CUSTOMER CHURN PREDICTION AND RETENTION STRATEGY DEVELOPMENT USING EXPLAINABLE MACHINE LEARNING

Angelo Baby

Phd Scholar, CHRIST University, Bangalore City/State, Karnataka, India. Email id: angelo.baby@res.christuniversity.in

Abstract

Customer churn is a severe problem for service companies since losses have a direct impact on profitability, revenue and long-term competitiveness. This research has created an interpretable machine learning model for customer churn prediction and designed a targeted retention strategy accordingly. The trained and evaluated four supervised learning models were Logistic Regression, Decision Tree, Random Forest and XGBoost and the following metrics were used for evaluation: Accuracy, Precision, Recall, Specificity, F1-Score, Balanced Accuracy, Matthews correlation coefficient, ROC-AUC and Precision–Recall AUC. XGBoost had the best overall predictive performance with ROC-AUC of 0.8545, PC-R AUC of 0.6724, PC-recall of 0.7941 and an F1 score of 0.6339 out of all the models evaluated. Explainability was added in the form of a SHapley Additive exPlanations (SHAP) analysis to highlight the key global and customer-level factors that affect churn. All the following were among the most influential: month-to-month contracts, short tenure, no dependents, higher monthly charges, fibre-optic internet service, electronic-check payment, no online security, and no online technical support. Customers were also separated into 4 groups: low risk, moderate risk, high risk, and critical risk according to their expected churn rates. Churn rate in the critical-risk segment was 71.74% while it was 2.17% in the low-risk segment. The results show the potential of explainable machine learning for enhancing customer churn prediction and managerial decisions by associating customer risk scores with interpretable churn drivers and value-based retention actions.

Keywords: *Customer churn; customer retention; explainable artificial intelligence; machine learning; customer segmentation; predictive analytics*

Received:08/04/2026 Revised: 11/05/2026 Acceptance: 19/05/2026 Publication: 26/05/2026

Introduction

Churn is when a customer no longer does business with a company or no longer uses its products or services. The cost of losing customers is a key strategic issue in telecommunications, banking, insurance, retail, e-commerce and other service industries, as customers are the lifeline of the business and their departure means the loss of revenue, market share and future profitability. Keeping customers is typically less expensive than constantly gaining new ones, so it's critical that you identify early on churn customers, a key part of customer relationship management. The analysis of churn has now shifted from a descriptive report to predicting the probability of a particular customer leaving a service contract (Ahn et al., 2020).

The telecom sector suffers from a high customer turnover rate because customers are easily tempted to switch providers due to the price, services, contract terms, technical issues or an offer from a competitor. With the growth of digital services and the extremely competitive environment, the need for proactive churn management is accentuated further. Basic statistical models and rule-based systems are not always effective in modelling the complex and non-linear relationships between the various demographic, contractual, behavioural, service-related and billing attributes. To better predict churn, patterns deep in large and multidimensional customer data have been increasingly identified using machine-learning techniques (Barsotti et al., 2024; Chang et al., 2024).

Logistic Regression, Decision Trees, Random Forests, Support Vector Machines, gradient-boosting models, and deep-learning models are the recent models used in churn prediction studies. The methods are variable in their predictive power, complexity of computation and interpretability. When there are several interdependent factors influencing customer behaviour, ensemble and deep-learning models can capture nonlinear interactions and perform better than traditional statistical models when prediction accuracy is important. (Khattak et al., 2023) But a high predictive accuracy is not enough to effectively manage churn. While a model can be effective at predicting whether a customer is likely to churn, it may not be clear why or what needs to be done to avoid customer loss.

This restriction has led to a growing interest in the field of explainable artificial intelligence, where the goal is to make decisions made by machine learning systems comprehensible to people. Explainable AI offers ways to detect variables that impact model predictions, assess the direction and magnitude of their impact, and explain why certain customers are assigned high or low churn-risk scores (Arrieta et al., 2020; Belle & Papantonis, 2021). In customer analytics, explainability is especially valuable as managers need to have clear and useful information to spend funds on retention, offer a discount, change contracts, and/or make some kind of customer-support personalization effort before them. If there are no explicit explanations, complex models can be hard to believe, validate and incorporate into organizational decision-making processes (Borrego-Díaz & Galán-Páez, 2022; Emmert-Streib et al., 2020).

There are two types of explainable machine-learning techniques, global and local. Global explanation methods are used to explain the variables that are typically significant across the entire customer base, while local explanation methods explain the variables that are why an individual customer. The explainable models can provide enhanced transparency, accountability, model validation, and communication among technical analysts and business decision makers (Islam et al., 2021; Linardatos et al., 2020). Among the available methods, Shapley Additive exPlanations (SHAP) have gained popularity since it offers theoretically grounded contributions of features, both for individual and aggregated predictions. SHAP is especially well-suited for tree-based models and may link to local explanations for customers to global patterns of model behavior (Lundberg et al., 2020).

However, there are significant challenges in bridging the gap between churn forecasting and creating effective retention strategies. One drawback of numerous studies is that they look at classification accuracy, model comparison, or feature ranking without paying much attention to the translation of the prediction results to targeted managerial actions. There are multiple reasons that customers churn; these reasons can be expensive charges, short tenures, weak service engagement, impractical contracts, poor technical support, or uncomfortable payment methods. It can thus be expensive and unproductive to apply the same retention offer to all high-risk customers. What would be more useful is for a model to include: a) predicted churn probability and b) the drivers of churn as communicated to the customer, so retention interventions could be customised to the specific context of each customer (Islam et al., 2022).

Therefore, this research combines explainable analysis and retention strategy development with machine learning prediction. The proposed approach involves comparing different classification models, measuring their performance using metrics appropriate to imbalanced data, and explaining the selected model using both global and local explanation techniques. The study also draws model-generated insights to develop risk-based and customer-specific retention recommendations. The research bridges the gap between predictive performance and management action by providing insights into the interpretability of machine learning results and how it can inform more actionable and transparent customer-retention decision-making.

Objectives

1. To create and evaluate machine learning models to accurately predict customer churn.
2. To understand the key churn drivers across the globe and customers based on explainable machine learning methods.

3. To develop personalized retention strategies based on the dominant factors influencing the individual customer decision by aligning churn risk with the factors.

2.1 Research Design
The study followed a quantitative predictive research design to construct a machine-learning framework that can predict customers who are likely to churn and create a strategy to retain them. Churning was considered a binary classification problem. The approach taken to the methodology was data preparation, exploratory analysis, machine-learning model development, performance evaluation, interpretation of the model, segmentation of customer risk, and formulation of a retention strategy.

2.2 Dataset Description
The Telco Customer Churn: IBM Dataset was used for the study (TanKY, 2020). The data includes 7,043 customer records and 33 variables related to customer demographics, services subscribed to, tenure, contract type, billing method, payment type, monthly charges, total charges, customer lifetime value, and churn behaviour. The dataset consisted of 1,869 churned customers and 5,174 retained customers with a churn rate of around 26.53%. The observations were of one telecommunications customer each.

2.3 Study Variables
Churn Value was the dependent variable, which was coded as 1 if a customer churned and 0 if they did not churn. Predictor variables were gender, senior-citizen status, partner status, dependents, tenure months, phone service, multiple lines, internet service, online security, online backup, device protection, technical support, streaming services, contract type, paperless billing, payment method, monthly charges, and total charges. The primary use of CLTV was to determine who the high-value customers were to use them for retention interventions. Variables that might lead to leakage of the target were removed from the training of the model. These comprised Churn Label, Churn Score, and Churn Reason. Other variables, such as CustomerID, Count, Country, and State, were also dropped as they were identifiers or constant variables that had little or no predictive value. Churn Reason was used only in model development to help formulate retention strategies.

2.4 Data Preprocessing
The data set was checked for any duplicated observations, inconsistent categories, missing values, and incorrectly formatted variables. The Total Charges variable was converted to numeric data, and any missing data for customers with no tenure was set equal to zero. The numerical encoding of binary categorical variables and one-hot encoding of categorical variables with multiple categories were performed. If the algorithm is sensitive to differences in measurement scale, continuous variables (e.g., tenure months, monthly charges, total charges) were standardized. A moderate customer churn ratio was fixed with class weight adjustment. The Synthetic Minority Oversampling Technique was used only on the training data to avoid information leakage when oversampling was used.

2.5 Exploratory Data Analysis
The features of the customers were summarized with descriptive statistics. Categorical variables were analyzed as frequencies and percentages, and continuous variables were analyzed as means, standard deviations, medians and interquartile ranges. Contract type, payment method, internet services, technical-support status, online-security status, and demographic groups were compared for churn rates. Patterns were identified, and potential churn determinants were identified using histograms, bar charts, boxplots and correlation analysis.

2.6 Data Partitioning and Model Development
A stratified sampling method was used to split the data into an 80% training set and a 20% testing set, while keeping the same percentages of churned and retained customers in each set. To select the models and optimize hyperparameters, we used stratified 5-fold cross-validation for the training set. The independent test set was only used as a final test of performance. Five machine-learning (ML) supervised models were trained and evaluated: Logistic Regression, Decision Tree, Random Forest, XGBoost, and CatBoost. Logistic regression was used as an interpretable baseline model, and the tree-based ensemble models for discovering nonlinear relationships and interactions among customer characteristics. Hyperparameters were optimized by random selection or grid search during the cross-validation process.

2.7 Model Performance Evaluation
Various models were assessed by accuracy, precision, recall, specificity, F1 score, balanced accuracy, Matthews correlation coefficient, receiver operating characteristic area under curve, and precision–recall curve area under curve. The confusion matrices were also analyzed to see what percentages of predictions were true positive, true negative, false positive and false negative. The importance of recall and precision recall area under the curve is highlighted because the failure to correctly identify an actual churner may mean a customer will be lost, who may

have been retained. The final model was chosen based on predictive performance, calibration, interpretability and the business of the model.

2.8 Explainable Machine-Learning Analysis
SHapley Additive exPlanations were used to interpret the best-performing model. Global SHAP analysis was used to determine variables whose overall effects on the customer churn were the largest, and the SHAP summary plots were used to show the direction and magnitude of the impact of each variable. To investigate the effect of key continuous variables (e.g., tenure and monthly charges) on churn probability, partial dependence plots were created. Waterfall plots for each local SHAP were created to interpret each customer's prediction. All of these plots showed the factors that increased or decreased the probability of churn for each customer. A month-to-month contract, for instance, or a high monthly charge, electronic-check payment, and no technical support might boost a customer's predicted churn risk.

2.9 Customer Segmentation and Retention Strategy Development
Low risk, moderate risk, high risk, and critical risk segments were defined by customers predicted churn risk probabilities. Customer lifetime value and the predicted probability were used to calculate the retention priority. High churn potential, high lifetime value customers were targeted for immediate personalized intervention. Retention actions were correlated with the key churn drivers determined by SHAP analysis. Customers who are impacted by high charges may be given a customised pricing plan, and customers with month-to-month contracts may be provided an incentive to sign up for longer contracts. Customers with no technical support or online security may be offered reduced service packages, while new customers may be geared up by means of enhanced onboarding and communication initiatives.

2.10 Ethical Considerations
No customer names or direct contact information were provided. Data were analyzed without identifying customers and only for academic purposes. Model outputs were an indication of churn risk, not a final verdict on customer behaviour. Fairness was evaluated by comparing the errors of the models for relevant demographic groups.

3. Results
3.1 Sample Characteristics and Churn Distribution
There were 7,043 telecommunications customers in the final analytical dataset. In these, 5,174 customers continued to be loyal to the company, and 1,869 customers churned. The churn rate observed was 26.54%, and the retention rate was 73.46%. The average customer's stay duration was 32.37 months, the average monthly cost was 64.76 monetary units, and the mean customer lifetime value was 4,400.30 monetary units, as shown in Table 1 below.

Table 1. General characteristics of the study sample

Characteristic	Value
Total customers	7,043
Retained customers	5,174
Churned customers	1,869
Retention rate	73.46%
Churn rate	26.54%
Mean tenure	32.37 months
Mean monthly charges	64.76
Mean customer lifetime value	4,400.30

The distribution in Figure 1 suggests a moderate class imbalance; around three-quarters of the observations were of retained customers. The imbalance was not extreme but warranted the use of class-weighted models and performance indicators other than the overall level of accuracy.

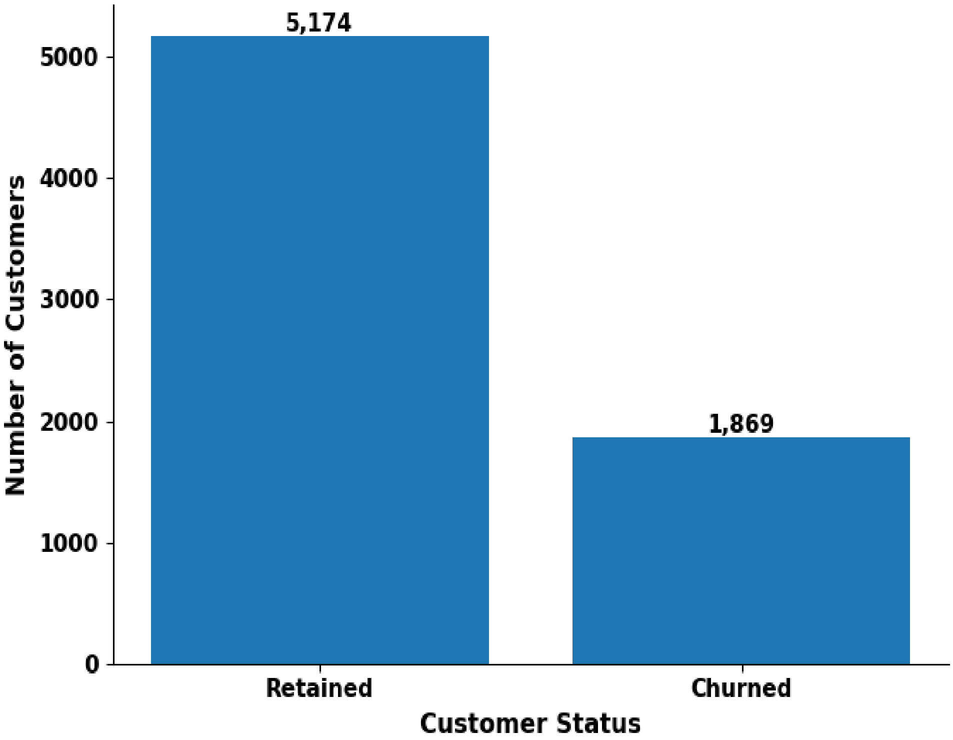

Figure 1. Distribution of retained and churned customers

3.2 Churn Patterns Across Customer Groups

Significant differences in churn were seen by contract type, payment method and category of internet service. Customers with the month-to-month contracts experienced the highest churn rate of 42.71%, whereas those with one year contract had 11.27% churn rate and only 2.83% churn rate for those with two years contract. The findings reveal that contractual commitment had a strong correlation with customer retention. The churn rate for electronic-check users was also quite high – 45.29% – compared with automatic bank transfer users and automatic credit-card payment and mailed check users. Fibre-optic customers also exhibited a relatively high churn rate, at 41.89%, compared to 18.96% for DSL and 7.40% for customers who do not use Internet service. The results for the whole group are displayed in Table 2.

Table 2. Churn rates across selected customer characteristics

Variable	Category	Customers	Churned customers	Churn rate
Contract	Month-to-month	3,875	1,655	42.71%
Contract	One year	1,473	166	11.27%
Contract	Two year	1,695	48	2.83%
Payment method	Bank transfer—automatic	1,544	258	16.71%
Payment method	Credit card—automatic	1,522	232	15.24%
Payment method	Electronic check	2,365	1,071	45.29%
Payment method	Mailed check	1,612	308	19.11%
Internet service	DSL	2,421	459	18.96%
Internet service	Fiber optic	3,096	1,297	41.89%
Internet service	No internet service	1,526	113	7.40%

The correlation between contract length and churning is even stronger, as shown in Figure 2. The churn rate also reduced significantly with the length of the contracted period, indicating customers with longer-term contractual commitments were much less likely to churn.

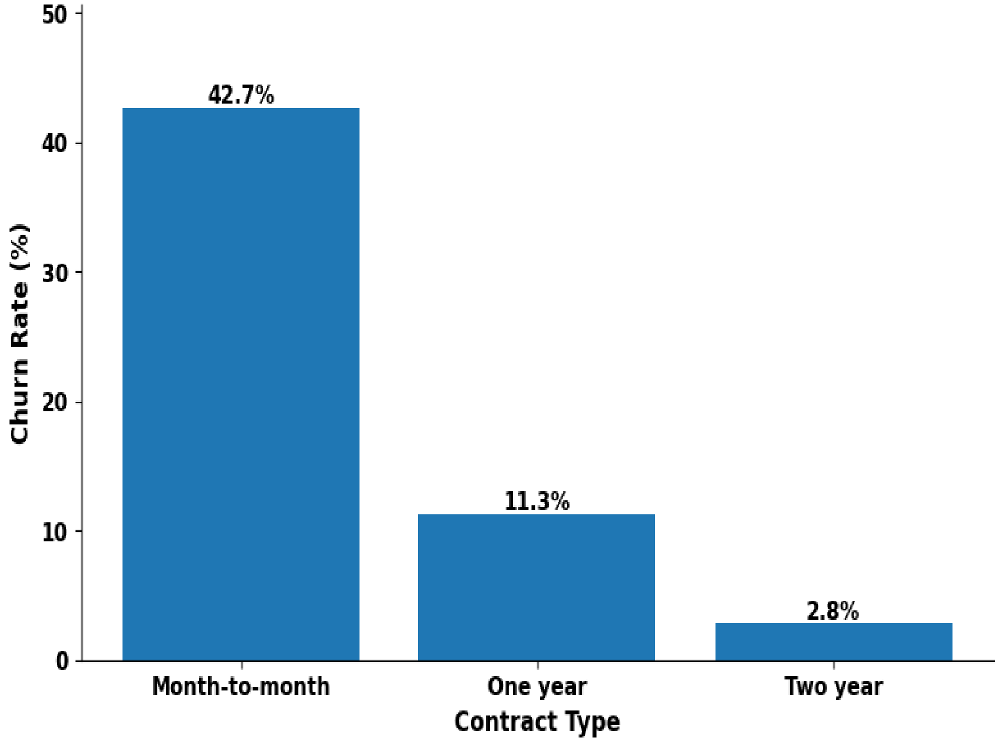

Figure 2. Churn rate according to contract type

3.3 Comparative Performance of Machine-Learning Models

The following machine-learning algorithms were tested on an independent test set: Logistic Regression, Decision Tree, Random Forest, XGBoost, and CatBoost. The accuracy, precision, recall, specificity, F1-score, balanced accuracy, Matthews correlation coefficient, ROC-AUC, precision–recall AUC, and Brier score were used to compare the models. XGBoost had the highest ROC-AUC of 0.8545 and the highest precision–recall AUC of 0.6724, as shown in Tab. 3. It earned the best recall-adjusted F1 score of 0.6339 and the best-balanced accuracy of 0.7686. The XGBoost model was able to accurately classify 79.41% of the churned customers. Random Forest had the highest overall accuracy (76.65%), the highest precision (54.44%) and the highest specificity (77.68%). But its recall was less 73.80% than of XGBoost. The CatBoost model showed competitive results with the ROC-AUC = 0.8513, precision-AUC = 0.6701. The precision–recall AUC values of Logistic Regression and Decision Tree were 0.6450 and 0.6170, respectively.

Table 3. Comparative performance of the machine-learning models

Model	Accuracy	Precision	Recall	Specificity	F1-score	Balanced accuracy	MCC	ROC-AUC	PR-AUC	Brier score
XGBoost	0.7566	0.5275	0.7941	0.7430	0.6339	0.7686	0.4842	0.8545	0.6724	0.1576
CatBoost	0.7637	0.5389	0.7594	0.7652	0.6304	0.7623	0.4787	0.8513	0.6701	0.1544
Random Forest	0.7665	0.5444	0.7380	0.7768	0.6266	0.7574	0.4736	0.8519	0.6667	0.1499
Logistic Regression	0.7445	0.5124	0.7727	0.7343	0.6162	0.7535	0.4570	0.8485	0.6450	0.1642
Decision Tree	0.7395	0.5060	0.7914	0.7208	0.6173	0.7561	0.4590	0.8414	0.6170	0.1674

Random Forest performed a bit better in terms of accuracy and probability calibration, but XGBoost was chosen as the top churn prediction model due to its highest ROC-AUC, precision–recall AUC, recall, F1-score, balanced accuracy, and Matthews correlation coefficient. It would be particularly critical in the context of churn management because it reduces the churned customers, who were not identified for intervention.

3.4 Receiver Operating Characteristic Analysis

The receiver operating characteristic (ROC) curves are plotted in Figure 3. The discriminatory performance of all five models was significantly better than random classification. XGBoost, Random Forest and CatBoost were the top three models with the largest area under the curve. The difference in ROC-AUC between the models was not very large, suggesting that all the models can differentiate between churned and retained customers. But XGBoost had the best true positive rate/false positive rate trade-off at other classification thresholds.

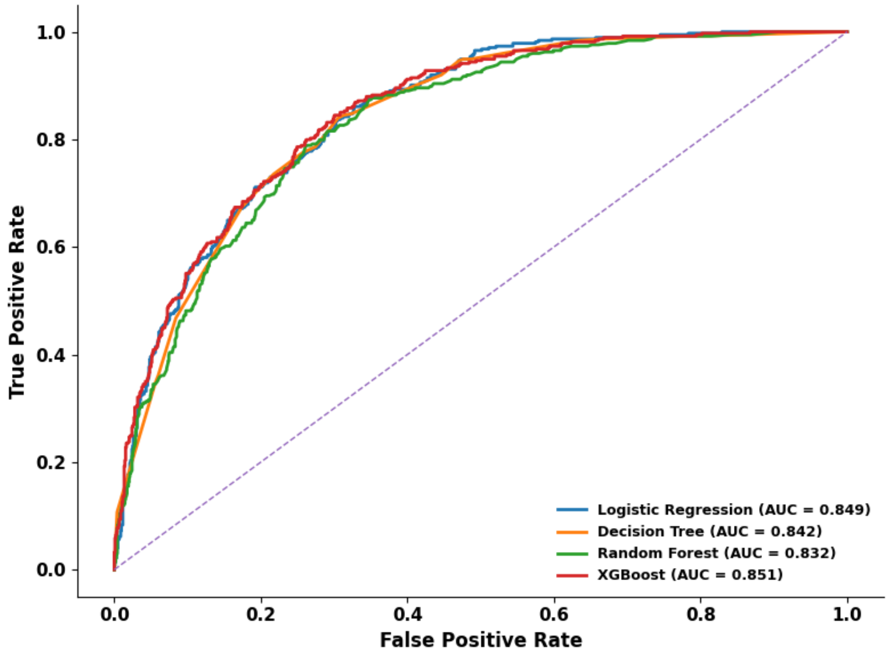

Figure 3. Receiver operating characteristic curves of the machine-learning models

3.5 Precision–Recall Analysis

To provide a more informative assessment in the context of the observed class imbalance, the precision–recall curves presented in Figure 4 are used. After CatBoost with the highest AUC of 0.6701, XGBoost has the highest precision–recall AUC of 0.6724, followed closely by the Random Forest with 0.6667. The better AUC for precision/recall of XGBoost showed that the model had a relatively high capability to pinpoint churners, while limiting the number of consumers classified as incorrect churn risks. The Decision Tree performed the worst in terms of precision–recall, indicating that its less complex structure could not capture all the complex interactions in customer behaviour.

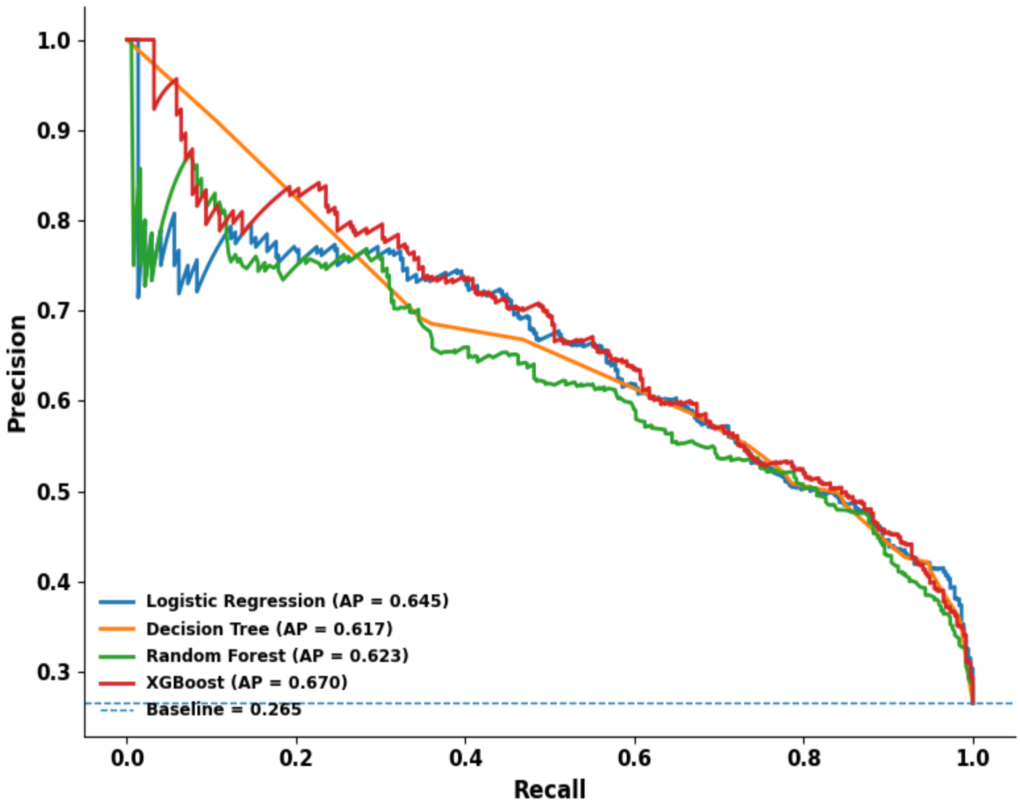

Figure 4. Precision–recall curves of the machine-learning models

3.6 Explainable Machine-Learning Results

To find the most important variables for churn predictions, SHAP analysis was applied to the selected XGBoost model. The results presented in Table 4 show that a month-to-month contract was the most influential predictor, with a mean absolute SHAP value of 0.7031. The second factor that was most important to the customers was tenure, followed by no dependents, monthly charges, fibre-optic internet service, no online security, total charges, electronic-check payment, and no technical support.

Table 4. Leading churn predictors identified through SHAP analysis

Rank	Predictor	Mean absolute SHAP value
1	Month-to-month contract	0.7031
2	Tenure months	0.5372
3	No dependents	0.4908
4	Monthly charges	0.2423
5	Fiber-optic internet service	0.2229
6	No online security	0.2189
7	Total charges	0.1919
8	Electronic-check payment	0.1860
9	No technical support	0.1708
10	Two-year contract	0.1297
11	No paperless billing	0.1122
12	No multiple lines	0.1088
13	No partner	0.0941
14	Has dependents	0.0821
15	No online backup	0.0688

The magnitude and direction of the relationships were shown in the SHAP summary plot in Figure 5. Generally, month-to-month contracts, shorter tenancy, higher monthly fees, fibre optics, electronic-check payment and no online security or technical support all increased churn risk predictions. Longer tenure and longer-term contracts, on the other hand, decreased the likelihood of churn. The results indicate that the XGBoost model was not sensitive to any factor. Rather, the combinations of contractual, financial, service, and household characteristics led to the emergence of churn risk.

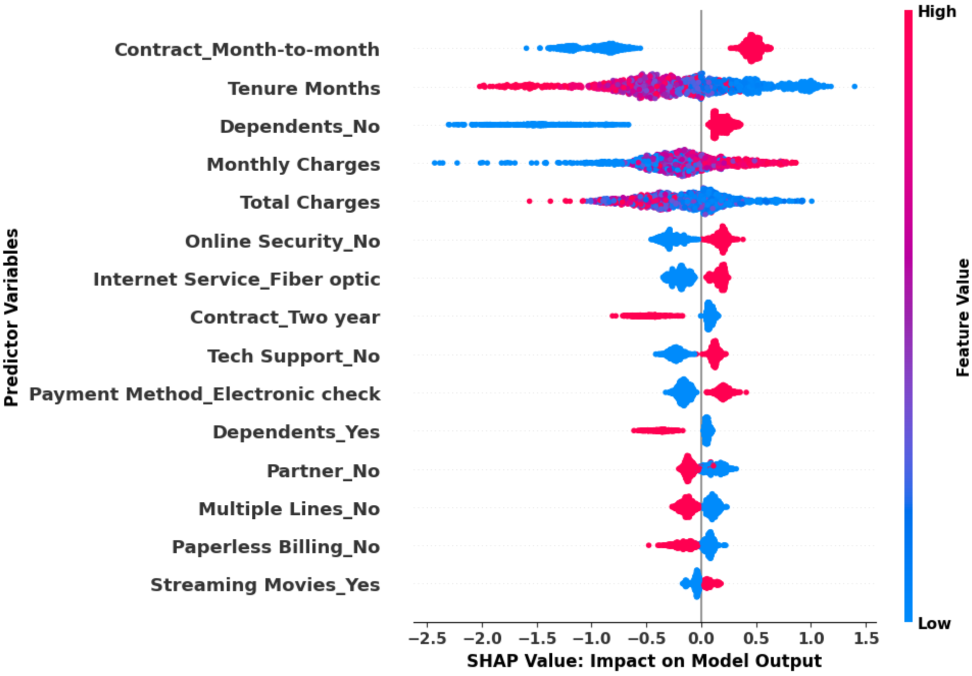

Figure 5. SHAP summary plot of the leading churn predictors

3.7 Churn-Risk Segmentation

Customers were categorized into four risk groups depending on the probability levels that they were predicted to churn. The low-risk group consisted of customers with a predicted probability of less than 25%, the moderate-risk group was 25% to above 50% and the high-risk group was 50% to above 75%, and the critical-risk group were those predicted at 75% or higher. 3,087 customers (43.83% of the total data) were classified as low risk, as presented in Table 5. Their average predicted churn rate was 7.61%, and the actual churn rate was just 2.17%. However, 1,451 customers were identified as critical risk. This segment accounted for 20.60% of the customers and had a customer churn rate of 71.74%.

Table 5. Predicted churn-risk segments and customer characteristics

Risk segment	Customers	Customer share	Mean predicted risk	Observed churn rate	Mean CLTV
Low	3,087	43.83%	7.61%	2.17%	4,640.99
Moderate	1,141	16.20%	37.06%	18.93%	4,374.51
High	1,364	19.37%	63.24%	39.96%	4,270.62
Critical	1,451	20.60%	85.59%	71.74%	4,030.40

The observed churn of each of the four groups of predicted risk points clearly shows that meaningful customer-risk rankings were achieved. The number of customers belonging to each segment is shown in Figure 6.

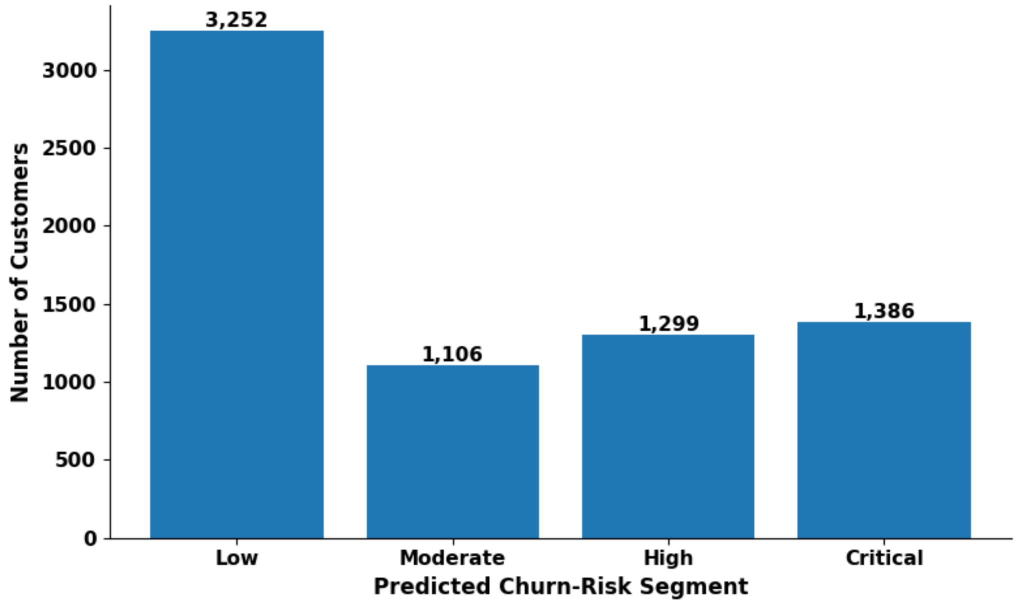

Figure 6. Distribution of customers across predicted churn-risk segments

3.8 Retention Strategy Classification

A risk-segmentation and SHAP were used to create specific retention actions. For customers in the critical-risk segment, immediate action was necessary, especially if the customers' lifetime values are high. Personal communication and service offers were required for customers in the high-risk group, while customers in the moderate-risk group were deemed to be suitable for preventive engagement. There is a need for regular surveillance and reinforcement of low-risk customers.

Table 6. Model-informed retention strategies for major churn drivers

Principal churn driver	Target customer group	Recommended retention strategy
Month-to-month contract	High- and critical-risk customers	Offer a discounted one-year or two-year contract conversion
Short customer tenure	Newly acquired high-risk customers	Strengthen onboarding, welcome communication, and early support
High monthly charges	Price-sensitive high-value customers	Provide personalized bundles, plan reviews, or limited discounts
Fiber-optic service	Dissatisfied fiber-optic customers	Conduct service-quality review and provide technical assistance
No online security	Customers without security services	Offer a free trial or discounted online-security package
Electronic-check payment	Electronic-check users	Encourage automatic payment through incentives
No technical support	Customers lacking support services	Provide temporary complimentary or discounted technical support
High risk and high CLTV	Most valuable at-risk customers	Assign personalized outreach and priority account management
High risk and low CLTV	Lower-value at-risk customers	Use automated, low-cost retention campaigns
Low risk and high CLTV	Valuable loyal customers	Provide loyalty rewards and relationship-building programmes

The outcomes showed that if a retention offer was implemented across the entire base, it would be inefficient since customers had various combinations of churn drivers. Predicted risk, customer lifetime value and individual SHAP explanations were integrated, enabling interventions to be tailored to the individual reason for a customer's risk of churning.

3.9 Summary of Principal Results

It was found that the use of the IBM Telco Customer Churn dataset provided good discriminatory accuracy for predicting customer churn. XGBoost was found to be the best model overall, with an ROC-AUC of 0.8545, a precision–recall AUC of 0.6724 and churn recall of 79.41%. The top churn influences included month-to-month contracts, short contract terms, no dependents, high monthly charges, fibre-optic service, no online security, electronic-check payment, and no technical support. The risk-segmentation analysis also revealed that the churn rate for the critical-risk group was 71.74% versus just 2.17% for the low-risk group. The results are also applicable in everyday practice to identify at-risk customers and comprehend what makes them tick in their likely future behaviour, to form differentiated retention strategies.

4. Discussion

The goal of the present study is to have an explainable machine-learning model to predict customer churn and formulate retention strategies. XGBoost had the best overall predictive performance with ROC-AUC of 0.8545, precision–recall AUC of 0.6724, recall of 0.7941 and F1 score of 0.6339 among the evaluated models. Random Forest achieved slightly higher accuracy and specificity, but XGBoost was chosen as the best model as it picked up a higher percentage of churners. It's significant to the retention business because if you don't identify a customer who later walks away, it can be more expensive than misidentifying a retained customer as high risk. It was therefore justified to use multiple evaluation indicators because, in imbalanced binary classification tasks, using accuracy as the only indicator can be misleading. Moreover, Matthews correlation coefficient and balanced accuracy offered a more accurate classification quality evaluation for both outcome classes (Chicco & Jurman, 2020).

The results of XGBoost were superior as previous research has shown that ensemble or hybrid learning is effective for churn prediction. Tree-based ensembles have the potential to capture nonlinear relationships, interaction effects and threshold effects that are not well represented by simpler linear models. In the past, similar evidence has demonstrated that carefully designed ensemble classifiers can outperform individual classifiers, without sacrificing interpretability completely (De Bock & De Caigny, 2021). Another application with good results in telecommunications churn prediction is the use of hybrid frameworks, which combine different preprocessing, feature-selection, and classification methods (Ouf et al., 2024). The results also show that model configuration is a key factor for performance, as the hyperparameter tuning process can significantly affect classification results, especially in complex machine-learning and deep-learning models (Domingos et al., 2021).

The month-to-month contracts, shorter tenure, no dependents, high monthly charges, fibre-optic internet service, no online security, electronic-check payment, and no technical support were found to be the top negative drivers for churn risk in the SHAP analysis. These findings show that the reasons for churning were not due to one specific attribute alone, but because of a few contractual, financial, service-related, and family-related benefits. Contract type has a significant effect, as customers who have only a small contractual obligation have more flexibility to switch providers. Likewise, low retention rates can be attributed to low customer attachment, less onboarding, or less embedded switching costs. The significance of customer profiling for these results aligns with previous studies demonstrating that customer segmentation, in addition to churn prediction, can provide more actionable insights than churn prediction alone (Geiler et al., 2022).

The adaptation of SHAP enhanced the usefulness of the predictive model by explaining overall patterns and individual customer predictions. This result is consistent with prior results, which demonstrate that an explainable-AI model can be used to convert complex churn models into meaningful business insights (Marín Díaz et al., 2022). However, the quality of the explanations should not be taken for granted only because an explanation method has been used. Systematic evaluation is still required to assess the reliability, robustness, and consistency of the explanatory outputs (Nauta et al., 2023). Structured ways of assessing the quality of the explanations, like the tools used in Quantus, could enhance the future assessment of the stability and faithfulness of the SHAP values (Hedström et al., 2023).

The risk-segmentation analysis also showed the operation value of the model. Customers identified as critical risk churned at 71.74% rate while low risk customers churned at 2.17% rate. This distinction is clear and implies that the model provided useful probability rankings, which could be used to inform differential retention strategies. You just need to find the right customers; however, for them to respond to intervention, you cannot identify them as likely churners. To overcome this limitation, uplift modelling identifies customers who can be persuaded to change their behavior in response to a retention action, but also customers who would stay or go without any actions taken (De Caigny et al., 2021). The current framework should be extended to include treatment-response data, and the incremental effect of retention offers should be evaluated experimentally in future research.

The research also correlated churn likelihood with the lifetime value of a customer, thereby prioritizing the

retention efforts. This approach is similar to profit-oriented churn modelling, which focuses on intervention cost and customer value in addition to predicting accuracy when deciding on retention (Maldonado et al., 2020). Recent studies also confirm that interpretable, profit-driven classifiers can enhance the economic value of churn predictions by giving higher weight to cases where it matters economically (Jiang et al., 2024). So, high-risk, high-value customers should be targeted in individual ways, while low-value customers can be better reached via automated and cost-effective campaigns.
There are some limitations in the study, however. The cross-sectional structure of the analysis is not able to reflect the changes in behaviour through time. When longitudinal data is available, time-varying indicators like recency and frequency scores as well as monetary indicators, may give better insights into the emerging churn risk (Mena et al., 2024). In addition, the proposed retention strategies were not tested interventions, but model-informed recommendations.

5. Conclusion

This research aims to create an explainable machine-learning model capable of predicting customer attrition and effectively using predictive analytics to guide retention efforts. This analysis showed that Logistic Regression, Decision Tree, Random Forest and XGBoost models have different levels of performance, with XGBoost being the best overall performer, especially in terms of ROC-AUC, precision–recall AUC, recall, F1-score, balanced accuracy and Matthews correlation coefficient. The results validated the effectiveness of ensemble learning for detecting high-risk customers. The explainability analysis also indicated that the month-to-month contract, short-term contract, high monthly charge, internet provider fibre-optic, electronic-check payment, online security lack, and technical support lack were the highest associated factors with churn. SHAP-based explanation helped to clarify the model by revealing both global churn patterns and customer-specific risk factors. It enabled the churn probabilities to be linked with real-life retention measures and not just viewed as standalone scores. Prioritising intervention through the segmentation of customers into low, moderate, high and critical-risk groups proved helpful. By integrating churn probability with the value of the customer, a more efficient allocation of resources could be achieved, especially toward the high-risk and high-value customers. In general, the findings of the study highlight the potential of explainable machine learning for effective, transparent, and insightful churn management. The retention actions proposed should be validated in future with the use of a longitudinal dataset, uplift modelling and/or controlled intervention experiment.

References

1. Ahn, J., Hwang, J., Kim, D., Choi, H., & Kang, S. (2020). A survey on churn analysis in various business domains. *IEEE access*, *8*, 220816-220839.
2. Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information fusion*, *58*, 82-115.
3. Barsotti, A., Gianini, G., Mio, C., Lin, J., Babbar, H., Singh, A., ... & Damiani, E. (2024). A decade of churn prediction techniques in the telco domain: a survey. *SN Computer Science*, *5*(4), 404.
4. Belle, V., & Papantonis, I. (2021). Principles and practice of explainable machine learning. *Frontiers in big Data*, *4*, 688969.
5. Borrego-Díaz, J., & Galán-Páez, J. (2022). Explainable Artificial Intelligence in Data Science: J. Borrego-Díaz, J. Galán-Páez. *Minds and Machines*, *32*(3), 485-531.
6. Chang, V., Hall, K., Xu, Q. A., Amao, F. O., Ganatra, M. A., & Benson, V. (2024). Prediction of customer churn behavior in the telecommunication industry using machine learning models. *Algorithms*, *17*(6), 231.
7. Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC genomics*, *21*(1), 6.
8. De Bock, K. W., & De Caigny, A. (2021). Spline-rule ensemble classifiers with structured sparsity regularization for interpretable customer churn modeling. *Decision Support Systems*, *150*, 113523.
9. De Caigny, A., Coussement, K., Verbeke, W., Idbenjra, K., & Phan, M. (2021). Uplift modeling and its implications for B2B customer churn prediction: A segmentation-based modeling approach. *Industrial Marketing Management*, *99*, 28-39.
10. Domingos, E., Ojeme, B., & Daramola, O. (2021). Experimental analysis of hyperparameters for deep learning-based churn prediction in the banking sector. *Computation*, *9*(3), 34.
11. Emmert-Streib, F., Yli-Harja, O., & Dehmer, M. (2020). Explainable artificial intelligence and machine learning: A reality rooted perspective. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(6), e1368.
12. Geiler, L., Affeldt, S., & Nadif, M. (2022). An effective strategy for churn prediction and customer profiling. *Data & Knowledge Engineering*, *142*, 102100.

13. Hedström, A., Weber, L., Krakowczyk, D., Bareeva, D., Motzkus, F., Samek, W., ... & Höhne, M. M. C. (2023). Quantus: An explainable ai toolkit for responsible evaluation of neural network explanations and beyond. *Journal of Machine Learning Research*, *24*(34), 1-11.
14. Islam, M. R., Ahmed, M. U., Barua, S., & Begum, S. (2022). A systematic review of explainable artificial intelligence in terms of different application domains and tasks. *Applied Sciences*, *12*(3), 1353.
15. Islam, S. R., Eberle, W., Ghafoor, S. K., & Ahmed, M. (2021). Explainable artificial intelligence approaches: A survey. *arXiv preprint arXiv:2101.09429*.
16. Jiang, P., Liu, Z., Abedin, M. Z., Wang, J., Yang, W., & Dong, Q. (2024). Profit-driven weighted classifier with interpretable ability for customer churn prediction. *Omega*, *125*, 103034.
17. Khattak, A., Mehak, Z., Ahmad, H., Asghar, M. U., Asghar, M. Z., & Khan, A. (2023). Customer churn prediction using composite deep learning technique. *Scientific Reports*, *13*(1), 17294.
18. Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2020). Explainable ai: A review of machine learning interpretability methods. *Entropy*, *23*(1), 18.
19. Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., ... & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature machine intelligence*, *2*(1), 56-67.
20. Maldonado, S., López, J., & Vairetti, C. (2020). Profit-based churn prediction based on minimax probability machines. *European Journal of Operational Research*, *284*(1), 273-284.
21. Marín Díaz, G., Galán, J. J., & Carrasco, R. A. (2022). XAI for churn prediction in B2B models: A use case in an enterprise software company. *Mathematics*, *10*(20), 3896.
22. Mena, G., Coussement, K., De Bock, K. W., De Caigny, A., & Lessmann, S. (2024). Exploiting time-varying RFM measures for customer churn prediction with deep neural networks. *Annals of Operations Research*, *339*(1), 765-787.
23. Minh, D., Wang, H. X., Li, Y. F., & Nguyen, T. N. (2022). Explainable artificial intelligence: a comprehensive review. *Artificial Intelligence Review*, *55*(5), 3503-3568.
24. Nauta, M., Trienes, J., Pathak, S., Nguyen, E., Peters, M., Schmitt, Y., ... & Seifert, C. (2023). From anecdotal evidence to quantitative evaluation methods: A systematic review on evaluating explainable ai. *ACM Computing Surveys*, *55*(13s), 1-42.
25. Ouf, S., Mahmoud, K. T., & Abdel-Fattah, M. A. (2024). A proposed hybrid framework to improve the accuracy of customer churn prediction in telecom industry. *Journal of Big Data*, *11*(1), 70.
26. TanKY. (2020). *Telco customer churn: IBM dataset* [Data set]. Kaggle. https://www.kaggle.com/datasets/yeanzc/telco-customer-churn-ibm-datase