

The Managerial Blind Spot Effect Explainable Machine Learning for Strategic Decision-Making

¹*Sai Srujan Reddy Tokala

¹Clark University, USA.

Abstract

Artificial Intelligence (AI) is used in strategic decision-making; however, it is not clear whether the use of explainable machine learning leads to more managerial judgment when prediction accuracy is high. This study analyzes the proposed 'Managerial Blind Spot Effect' by combining predictive modelling with behavioural analysis applied to the IBM Telco Customer Churn dataset and the AI Decision Dependence and Cognitive Caution dataset. Several models, including Explanatory Boosting Machine (EBM), XGBoost, Random Forest and Neural Network, were tested, and the explanation produced by the native model EBM was enhanced by SHAP-based explanation. The behavioural results revealed that both Trust in AI and Cognitive Caution significantly contributed to an increase in AI Dependence. EBM delivered the best predictive performance with intrinsic transparency. The results highlight the importance of using explainable AI prediction models in conjunction with managerial judgment in dashboards.

Keywords: Explainable Artificial Intelligence, Strategic Decision-Making, Managerial Blind Spot Effect, AI Dependence, Customer Churn Prediction.

Introduction

Organizations have embraced Artificial Intelligence (AI), as it plays a crucial role in decision-making processes (Shrestha et al., 2019). AI can help businesses navigate the vast amounts of information generated by their operations and convert that data into valuable information that can guide their decision-making for both strategy and operations (Rane et al., 2024). In the finance sector, CRM, healthcare, and supply chain management, machine learning models are increasingly being leveraged to deliver more predictive analysis, automate simple analysis and improve the organization's responsiveness (Khan, 2025). The suggestions these predictive systems give are also now playing a role in the management workflow and are making decisions that were previously made by the experts. The growing utility of predictive analytics also highlighted a lack of transparency, accountability and managerial knowledge of algorithmic recommendations, as the number of complex machine learning models keeps increasing (Busuioc, 2021; Olayinka, 2019).

To address these concerns, Explainable Artificial Intelligence (XAI) has begun to emerge as one of the most

significant research areas for making machine learning systems more interpretable (Xu et al., 2019). The methods of explainability give explanations that are comprehensible to the user of the model, and can be employed to uncover what is driving the model's prediction (Hassija et al., 2024; Linardatos et al., 2020). One of the most popular post-hoc explanation approaches was known as SHapley Additive exPlanations (SHAP), which is used extensively to explain black-box models (Kamath & Liu, 2021; Tan et al., 2023). There has been growing attention to intrinsically interpretable models, like the Explainable Boosting Machine (EBM), that predict without the need for another explainability algorithm. Explainability is becoming a necessary and integral part of a trustworthy AI decision-making system, it is not just a desirable, but useful, analytical tool (Chinnaraju, 2025).

Although these progressions have been made, there is still fundamental issues that are not fully clarified. Existing research on explainable AI concentrates primarily on technical issues such as the quality of explanations, the importance of features and predictability. Less focus has been placed on the behavioural implications of explainable

Sai Srujan Reddy Tokala
Clark University, USA.
Email: saisrujan816@gmail.com

Received: 2-Sep-2026

Revised: 14-Sep-2026

Accepted: 8-Oct-2026



© 2026 Copyright by the Authors.

Licensed as an open access article using a [CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/).

AI, particularly how managers understand, trust, and ultimately act on AI-recommended actions. Indeed, often decisions in organizations were taken under pressure due to which a manager follows the algorithmic advice given without critically verifying the reasons for the decision or evaluating other choices (Peeters, 2020). But, explainability is not enough to reduce the risk of using the AI inappropriately when the user takes the explanation as an indication of correctness, instead of as information for analysis (Bauer et al., 2023).

The phenomenon in this study is called the Managerial Blind Spot Effect the more a manager relies on AI's recommendations, the less they rely on explanations meant to guide them toward making their own decisions (Chen et al., 2023). The proposed perspective suggests that the explainability of the decision might also be beneficial to managerial behavior, but the technical characteristics of the AI system, and the cognitive characteristics of the individual, also influence managerial behavior (Bankins et al., 2024). With the increasing use of explainable AI dashboards by more organizations to inform strategic decisions on customer retention, financial forecasting, operational planning and resource management, it is becoming more critical to understand this interaction.

This phenomenon is studied by an empirical design with a dual-dataset. The first dataset is the IBM Telco Customer Churn dataset for which four machine learning models, namely, Explainable Boosting Machine (EBM), XGBoost, Random Forest, and Neural Network, are evaluated for predictive performance. Native explainability produced by EBM is then contrasted with post-hoc SHAP-based explanations to check the difference between explainability approaches in terms of both explainability and managerial transparency. The second dataset, the AI Decision Dependence and Cognitive Caution dataset, is used to study the behavioural aspects of AI-driven decision making using the relationships between AI Dependence, Cognitive Caution, and Trust in AI. Predictive modelling and behavioural analysis used together to study explainability from a technical and managerial point of view (Vudugula et al., 2023).

The research contributes to the literature in three major ways. First, it moves explainable AI from a model-centered approach to a decision-centered approach, and asks whether explainability enhances managerial understanding instead of simply helping to understand the model. Second, it embeds intrinsic explainability, post-hoc explainability, and behavioural decision analysis

into a single empirical framework to allow for an all-encompassing evaluation of explainable machine learning dashboards. Third, it brings empirical evidence in favor of the proposed Managerial Blind Spot Effect, showing that although models that are explainable do not negatively impact the predictive ability of AI. Having an explainable model also positively influences the manager's trust in the model and their willingness to critically assess the algorithmic recommendations for improved organizational decision making (Patil, 2024). As a whole, the works supply a comprehensive look at the domain of explainable business analytics, highlighting that AI systems need to be created using not just accurate and explainable models, but additionally with human understanding.

2. Literature Review and Research Gap

2.1 Explainable AI in Strategic Decision-Making

AI has transformed how business decisions are made, enabling managers to analyze massive amounts of data, identify trends, and make informed decisions for the future of their business (Rane et al., 2024). In various industries, AI models are being used to support decision-making processes and enhance efficiency, such as in customer analytics, healthcare, financial forecasting, and supply chain management (Samara et al., 2024; Selvarajan, 2021). When AI advice is part of strategic decisions, there are issues of accountability, trust, and managerial acceptance due to the lack of interpretability (Ferdousi et al., 2026).

To address these issues, a new study Explicable Artificial Intelligence (XAI) has emerged to create more transparent predictive models (Hassija et al., 2024). Explainability enables decision makers to grasp why a model is giving a particular recommendation, and thus assists them in making informed, accountable decisions (Mathew et al., 2025). Consequently, explainability has become an important consideration in building AI systems that can be trusted (Chamola et al., 2023).

2.2 Native and Post-hoc Explainability

In general, there are two types of explainability approaches, namely, intrinsic and post-hoc explainability. Intrinsically interpretable means that the meaning of the model is built into the learning process, so that the learner can grasp how predictions are being made without having to perform further analyses (Murdoch et al., 2019). Explainable Boosting Machine (EBM) is one of these models that has attracted attention due to its high

predictiveness, besides its native feature contribution analysis, which makes it attractive to decision support applications where transparency is a key asset (Fajar et al., 2024).

Moreover, post-hoc explanation techniques, on the other hand, aim to explain the complex black-box models after the models have been developed (Guidotti et al., 2018). One of the most popular post-hoc models is SHAP, which measures the effect of each feature in a specific prediction through cooperative game theory (Li et al., 2024). While SHAP has significantly enhanced the explainability of black-box algorithms, its explanations are still reliant on external analysis of an already-trained model (Linardatos et al., 2020). This means that in many cases, organizations are now faced with a practical dilemma: when predictive performance is similar, is native explainability more useful in guiding decisions?

2.3 Behavioural Perspectives on AI-Assisted Decision-Making

As the technical dimension is traditionally explored in terms of explainability, the decision-making process of organizations is essentially behavioural (Çeri & Erhan, 2025). AI's recommendations are not typically judged independently of their predictive capabilities, but rather based on trust, past experience, mental orientation, and the organization's context. Trust in AI has always been deemed a critical factor in technology adoption, as people are more likely to use the recommendations of AI to make their decisions if they trust them (Afroogh et al., 2024).

Simultaneously, as the use of AI becomes more prevalent, there are growing worries about algorithmic reliance. Some level of dependence can help to increase efficiency and decision-making quality, but excessive reliance can limit independent assessment and make it easier to fall victim to algorithmic mistakes or algorithmic bias (Zerilli et al., 2019). The cognitive caution helps to balance this by prompting the users to critically analyze the AI recommendations before implementing them. However, the combined effect of trust, cognitive caution, and reliance on AI in explainable decision support systems is not well studied empirically (Wang & Yin, 2021). Thus, the knowledge of how these behavioural factors relate to explainable AI is therefore a key research area in business analytics.

2.4 Research Gap

Although explainable AI has seen significant

development, there are still several key research gaps that need to be addressed. To begin with, past research has tended to assess the explainability of an explanation based on technical performance measures, but comparatively few studies have focused on the outcome of a decision when behavioural measures are relevant. Second, very few studies test the performance of models that are intrinsically explainable against the performance of models that are generated via post-hoc explanation techniques, in a single empirical setting, and have a similar predictive performance. Third, the link between explainability and managerial decision behavior is still poorly understood, especially the impact of trust and cognitive evaluation on trust in AI recommendations.

In addition, the majority of explainable AI research primarily focuses on either predictive modelling or on human factors, with little consideration for the combination of both (Islam et al., 2021). Thus, there is little empirical evidence that explainable machine learning dashboards actually improve strategic decision-making or that they just boost confidence in algorithmic recommendations.

In this regard, the present study aims to fill these gaps by combining machine learning evaluation and behavioural analysis by using the dual-dataset research design. The study explores the Managerial Blind Spot Effect as a perspective on explainable AI from a decision-centric lens and compares the Explainable Boosting Machine with well-known black-box algorithms, native explainability and post-hoc explainability, and investigates the relationship between Trust in AI, Cognitive Caution, and AI Dependence. The integrated system goes beyond the mere focus of predictability in the field of explainable AI to the understanding and analysis of the impact of transparency on managerial judgment and strategic decision-making.

3. Proposed Framework

The research framework proposed for the study of the Managerial Blind Spot Effect is shown in Figure 1. The framework combines predictive modelling with behavioural decision analysis to assess whether explainable machine learning dashboards aid in strategic decision-making while retaining the same predictive accuracy provided by various algorithms. While existing explainable AI research mainly addresses the explainability of models or model quality, the proposed approach takes a decision-centric perspective with technical explainability coupled with managerial cognition.

The framework is made up of two analytic layers that are

complementary. The first layer corresponds to the machine learning component in which we are using the IBM Telco Customer Churn dataset to build predictive models using Explainable Boosting Machine (EBM), XGBoost, Random Forest, and Neural Network algorithms (Joy et al., 2024). The performance of the model is assessed by Accuracy, Precision, Recall, F1-score, ROC-AUC, Brier Score, Confusion Matrix, and Calibration curve. The Explainable Boosting Machine is the main explainable model, which offers native feature importance and prediction mechanisms, and SHAP explanations are generated for XGBoost to give a post-hoc explanation framework for comparison.

The second layer is the managerial decision layer, in which the outputs of the models are actually explained to include in the decision-making processes of human participants. The framework does not necessarily mean that clear predictions are better predictors of organizational outcomes; it acknowledges that there are behavioral characteristics that are also useful predictors of organizational outcomes. Therefore, the AI Decision Dependence and Cognitive Caution dataset is applied to the examination of the relationships between AI Dependence, Cognitive Caution, and Trust in AI in terms of reliability,

correlation, and regression analysis. These constructs can reveal the interpretation and reliance on the AI-generated recommendations by managers, even if the advice is technically incorrect.

These two analytical layers are proposed to be integrated to become the basis of the proposed “Managerial Blind Spot Effect”. The framework clearly asserts that explainability is not always the solution to the problem of inappropriate reliance on AI. Instead, strategic decision quality is the outcome of a combination of transparency in machine learning models and sound business sense. Therefore, explainable dashboards should not only provide accurate forecasts and explanations but should also foster a critical evaluation of AI recommendations before organizational decisions are made (Akter & Kudapa, 2024).

In sum, the proposed framework highlights the technical and behavioural nature of explainability and shows how explainable models are vital for trustworthy decisions in a way that relies on the interplay between predictive performance, explainability, and managerial cognitive processing. This holistic view offers the conceptual basis for the empirical method used and the analyses to follow in this study.

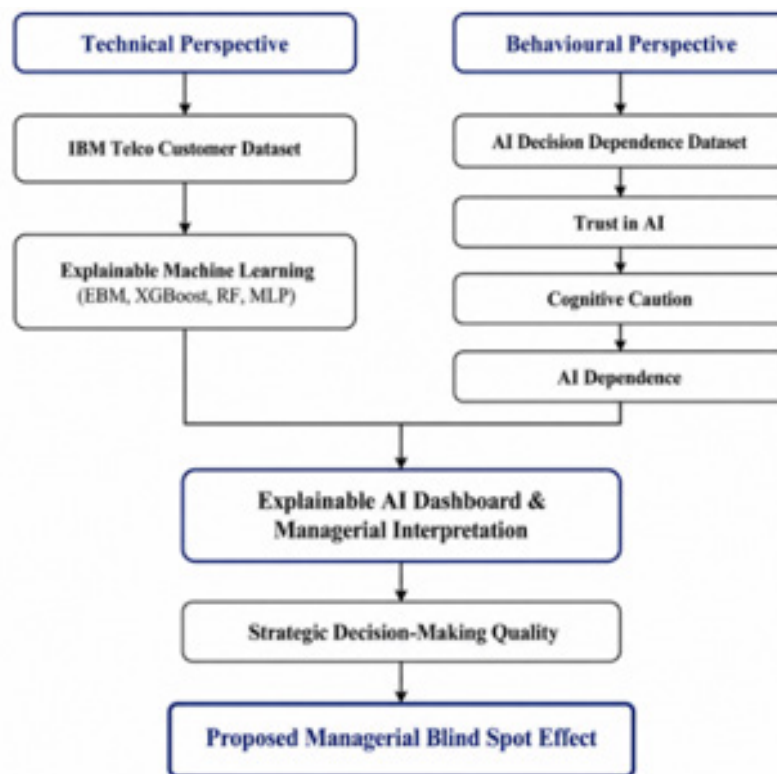


Figure 1. Proposed research framework

4. Methodology

4.1 Research Design

The two-stage quantitative research design was used in this study to assess whether explainable machine learning dashboards enhance the managerial decision-making process when the prediction accuracy is close across different machine learning algorithms used. The proposed Managerial Blind Spot Effect is explored in the framework that consists of a technical machine learning layer and a behavioural analysis layer. The first stage created customer churn prediction models using the Telco Customer Churn dataset from IBM, and the second stage examined the relationships between the variables of AI dependence, cognitive caution, and trust in AI using survey-based behavioural data.

4.2 Datasets

Two sets of independent data were used. The first was the dataset of 7,043 customer records from IBM's Telco Customer Churn database, which included demographic, contractual, service, and billing information. The binary target variable was the status of customer churn. The remaining features after discarding identifier variables, redundant attributes, and potential leakage variables were used for predictive modelling.

The second dataset was the AI Decision Dependence and Cognitive Caution 2025 dataset, comprising 201 survey responses of behavioural traits related to AI-informed decision making. Three constructs were examined: AI Dependence, Cognitive Caution, and Trust in AI. The data was analyzed to evaluate the behaviours that enable managerial reliance on AI recommendations.

4.3 Data Pre-processing

For the IBM Telco dataset, variables that could potentially lead to data leakage were removed from the dataset prior to analysis. Categorical items that had no value were replaced with the most frequent category, and numerical items were replaced with a median value. Categorical attributes were coded according to the requirements of the learning algorithms, and numerical variables were retained in their original range. The data was divided into two sets, training and test, with stratified random sampling with a fixed random state value of 42 to maintain class distribution.

For the composite scores of AI Dependence and Cognitive Caution, items that were not completed were dropped prior to calculating the composite score, which was computed as

the mean of the items. The construct of trust observed was then incorporated into further statistical analyses.

4.4 Machine Learning Models

The same experimental conditions were used for the four supervised classification methods. The major model used was a generalized additive model, specifically an Explainable Boosting Machine (EBM), which is intrinsically interpretable with comparable predictive capabilities that give transparent feature contributions. As EBM is the native explainability, it was suggested to be the model for the explainable dashboard.

Three widely used algorithms were chosen to compare the results: Extreme Gradient Boosting (XGBoost), Random Forest and Multi-Layer Perceptron (MLP) Neural Network. XGBoost was an up-to-date gradient boosting model; Random Forest was a baseline model for ensemble learning; a neural network was a deep learning tabular classification model. All models were trained and tested on the same data partitions for a fair comparison.

4.5 Explainability Analysis

Two approaches to explainability were discussed: intrinsic and post-hoc. For native explainability, global feature importance and additive feature contribution functions were derived from the model using the Explainable Boosting Machine (XBMT). To compare the intrinsic transparency with the post-hoc interpretation, SHAP values were computed for the XGBoost model. The fusion of EBM and SHAP allowed for comparing transparent machine learning models and externally explained black-box algorithms in a managerial decision-support setting.

4.6 Behavioural Analysis

Behavioural analysis was used to assess the reliability and relationships between the survey constructs. Cronbach's alpha was used to evaluate internal consistency, and descriptive statistics were used to summarize the central tendencies of AI Dependence, Cognitive Caution, and Trust in AI. All these constructs were analyzed using Pearson correlation analysis. Finally, multi-linear regression analysis was conducted to explore the behavioural factors that predict AI reliance when making decisions through AI and Cognitive Caution as the predictors.

4.7 Evaluation Metrics

The markers used for evaluation of model

performance included: Accuracy, Precision, Recall, F1-score, and Receiver Operating Characteristic Area Under the Curve (ROC-AUC). A Confusion Matrix was used to further explore the performance of the best model used for classification. To assess probability calibration, the Brier Score and Calibration Curve were chosen, as they are crucial for managerial dashboards. To determine the most important features influencing the customer churn predictions, explainability was done using EBM's built-in feature importance and SHAP visualizations.

4.8 Implementation Environment

In Google Colaboratory, Python was used to

perform all analyses. Pandas, NumPy, Scikit-learn, Statsmodels, and Matplotlib were used for data pre-processing and statistical analyses. An explainable boosting machine was applied using the InterpretML library, and SHAP was used for generating explanations of the XGBoost model. This was a fully reproducible process for predictive modelling and for behavioural analysis.

5. Results

5.1 IBM Telco Dataset Characteristics

After pre-processing the data and removing leakage, we had a dataset of 7043 customers for IBM Telco Customer Churn. Customer churn was used as the binary

Table 1. Descriptive Characteristics of the IBM Telco Customer Churn Dataset

Characteristic	Value
Total observations	7,043
Target variable	Churn Value
Total predictive features	23
Training samples (80%)	5,634
Testing samples (20%)	1,409
Churn class	Binary (0 = No, 1 = Yes)
Machine learning task	Binary classification
Non-churn customers	5,174 (73.5%)
Churn customers	1,869 (26.5%)

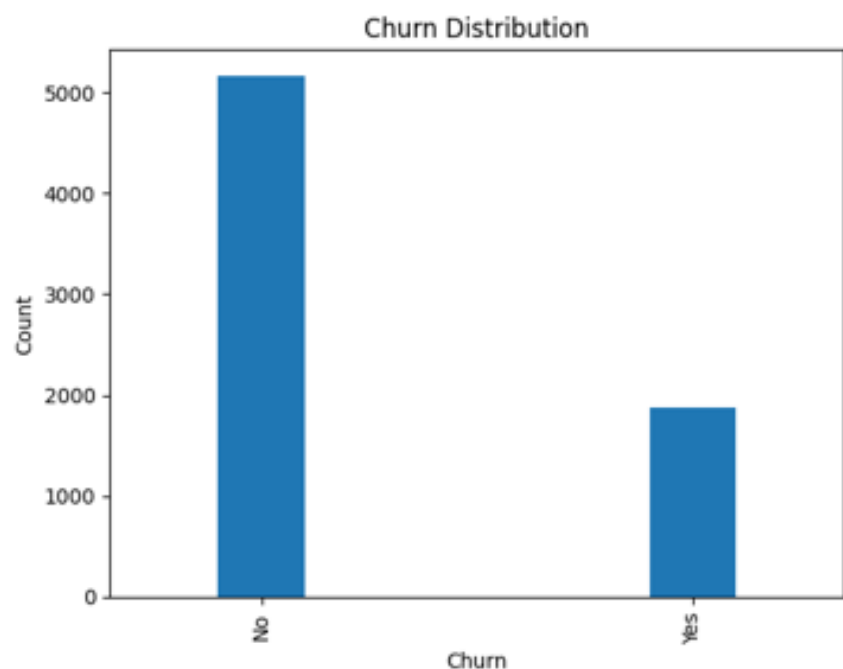


Figure 2. Distribution of Customer Churn

outcome variable, and a set of demographics, contractual, service, and billing attributes was included as predictive features. The data characteristics are summarized in Table 1, while the distribution of the churn is shown in Figure 2. The distribution shows that the majority class is the non-churn customers, which is a moderately imbalanced classification problem. Thus, the metrics assessed for model evaluation were not limited to accuracy, but also precision, recall, F1-score, ROC-AUC, and probability calibration.

5.2 Comparative Performance of Machine Learning Models

For the predictive performance, the results of

the Explainable Boosting Machine (EBM), XGBoost, Random Forest and Neural Network are summarized in Table 2. EBM has the greatest overall performance with an accuracy of 80.55%, a precision of 65.53%, a recall of 56.42%, an F1-score of 60.63%, an ROC-AUC of 0.8578, and the lowest Brier score (0.1315) among all the evaluated models. XGBoost (79.84%) and ROC-AUC (0.8489) achieved the second-highest accuracy, while Random Forest and Neural Network gave relatively poor prediction capabilities. Results indicate that EBM has competitive predictive ability and also outperforms traditional black-box models, and that it is inherently interpretable. The result showed its applicability as the main explainable dashboard model to conduct further analyses.

Table 2. Comparative Performance of Machine Learning Models

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC	Brier Score
Explainable Boosting Machine	0.8055	0.6553	0.5642	0.6063	0.8578	0.1315
XGBoost	0.7984	0.6424	0.5428	0.5884	0.8489	0.1356
Random Forest	0.7935	0.6370	0.5160	0.5702	0.8399	—
Neural Network	0.7573	0.5457	0.5107	0.5276	0.8077	—

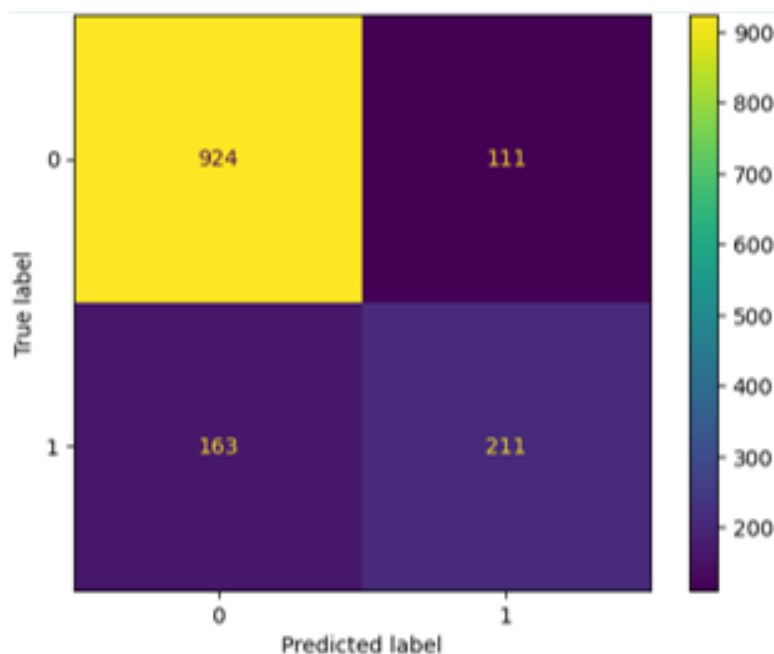


Figure 3. Confusion Matrix of the Explainable Boosting Machine

5.3 Reliability of the Explainable Boosting Machine

5.3.1 Classification Performance

The Explainable Boosting Machine classification accuracy was further tested with the confusion matrix in

Figure 3. The model was able to correctly identify most of the churn customers and non-churn customers, and it had a few false classifications. There were still some customers who were falsely classified as not churning, but the

overall pattern of classification shows balanced predictive behavior. Managerial implications from these findings are that EBM can be used to make good customer retention decisions and identify when further managerial oversight is needed.

5.3.2 Probability Calibration

In addition to the classification accuracy, the reliability of the predicted probabilities was evaluated with

the calibration curve in Figure 4. The calibration curve is fairly close to the line of reference, showing good agreement between the predicted and observed churn probabilities. In line with the above, EBM performed best on the Brier score out of the models tested (Brier score 0.1315), which indicates its ability to calibrate the probabilities well. Based on these results, EBM could be applied to allow precise risk assessment for decision-making at dashboards.

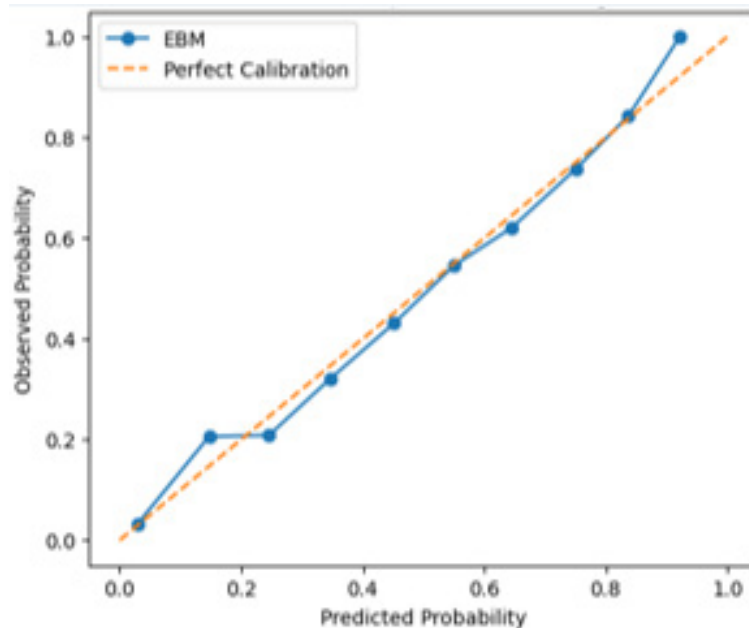


Figure 4. Calibration Curve of the Explainable Boosting Machine

5.4 Native and Post-hoc Explainability Analysis

5.4.1 Native Explainability Using Explainable Boosting Machine

Table 3 displays the results for the native EBM explainability. The most important variables were Dependents, Contract, Tenure Months, Internet Service,

Total Charges, Monthly Charges, Online Security, and Tech Support. The key findings were that the major factors were Dependents (0.534) and Tenure Months (0.402), indicating that commitment and service relationships are the most important factors to predict churn.

The results are applicable directly to management as they

Table 3. Top Feature Importance Identified by the Explainable Boosting Machine

Rank	Feature	Importance
1	Dependents	0.5337
2	Contract	0.5337
3	Tenure Months	0.4016
4	Internet Service	0.3002
5	Total Charges	0.2906
6	Monthly Charges	0.2332
7	Online Security	0.2093
8	Tech Support	0.1855

pinpoint the characteristics that are most associated with customer churn. Short-term customers and those with less service support and long-term engagement seem to be significantly more at risk of churning. In contrast to traditional black box models, EBM builds these explanations as part of the model-building process so that they can be interpreted transparently without the need for extra explanation methods.

5.4.2 Post-hoc Explainability Using SHAP

To compare intrinsic and post-hoc explainability, SHAP explanations were generated for the XGBoost model. In Figure 5, the SHAP summary plot revealed that the top factors that contributed to predicting churn were: contract type, tenure, monthly charge, internet service, online security, and technical support. The results are very similar to those derived by EBM's variables, showing consistency across native vs. post hoc explanation strategies.

In Figure 6, a SHAP waterfall plot is another representation that shows the combined effect of several customer

features on the probability of churn for a given prediction. EBM offers the same explanatory power directly through the modelling process as SHAP does in terms of explaining individual decisions made by a black-box model. Combined, these complementary analyses prove that explainable machine learning dashboards can increase transparency without compromising predictive accuracy.

5.5 Behavioural Evidence Supporting the Managerial Blind Spot Effect

5.5.1 Reliability Analysis

Table 4 indicates that there was good internal consistency for both of the behavioural constructs with a reliability analysis. The Cronbach's alpha for AI Dependence was 0.898, and for Cognitive Caution was 0.883, which was above the threshold of acceptable reliability. These results validate the appropriateness of both constructs for statistical analysis.



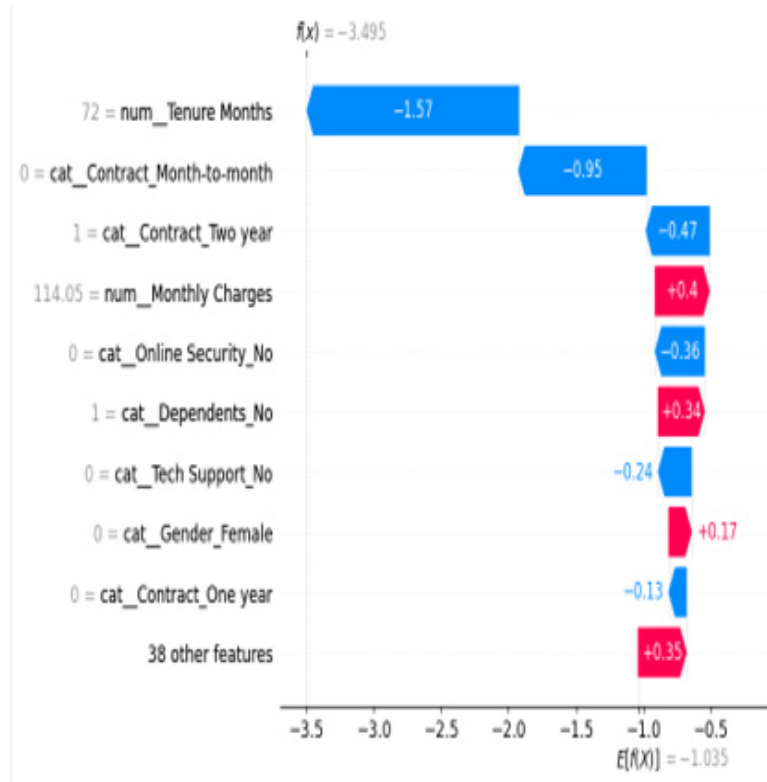


Figure 6. SHAP Waterfall Plot Explaining an Individual Prediction

Table 4. Reliability Analysis of Behavioural Constructs

Construct	Number of Items	Cronbach's Alpha
AI Dependence	7	0.8978
Cognitive Caution	7	0.8833

5.5.2 Behavioural Descriptive Statistics

The descriptive statistics are presented in Table 5. The mean score for AI Dependence was 3.77, for Cognitive Caution it was 3.83, and for Trust in AI it was the highest mean of 4.23. This observed variability

suggests a meaningful difference in respondents' perceptions of AI-supported decision-making, warranting further investigation of the interrelationships of the various variables.

Table 5. Descriptive Statistics of Behavioural Constructs

Construct	N	Mean	Standard Deviation	Minimum	Maximum
AI Dependence	201	3.765	1.476	1.000	6.429
Cognitive Caution	201	3.831	1.506	1.000	7.000
Trust in AI	201	4.234	1.732	1.000	7.000

5.5.3 Correlation Analysis

As shown in the correlation matrix presented in Table 6, there are strong positive correlations between the behavioural constructs. The results indicated that there was a significant correlation between AI Dependence and

Cognitive Caution ($r = 0.883$) and between AI Dependence and Trust in AI ($r = 0.720$), and a moderate positive correlation between Cognitive Caution and Trust in AI ($r = 0.614$). The results showed that an increase in trust in AI is correlated with an increase in AI-supported decisions,

and reflective evaluation is still positively related to AI dependence.

5.5.4 Regression Analysis

As shown in the results of multiple regression reported in Table 7, the model was found to have a significant prediction of AI Dependence ($R^2 = 0.830$, Adjusted $R^2 = 0.829$, $F = 484.8$, $p < 0.001$). The Trust in AI ($\beta = 0.244$, $p < 0.001$) and Cognitive Caution ($\beta =$

0.693, $p < 0.001$) constructs were both significant positive contributors to AI Dependence, with the latter being the stronger predictor. The results showed that, besides having trust in the AI system, reflective judgment of AI-generated recommendations also affects managerial trust. Altogether, the behavioural results support the machine learning findings: explainable prediction models need to be supported by informed managerial judgment and illustrate the proposed Managerial Blind Spot Effect.

Table 6. Pearson Correlation Matrix

Variable	AI Dependence	Cognitive Caution	Trust in AI
AI Dependence	1.000	0.883	0.720
Cognitive Caution	0.883	1.000	0.614
Trust in AI	0.720	0.614	1.000

Table 7. Multiple Regression Results Predicting AI Dependence

Predictor	Coefficient (β)	Std. Error	t-value	p-value
Constant	0.078	0.128	0.614	0.540
Trust in AI	0.244	0.032	7.711	<0.001
Cognitive Caution	0.693	0.036	19.087	<0.001

Note: R^2 (0.830), Adjusted R^2 (0.829), F-statistic (484.80) & Model Significance ($p < 0.001$)

6. Discussion

6.1 Explainable Machine Learning and Predictive Performance

Data for this study clearly demonstrates that explainability is possible without sacrificing predictive capabilities. The Explainable Boosting Machine (EBM) algorithm was also found to be the best-performing algorithm with all of the major evaluation methods, in addition to being intrinsically transparent. EBM is an integrated modelling approach, in contrast to the black-box approaches, which do not call on external interpretation approaches to predict and explain. The results indicate transparency is not a compromise on predictive power, but sometimes it is a necessity, as in a business decision-making application.

The excellent calibration performance also illustrates the importance of EBM. Accurate probability estimates help prioritize customer retention programs and move customers beyond the “churn”/“nonchurn” bucket into a probability classification. The aim is to make correct predictions and also to give back reliable information regarding the likelihood of each of the predictions, which can assist in the process of allocation of resources and

strategic planning.

6.2 Native versus Post-hoc Explainability in Managerial Dashboards

The comparison between intrinsic and post hoc explainability is one of the contributions of this study. The key drivers of churn found by native explanations produced by EBM were the contractual commitment, customer tenure, service usage, and billing characteristics. These explanations are directly incorporated into the predictive model, that is, without any additional analysis. For managers, they can see the contribution of each variable in the model to the prediction result through the process.

The feature importance that we got from the XGBoost model was also similar to the SHAP analysis, meaning that the churn drivers we identified are consistent across the modelling approaches. SHAP explanations are still based on a black-box model and are produced after modeling. EBM, by contrast, has feature contribution as an integral part of the prediction process and is transparent. Explainability is an important distinction, especially in managerial settings where it is not a stand-alone analytical activity, but a part of any prediction. This is therefore an

example of how EBM and SHAP can be used in tandem to create explainable dashboards that can be both predictive and decision-supportive.

6.3 Behavioural Evidence for the Managerial Blind Spot Effect

The behavioural analysis applies the technical findings by proactively demonstrating that explainability is not the only driver of the extent to which managers use AI. Both predictors are positive, and the regression results indicate that Trust in AI and Cognitive Caution are positively influencing AI Dependence, with the latter being more influential. The findings suggest that the more time managers dedicate to discussing the AI suggestions, the more likely they are to take AI-generated suggestions into account when making decisions. Relying on AI does not appear to be an uncritical acceptance, but rather a process of trust in the AI systems' functionality and consideration of the outcomes.

The findings confirm and replicate the hypothesized Managerial Blind Spot Effect. Managers are increasingly adopting AI-generated recommendations, potentially leading to increasingly algorithmically driven strategic decisions. Explainable dashboards can mitigate this risk by providing managers with the reasoning behind predictions, which enables them to make a decision rather than relying on the accuracy of the prediction and following blindly. Thus, explainability is an approach to making it possible for managers to monitor the situation without losing the benefits of AI-backed decision-making (Wolf, 2019).

6.4 Theoretical and Practical Implications

The overall contribution of the paper to the field of explainable AI and business analytics is that the authors demonstrate that by using intrinsically interpretable models, one might be able to achieve the same or better predictive accuracy than the widely used black-box models. Combining native explainability, post-hoc interpretation, and even behavioural evidence, the study goes beyond accuracy prediction to the quality of decision and managerial insight. The challenge for the business world is to select a model that can make accurate predictions and give insightful explanations. Native explainability empowers managers to understand the business context behind factors influencing customer attrition, and behavioural evidence underscores the importance of questioning AI recommendations. Based on these findings, it can be concluded that an AI-based decision-making

support system must not only increase the accuracy of predictions but also create trust and accountability among managers and ensure the quality of strategic decisions.

7. Limitations

There are a couple of limitations in this study. First, the analysis employed a cross-sectional survey dataset that only offers descriptive or statistical inferences of the relationships between AI Dependence, Trust in AI and Cognitive Caution rather than causal inferences. Second, the machine learning model was tested with just one data set of the customer churn prediction, which means that the model may behave differently in other domains or contexts of an organization. Third, the Explainable Boosting Machine is a model with an intrinsic explainability, while this study did not directly witness real managerial interactions with explainable dashboards in the field. Therefore, the proposed Managerial Blind Spot Effect is not proposed in the context of controlled managerial experiments but rather by integrating predictive modelling and behavioural evidence.

8. Future Research

More research is required to validate the Managerial Blind Spot Effect by controlling and testing the manager's decision-making process for an "explainable" and "non-explainable" AI dashboard context. Generalization of the framework to other application domains, like healthcare, financial risk assessment, and supply chain management, would further explore the generalizability. Other factors that might affect algorithmic decision-making, such as algorithmic literacy, risk perception, organizational experience, and decision accountability, can also be explored in future studies. Longitudinal studies may further examine the relationship between the development of trust, cognitive vigilance, and reliance on AI over time as managers become more experienced with the use of explainable AI.

9. Conclusion

In this study, we investigated whether machine learning dashboards that explain models had a positive impact on strategic decisions, while predictive accuracy was similar for different models. The study combined predictive modelling with behavioural analysis, leveraging to present a decision-centric view of explainable AI, the "Managerial Blind Spot Effect. The results of the empirical study showed that the Explainable Boosting

Machine achieved the best predictive performance without compromising explainability, as it already offers native model transparency without sacrificing predictive capability. Complementary SHAP analysis also revealed the power of the factors affecting churn prediction.

The behavioural analysis extended these technical findings by showing that both the trust in AI and the cognitive caution have significant positive effects on AI dependence, indicating that AI dependence is not just based on confidence in intelligent systems but also on reflective consideration of the recommendations produced by the AI. The results suggest that explainability is not sufficient for improved decision making, but rather a mix of human management and transparent algorithms is needed for successful AI-enabled decision-making.

Overall, the study makes a contribution to explainable business analytics by linking the intrinsic explainability, post-hoc explainability, and the behavioural decision factors into a single empirical framework. The results offer actionable insights for organizations aiming to implement AI dashboards that are both accurate and trustworthy, offering predictive insights while maintaining transparency, accountability, and human oversight for informed and responsible strategic decision-making.

Reference

Afroogh, S., Akbari, A., Malone, E., Kargar, M., & Alambeigi, H. (2024). Trust in AI: progress, challenges, and future directions. *Humanities and Social Sciences Communications*, 11(1), 1568.

<https://doi.org/10.1057/s41599-024-04044-8>

Akter, M., & Kudapa, S. P. (2024). A comparative analysis of artificial intelligence-integrated bi dashboards for real-time decision support in operations. *International Journal of Scientific Interdisciplinary Research*, 5(2), 158-191.

<https://doi.org/10.63125/47jjv310>

Bankins, S., Ocampo, A. C., Marrone, M., Restubog, S. L. D., & Woo, S. E. (2024). A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. *Journal of organizational behavior*, 45(2), 159-182.

<https://doi.org/10.1002/job.2735>

Bauer, K., Von Zahn, M., & Hinz, O. (2023). Expl (AI) ned: The impact of explainable artificial intelligence on users' information processing. *Information systems research*, 34(4), 1582-1602.

<https://doi.org/10.1287/isre.2023.1199>

Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. *Public administration review*, 81(5), 825-836.

<https://doi.org/10.1111/puar.13293>

Çeri, Ş. Ö., & Erhan, T. (2025). A Conceptual Study on the Effects of Artificial Intelligence in Managerial Decision-Making. *European Journal of Digital Economy Research*, 6(1), 5-21.

Chamola, V., Hassija, V., Sulthana, A. R., Ghosh, D., Dhingra, D., & Sikdar, B. (2023). A review of trustworthy and explainable artificial intelligence (XAI). *IEEE Access*, 11, 78994-79015.

<https://doi.org/10.1109/ACCESS.2023.3294569>

Chen, V., Liao, Q. V., Wortman Vaughan, J., & Bansal, G. (2023). Understanding the role of human intuition on reliance in human-AI decision-making with explanations. *Proceedings of the ACM on human-computer interaction*, 7(CSCW2), 1-32.

<https://doi.org/10.1145/3610219>

Chinnaraju, A. (2025). Explainable AI (XAI) for trustworthy and transparent decision-making: A theoretical framework for AI interpretability. *World Journal of Advanced Engineering Technology and Sciences*, 14(3), 170-207.

<https://doi.org/10.30574/wjaets.2025.14.3.0106>

Fajar, A., Yazid, S., & Budi, I. (2024). Enhancing phishing detection through feature importance analysis and explainable AI: A comparative study of CatBoost, XGBoost, and EBM models. *arXiv preprint arXiv:2411.06860*.

Ferdousi, J., Shokran, M., & Islam, M. S. (2026). Designing Human-AI Collaborative Decision Analytics Frameworks to Enhance Managerial Judgment and Organizational Performance. *Journal of Business and Management Studies*, 8(1), 01-19.

<https://doi.org/10.32996/jbms.2026.8.1.1>

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys (CSUR)*, 51(5), 1-42.

<https://doi.org/10.1145/3236009>

Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., & Hussain, A. (2024). Interpreting black-box models: a review on explainable artificial intelligence. *Cognitive Computation*, 16(1), 45-74.

<https://doi.org/10.1007/s12559-023-10179-8>

Islam, S. R., Eberle, W., Ghafoor, S. K., & Ahmed, M. (2021). Explainable artificial intelligence approaches: A survey. *arXiv preprint arXiv:2101.09429*.

Joy, U. G., Hoque, K. E., Uddin, M. N., Chowdhury, L., & Park, S.-B. (2024). A big data-driven hybrid model for enhancing streaming service customer retention through churn prediction integrated with explainable AI. *IEEE Access*, 12, 69130-69150.

<https://doi.org/10.1109/ACCESS.2024.3401247>

Kamath, U., & Liu, J. (2021). Post-hoc interpretability and explanations. In *Explainable Artificial Intelligence: An Introduction to Interpretable Machine*

Learning (pp. 167-216). Springer.

https://doi.org/10.1007/978-3-030-83356-5_5

Khan, M. N. M. (2025). Artificial intelligence driven big data and business analytics: A comprehensive review of multi-sectoral applications in healthcare, finance, supply chain, and organizational innovation. *Pacific Journal of Business Innovation and Strategy*, 2(4), 122-137.

<https://doi.org/10.70818/pjbis.v02i04.0130>

Li, M., Sun, H., Huang, Y., & Chen, H. (2024). Shapley value: from cooperative game to explainable artificial intelligence. *Autonomous Intelligent Systems*, 4(1), 2.

<https://doi.org/10.1007/s43684-023-00060-8>

Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2020). Explainable ai: A review of machine learning interpretability methods. *Entropy*, 23(1), 18.

<https://doi.org/10.3390/e23010018>

Mathew, D. E., Ebem, D. U., Ikegwu, A. C., Ukeoma, P. E., & Dibiaezue, N. F. (2025). Recent emerging techniques in explainable artificial intelligence to enhance the interpretable and understanding of AI models for human. *Neural Processing Letters*, 57(1), 16.

<https://doi.org/10.1007/s11063-025-11732-2>

Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Interpretable machine learning: definitions, methods, and applications. *arXiv preprint arXiv:1901.04592*.

Olayinka, O. H. (2019). Leveraging predictive analytics and machine learning for strategic business decision-making and competitive advantage. *International Journal of Computer Applications Technology and Research*, 8(12), 473-486.

Patil, D. (2024). Explainable Artificial Intelligence (XAI) for industry applications: Enhancing transparency, trust, and informed decision-making in business operation. *Trust, And Informed Decision-Making In Business Operation* (December 03, 2024).

<https://doi.org/10.2139/ssrn.5057402>

Peeters, R. (2020). The agency of algorithms:

Understanding human-algorithm interaction in administrative decision-making. *Information Polity*, 25(4), 507-522.

<https://doi.org/10.3233/IP-200253>

Rane, N. L., Paramesha, M., Choudhary, S. P., & Rane, J. (2024). Artificial intelligence, machine learning, and deep learning for advanced business strategies: a review. *Partners Universal International Innovation Journal*, 2(3), 147-171.

<https://doi.org/10.2139/ssrn.4835661>

Samara, F. Y. A., Taha, A. H. A., Massa, N. M., Jamie, T. N. A., Harara, F. E., Abu-Nasser, B. S., & Abu-Naser, S. S. (2024). The role of AI in enhancing business decision-making: Innovations and implications.

Selvarajan, G. (2021). Leveraging AI-enhanced analytics for industry-specific optimization: A strategic approach to transforming data-driven decision-making. *International Journal of Enhanced Research In Science Technology & Engineering*, 10(1), 78-84.

Shrestha, Y. R., Ben-Menahem, S. M., & Von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California management review*, 61(4), 66-83.

<https://doi.org/10.1177/0008125619862257>

Tan, S., Hooker, G., Koch, P., Gordo, A., & Caruana, R. (2023). Considerations when learning additive explanations for black-box models. *Machine Learning*, 112(9), 3333-3359.

<https://doi.org/10.1007/s10994-023-06335-8>

Vudugula, S., Chebrolu, S. K., Bhuiyan, M., & Rozony, F. Z. (2023). Integrating artificial intelligence in strategic business decision-making: A systematic review of predictive models. *International Journal of Scientific Interdisciplinary Research*, 4(1), 01-26.

<https://doi.org/10.63125/s5skge53>

Wang, X., & Yin, M. (2021). Are explanations helpful? a comparative study of the effects of explanations in ai-assisted decision-making. *Proceedings of the 26th International Conference on Intelligent User Interfaces*,

<https://doi.org/10.1145/3397481.3450650>

Wolf, C. T. (2019). Explainability scenarios: towards scenario-based XAI design. *Proceedings of the 24th International conference on intelligent user interfaces*,

<https://doi.org/10.1145/3301275.3302317>

Xu, F., Uszkoreit, H., Du, Y., Fan, W., Zhao, D., & Zhu, J. (2019). Explainable AI: A brief survey on history, research areas, approaches and challenges. *CCF international conference on natural language processing and Chinese computing*,

https://doi.org/10.1007/978-3-030-32236-6_51

Zerilli, J., Knott, A., Maclaurin, J., & Gavaghan, C. (2019). Algorithmic decision-making and the control problem.

<https://doi.org/10.1007/s11023-019-09513-7>