International Journal of Commerce and Management Research
www.managejournal.com
ISSN: 2455-1627
Received: 25-08-2026, Accepted: 24-09-2026, Published: 09-10-2026
Volume 12, Issue 4, 2026, Page No. 128-130

An explainable AI-based customer churn prediction and retention decision support system

S Inbaroja[1], Dr. Germanus Alex[2]
[1] Research Scholar, St. Xavier's College Autonomous, Palayamkottai, Tamil Nadu, India
[2] Research Guide, Government Arts and Science College, Nagercoil, Tamil Nadu, India

Abstract
Customer churn is a very big challenge for organizations because the loss of already available customers can negatively affect revenue, customer lifetime value, and long-term business growth. The increasing the already available customer demographic, transactional, behavioral, and service-related data has created opportunities for machine learning techniques to identify customers who are dislike a product or service. However, conventional machine learning models often focus on predictive performance and may provide limited insight into the reasons behind individual predictions. This lack of interpretability can reduce user trust and make it difficult for organizations to convert churn predictions into effective retention strategies. This paper proposes an Explainable Artificial Intelligence (XAI)-based customer churn prediction and retention decision support framework that combines machine learning prediction, model explainability, customer risk profiling, and retention recommendations.
The proposed approach evaluates suitable machine learning algorithms for predicting customer churn and employs SHapley Additive exPlanations (SHAP) to identify the contribution of individual customer features to model predictions. The system provides both global explanations, which identify the most influential factors across the customer population, and local explanations, which describe the factors contributing to an individual customer's churn risk. Based on predicted churn probability and influential characteristics, customers can be categorized according to their level of risk and appropriate retention strategies can be considered. The proposed framework aims to bridge the gap between predictive analytics and practical customer relationship management by transforming model predictions into interpretable and actionable decision support. The approach can assist organizations in identifying high-risk customers, understanding major churn drivers, prioritizing retention efforts, and improving data-driven customer management.

Keywords: Customer churn prediction, explainable artificial intelligence, machine learning, shap, customer retention, decision support system, predictive analytics, customer relationship management

Introduction

Customer retention is an important concern for organizations operating in highly competitive markets. Customers may discontinue a service because of dissatisfaction, pricing issues, poor service quality, inadequate customer support, changes in personal requirements, or the availability of alternative products and services. Customer churn can result in revenue loss and can increase the resources required to acquire new customers. Consequently, organizations increasingly seek methods for identifying customers who are likely to churn before the actual churn event occurs.
Traditional churn-management approaches generally rely on historical reports, manually designed rules, customer surveys, and reactive customer-service activities. Although these methods can provide useful information, they may be unable to capture complex relationships among the large number of variables associated with customer behavior. Machine learning provides an opportunity to overcome this limitation by learning patterns from historical customer data and generating predictions for customers who may be at risk of leaving.
Various machine learning algorithms have been applied to customer churn prediction, including Logistic Regression, Decision Trees, Random Forest, Support Vector Machines, Gradient Boosting, and XGBoost. These methods can identify nonlinear relationships and interactions between customer characteristics and churn outcomes. However, predictive accuracy alone does not necessarily make a churn model useful for business decision-making. Complex models can operate as black boxes, producing a churn probability without explaining the factors responsible for the prediction.
Interpretability is particularly important in customer retention applications. For example, if a model predicts that a customer has a high probability of churn, a retention manager needs to understand whether the prediction is associated with factors such as contract type, customer tenure, monthly charges, payment behavior, service usage, or customer-support interactions. Such information can help the organization select a more appropriate retention strategy.
Explainable Artificial Intelligence has emerged as an important research area for addressing the lack of transparency in machine learning systems. XAI techniques aim to provide understandable information about how a model produces its predictions. SHAP is one of the prominent explanation methods and is based on Shapley values from cooperative game theory. Lundberg and Lee introduced a unified framework for interpreting model predictions using additive feature attribution.
Another widely known method is Local Interpretable Model-Agnostic Explanations (LIME), which explains individual predictions by constructing an interpretable approximation around the prediction of interest. Ribeiro *et al*. demonstrated the use of LIME for explaining predictions generated by different machine learning classifiers.
Recent research has increasingly combined machine learning-based churn prediction with explainability. Studies

have investigated SHAP-based feature analysis to identify important churn drivers and improve the transparency of predictive models. Recent work has also explored the integration of explainability with customer segmentation and retention-oriented decision-making.

Despite these developments, there remains a need for frameworks that connect prediction, explanation, and retention decision-making. A conventional churn model may identify a customer as high risk but provide no explanation, while an explainability system may identify influential features without translating them into practical retention support. Therefore, this research focuses on developing an Explainable AI-based customer churn prediction and retention decision support framework that integrates these components into a unified process.

The proposed approach aims to predict customer churn, identify the factors responsible for the prediction, classify customers according to risk, and provide retention-oriented decision support. The framework is intended to improve the transparency and practical usefulness of machine learning-based customer churn prediction while supporting more targeted customer retention strategies.

Customer Churn Prediction Using Machine Learning

Customer churn prediction can be formulated as a supervised classification problem in which historical customer characteristics are used to predict whether a customer will remain with or leave an organization. The input data can contain demographic information, transaction history, customer tenure, contract characteristics, service usage, payment information, and other behavioral attributes. The target variable generally represents whether the customer has churned.

Machine learning algorithms can identify relationships between these characteristics and historical churn outcomes. Logistic Regression provides an interpretable statistical baseline, while Decision Trees can represent predictions using hierarchical decision rules. Random Forest uses an ensemble of decision trees to improve predictive robustness, whereas Gradient Boosting and XGBoost can capture more complex nonlinear patterns.

The effectiveness of a churn prediction model should not be evaluated using accuracy alone. Precision, Recall, F1-score, ROC-AUC, and PR-AUC provide additional information about the model's ability to identify customers who are genuinely at risk. Recall is particularly relevant when failing to identify a potential churner can result in a missed retention opportunity.

Explainable Artificial Intelligence for Churn Prediction

Machine learning models can produce highly accurate predictions while remaining difficult for non-technical users to interpret. Explainable Artificial Intelligence addresses this issue by providing information about the relationship between input features and model predictions.

Explainability can be considered at two levels. Global explanations describe the overall behavior of a model and identify features that are important across the complete dataset. Local explanations focus on a single customer and identify the factors responsible for that particular prediction.

In customer churn prediction, global explanations can help organizations understand general churn patterns, while local explanations can help retention personnel understand why a particular customer has been assigned a high churn probability. This distinction is important because the factors influencing churn may differ between individual customers.

SHAP-Based Feature Explanation

SHAP, or SHapley Additive exPlanations, provides a framework for assigning contribution values to individual features. The method is based on the Shapley value concept from cooperative game theory and estimates how each feature contributes to a prediction relative to a baseline.

For a given prediction, the SHAP representation can be expressed conceptually as:

$$f(x) = \varphi_0 + \Sigma\varphi_i$$

where $f(x)$ represents the model prediction, φ_0 is the baseline value, and φ_i represents the contribution of feature i.

A positive SHAP value generally indicates that a feature increases the predicted probability of the target class, while a negative value indicates that it moves the prediction toward the opposite class. This allows the proposed system to identify both the most important churn factors at the population level and the specific factors influencing an individual customer's prediction.

The original SHAP research proposed a unified framework for interpreting model predictions and established important theoretical properties for additive feature attribution.

Customer Risk Profiling and Retention Decision Support

Prediction alone does not necessarily provide sufficient information for effective customer retention. Therefore, the proposed framework converts predicted churn probabilities into customer risk categories. Customers with relatively low predicted risk can receive routine engagement, whereas customers with moderate or high risk can be prioritized for preventive or targeted retention activities.

The risk classification can be combined with SHAP explanations to support more personalized decision-making. For example, if service-related factors contribute strongly to a customer's predicted churn risk, proactive service support may be considered. If contract-related characteristics contribute significantly, an appropriate plan review may be considered. Similarly, low customer engagement may motivate targeted communication or engagement activities.

The purpose of the decision-support component is not to claim that a particular intervention will definitely prevent churn. Instead, it provides a structured basis for human decision-makers to select and prioritize possible retention actions.

The overall decision process can therefore be represented as:

Customer Data → Churn Prediction → Explainability → Risk Profiling → Retention Decision

Proposed Research Framework

The proposed framework integrates data preprocessing, machine learning, explainable AI, and retention decision support. Initially, customer data are cleaned and transformed into a suitable format for model training. Categorical variables are encoded, numerical variables are processed as required, and class imbalance is addressed where necessary.

Multiple machine learning models can then be trained and evaluated. The best-performing model is selected based on

appropriate predictive metrics. SHAP is subsequently applied to the selected model to generate global and local explanations.

The churn probability and explanation information are then passed to the risk-profiling component. Customers are assigned risk categories based on validated thresholds, and retention-oriented recommendations are generated according to the customer's risk level and influential characteristics.

This framework provides an integrated pathway from prediction to explanation and decision support rather than treating these functions as independent components.

Advantages of the Proposed Approach

The proposed approach provides several advantages over prediction-only churn systems. It enables early identification of customers who may be at risk, supports comparison of different machine learning algorithms, provides interpretable explanations for predictions, identifies important churn-related factors, and supports risk-based prioritization of retention resources.

The integration of explainability can also improve communication between technical and business users. Data scientists can use SHAP outputs to analyze model behavior, while business users can use customer-level explanations to understand the reasons associated with individual churn predictions.

However, explainability should not be interpreted as causal evidence. A feature that contributes strongly to a prediction is not necessarily a factor that can be changed to prevent churn. Therefore, retention recommendations should be validated through appropriate business experiments.

Future Work

Future research can extend the proposed framework by incorporating temporal customer behavior. Rather than treating customer information as a static snapshot, future models can analyze sequences of transactions, service interactions, complaints, payments, and engagement activities to identify changes in churn risk over time.

Customer lifetime value can also be integrated with churn probability to improve retention resource allocation. A customer with a high probability of churn may not necessarily require the same level of intervention as another customer who has both high churn probability and high lifetime value.

Another important direction is the integration of causal inference and uplift modeling. Conventional churn prediction answers the question of which customers are likely to churn, whereas uplift modeling can help determine which customers are most likely to respond positively to a particular retention intervention. This could allow the system to recommend not only which customers require attention but also which type of intervention may have the greatest expected benefit.

Future studies can also incorporate real-time customer data, automated model monitoring, concept-drift detection, and periodic retraining. These capabilities would enable the system to adapt to changing customer behavior and market conditions.

Human-centered evaluation is another important research direction. Future studies should investigate whether business users understand SHAP explanations correctly, whether explanations improve decision quality, and whether users develop an appropriate level of trust in the predictions.

Finally, real-world experimental evaluation should be performed to determine whether the recommendations generated by the system actually improve customer retention and profitability. Controlled retention campaigns or A/B testing can provide stronger evidence than predictive performance alone.

Conclusion

This paper presents an Explainable AI-Based Customer Churn Prediction and Retention Decision Support System that integrates machine learning, SHAP-based explainability, customer risk profiling, and retention-oriented decision support. The proposed approach addresses an important limitation of conventional churn prediction systems, which often focus on predicting whether a customer will leave without sufficiently explaining the reasons behind the prediction.

By incorporating explainability, the proposed framework can identify the factors that contribute to both overall churn patterns and individual customer predictions. The integration of these explanations with risk profiling provides a basis for more targeted and informed retention decision-making.

The proposed framework has potential applications in telecommunications, banking, insurance, retail, e-commerce, subscription services, and other industries where customer retention is strategically important. Although predictive performance remains an important consideration, the practical value of the system depends on the ability to connect predictions with understandable explanations and effective retention actions.

Future integration of customer lifetime value, causal inference, uplift modeling, temporal behavior, and real-world intervention evaluation can further improve the framework. Overall, the proposed approach demonstrates how explainable machine learning can help bridge the gap between customer churn prediction and practical retention decision support.

References

1. SM Lundberg, SI Lee. A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems, 2017, 30.
2. MT Ribeiro, S Singh, C Guestrin. “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,2016:1135–1144.
3. A El Attar, M El-Hajj. Explainable AI-Driven Customer Churn Prediction: A Multi-Model Ensemble Approach with SHAP-Based Feature Analysis. Frontiers in Artificial Intelligence, 2026.
4. A Data-Driven Approach with Explainable Artificial Intelligence for Customer Churn Prediction in the Telecommunications Industry. Results in Engineering,2025,26:104629.
5. Özkurt C. Transparency in Decision-Making: The Role of Explainable AI (XAI) in Customer Churn Analysis. Information Technology in Economics and Business, 2025, 2(1).
6. Add the authoritative citation for the exact customer churn dataset used in the study.
7. Add the original reference for the class-imbalance technique used in the experiment, if applicable.
8. Add additional recent peer-reviewed studies specifically related to customer churn prediction, explainable AI, and retention decision support.