought leadership across disciplines

that grounds it in wisdom and ethical reasoning is essential to imbue these transformative tools with a moral compass.

Future Work – Disambiguating Noise and Bias when working with Decision Trees

While this paper attempts to identify the problems and challenges in keeping Decision Tree models trustworthy, it would be valuable to introduce the notions of predictive and evaluative judgement of decision tree models to the picture. The purpose would be to add a more focused quantitative angle which would help set an objective framework when assessing and managing the risks of using AI models is uncertain [17].

Author Contributions: Writing-original draft PFD. writing-review and editing PFD, JvdK

Conflicts of Interest: Author declares no conflict of interest.

References:

- [1] Daniel Kahneman, 2011. "Thinking fast and slow"
- [2] "Dynamic models of choice" (Busemeyer & Diederich, 2002) - Reviews different dynamic models of decision-making including connections between decision field theory and prospect theory.
- [3] Darveau, P., 1993. C programming in programmable logic controllers.
- [4] Darveau, P., 2020. Software UX enhancement using AI and NLP
- [5] "Advances in prospect theory: Cumulative representation of uncertainty" (Tversky & Kahneman, 1992) - Seminal paper by Kahneman and Tversky extending prospect theory to uncertain choices.
- [6] https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=prospect+theory&btnG=
- [7] https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=kahneman+and+tversky&btnG=
- [8] https://scholar.google.com/scholarhl=en&as_sdt=0%2C5&q=uncertainty+prospect+theory&btnG=
- [9] https://scholar.google.com/scholarhl=en&as_sdt=0%2C5&q=kahneman+tversky+human+judgment&btnG=

- [10] "A quantum theoretical explanation for probability judgment errors" (Busemeyer et al., 2011) - Proposes using quantum probability as a framework to integrate existing decision-making theories.
- [11] "Decision field theory: An integrative framework for behavioral decision making" (Busemeyer & Johnson, 2004) - Provides a comprehensive overview of decision field theory and relates it to other decision-making models
- [12] Darveau, P. Prognostics and Availability for Industrial Equipment Using High Performance Computing (HPC) and AI Technology. Preprints 2021, 2021090068 .
- [13] Darveau, P., 2020. Production Alarms using Machine Learning
- [14] Lusty, A., 2021. *Understanding decision making in robotic surgery: a knowledge gap survey and cognitive task analysis of robotic prostatectomy* (Doctoral dissertation, Université d'Ottawa/University of Ottawa).
- [15] Szilagyi, K., 2022. *Artificial Intelligence & the Machineation of the Rule of Law* (Doctoral dissertation, Université d'Ottawa/University of Ottawa).
- [16] Burke, M., 2015. *Frege, Hilbert, and Structuralism* (Doctoral dissertation, Université d'Ottawa/University of Ottawa).
- [17] "Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment" (Busemeyer & Townsend, 1993) - Discusses integrating decision field theory and prospect theory for modeling preferences.