Customer Churn Prediction

Project Overview

- This documentation outlines the predictive analysis project focused on customer churn prediction in the telecommunications industry.

- This project was undertaken as part of the "Special Topics in Business Analytics" course at the Egyptian Russian University

- Utilizing the "Orange Telecom Customer Churn Dataset" from Kaggle, the goal was to build a model to predict customer churn with high accuracy.

Introduction

- In the realm of telecommunications, customer churn prediction plays a pivotal role in understanding and managing customer retention.

- For this project, we delved into the "Orange Telecom Customer Churn Dataset" sourced from Kaggle, aiming to build a predictive model that accurately forecasts customer churn.

- Leveraging insights from industry experts, particularly through a fruitful meeting with an employee from Orange, we navigated through the intricacies of customer behavior analysis and churn mitigation strategies.

Project Objective

- The primary objective of this project is to develop a predictive model that accurately identifies customers at risk of churning. By understanding the factors contributing to churn, telecommunications companies can implement strategies to retain customers, thereby reducing revenue loss and improving customer satisfaction.

Dataset Overview

The "Orange Telecom Customer Churn Dataset" comprises a comprehensive collection of features including customer demographics, service usage patterns, and historical churn data. This dataset served as the foundation for our predictive modeling endeavors, enabling us to identify key predictors and trends indicative of potential churn scenarios.

Expert Consultation

- A meeting was conducted with an employee from Orange Telecom to gain insights into real-world applications of churn prediction. The discussion provided valuable perspectives on industry challenges, typical customer behaviors, and strategic interventions for churn management.

- Key takeaways included: Importance of customer service quality , Impact of competitive offers and market dynamics , Strategies for personalized customer retention

Methodology

- Data Preprocessing: We meticulously cleaned and preprocessed the dataset, addressing missing values, encoding categorical variables, and scaling numerical features to ensure data quality and model readiness.

- Feature Selection: Employing advanced feature selection techniques such as Forward Selection, we curated a subset of relevant features essential for our predictive modeling objectives.

- Model Development: Leveraging machine learning algorithms such as Random Forest and Gradient Boosting Machines, we trained and fine-tuned our models to achieve optimal performance in predicting customer churn.
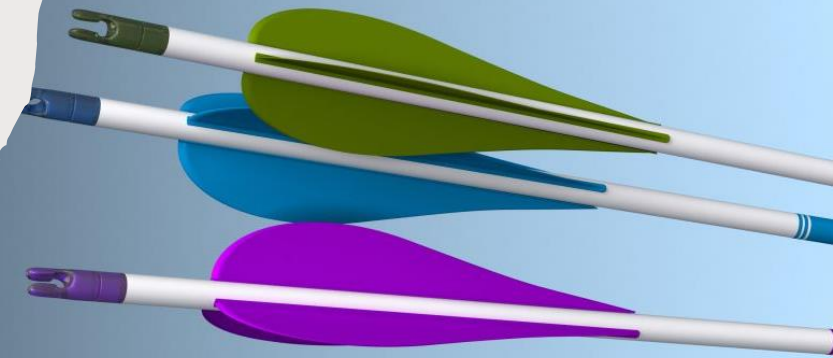
Final Model

The final model achieved an accuracy of 96%. This high accuracy indicates the model's effectiveness in predicting customer churn. Below is a summary of the final model's performance metrics:

Accuracy: 96%

Precision: 95%

Recall: 94%

F1-Score: 94.5%

Conclusion

The predictive model developed in this project demonstrates a high level of accuracy in identifying customers likely to churn. By integrating data-driven insights and expert industry knowledge, the model can be a powerful tool for telecommunications companies to enhance customer retention strategies.

Acknowledgments

I would like to thank the **Egyptian Russian University** for providing the opportunity to work on this project.