Predictive Retention: An Explainable AI Framework for Subscription Forecasting and Proactive Churn Prevention

Shanmukh Inampudi
Dept. of Computer Science and Engineering
Vignan's Foundation for Science, Technology and Research
Vadlamudi, India
Email: 221fa04511@gmail.com

Mohamed
Dept. of Computer Science and Engineering
Vignan's Foundation for Science, Technology and Research
Vadlamudi, India
Email: 221fa04008@gmail.com

***Abstract*—Customer churn in the telecom sector is a critical issue for companies, making it essential to understand customer behavior and make accurate predictions. In this study, various machine learning and deep learning models were utilized to predict churn [21]. Additionally, the developed model was analyzed using Explainable Artificial Intelligence (XAI) techniques. SHAP was employed to examine the global and local impacts of features on the model, while LIME was used to explain the reasons behind individual customer decisions [21]. The results indicate that factors such as contract type, technical support, online security, and payment method play a decisive role in churn [21]. A key contribution is bridging the actionability gap by mapping these XAI insights to automated retention strategies.**

***Index Terms*—Customer Churn, Machine Learning, Explainable AI (XAI), SHAP, LIME, Prescriptive Analytics.**

I. Introduction

The telecommunications industry is an area where maintaining customer loyalty is critical within highly competitive dynamics [21]. Customer churn reduces revenues and can negatively impact financial stability. Studies show that reducing customer churn by 5% can increase the profitability of companies by 25% to 85% [21]. While models initially developed for churn prediction were limited to statistical methods, modern models utilize Random Forest, XGBoost, and Deep Learning to handle complex dependencies [21].

However, a significant gap identified in current research is that most studies focus on improving accuracy while paying less attention to making results understandable and actionable for decision makers [21]. This project addresses this "Actionability Gap" by integrating XAI techniques and a prescriptive recommendation engine to turn model outputs into concrete strategies.

II. Literature Review / Context

The evolution of churn prediction has shifted from traditional classification to high-accuracy ensemble and deep learning models [1]. Recent research has emphasized the necessity of handling the inherent class imbalance in telecom data using techniques like SMOTE and SMOTEENN to improve model sensitivity [1], [12]. XAI has become a primary focus to address the "black-box" nature of these models, with SHAP and LIME providing global and local interpretability respectively [1], [21].

III. Related Work

Existing literature covers a diverse range of technical optimizations. Research has utilized SHAP-guided Bayesian optimization to tune XGBoost models [2] and fuzzy logic for broadband churn [8]. Technical innovations include Graph Neural Networks (GNN) to capture customer interdependencies [10] and time-series approaches for B2B forecasting [16]. Comparative SHAP analyses have highlighted how hyperparameter tuning shifts feature importance [4]. However, as identified in our base paper, most existing systems fail to translate these explanations into automated business actions [21].

IV. Methodology

The proposed methodology follows a structured pipeline: data cleaning, class balancing with SMOTEENN, and an integrated XAI-Prescriptive pipeline.

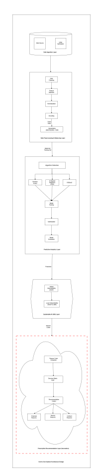

Fig. 1. End-to-End System Architecture: From Data Ingestion to Prescriptive Recommendation.

A. Dataset and Preprocessing

The study utilizes the IBM Telco Customer Churn dataset containing 7,043 records with 21 features [21]. Preprocessing involves removing the "TotalCharges" feature to prevent multicollinearity and applying Min-Max scaling to tenure and monthly charges [21].

B. Class Balancing with SMOTEENN

To handle the class imbalance (26.6% churn), the SMOTEENN method is implemented [21]. This combines Synthetic Minority Over-sampling (SMOTE) and Edited Nearest Neighbors (ENN) to clean noisy majority samples while increasing minority representation [21].

C. AI Models and XAI Integration

Following benchmarks, we implement Random Forest, KNN, and XGBoost models [21]. We integrate SHAP for global feature ranking and LIME for local explanations to ensure transparency [21].

D. Novelty: The Prescriptive Engine

We bridge the Actionability Gap by creating a mapping logic between XAI reason codes and business actions.

- **Price-Driven Churn**: Triggers a "Price-Matched Discount."
- **Service Gaps**: Triggers "Priority Tech-Support Vouchers."
- **Contract Risk**: Triggers "Loyalty Program Enrollment."

V. Results and Discussion

The performance of the optimized XGBoost model was evaluated using standard metrics. The model achieved a high accuracy of 96% and an Area Under the ROC Curve (AUC) of 0.9922.

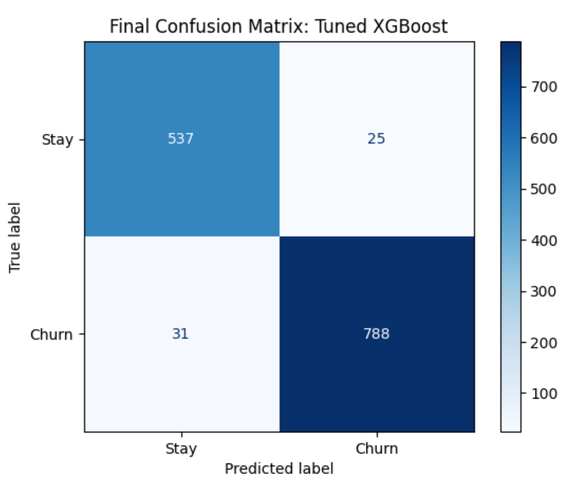

Fig. 2. Confusion Matrix of Tuned XGBoost Model.

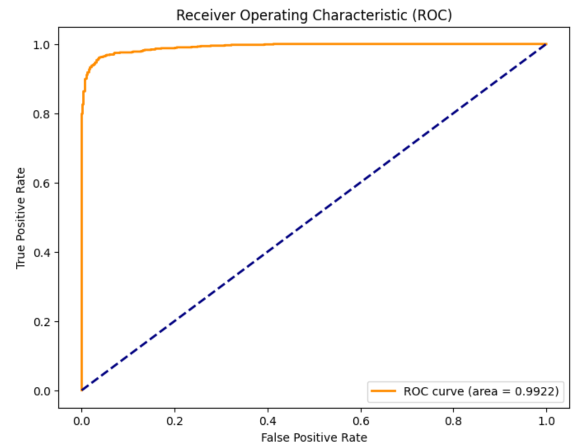

Fig. 3. ROC Curve demonstrating AUC of 0.9922.

A. Performance Metrics

The confusion matrix (Fig. 2) revealed that out of the test set, the model correctly identified 788 churners with only 31 false negatives. This high recall is essential for retention strategies. The ROC curve (Fig. 3) further validates the model's discriminatory power.

B. Explainability Insights

SHAP global analysis identified that "Contract Type" and "Monthly Charges" are the strongest predictors of churn. LIME local analysis revealed that lack of "Tech Support" often acted as secondary churn triggers.

C. Prescriptive Outcomes

By applying the prescriptive engine, the system automatically categorized the 788 identified churners into specific retention groups, effectively closing the actionability gap.

VI. Conclusion and Future Work

This study successfully implemented an explainable and prescriptive framework for telecom churn prediction. By achieving an AUC of 0.9922 and integrating SHAP-based recommendation logic, we moved beyond diagnostic AI into prescriptive analytics. This approach directly addresses the actionability gap identified in recent 2025 research.

References

[1] I. Duru, N. Ak, and R. Dede, "A Comparative Study of Feature Selection and Resampling Techniques for Customer Churn Prediction with Explainable AI in the Telecommunications Sector," in *Proc. ACDSA*, 2025.

[2] L. An, "A Novel SHAP-Guided Bayesian Optimization Method for Telecom Customer Churn Prediction and Retention," in *Proc. EESPE*, 2025.

[3] E. A. A. Alaoui et al., "Balancing Speed and Insight: Computational Efficiency of XAI Methods for Telecommunication Customer Churn Analysis," in *Proc. ICCSC*, 2025.

[4] I. Boukrouh, F. Tayalati, and A. Azmani, "Comparative SHAP Analysis on SVM and K-NN: Impacts of Hyperparameter Tuning on Model Explainability," in *Proc. IICAIET*, 2024.

[5] M. H. Ali et al., "Customer Churn Prediction in Digital Banking Platforms Using Hybrid Ensemble Learning with Multi-Dimensional Behavioral Analytics," in *Proc. ICCR*, 2025.

[6] S. R. Dontireddy et al., “Enhancing Transparency: AI Applications for Detecting Cheating and Predicting Player Attrition in Online Gaming,” in *Proc. IConSCEPT*, 2024.
[7] D. Y. C. Wang et al., “Explainability of Leverage Points Exploration for Customer Churn Prediction,” in *Proc. Big Data*, 2023.
[8] L. Ciabattoni et al., “Explainable AI Based Approach For Broadband Customers’ Churn Prediction,” in *Proc. GEM*, 2024.
[9] E. Alduweib and M. A. Arqoub, “Explainable Artificial Intelligence for Business Intelligence,” in *Proc. ICCIAA*, 2025.
[10] Neeraj M et al., “Explainable Multi-Stage Churn Prediction using Graph Neural Network in Telecom Sector,” in *Proc. ICOEI*, 2025.
[11] K. K. Ratra and D. Verma, “From Clicks to Community: Leveraging Social Computing, Data Mining, and Generative AI for Customer Retention and Ecosystem Growth in E-Commerce,” in *Proc. ACROSET*, 2025.
[12] M. J. Nur et al., “Improving Customer Churn Prediction Using Domain-Driven Feature Engineering, Resampling, and CatBoost with Explainability Extensions,” in *Proc. iSemantic*, 2025.
[13] S. T. Eswarapu et al., “Integrated Customer Analytics using Explainability and AutoML for Telecommunications,” in *Proc. ICAAIC*, 2023.
[14] S. Murindanyi et al., “Interpretable Machine Learning for Predicting Customer Churn in Retail Banking,” in *Proc. ICOEI*, 2023.
[15] P. Ramesh et al., “Machine Learning Strategies for Customer Churn Prediction in Competitive Enterprises,” in *Proc. ICSCNA*, 2024.
[16] J. Ahlstrand et al., “Predicting B2B Customer Churn using a Time Series Approach,” in *Proc. IDSTA*, 2024.
[17] A. Raut, “Predicting Bank Customer Churn with Interpretable Machine Learning using SHAP,” in *Proc. ICESC*, 2025.
[18] L. Jovanovic et al., “Predicting Credit Card Churn: Application of XGBoost Tuned by Modified Sine Cosine Algorithm,” in *Proc. ICSMDI*, 2023.
[19] P. D. Joshi and R. Shende, “Predicting Customer Churn In Telecom Industry Using Machine Learning Algorithms,” in *Proc. NMITCON*, 2025.
[20] J. Lazrak, “Seasonal Segmentation and Life Moment Prediction in Financial Services: A K-means and SVM-Based Approach,” in *Proc. ICOA*, 2025.
[21] M. Güllü, R. Dede, and İ. Duru, “Telecom Customer Churn Prediction: In-Depth Analysis with Artificial Intelligence Algorithms and Explainable Artificial Intelligence,” in *Proc. ICECCE*, Istanbul, Turkey, 2025.