Telecom Customer Churn Prediction

Abstract

Customer churn, or customer attrition, is a critical issue in the telecom industry, leading to significant revenue losses and increased customer acquisition costs. To address this challenge, machine learning (ML) techniques can be leveraged to analyze customer behavior and predict churn probability. This project aims to develop an accurate and efficient ML-based model for telecom customer churn prediction.

The study involves data collection, preprocessing, feature engineering, and model training using multiple algorithms, including Decision Tree, Random Forest, Gradient Boosting, and K-Nearest Neighbors (KNN). The models are evaluated based on performance metrics such as accuracy, precision, recall, F1-score, and AUC-ROC to determine the best-performing approach. Additionally, hyperparameter tuning techniques like Grid Search and Random Search are applied to enhance model efficiency.

A comprehensive literature review highlights the existing research on telecom churn prediction, identifying key challenges such as class imbalance, feature selection complexity, lack of real-time prediction, and model interpretability issues. The study also proposes solutions, including the use of Synthetic Minority Oversampling Technique (SMOTE), automated feature selection, real-time model deployment, and explainable AI techniques like SHAP and LIME.

By developing a user-friendly predictive system, this project provides telecom companies with actionable insights to identify at-risk customers and implement data-driven retention strategies. The proposed approach aims to enhance customer satisfaction, minimize churn, and improve overall business profitability in the competitive telecom sector.