TELECOM CUSTOMER CHURN PREDICTION AND PREVENTION

Ms. Sreelakshmi K R and Ms. Anju L

***Abstract*— Customer churn stands as a critical concern within firms, amplified by the intensifying competition, the amplified significance of marketing tactics, and the increasingly discerning nature of customers in recent times. In the swiftly evolving telecom sector, customer defection remains pivotal. The act of transitioning from one telecom service provider to another is driven by various factors such as superior services, competitive rates, or enticing advantages offered by rival firms during customer acquisition. The principal goal of this project centers on delving into the forecast of customer churn within the telecom domain, leveraging advanced machine learning technologies on extensive data platforms. Machine learning methodologies are applied to gauge the likelihood of customers churning, construct predictive models to pinpoint those at heightened risk, and delineate the primary indicators influencing churn. Additionally, the project aims to devise preventive strategies aimed at circumventing customer defection.**

I. INTRODUCTION

Acquiring new customers is considerably more costly compared to retaining existing ones, with the expense of acquiring a new customer being six to seven times greater than retaining an already existing one. Customers hold a pivotal position as the primary contributors to profits across industries. Hence, companies prioritize retaining their current customer base more than ever. Successfully reducing customer churn relies on accurately predicting customer behavior and comprehensively understanding the internal factors influencing attrition within a firm. By distinguishing between customers who continue their engagement and those who discontinue, it becomes feasible to forecast potential churn occurrences. In the realm of telecommunications, specialized telecom analytics stands tailored to meet the unique demands of this sector.

Telecom analytics predominantly focuses on amplifying profits, curtailing expenses, and mitigating fraud. The core objective revolves around forecasting and optimizing across multiple dimensions. The telecom sector often grapples with customer churn, impacting revenues when customers transition from one service provider to another. To expand their revenue streams, telecom companies must not only attract new customers but also mitigate terminations (churn). Churn analysis reveals that customers terminate their contracts due to various reasons, including superior price offers, enticing packages, unsatisfactory service encounters, and personal circumstances changes.

This initiative, named ”Telecom Customer Churn Prediction and Prevention,” aims to devise a predictive model to assess customer details and anticipate potential churn from the service. Leveraging diverse machine learning algorithms like Logistic Regression, Random Forest, ADA Boost, and XG Boost, the project seeks to implement and fine-tune these models to attain optimal predictive accuracy. The system not only expedites decision-making but also provides effective strategies to prevent customer churn.

II. System Architecture

The process for Customer Churn Prediction and Prevention using a logistic regression model in a telecom setting involves data collection, preprocessing, model development, deployment, and strategy implementation.
Customer data, including demographics and usage patterns, is gathered and processed to train a logistic regression model. Once deployed, this model predicts the likelihood of customer churn. Strategies for retaining customers are formulated based on these predictions, allowing targeted interventions. Continuous monitoring and adaptation ensure the system remains effective, providing a structured framework to proactively manage churn and enhance customer retention in the telecom industry.

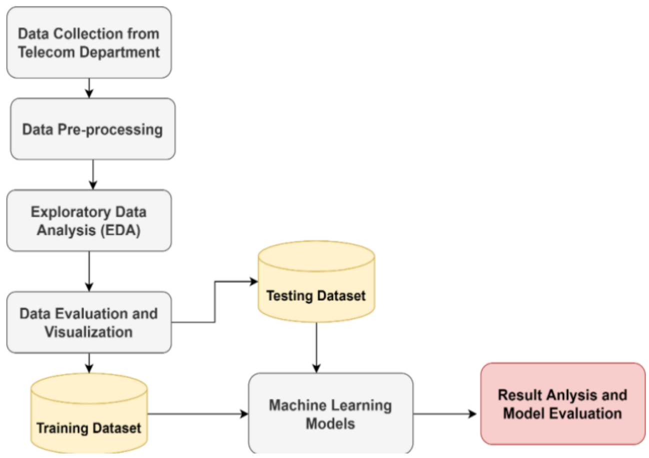

Fig. 1. Flow diagram of Proposed Work [1]

III. Methodology

A. Data Processing

The Data Processing module encompasses the aggregation and cleansing of data, integrating information regarding customer demographics, service usage, and churn history into a cohesive dataset. Data cleaning procedures, including handling missing values, outliers, and ensuring data consistency, are imperative to fortify the dataset's quality. Concurrently, Feature Engineering within this module enriches the dataset by creating new attributes or transforming existing ones, thereby enhancing the predictive power of subsequent models. Furthermore, the Data Splitting process stratifies the dataset into segments for distinct purposes, facilitating effective model training and validation.

B. Data Exploration

The churn rate graph, extracted from the dataset for the project, illustrates how customers' engagement changes over time. Concurrently, the graph plotting the correlation between monthly charges and total charges manifests a direct proportionality, indicating that as the monthly bills increase, so do the total charges incurred by customers[2].Moreover, Figure 2, prominently featuring the rate of churn, stands as a pivotal representation within the dataset analysis. This figure sheds light on the dynamics of customer attrition, delineating the frequency or percentage of customer departures from the service or product over a specified duration.

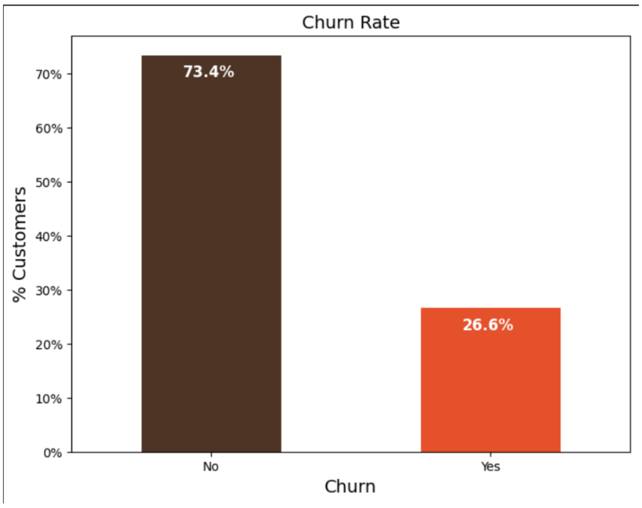

Fig. 2. Customer Churn Rate Visualization

IV. RESULT ANALYSIS

An experimental setup was established to create a Telecom Customer Churn Prediction model using Logistic Regression as the primary algorithm. This model aimed to predict customer churn based on various input features extracted from telecom service usage data.
The model was capable of handling real-time data from the telecom company's customer database or pre-recorded datasets. It processed this information using Logistic Regression to predict customer churn. Assessing individual customer instances, it provided predictions and associated probabilities, categorizing each instance as either 'churn' or 'no churn'. Following thorough testing and evaluation, the Logistic Regression model achieved an accuracy rate of 85 percentage.

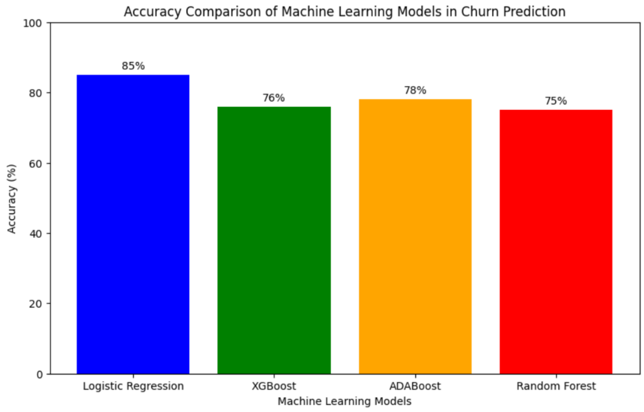

Fig. 3. Customer Churn Rate Visualization

V. CONCLUSION

The Telecom Customer Churn Prediction and Prevention project, centered around the utilization of Logistic Regression as the primary predictive model, has yielded considerable insights and demonstrated substantial potential in identifying customers at risk of churning within the telecom service domain. Leveraging Python programming for model development and analysis, this project aimed to proactively forecast potential churners based on a diverse array of telecom service usage data. The deployed model successfully categorized customers into churn and non-churn segments, enabling strategic initiatives for customer retention and business sustainability.
The analysis uncovered areas for potential enhancement to optimize the model's performance and mitigate potential revenue loss resulting from false predictions. Strategies aimed at minimizing false negatives, which could lead to losing valuable customers, while simultaneously enhancing precision, are pivotal for refining the model's predictive efficacy.

VI. FUTURE SCOPE

Introducing sentiment analysis or natural language processing techniques to assess customer feedback allows for a deeper understanding of customers' sentiments, preferences, and concerns. Additionally, establishing an interactive feedback loop integrating machine learning models can dynamically adapt to evolving customer behaviour and preferences. Moreover, integrating feedback analytics into the churn prediction model could offer invaluable insights.

VII. REFERENCES

[1] Prediction of Customer Churn in Telecom Industry: A Machine Learning Perspective by Lopamudra Hota1, Prasant Kumar Dash jer2 - Department of Computer Science and Engineering, National Institute of Technology, Rourkela, India*

[2] Hota, L., Dash, P. "Comparative Analysis of Stock Price Prediction by ANN and RF Model", Computational Intelligence and Machine Learning, 2.1, (2020), 1-9.

[3] Labhsetwar, Shreyas Rajesh. "Predictive Analysis of Customer Churn in Telecom Industry using Supervised Learning." ICTACT Journal on Soft Computing 10.2 (2020): 2054-2060..