Applied Mathematics and Computational Intelligence
Volume 15, No. 4, 2026 [89-106]

Bridging Machine Learning and Strategic Decision-Making: A Framework for Customer Churn in Banking

Dong Zhang[1], Wei Chien Ng[1]*, Linlin Liu[1], Yu Qing Soong[1,2]

[1]School of Management, 11800, Universiti Sains Malaysia, Pulau Pinang, Malaysia
[2]Department of Accountancy and Business, Tunku Abdul Rahman University of Management and Technology, Penang Branch, 11200 Tanjung Bungah, Pulau Pinang, Malaysia

* Corresponding author: ngweichien@usm.my

Received: 9 September 2025
Revised: 20 October 2025
Accepted: 3 March 2026

ABSTRACT

Customer churn management has become increasingly important in modern customer relationship strategies. Among the main applications of machine learning and artificial intelligence in this field are customer churn prediction and model interpretation. Research on prediction models is already well developed, and work on interpretation methods to address specific business problems is steadily growing. However, major challenges remain, such as overly complex models, unverifiable recommendations, and high technical barriers for users. This study examines customer churn in the banking industry as a case example. It introduces the Simplicity, Goal orientation, and Verifiable solutions (SGV) framework, which combines simplified model construction, goal-oriented interpretation, and verifiable recommendations. The framework is validated using SHAP explanations. The proposed interpretation method analyzes and visualizes the relationship between the target outcome and key variables through simulated data, providing an integrated link between interpretation, findings, and actionable strategies. SGV framework offers a practical toolkit for practitioners across industries. It enables effective use of machine learning and AI for in-depth data analysis, reduces the entry barrier for non-specialists, and ultimately contributes to improved productivity.

Keywords: Customer churn prediction, Predictive framework, Banking industry, SGV framework, Big Data Analytics Capability

1 INTRODUCTION

Customer churn is a challenge across industries, particularly in high-tech, e-commerce, and financial services [1-3]. In streaming service industry, for instance, Netflix experiences a churn rate of 9%, while Hulu has reported a staggering customer churn rate up to 50% [4]. Currently, with the rising of anti-globalization and protectionism, the situation of customer retention for all business is becoming increasingly challenging.

Customer churn always causes great economic loss to enterprises worldwide. On one hand, it will directly lead to the shrinking business scale and eventually result in revenues dropping [5]. On the other hand, it will eventuate in surging marketing expenses, as the cost of bringing in a new client is often five to six times more than keeping an existing one [6]. Moreover, the lost customers are more likely to turn to their competitors, which will impose heavier competitive pressure and potentially result in a loss of market share [7]. If new customers fail to offset the losses, particularly for asset-light companies, the firm may face asset erosion or even bankruptcy [8]. In the business to business (B2B) context, customer churn can also have profound implications. Global B2B distributors have reported annual revenue decline ranging from 4.33% to 15.07%, due to customer churn [9].

Thus, understanding the causes of churn, predicting churn probability, and implementing appropriate mitigation strategies are essential for business development in the long run. This is particularly true for the financial industry, a key pillar of the global economy, which bears responsibilities such as investment intermediation, financing, resource allocation, money circulation, and risk management. Moreover, the global financial industry has sourced about 410 trillion dollars in 2023, which brought over 6.8 trillion dollars revenue [10].

2 LITERATURE REVIEW

2.1 Data Analysis Based Customer Churn Management

Data analysis is becoming the core basis for strategic decision-making in all sectors. By analyzing market data and applying customer behaviour insights, enterprises can be more efficient in knowing about industry issues and capturing customer demand [2]. Customer churn analysis is one of the important applications. Machine learning algorithms have been widely applied to customer churn prediction (CCP) and continue to play a significant role. These methods effectively leverage large scale data resources to enhance both the accuracy and farsightedness of churn prediction, while also offering high computational efficiency. Traditional models typically treat churn prediction as a binary classification problem [11]. Through processes such as data preprocessing, model training, prediction, evaluation, and optimization, basic models such as Logistic Regression and K-Nearest Neighbors (KNN) have achieved 89.58% and 95.79% accuracy, respectively, in predicting airline customer churn [12]. Ensemble learning models like Random Forest (RF) and XGBoost have pushed the accuracy further to 99.95% and 99.99%, respectively [12]. Recently, more advanced approaches, including Neural Networks and Deep Learning, have continued to emerge, further improving prediction performance.

Currently, effective customer churn research frameworks primarily contain two components: (i) CCP models, which aim to predict whether a particular customer is likely to churn; and (ii) analysis of influencing factors (AIF) models, which are mainly used to identify the underlying causes of customer churn [13]. Within the CCP domain, improving accuracy has always been the main theme. [14], others rely on ensemble learning approaches [12]. Meanwhile, some researchers specialize in hyperparameter optimization [15], and some do well in deep learning models [16]. Moreover, the latest studies have introduced unconventional and more complex algorithms such as genetic algorithm (GA) and optimized brief rule based (BRB) algorithm [17]. These advancements have significantly improved CCP performance.

Although CCP models are still the mainstream of related research, there is a growing number of works focusing on churn reasoning. These studies typically leverage various interpreting techniques to infer the primary factors behind customer churn [18]. By interpreting these models, practitioners can better understand why specific customers are at risk of churning and take prompt measures accordingly. Furthermore, some researchers have shifted the focus from churn propensity to profit maximization, adopting a managerial perspective to maximize the profit [19]. This profit-driven approach is especially suitable for heterogeneous customer structure, where the potential economic value of individual customers varies significantly [20]. In fact, some studies have found a negative correlation between churn propensity and customer value. In other words, high-value customers are generally less likely to churn [21]. This phenomenon is more common in B2B relationships, where new modeling approaches continue to emerge.

2.2 Research Gaps

However, both the traditional churn propensity focused frameworks and the profit maximization driven approaches have notable limitations. Firstly, excessive hyper-parameters tuning, sophisticated modeling and cumbersome formulas can increase the risk of over-fitting [22]. Second, profit-driven models may introduce bias by overemphasizing high net value customers, thereby reducing the accuracy of churn prediction for ordinary ones, and diminishing the objectivity of subsequent analysis. Moreover, the pursuit of higher predictive accuracy often leads to increasingly complicated models, which raise implementation costs and diminish practicality, even an Apple of Sodom. Finally, current research primarily focuses on churn propensity prediction and partly on churn reasoning. However, few studies have provided practical and verifiable strategies for reducing customer churn. The comparison of more research was shown in Table 1, which highlights the main gaps and limitations of previous studies. Churn prediction and analysis models should keep focusing on their core application value, namely, reducing customer churn.

Table 1: A Comparison of Related Studies

Refs.	Domains	Methods	Findings	Gaps
[12]	Telecom	Logistic Regression, RF, XGBoost, KNN	The ensemble algorithms perform better than the traditional ones in CCP.	Lack of the model explanation and corresponding recommendations.
[14]	Telecom	Decision Tree and RF	RF model achieved higher accuracy than decision tree. Up-sampling could enhance the performance.	More machine learning algorithms are needed.

[21]	B2B	AdaBoost, LASSO, RF, Support Vector Machine (SVM), XGBoost	This study takes customer life value into account, and construct model aiming to maximize the expected profit, which reached the integration of model aims and business goals.	Overemphasizing the characteristics of high value customers may lead to more bias, as the high value customers are just a small proportion.
[23]	B2B	Logistic Regression, Decision Tree, Naive Bayes, B2B-ARM	B2B-ARM achieved a similar performance with other models, and it distinguished by only focusing on the workable features.	Too many features are contained by the model, making it a time-consuming method.
[24]	Aviation	Logistic Regression, RF, SVM, KNN	RF is selected as the champion model. Entertaining services plays a more important role than utilitarian function in aviation customer churn management.	The work is limited in aviation industry, and difficult to transform to other areas.
[25]	Internet	RSF and Cox Regression	Lower prices or offering discounts do little to the customer churn in internet service. Higher internet speed and more interaction will contribute to reduce churn.	This research is based on a specific region and a particular customer group, which will largely limit its generalization.
[26]	Tourism	SVC, Logistic Regression, Decision Tree, RF, Neutral Networks, AdaBoost, XGBoost, LightGBM	This study builds the model by combing the customer life value with incentive cost, which successfully caught the core of business performance. By assigning different weights to the features, PAM achieved dynamic analyzing capability.	This study set too many prerequisites related to the business domain, which in turn has restricted the generalization of the model.

In response, this study proposed the Simplicity, Goal orientation, and Verifiable solutions (SGV) framework, a practical and generalizable framework for customer churn management that emphasizes simplicity, goal orientation, and verifiable solutions. It aims to deliver: (1) a simplified prediction process, (2) goal-driven cause analysis focused on reducing churn rate, and (3) actionable recommendations with measurable impact, thus establishing a practical and interdisciplinary framework for churn reduction.

3 METHODOLOGY

3.1 Research design

The SGV framework proposed in this study is a predictive analytics framework, which is designed to support practitioners across industries in predicting customer churn, analyzing its underlying causes, and providing corresponding recommendations. This framework is based on the BDAC theory and integrates predictive analytics methodologies [27]. It can be applied in various business scenarios involving customer relationship management. As illustrated in Figure 1, the framework consists of three main stages: model construction and evaluation, goal-oriented model interpretation and feasible suggestions.

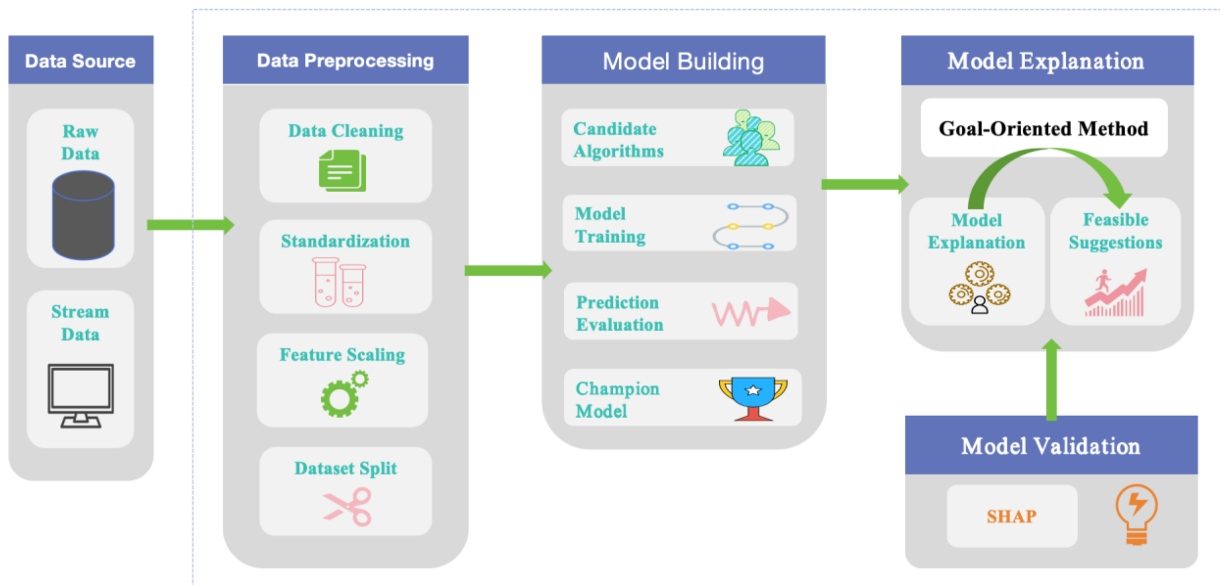

Figure 1: Overview of the SGV Framework

3.2 Data Collection

The dataset used in this study is obtained from the Kaggle community, titled 'A Synthetic Dataset for Predicting Customer Churn in the Banking Sector'. This dataset consists of 23 features across five dimensions: customer demographics, debt status, account information, customer activity, and customer satisfaction, along with a target indicating whether the customer has churned. It is well suited for research on CCP and related topics. Table 2 provides detailed descriptions of all the features.

Table 2: Summary of Attributes used in the Dataset

Column	Description	Data Types
Customer_ID	Unique identifier for each customer	Object
Age	The customer's age	Integer
Gender	The gender of the customer	Object
Is_Employed	Whether the customer is employed	Bool
Marital_Status	The marital status of the customer	Object
Annual_Income	The total annual income of the customer	Float
Region	The geographical region of the customer	Object
Loan_Amount	Amount of any loan taken by the customer	Float
Loan_Type	The type of the customer loan	Object
Credit_Score	The customers' credit score, indicating their financial health	Integer
Account_Open_Date	The date the customer's account was opened	Date
Account_Type	Type of account (hecking, Savings, Investment)	Object
Account_Balance	The balance in the customer's account	Float
Branch	Location of the branch where the account is held	Object
Last_Transaction_Date	Date of the last transaction made by the customer	Date
Transaction_Amount	Amount of the last transaction	Float
Transaction_Type	Type of the last transaction (Deposit, Withdrawal, etc.)	Object
Number_of_Transactions	Total number of transactions in the past year	Integer
Account_Activity_Trend	Whether the account activity is Increasing, Decreasing, or Stable	Categorical
Customer_Service_Interactions	Number of interactions with customer service	Integer
Recent_Complaints	Number of complaints made by the customer	Integer
Change_in_Account_Balance	Subtotal of account balance over last year	Float
Customer_Satisfaction_Score	Customer satisfaction rating (1-5), with lower scores indicating higher churn risk.	Integer
Churn_Label	Whether the customer has churned or not	Target

3.3 Data Preprocessing

Data preprocessing is considered as the foundation for effective data analysis. Raw datasets often contain various noises such as missing values, outliers, incorrect records, redundant variables, inconsistent data types, and large differences in feature scales. These types of noises not only increase the computational and time cost of analysis but can significantly reduce model performance and stability [28]. Therefore, precise and appropriate preprocessing is essential to ensure the quality of machine learning models.

3.3.1 Detection and Handling Outliers

For numerical features, outliers were always identified using the IQR $\pm$ 1.5 $\times$ IQR rule. Detected outliers were then capped at the upper or lower boundaries to minimize their influence on model training. An exception was made for the feature ‘Account_Balance’, as positive outliers in this variable most likely represent high net worth customers in the context of banking customers churn prediction. As this information is potentially crucial for the model’s predictive capability, no outlier treatment was applied to this feature.

3.3.2 Feature Encoding

The variable ‘Customer_ID’ is just used as a unique identifier for each customer and is excluded from model training. For numerical variables such as ‘Age’, ‘Annual_Income’, ‘Account_Balance’, ‘Loan_Amount’, ‘Credit_Score’, ‘Number_of_Transactions’, ‘Customer_Service_Interactions’, ‘Last_Transaction_Amount’, ‘Recent_Complaints’, ‘Change_in_Account_Balance’, and ‘Customer_Satisfaction_Score’, those can be used directly in model training and testing. For the Boolean variable ‘Is_Employed’, although it’s not necessary to encode, its values ‘True’ and ‘False’ were converted to 1 and 0, respectively. The feature ‘Account_Activity_Trend’, being an ordinal variable, was encoded using Label Encoding. The variables including ‘Gender’, ‘Marital_Status’, ‘Region’, ‘Account_Type’, and ‘Last_Transaction_Type’ are nominal and without order, thus were encoded using One-Hot Encoding. The variable of ‘Branch’ contains a large number of unique values, making it unsuitable for One-Hot Encoding. Therefore, Frequency Encoding was applied. The date-type variables ‘Account_Open_Date’ and ‘Last_Transaction_Date’ can never be directly accepted by the model. Hence, they were transformed into quantitative time based features: ‘Account_Age_Days’ and ‘Days_Since_Last_Transaction’. The feature ‘Account_Age_Days’ reflects the number of days from account opening to the current date (June 1, 2025), serving as an important indicator of customer loyalty. Another derived feature, ‘Days_Since_Last_Transaction’, measures the inactivity period since the last transaction up to June 1, 2025. This variable is particularly important for predicting potential churn behavior, as it captures customer engagement levels. Since this feature was not available in the original dataset, it was constructed manually based on the background of bank customer management.

3.3.3 Feature Scaling

Feature scaling is not required for tree-based models such as RF, XGBoost, and CatBoost, as these algorithms are insensitive to the scale of input features. However, feature scaling is essential for models that rely on distance metrics or mathematical functions, including SVM, Logistic Regression, and Neural Networks [29,30].

Numerical features in the datasets, such as 'Age', 'Annual_Income', and 'Transaction_Amount', as well as ordinal categorical variables encoded via Label Encoding, were executed feature scaling with the StandardScaler method, to ensure features at different magnitudes have equal contributions to model training.

3.3.4 Feature Selection

Different machine learning algorithms have varying requirements regarding feature selection. Algorithms such as Logistic Regression, SVM, and Neural Networks are more sensitive to irrelevant or redundant features and always benefit significantly from feature selection, both in improving performance and reducing the risk of over-fitting. In this study, we applied univariate feature selection using SelectKBest from Scikit-learn, which selects features based on univariate statistical tests, to enhance model performance and generalization. In contrast, tree-based algorithms such as RF, XGBoost, and CatBoost are inherently robust to noise and high dimensional data. These models can effectively handle large numbers of features due to their inherent feature importance mechanisms. Moreover, the datasets used in this study contains fewer than 30 features, making it unnecessary for these algorithms to perform feature selection.

3.4 Machine Learning Algorithms

This study explores six commonly used machine learning models, aiming to identify the most effective algorithm through a practical and efficient approach. SVM is a widely used binary classification algorithm that works by mapping data into a high dimensional feature space and identifying the optimal hyperplane that maximizes the margin between classes. The use of Kernel functions enables SVM to perform both linear and nonlinear classification. Its reliance on a small subset of support vectors helps to reduce the risk of over-fitting and enhance robustness [31].

RF is an ensemble learning method that builds multiple decision trees during training. Each tree outputs a classification result, and the final prediction is determined by majority voting (bagging). Due to the randomness in both sample selection and feature selection, Random Forest is highly stable and effective in handling noisy data [32]. Logistic Regression: Logistic Regression is similar to linear regression but is specifically designed for binary classification. It estimates the probability of an event occurring based on input features. Its relatively simple training process and high computational efficiency make it a popular choice. Additionally, regularization techniques can be applied to reduce the risk of over-fitting [33].

XGBoost is a powerful gradient boosting algorithm. It supports parallel processing and is capable of handling large and complex datasets. It requires minimal data preprocessing and is equipped with inherent regularization to reduce over-fitting, making it highly suitable for practical applications [34]. CatBoost is a gradient boosting algorithm, which is specifically designed for handling categorical features. It automatically processes categorical variables, saving the procedure of manual encoding. With the advantages of strong predictive performance robust model, CatBoost is suitable for large scale and complex classification tasks [35]. Neural Networks: Neural Networks are statistical learning models, which are inspired by the structure of the human brain. As the foundation of deep learning, they are capable of learning complex, nonlinear relationships between features. Neural networks have been widely applied in areas such as natural language processing and image recognition, particularly where traditional models fall short [36].

To enhance the adaptability of the SGV framework, this study adopts Optuna, an automatic hyper-parameter tuning tool. By setting the parameter range, it can easily and efficiently generate the optimal parameters, which will help to improve model performance [37]. Compared to traditional methods, such as Grid Search and Random Search, Optuna offers greater efficiency and flexibility, enabling better model performance through intelligent and automated tuning [38].

3.5 Prediction and Evaluation

This study implemented six widely used machine learning models to predict customer churn, covering traditional algorithms, ensemble methods, and deep learning techniques. To ensure the repeatability of the methodology, all models were developed using Python. Given the imbalanced distribution of the target variable, where the proportion of positive supports is approximately 75.28%, and negative supports (churned customers) around 24.72%, we employed a combination of evaluation metrics, which includes Accuracy, Precision, Recall, F1-score, Confusion Matrix, and the AUC (area under the ROC curve), to comprehensively assess model performance and identify the champion model.

The dataset contains 5,000 records, which were divided into a training set (80%) and a testing set (20%). To address the issue of class imbalance, we applied a twofold approach. Synthetic Minority Over-sampling Technique (SMOTE) was used at the data level to balance the training set. The 'class_weight' parameter was set at the algorithm level (for applicable models) to give higher weight to the minority class. This combined strategy effectively mitigated the adverse impact of imbalance samples on model performance.

3.6 Interpretative Analysis and Recommendation

This study originally proposed a goal-oriented model interpretation method and compared it with the traditional Shapley Additive exPlanations (SHAP) approach for validation. SHAP is an important model interpretation method, which is based on game theory to calculate the Shapley value of each feature and then explain the contribution of each feature to the model [39].

The main idea of the goal-oriented model interpretation method is to reversely deduce the number of positive and negative support in the target data set, based on the performance metrics of the selected champion model, such as Precision and Recall, consequentially to obtain the estimated 'customer churn rate' or 'customer retention rate'. Then, by adjusting the values of important features in the data set, we can simulate the effect of adopting the suggestions or stream data and meanwhile calculate and record the changes in the 'estimated customer churn rate'. Finally, when each feature changes in a positive or negative direction, the 'estimated customer retention rate' is visualized, so as to intuitively obtain the 'optimal solution' that can efficiently reduce the 'customer churn rate'.

On the one hand, this method can provide the deduced outputs when each feature value is changed. One the other hand, the difference of the outputs also reflects the different efficiencies to impact the target. So we can easily get the most effective feature to achieve the business goals.

Although this method holds those advantages, it relies on the assumption that the model's performance on the test set being a perfect proxy for the population. The specific design is as follows.

Step 1: Deduce the 'estimated customer churn rate'. Figure 2 demonstrated the derivation of the 'estimated customer churn rate'.

Step 2: Simulate the stream data. Since it's difficult to obtain the perfect stream data to analyze the exact relationship between the target and important features, we just perform a simulation. So, We sequentially adjust the values of the important features in the test datasets, with a variation range of ±100% and a step size of 1%. For each adjustment, we use the champion model to make predictions and record the number of predicted positives (PP) and predicted negatives (PN). Basing on Step 1, we can easily derive the corresponding 'Churn rate' and 'Retention rate'.

Step 3: Visualize the variation of targets as the feature values change. Setting the number of changes in feature values as the X-axis and the calculated 'Churn rate' or 'Retention rate as the Y-axis, we can plot the trend of target for every feature as its value changes."

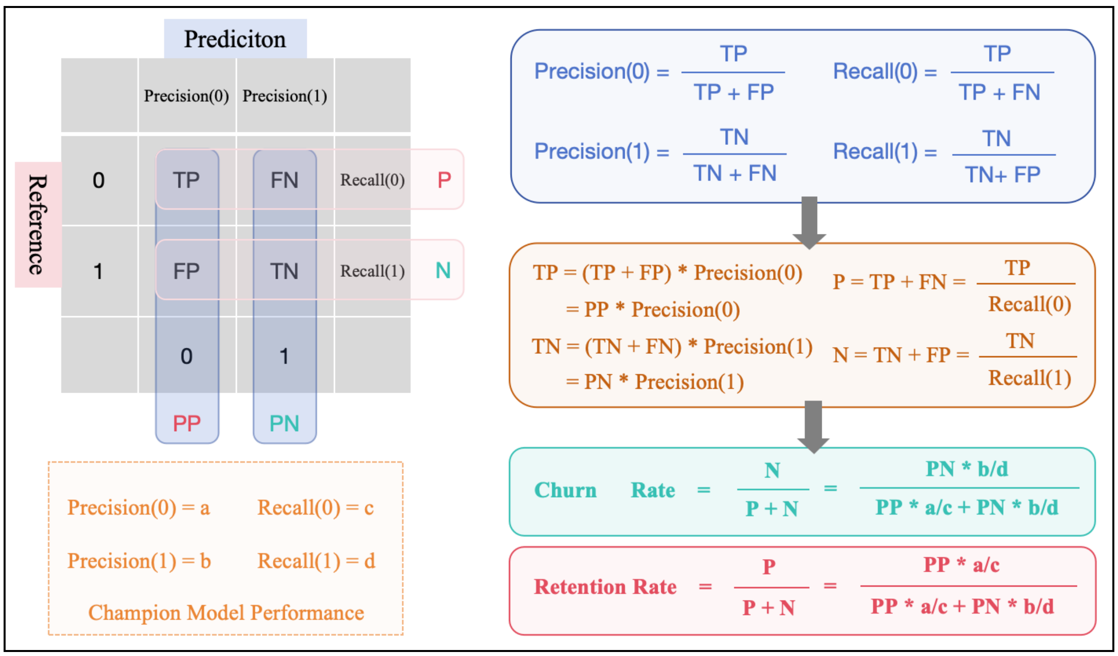

Figure 2: Derivation of the Estimated Customer Churn Rate

4 RESULTS AND DISCUSSION

4.1 Model Performance Evaluation

This study evaluates candidate models based on a set of comprehensive metrics including Accuracy, Precision, Recall, F1 Score, and AUC. By comparing these indicators, the Random Forest model was selected as the champion model due to its superior and consistent performance, with the highest accuracy of 0.857, as well as the outstanding F1 Score and AUC. Table 3 presents the comparative performance metrics of all evaluated models.

Table 3: Performance Comparison of Machine Learning Models

Classifier	Accuracy	Precision		Recall		F1 Score		AUC
		0	1	0	1	0	1	
SVM	0.819	0.91	0.61	0.84	0.75	0.87	0.68	0.891
Logistic Regression	0.820	0.92	0.61	0.83	0.80	0.87	0.69	0.901
Random Forest	0.857	0.88	0.76	0.93	0.63	0.91	0.69	0.903
XGBoost	0.851	0.90	0.70	0.90	0.67	0.90	0.70	0.906
CatBoost	0.849	0.89	0.71	0.91	0.66	0.90	0.69	0.893
Neutral Networks	0.795	0.93	0.56	0.79	0.82	0.85	0.67	0.890

4.2 Goal-Oriented Model Interpretation and Recommendations

Firstly, we applied the SHAP model interpretation method to analyze feature importance in the champion model. Based on the results, as illustrated in Figure 3, the finds were summarized as follows: The first five influential features on customer churn are: 'Recent Complaints', 'Customer Satisfaction Score', 'Customer Service Interactions', 'Number of Transactions', and 'Account Balance'. 'Recent Complaints' shows a strong positive impact on churn when the feature values are high (represented in red), and a strong negative impact when values are low (green). This indicates that a higher number of complaints significantly increases the likelihood of customer churn. 'Customer Satisfaction Score' has a strong negative relationship with churn when the score is high. This implies that improving customer satisfaction scores can greatly reduce churn probability. 'Customer Service Interactions' demonstrates a moderate positive correlation with churn risk, while 'Number of Transactions' and 'Account_Balance' show a moderate negative correlation with churn likelihood.

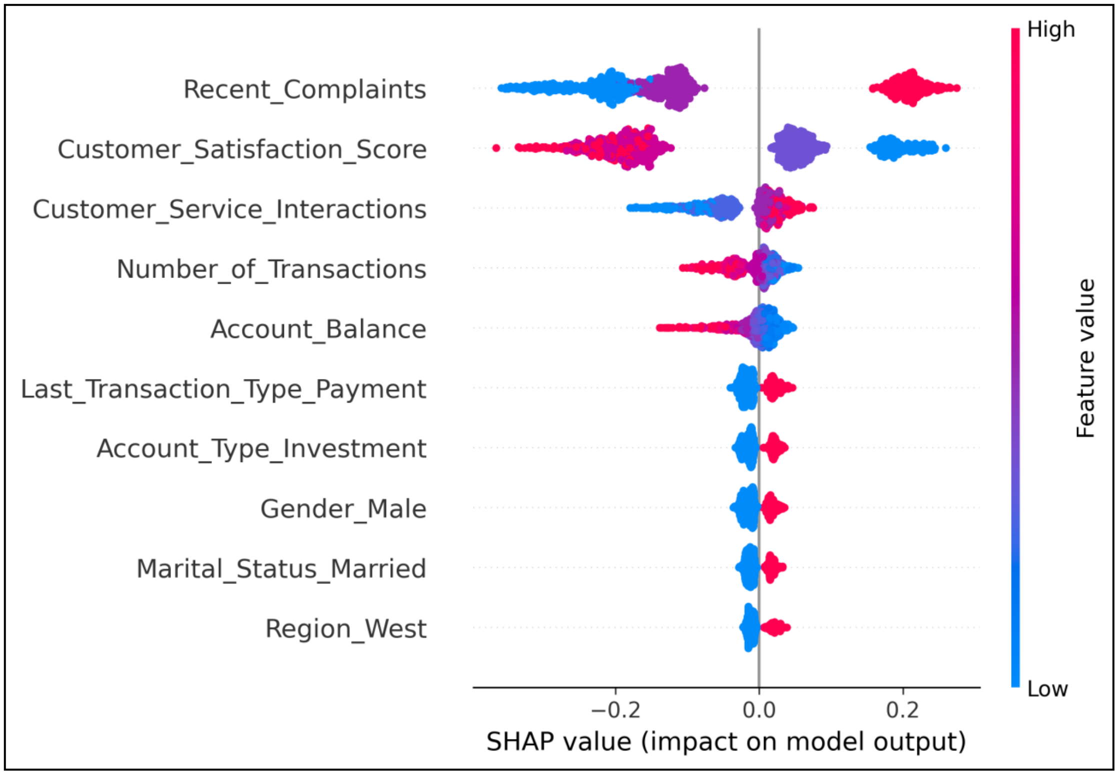

Figure 3: SHAP Bee Swarm Plot of Customer Churn

Next, we applied the goal-oriented model explanation method to analyze the importance of the five key features: 'Recent Complaints', 'Customer Satisfaction Score', 'Customer Service Interactions', 'Number of Transactions', and 'Account Balance'. The discoveries are displayed in Figure 4.

Based on the outputs, we can intuitively get the idea that the increase in 'Recent Complaints' leads to a substantially rise in the customer churn rate, while an increase in 'Customer Satisfaction Score' results in a severe drop of the churn rate. In addition, 'Number of Transactions' and 'Account_Balance' are both negatively correlated with churn rate, indicating that improving these business indicators could steadily diminish the customer churn rate. These discoveries are consistent with the conclusions derived from the SHAP model interpretation. Moreover, through the analysis using the goal-oriented' model interpretation method, the following additional insights were disclosed. When the feature values change within a ±50% range, 'Recent Complaints' could enormously increase or decrease the customer churn rate, while 'Customer Satisfaction Score' has the reverse effect. This highlights that reducing customer complaints and enhancing customer satisfaction are the most effective strategies for reducing churns, offering practical guidance to practitioners. As 'Number of Transactions' and 'Account_Balance' increase, the churn rate shows a continuous and moderate decline, indicating that enhancing a customer's account balance and transaction frequency, which reflecting customer engagement and usage intention, could consistently lower churn rates in a long run, which introduced a strategic ameliorating measure for customer relationship management.

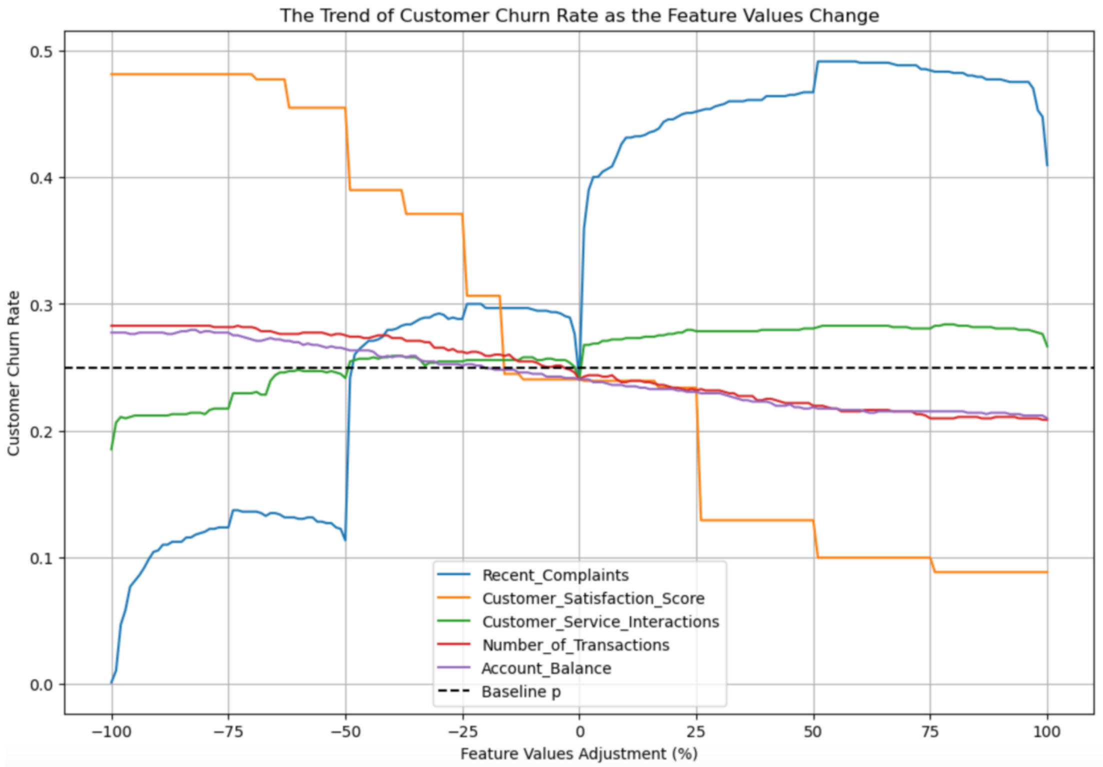

Figure 4: Goal-Oriented Model Interpretation

4.3 Strengths and Limitations of Goal-Oriented Model Interpretation

By implementing a comparative analysis, it becomes evident that the goal-oriented model interpretation method proposed in this study offers distinct advantages. It not only clearly illustrates the relationship between individual features and customer churn rate, but also visually demonstrates how changes in feature values affect the trajectory of the target variable. Furthermore, by analyzing each variable independently, this method enables a detailed examination of the marginal impact of feature changes on the target variable, thereby allowing practitioners to assess the input and output trade offs in diminishing customer churn.

Moreover, by conducting comparative analyses across multiple variables, this method facilitates the identification of the optimal feature adjustment strategy, or the 'best solution', to maximize customer retention or minimize churn. Although the 'ARM' model has taken the incentive cost into account, which makes it closer to the real-world question [26]. The innovative aspect of the goal-oriented explanation method lies in its integrated nature, where recommendations are not only the outcome of the analysis but also inherently verified by the analytical process itself. The method is rooted in the stable performance metrics of the selected champion model. Its procedure involves systematically modifying the values of individual variables in the prediction datasets, using the model to predict outcomes for each scenario, and reversely inferring the distribution of the target variable. This process also enables the model of demonstrating how feature variations directly influence the target outcome. The goal-oriented nature of this approach ensures that the conclusions, actionable recommendations, and their validation form a coherent and unified analytical framework.

4.4 Exploration of Generalizability

Selecting the bank CCP as a case study, this research demonstrates the complete analytical workflow, from data preprocessing, model construction, evaluation, and interpretation to the formulation and validation of actionable recommendations, summarized as the SGV framework. The proposed SGV framework is a predictive analytical approach characterized by three core components: Simplified predictive modeling, Goal-oriented model interpretation, and Verifiable recommendations. Compared to 'B2B-ARM' [23], which also achieved excellent performance in customer churn analysis, the simplified predictive modeling approach reduces modeling complexity by employing standard preprocessing techniques and automated hyperparameter optimization, ultimately yielding a model that is both efficient and robust. The goal-oriented model interpretation method ensures an organic integration of findings, recommendations, and validation within a single coherent framework.

Hence, the SGV framework provides a standardized, low barrier, and highly efficient solution for practitioners across industries and domains, which could further be extended to other fields. It is particularly well suited for professionals with limited backgrounds in data science or engineering, enabling them to leverage machine learning models and AI tools effectively in their specific contexts. This framework empowers users to extract deep value from data and convert those insights into tangible productivity gains within their industries [40].

5 CONCLUSION

Based on the example of customer churn analysis in the banking industry, this research applied the SGV framework to perform data preprocessing, model development, and model interpretation, in which the Random Forest was selected as the champion model. Meanwhile, the goal-oriented model explanation method was established and validated with SHAP, a popular model interpretation method. In addition, by performing the goal-oriented method, it's proved that customer churn is highly sensitive to complaint frequency and satisfaction scores, and that enhancing customer engagement can steadily and sustainably improve retention rates. Hence, the timely solution to handling customer churn is to avoid customer complaints and improve customer satisfaction. Enhancing customer engagement and usage intent, which includes scale up customer's account balance and transaction frequency, could boost customer retention.

Moreover, this study demonstrates that the SGV framework, characterized by its simplicity and efficiency, could empower practitioners across various industries to leverage machine learning algorithms and AI tools for deep data value extraction, and ultimately, will contribute to enhanced productivity in their respective fields. The SGV framework not only provides a new approach and methodology for digital transformation and AI empowerment across sectors but provides a novel toolkit for the evolution of analytics.

Despite its contributions, this research still has several limitations. Firstly, the theoretical foundation of the goal-oriented model interpretation method relies on reversing the estimation of positive and negative sample distributions, based on model performance metrics. This process introduces estimation errors which are difficult to quantify. Besides, the current version of the goal-oriented interpretation approach is limited to qualitative analysis. It does not yet support quantitative modeling of the relationships between features and the target. Moreover, the dataset applied in this research is a synthetic one, which may impair the framework's performance with real-word data. Finally, the SGV framework is primarily designed for users without professional data engineering skills, for whom simplicity and efficiency are preferred. As such, data preprocessing, feature engineering, and hyperparameter optimization are largely dependent on automated tools, which may limit the flexibility or fine tuning of models in highly complex scenarios.

Upon the limitations identified earlier, future research will primarily focus on the following directions. Optimizing the model interpretation approach by developing quantitative methods to explore the relationships between individual features and the target variable, which will give more precise explanation. Including sensitivity and behaviour analysis to interpret customer churn in a more quantitative and precise way. Evaluating the applicability of the SGV framework across different domains or using cases, to refine and promote this approach more broadly. Integrating advanced algorithms or deep learning models into the SGV framework, which could potentially enhance the effectiveness in both model performance and interpretation.

REFERENCES

[1] V. Chang, K. Hall, Q. A. Xu, F. O. Amao, M. A. Ganatra, and V. Benson, "Prediction of customer churn behavior in the telecommunication industry using machine learning models," *Algorithms*, vol. 17, no. 6, p. 231, May 2024, doi:10.3390/a17060231.

[2] O. S. Owolabi, P. C. Uche, N. T. Adeniken, O. Efijemue, S. Attakorah, O. G. Emi-Johnson, E. Hinneh *et al.*, "Comparative analysis of machine learning models for customer churn prediction in the U.S. banking and financial services: Economic impact and industry-specific insights," *Journal of Data Analysis and Information Processing*, vol. 12, no. 3, pp. 388–418, Jun. 2024, doi:10.4236/jdaip.2024.123021.

[3] J. Shobana, C. Gangadhar, R. K. Arora, P. N. Renjith, J. Bamini, and Y. D. Chincholkar, "E-commerce customer churn prevention using machine-learning-based business intelligence strategy," *Measurement: Sensors*, vol. 27, p. 100728, Jun. 2023, doi:10.1016/j.measen.2023.100728.

[4] A. Caldwell, "Losing customers? Calculate why," *Oracle NetSuite*, Jan. 27, 2021. [Online]. Available: https://www.netsuite.com/portal/resource/articles/human-resources/customer-churn-analysis.shtml. [Accessed: Aug. 21, 2025].

[5] J. Brereton, "The cost of customer churn: Understanding the financial impact," *LaunchNotes*, Nov. 22, 2022. [Online]. Available: https://www.launchnotes.com/blog/the-cost-of-customer-churn-understanding-the-financial-impact. [Accessed: Aug. 21, 2025].

[6] K. Peng, Y. Peng, and W. Li, "Research on customer churn prediction and model interpretability analysis," *PLOS ONE*, vol. 18, no. 12, p. e0289724, Dec. 2023, doi:10.1371/journal.pone.0289724.

[7] A. K. Ahmad, A. Jafar, and K. Aljoumaa, "Customer churn prediction in telecom using machine learning in big data platform," *Journal of Big Data*, vol. 6, no. 1, Mar. 2019, doi:10.1186/s40537-019-0191-6.

[8] P. Kündig, "Alternative data for restaurant bankruptcy prediction," *Swiss Journal of Economics and Statistics*, vol. 161, no. 1, p. 6, 2025, doi:10.1186/s41937-024-00141-7.

[9] Zilliant, "2022 Global B2B Benchmark Report: Distribution Edition (P7)," 2022. [Online]. Available: https://zilliant.com/images/hubs/Downloadable-Resources/distribution-edition-2022-global-b2b-benchmark-report.pdf. [Accessed: Aug. 21, 2025].

[10] McKinsey & Company, "McKinsey's global banking annual review," Oct. 17, 2024. [Online]. Available: https://www.mckinsey.com/industries/financial-services/our-insights/global-banking-annual-review. [Accessed: Aug. 21, 2025].

[11] V. Agarwal, S. Taware, S. A. Yadav, D. Gangodkar, A. Rao, and V. K. Srivastav, "Customer churn prediction using machine learning," in *Proc. IEEE Xplore*, Oct. 2022. doi:10.1109/9988187.

[12] P. Boozary, S. Sheykhan, H. GhorbanTanhaei, and C. Magazzino, "Enhancing customer retention with machine learning: A comparative analysis of ensemble models for accurate churn prediction," *International Journal of Information Management Data Insights*, vol. 5, no. 1, p. 100331, Feb. 2025, doi:10.1016/j.jjimei.2025.100331.

[13] S. S. Poudel, S. Pokharel, and M. Timilsina, "Explaining customer churn prediction in telecom industry using tabular machine learning models," *Machine Learning with Applications*, vol. 17, p. 100567, Jun. 2024, doi:10.1016/j.mlwa.2024.100567.

[14] S. K. Wagh, A. A. Andhale, K. S. Wagh, J. R. Pansare, S. P. Ambadekar, and S. H. Gawande, "Customer churn prediction in telecom sector using machine learning techniques," *Results in Control and Optimization*, vol. 14, p. 100342, Nov. 2023, doi:10.1016/j.rico.2023.100342.

[15] V. Gkonis and I. Tsakalos, "Deep dive into churn prediction in the banking sector: The challenge of hyperparameter selection and imbalanced learning," *Journal of Forecasting*, Sep. 2024, doi:10.1002/for.3194.

[16] S. Arefin, R. Parvez, T. Ahmed, F. Jahin, F. Sumaiya, and M. Ahsan, "Enhancing bank consumer retention: A comprehensive analysis of churn prediction using machine-learning and deep learning techniques," *Lecture Notes in Networks and Systems*, vol. 1075, pp. 265–281, Oct. 2024, doi:10.1007/978-981-97-6106-7_15.

[17] L. Kun, H. Alli, and A. Azlin, "GA optimization-based BRB AI reasoning algorithm for determining the factors affecting customer churn for operators," *Social Sciences & Humanities Open*, vol. 10, p. 100944, Jan. 2024, doi:10.1016/j.ssaho.2024.100944.

[18] A. D. Caigny, K. W. D. Bock, and S. Verboven, "Hybrid black-box classification for customer churn prediction with segmented interpretability analysis," *Decision Support Systems*, vol. 181, p. 114217, Jun. 2024, doi:10.1016/j.dss.2024.114217.

[19] B. Janssens, M. Bogaert, A. Bagué, and D. V. D. Poel, "B2Boost: Instance-dependent profit-driven modelling of B2B churn," *Annals of Operations Research*, vol. 341, Mar. 2022, doi:10.1007/s10479-022-04631-5.

[20] F. E. Usman-Hamza, A. O. Balogun, R. T. Amosa, L. F. Capretz, H. A. Mojeed, S. A. Salihu, A. G. Akintola, and M. A. Mabayoje, "Sampling-based novel heterogeneous multi-layer stacking ensemble method for telecom customer churn prediction," *Scientific African*, vol. 24, p. e02223, Jun. 2024, doi:10.1016/j.sciaf.2024.e02223.

[21] S. Rahman, B. Janssens, and M. Bogaert, "Profit-driven pre-processing in B2B customer churn modeling using fairness techniques," *Journal of Business Research*, vol. 189, p. 115159, Dec. 2024, doi:10.1016/j.jbusres.2024.115159.

[22] C. Aliferis and G. Simon, "Overfitting, underfitting and general model overconfidence and under-performance pitfalls and best practices in machine learning and AI," in *Computers in Health Care*, pp. 477–524, Jan. 2024, doi:10.1007/978-3-031-39355-6_10.

[23] E. Guliyev, J. S. Ramirez, A. D. Caigny, and K. Coussement, "Improving B2B customer churn through action rule mining," *Industrial Marketing Management*, vol. 125, pp. 1–11, Feb. 2025, doi:10.1016/j.indmarman.2024.12.005.

[24] Q. Li, Y. Li, and R. Jing, "A variant bagging forecasting framework for customer churn in airline," *Journal of Air Transport Management*, vol. 125, p. 102795, Apr. 2025, doi:10.1016/j.jairtraman.2025.102795.

[25] K. A. Pflughoeft, N. T. Butz, and A. Corbley, "Customer churn prediction for fixed wireless access: The case of a regional internet service provider," *Telecommunications Policy*, vol. 49, no. 4, p. 102929, Mar. 2025, doi:10.1016/j.telpol.2025.102929.

[26] Y. Feng, Y. Yin, D. Wang, J. Ignatius, T. C. E. Cheng, M. Marra, and Y. Guo, "Enhancing e-commerce customer churn management with a profit- and AUC-focused prescriptive analytics approach," *Journal of Business Research*, vol. 184, p. 114872, Nov. 2024, doi:10.1016/j.jbusres.2024.114872.

[27] K. Lepenioti, A. Bousdekis, D. Apostolou, and G. Mentzas, "Prescriptive analytics: Literature review and research challenges," *International Journal of Information Management*, vol. 50, no. 1, pp. 57–70, 2020, doi:10.1016/j.ijinfomgt.2019.04.003.

[28] S. Gupta and A. Gupta, "Dealing with noise problem in machine learning datasets: A systematic review," *Procedia Computer Science*, vol. 161, pp. 466–474, 2019, doi:10.1016/j.procs.2019.11.146.

[29] D. K. Thara and B. G. PremaSudha, "Auto-detection of epileptic seizure events using deep neural network with different feature scaling techniques," *Pattern Recognition Letters*, vol. 128, pp. 544–550, Dec. 2019, doi:10.1016/j.patrec.2019.10.029.

[30] D. U. Ozsahin, M. T. Mustapha, A. S. Mubarak, Z. S. Ameen, and B. Uzun, "Impact of feature scaling on machine learning models for the diagnosis of diabetes," in *Proc. Int. Conf. Artificial*

Intelligence in Everything (AIE), Aug. 2022, doi:10.1109/aie57029.2022.00024.

[31] F. Yang, Z. Xu, H. Wang, L. Sun, M. Zhai, and J. Zhang, "A hybrid feature selection algorithm combining information gain and grouping particle swarm optimization for cancer diagnosis," *PLOS ONE*, vol. 19, no. 3, p. e0290332, Mar. 2024, doi:10.1371/journal.pone.0290332.

[32] P. Lan, Q. Shi, P. Zhang, Y. Chen, R. Yan, X. Hua, Y. Jiang, J. Zhou, and Y. Yu, "Core genome allelic profiles of clinical *Klebsiella pneumoniae* strains using a random forest algorithm based on multilocus sequence typing scheme for hypervirulence analysis," *The Journal of Infectious Diseases*, vol. 221, no. Suppl. 2, pp. S263–S271, Mar. 2020, doi:10.1093/infdis/jiz562.

[33] J.-H. Chen, C.-S. Chen, M.-F. Huang, and H.-C. Lin, "Estimating the probability of rare events occurring using a local model averaging," *Risk Analysis*, vol. 36, no. 10, pp. 1855–1870, Feb. 2016, doi:10.1111/risa.12558.

[34] W. He, H. He, F. Wang, S. Wang, R. Li, J. Chang, and C. Li, "Rapid and noninvasive characterization of bananas by hyperspectral imaging with extreme gradient boosting (XGBoost)," *Analytical Letters*, vol. 55, no. 4, pp. 620–633, Apr. 2022, doi:10.1080/00032719.2021.1952214.

[35] N.-H. Nguyen, K. T. Tong, S. Lee, A. Karamanli, and T. P. Vo, "Prediction of compressive strength of cement-based mortar containing metakaolin using explainable categorical gradient boosting model," *Engineering Structures*, vol. 269, p. 114768, Oct. 2022, doi:10.1016/j.engstruct.2022.114768.

[36] C. Rohlfs, "Generalization in neural networks: A broad survey," *Neurocomputing*, vol. 611, p. 128701, Oct. 2024, doi:10.1016/j.neucom.2024.128701.

[37] S. Dhanka and S. Maini, "HyOPTXGBoost and HyOPTRF: Hybridized intelligent systems using Optuna optimization framework for heart disease prediction with clinical interpretations," *Multimedia Tools and Applications*, vol. 83, Feb. 2024, doi:10.1007/s11042-024-18312-x.

[38] H. Almarzooq and U. B. Waheed, "Automating hyperparameter optimization in geophysics with Optuna: A comparative study," *Geophysical Prospecting*, vol. 72, no. 5, Feb. 2024, doi:10.1111/1365-2478.13484.

[39] P.-Y. Tseng, Y.-T. Chen, C.-H. Wang, K.-M. Chiu, Y.-S. Peng, S.-P. Hsu, K.-L. Chen, C.-Y. Yang, and O. K.-S. Lee, "Prediction of the development of acute kidney injury following cardiac surgery by machine learning," *Critical Care*, vol. 24, no. 1, p. 478, Jul. 2020, doi:10.1186/s13054-020-03179-9.

[40] E. SY, T.-C. Peng, H.-Y. Lin, S.-H. Huang, Y.-C. Chang, and C.-P. Chung, "Ensemble BERT techniques for financial sentiment analysis and argument understanding with linguistic features in social media analytics," *Journal of Information Science and Engineering*, vol. 41, no. 3, pp. 579–599, 2025, doi:10.6688/JISE.202505_41(3).0004.