International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 4, July-Aug-2026, ISSN (Online): 2395-566X

Customer Churn Prediction and Explainable Analysis in Telecom: An XGBoost and SHAP-based Framework with Interactive Tableau Visualization

Komal[1], Balvir Singh Thakur[1]
[1]Department of Computer Science and Engineering, University Institute of Technology, Himachal Pradesh University, Shimla, India

Abstract — A major challenge for the telecommunication industry is the frequent customer switching, resulting in substantial revenue loss and increased customer acquisition costs. Most of the existing studies on churn prediction focus on the predictive accuracy and pay less attention to the interpretability of the model and how to convert the model output to business-facing visual tools. In this study, an integrated framework is proposed with XGBoost for churn classification, SHAP (Shapley Additive explanations) for global and local interpretability and Tableau for the interactive dashboard visualization. The Telco Customer Churn dataset with 7,043 customer records and 21 features is publicly available and is used for experimentation. The accuracy of the XGBoost model on a held-out test set of 1,409 customers is 80.41%, with better results on the majority (retained) class than on the minority (churned) class, which is consistent with the moderate imbalance between classes in the dataset. This outcome will be contextualized through baseline comparison against Logistic Regression and Random Forest. SHAP analysis reveals that contract type, tenure, monthly charges, total charges and internet service type are the dominant churn drivers globally, and local SHAP explanations illustrate how these factors combine for individual customers. In addition to the SHAP-ranked churn drivers, an interactive Tableau dashboard was created to visualize churn distribution by contract type, payment method, and internet service type for business-facing exploration of these insights. The study concludes that the combination of predictive accuracy with explainability and visualization leads to a more actionable churn-management framework than accuracy-oriented approaches alone.

Keywords— Customer churn prediction; XGBoost; Explainable artificial intelligence; SHAP; Tableau; Telecommunications; Machine learning.

I. INTRODUCTION

Customer retention is a critical issue in the telecommunication industry due to intense market competition and availability of multiple service providers. Customers often switch from one provider to another in search of better pricing, better service quality or better offers. This phenomenon is called ‘customer churn’, and it is directly impacting the profitability of telecom companies, as it is much more expensive to acquire new customers than to keep the existing ones [1].

With the exponential growth of digital data, telecom companies now collect vast amounts of customer data such as usage patterns, billing details, contract type, and service preferences. The data provides an opportunity to use machine-learning techniques to identify possible churners. When churn can be predicted early enough, companies receive a real window to act — offering a better deal, addressing a complaint, or simply reaching out before the customer decides to leave for good [2], [3], [4].

Among various machine-learning methods, gradient-boosting algorithms such as XGBoost have shown good performance in classification problems due to their high predictive accuracy and efficiency in handling structured, tabular data [4]. However, predictive accuracy alone is often not sufficient for real-world business applications, as decision-makers also need to understand the rationale behind a prediction before they are able to act on it.

To overcome this limitation, explainable artificial intelligence (XAI) methods such as SHAP have been proposed for churn analysis [1], [5]. SHAP quantifies the contribution of each feature to the output of a model. We use it to point out the overall drivers of churn (global explanation) and to explain a single customer’s churn (local explanation). Additionally, interactive visualization using tools like Tableau can help make these insights available to business stakeholders who may not be directly involved with the underlying model.

Although many studies have individually investigated XGBoost and SHAP for telecom churn prediction [1], [4]–[7], few studies have presented a single end-to-end framework that integrates forecasting, dual-level (global and local) explainability, and interactive dashboard-based visualization. The present study fills this gap. The rest of this paper is organized as follows. Section 2 covers methodology, including dataset description, preprocessing, model design and evaluation criteria. In Section 3 the experimental results obtained so far are given. Section 4 presents findings, limitations and future directions. The paper concludes with Section 5.

II. METHODOLOGY

This study uses a comparative, experimental design to evaluate the predictive performance of XGBoost compared to classical baseline classifiers and to evaluate the interpretability and visual usability of the resulting predictions.

1. Research Questions

The following research questions guide the experimental design:

- RQ1: Does XGBoost perform better than classical baseline classifiers (Logistic Regression and Random Forest) in predicting customer churn in the telecom industry?
- RQ2: According to SHAP, which customer attributes are important to churn at global (population) and local (individual customer) level?
- RQ3: How to convert model predictions and SHAP-based explanations into an interactive decision-support dashboard for non-technical stakeholders?

2. Dataset Description

The Telco Customer Churn dataset (IBM Sample Data Set) used in this study is a publicly available dataset of 7043 customer records with 21 attributes that comprise demographic information, account information and service subscriptions. The target variable Churn tells us whether the customer left the service and is converted from original Yes/No labels to a binary 1/0 value. The target variable analysis reveals a class distribution of 73.46% retained customers and 26.54% churned customers, indicating a moderate class imbalance that is retained in the train-test split. The main features of the dataset are summarized in Table 1.

Table 1. Summary of the Telco Customer Churn dataset

Attribute	Value	Attribute	Value
Total records	7,043	Total features	21
Churn rate	26.54%	Retained rate	73.46%
Categorical features	17	Numerical features	4
Target variable	Churn (Yes/No)	Source	IBM Sample Telco Dataset
Training set size	5,634 (80%)	Test set size	1,409 (20%)

3. Data Preprocessing

The preprocessing of data is performed in Python with the help of Pandas library. The Churn column is mapped from binary label 1/0 to Yes/No. The TotalCharges field is an object/string type and is converted to a numeric type and invalid or blank entries are coerced to missing values. There are 21 attributes and 12 predictor features are selected for modeling, they are tenure, MonthlyCharges, TotalCharges, SeniorCitizen, Contract, InternetService, PaymentMethod, OnlineSecurity, TechSupport, PaperlessBilling, StreamingTV and StreamingMovies. Among these, categorical features are converted to numerical form using one-hot encoding (pandas get_dummies with the first category dropped to prevent multicollinearity). The generated feature matrix is split into train and test partitions at a ratio of 80:20 and a fixed random seed (random_state = 42). This results in a train set of 5,634 records and a test set of 1,409 records, while preserving the original class ratio in both partitions.

4. Exploratory Data Analysis

Exploratory Data Analysis (EDA) with Pandas, Matplotlib and Seaborn. The analysis includes: bar chart of churn count between retained and churned customers; stacked histogram of tenure distribution by churn status; boxplot of monthly charges by churn status; count plots of churn by payment method, internet service type, contract type, senior-citizen status, online security and tech-support subscription; and Pearson correlation heatmap between tenure, monthly charges, total charges, senior-citizen status and churn. These visualizations are useful to identify the variables that visually differentiate the churned customers from the retained customers prior to model training and to inform the feature set used for forecasting.

5. Predictive Modeling

We plan to apply three classification models on the same one-hot-encoded feature set as described in Section 2.3, namely Logistic Regression as interpretable linear baseline, Random Forest as nonlinear ensemble baseline, and XGBoost as the primary predictive model. The current phase implements the classifier using the XGboost Python library with 200 estimators, a learning rate of 0.1, a maximum depth of 6 and a fixed random seed (random_state = 42), trained on the training partition containing 5,634 records and evaluated on the held-out test set composed of 1,409 records. The Logistic Regression and Random Forest baselines will be implemented to close the comparison and directly answer RQ1.

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 4, July-Aug-2026, ISSN (Online): 2395-566X

6. Explainability using SHAP
SHAP (SHapley Additive EXPlanations) is applied on the trained XGBoost model using shap.TreeExplainer to quantify the contribution of each feature to individual predictions based on cooperative game-theoretic Shapley values [5]. A global SHAP bar plot and summary (beeswarm) plot rank features by their overall mean absolute contribution to the output of the model. A local SHAP waterfall plot is used to explain an individual prediction at the customer level. This two-level explanation enables identifying both population-wide churn drivers and customer-specific reasoning.

7. Visualization using Tableau
SHAP derived global feature importances and model outputs were exported from the notebook environment and visualized in an interactive Tableau dashboard. The dashboard shows the churn distribution by contract type, payment method, and internet service type, as well as a ranked bar chart of the top churn drivers based on SHAP values. This provides business stakeholders the ability to explore the key patterns behind churn without requiring them to interact with the underlying model directly.

8. Evaluation Criteria
Model performance is measured using classification accuracy, precision, recall, F1-score and a confusion matrix. AUC-ROC is planned as an additional metric once all three models have been compared. Together, these metrics capture overall correctness and the trade-off between false positives and false negatives, which is especially relevant in the case of churn prediction, where a false negative, i.e. failing to identify a churning customer, entails a direct business cost.

III. RESULTS

In this section we present the results obtained so far. The completed experiments constitute the basis of the XGBoost classification results, confusion matrix, the findings of SHAP-based feature importance (Sections 3.1–3.2), and the Tableau dashboard (Section 3.4). As a follow-up experiment, we plan to include comparative baselines against Logistic Regression and Random Forest, which are not yet included.

1. Classification Performance
Table 2 presents the classification performance of the XGBoost model on the held-out test set. The model achieved an overall accuracy of 80.41% on the test set of 1,409 records. We plan to conduct a direct response to RQ1 by means of a follow-up experiment against the Logistic Regression and Random Forest baselines, but this is not available at the time of writing.

Table 2. Classification performance of the compared models on the test set

Model	Acc.	Prec.	Rec.	F1	AUC
XGBoost (Proposed)	80.41 %	80% (wtd.)	80% (wtd.)	0.80 (wtd.)	[Pending]

Since the dataset is slightly imbalanced (73.46% stayed vs. 26.54% churned), overall accuracy can be misleading. Therefore, Table 3 presents the precision, recall and F1-score of XGBoost for each class. The model performs well on the majority (Stayed) class but relatively poorly on the minority (Churned) class, correctly identifying 203 of 373 actual churners (recall = 0.54) with a precision of 0.66 for churn predictions.

Table 3. XGBoost per-class performance on the test set

Class	Precision	Recall	F1-score
Stayed (0)	0.85	0.90	0.87
Churned (1)	0.66	0.54	0.60

The corresponding confusion matrix is shown in Table 4. Of the 373 customers who churned, 203 were correctly identified (true positives), and 170 were missed (false negatives). Of the 1,036 customers who stayed, 930 were correctly identified (true negatives) and 106 were incorrectly flagged as at risk of churn (false positives).

Table 4. Confusion matrix for the XGBoost model on the test set

Actual \ Predicted	Stayed (0)	Churned (1)
Stayed (0)	930 (TN)	106 (FP)
Churned (1)	170 (FN)	203 (TP)

When the Logistic Regression and Random Forest baselines are completed, this section will also report whether the accuracy, precision, recall and F1-score of the XGBoost are higher than the baseline models, answering RQ1 directly.

2. Global and Local SHAP Feature Importance
Table 5 shows the features with the highest mean absolute SHAP value from the global explanation of the XGBoost model, obtained from the SHAP bar and summary plots generated using Shap.TreeExplainer, and the direction of their impact on churn probability. The most significant predictors of retention were having a two-year contract and longer tenure, whereas higher monthly charges, fiber-optic internet service, and electronic-check payment were linked to higher churn risk

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 4, July-Aug-2026, ISSN (Online): 2395-566X

Table 5. Top features ranked by mean absolute SHAP value (global explanation)

Rank	Feature	Direction of effect (from SHAP)
1	Contract (Two year)	Presence strongly decreases churn probability
2	Tenure	Longer tenure decreases churn probability
3	Monthly Charges	Higher charges increase churn probability
4	Total Charges	Interacts with tenure; lower values associate with higher churn
5	Internet Service (Fiber optic)	Presence increases churn probability
6	Contract (One year)	Presence decreases churn probability
7	Payment Method (Electronic check)	Presence increases churn probability
8	Paperless Billing (Yes)	Presence mildly increases churn probability

The local SHAP waterfall explanation for a representative test customer shows how these global patterns combine for an individual customer: for a customer with only one month of tenure, low total charges, a month-to-month contract, and electronic cheque payment, SHAP assigns a large positive contribution to churn risk from the short tenure (+0.70) and low total charges (+0.66), partially offset by the lack of fiber-optic internet service (−0.49), for a final predicted churn score of f(x) = 0.58. This customer-level breakdown demonstrates how the framework can generate personalized, human-interpretable explanations for a predicted churn risk, as opposed to only providing a population-level explanation.

3. Computational Characteristics

The approximate training time, inference latency and qualitative interpretability of the XGBoost model are shown in Table 6. Training on the 5,634 training set records finishes in seconds on standard hardware, with low inference latency, as expected for XGBoost's general suitability for near-real-time churn-scoring pipelines. Comparable figures for Logistic Regression and Random Forest baselines will be added once those experiments are complete.

Table 6. Training time, inference latency, and interpretability of the compared models

Model	Training time	Inference latency	Interpretability
XGBoost + SHAP	Low (seconds)	Low	Moderate (via SHAP)

4. Dashboard-based Visualization

Results of the churn analysis were presented in a business-facing, non-technical format using an interactive Tableau dashboard. Fig. 1 shows the composite dashboard view, arranged as a 2x2 grid of linked charts.

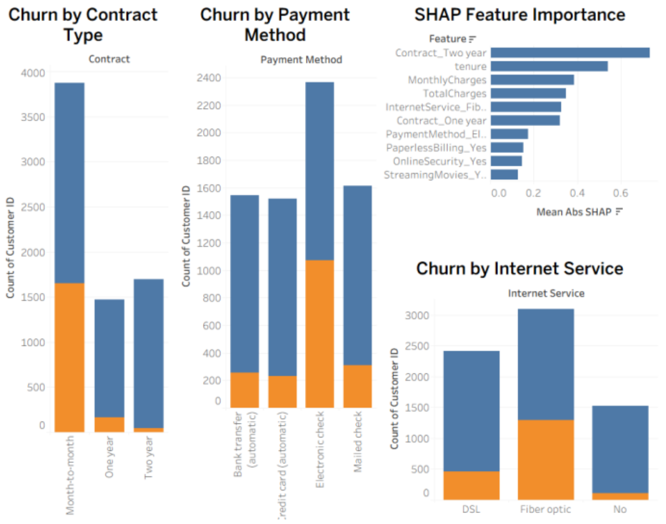

Fig. 1. Tableau dashboard summarizing churn distribution by contract type, payment method, and internet service type, alongside the top SHAP-ranked churn drivers.

The top left panel (Churn by Contract Type) indicates that the vast majority of churn observed is from customers with month-to-month contracts, while customers with one-year and especially two-year contracts churn much less frequently. In the top-centre panel (Churn by Payment Method), it is apparent that customers that have paid through electronic cheque have a much higher proportion of churn than those that have paid through bank transfer, credit card or mailed-cheque payment. The bottom right panel (Churn by Internet Service) shows that customers with fiber-optic internet service churn significantly more than customers with DSL or no internet service. The top-right panel shows the top nine features ranked by SHAP (Contract_Two year, tenure, MonthlyCharges, TotalCharges, InternetService_Fiber optic, Contract_One year, PaymentMethod_Electronic check, PaperlessBilling_Yes, and OnlineSecurity_Yes) in the same order as the global SHAP importance ranking reported in Section 3.2, thus providing a direct visual link between the model-level explanation and the business-facing churn patterns. These four combined views complement SHAP-based results with descriptive and aggregate visualization, and reveal that the same drivers of churn identified by explainable-AI are also visible in simple customer segment analysis.

IV. DISCUSSION

The findings presented so far provide a partial but informative answer to the research questions of the study. For RQ2, the SHAP analysis clearly indicates contract type, tenure, monthly charges, total charges, and internet-service type as the major drivers of churn for the XGBoost model, which is consistent with the results reported in the prior studies on the same dataset [1], [6], [7]. This convergence across independent studies builds confidence in the business interpretation underneath: customers on short tenures, month-to-month contracts, fiber-optic internet, and electronic cheque payment are the highest-risk segment for targeted retention offers.

To give a definite answer to RQ1, we need the Logistic Regression and Random Forest baselines, which are not finished yet. The accuracy of the XGBoost model was 80.41%, which is in line with accuracies reported in the literature for this data set, which typically range from 80% to 96% depending on feature engineering and model tuning [1]–[3]. The class-imbalance-sensitive metrics in Table 3 show that the model as it is now, without tuning, is much better at predicting customers that will stay than customers that will churn (recall of 0.90 versus 0.54, respectively). The asymmetry is a significant limitation that needs to be addressed to make the model business-ready, for instance through class weighting, resampling (e.g., SMOTE) or threshold tuning, because the business cost of missing an actual churner (a false negative) is usually higher than the cost of a false retention alarm.

For RQ3, the local SHAP waterfall explanation confirms that individual, customer-level explanations can be generated in parallel with aggregate predictions, and the interactive Tableau dashboard (Section 3.4) translates these results into a non-technical, business-facing view, corroborating the SHAP-based global drivers with descriptive segment-level charts. This addresses an area noted in prior literature as relatively under explored: the interactive, dashboard-based communication of churn insights [7].

Some limitations of this study are at this stage. It is based on a single, relatively small, static dataset (7,043 records) and therefore limits the extent to which findings may be generalized to other telecom operators or to real-time, streaming data conditions. The XGBoost model has not been systematically tuned for hyper-parameters and class-imbalance handling yet, nor been compared yet against Logistic Regression, Random Forest or other gradient-boosting frameworks such as LightGBM or CatBoost. The interpretability assessment of SHAP is based on established theoretical properties rather than a formal user study with business stakeholders and the Tableau dashboard, while functional, shows aggregate segment-level views and not individual customer-level SHAP explanations.

In future work we will finish the Logistic Regression and Random Forest baselines and the AUC-ROC comparison, address the class imbalance observed in the churn-class recall, expand the Tableau dashboard with customer-level drill-down views, and if time permits compare SHAP with alternative explanation methods such as LIME.

V. CONCLUSION

We propose an integrated framework for telecom customer-churn management where XGBoost is used for prediction, SHAP for global and local explainability and Tableau for interactive visualization. At the current stage, the XGBoost model was trained and evaluated on the Telco Customer Churn dataset, achieving 80.41% accuracy on a held-out test set. SHAP analysis identified contract type, tenure, monthly charges, total charges, and internet service type as the dominant churn drivers at the global and individual-customer level. An interactive Tableau dashboard is built to show these drivers and churn distribution by contract type, payment method and internet service type. Baseline comparison against Logistic Regression and Random Forest remains to be completed. When completed, the framework is expected to support the assertion that accurate prediction of churn alone is not sufficient for an effective retention strategy; interpretability and easy visualization are also important for translating predictive models into real business actions.

REFERENCES

1. D. Asif, M. S. Arif, A. Mukheimer, "A data-driven approach with explainable artificial intelligence for customer churn prediction in the telecommunications industry", Results in Engineering 26, 104629 (2025). https://doi.org/10.1016/j.rineng.2025.104629
2. S. S. Poudel, A. Alzubaidi, G. An, and J. Zhang, Explaining customer churn prediction in the telecom industry using tabular machine learning models, Machine Learning with Applications (2024).

3. V. Chang, K. Hall, Q. A. Xu, F. O. Amao, M. A. Ganatra, V. Benson, Prediction of customer churn behaviour in the telecommunication industry using machine learning algorithms, Algorithms 17(6), 231 (2024).
4. T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system, in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, 785–794 (2016). https://doi.org/10.1145/2939672.2939785
5. S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in Adv. Neural Inform. Process. Syst. 30, 4765–4774 (2017).
6. 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 9, 1748799 (2026).
7. K. Peng, Y. Peng, Research on telecom customer churn prediction based on GA-XGBoost and SHAP, J. Comput. Commun. (2022).
8. IBM, Telco Customer Churn dataset, accessed 18 July 2026. https://www.kaggle.com/datasets/blastchar/telco-customer-churn