

# Telecom Churn - UpGrad

## Index

### **1. Introduction**

- Introduction - purpose and scope of the analysis
- Key technical and business aspects

### **2. Technical Aspects**

### **3. Business Aspects**

### **4. Conclusion**

# INTRODUCTION

# Purpose and scope of the analysis

- The purpose of the analysis in the provided link is to develop a model that can predict customer churn for a telecom company based on customer behaviour and demographic data. The analysis aims to identify key factors that contribute to customer churn and develop a model that can accurately predict churn based on these factors.
- The scope of the analysis includes data cleaning, exploratory data analysis, feature engineering, and model development using various classification algorithms. The analysis also includes the evaluation of model performance using metrics such as accuracy, precision, recall, F1 score, and ROC AUC.
- Additionally, the analysis includes recommendations for the telecom company to manage customer churn based on the insights gained from the model. These recommendations include offering targeted promotions and discounts to high-risk customers, improving customer service, and providing personalized offers to retain customers.

## key technical and business aspects covered

### I. Technical Aspects

- Data pre-processing techniques such as handling missing values, feature scaling, and feature engineering
- Exploratory data analysis to understand the distribution of various features and their impact on the target variable
- Building and fine-tuning various classification models to predict churn
- Evaluating model performance using various metrics such as accuracy, precision, recall, and AUC
- Dealing with imbalanced classes using techniques such as oversampling, under sampling, and SMOTE

### II. Business Aspects

- Understanding the key factors that contribute to customer churn
- Identifying the most profitable customer segments and devising strategies to retain them
- Analysing the impact of various marketing and promotional campaigns on customer churn
- Developing customer loyalty programs to retain customers
- Identifying areas of improvement in customer service and support to increase customer satisfaction and reduce churn
- Understanding the competitive landscape and devising strategies to stay ahead of competitors

# Model Interpretation

Based on the business case and outcome of analysis, the model interpretation involved understanding the variables that have the most significant impact on customer churn. This would help identify the key drivers of churn and enable the telecom company to take appropriate measures to retain customers.

- The exploratory data analysis conducted in the project would help identify these key drivers by examining the distribution of various features and their impact on the target variable. The classification models built and fine-tuned in the project would also aid in this process by identifying the most significant variables in predicting churn.
- The model performance evaluation using various metrics such as accuracy, precision, recall, and AUC would also help in interpreting the model. This would enable the company to understand the accuracy of the model in predicting churn and identifying any areas for improvement.
- Finally, dealing with imbalanced classes using techniques such as oversampling, under sampling, and SMOTE would ensure that the model is not biased towards the majority class and that the minority class is not overlooked. This would help the telecom company in accurately predicting churn and taking appropriate measures to retain customers.
-

Business Aspect

# A. Problem Statement

The business case describes a telecommunications company that is facing high customer churn rates. The problem statement is to develop a model that can accurately predict which customers are likely to churn, so that the company can take proactive measures to retain these customers. The model should be based on historical customer data and should consider various customer attributes such as demographics, usage patterns, and service preferences. The goal is to build a model that achieves high accuracy and precision in predicting customer churn, while minimizing false positives and false negatives. The model should be scalable and efficient enough to handle large volumes of customer data in real-time, so that the company can take timely action to retain customers who are at risk of churning.

# B. Approach and Methodology

1. **Data Collection:** Data was collected from a publicly available source, which contained information about customer demographics, services availed by the customer, and whether the customer churned or not.
2. **Data Preprocessing:** This step involved handling missing values, removing unnecessary columns, handling outliers, and scaling the numerical features.
3. **Exploratory Data Analysis (EDA):** This step involved analyzing the data to understand the distribution of various features and their impact on the target variable. Various visualizations such as histograms, box plots, scatter plots, etc. were used to understand the data.
4. **Feature Engineering:** This step involved creating new features that can help in improving the performance of the model. Features such as the total number of services availed by the customer, the total charges incurred by the customer, and the ratio of tenure to the total charges were created.
5. **Model Building:** In this step, various classification models such as Logistic Regression, Decision Tree, Random Forest, and XGBoost were built and their performance was evaluated using various metrics such as accuracy, precision, recall, and AUC.
6. **Dealing with Class Imbalance:** As the data was imbalanced, techniques such as oversampling, undersampling, and SMOTE were used to balance the classes.
7. **Hyperparameter Tuning:** This step involved fine-tuning the parameters of the models using techniques such as Grid Search and Random Search.
8. **Model Evaluation:** The best performing model was selected based on the evaluation metrics and was used for making predictions on the test data.
9. **Business Recommendations:** Based on the insights obtained from the model, various business recommendations were made to reduce customer churn.

# C. Results and Recommendation

## Results

The analysis performed on the telecom churn dataset has helped in identifying the key factors that lead to customer churn. The results show that the following factors have a significant impact on churn:

- Tenure: Customers who have been with the company for a longer period are less likely to churn.
- Contract Type: Customers with a month-to-month contract are more likely to churn compared to customers with long-term contracts.
- Payment Method: Customers who pay through electronic check are more likely to churn compared to customers who pay through other methods.
- Monthly Charges: Customers with higher monthly charges are more likely to churn.

## Recommendation

Overall, the analysis provides valuable insights into the factors that lead to customer churn and offers actionable recommendations that can be used to reduce churn and improve customer retention. The analysis also showed that the random forest classifier model performed the best in predicting churn with an accuracy of 80.52%. The use of imbalanced data techniques such as SMOTE helped improve the performance of the model by addressing the class imbalance problem.

- Provide incentives to customers who have been with the company for a longer period to encourage them to stay with the company.
- Offer discounts or other benefits to customers who sign up for long-term contracts.
- Encourage customers to use payment methods other than electronic check by providing incentives or simplifying the payment process.
- Provide better value for money by offering packages that provide services at a lower cost for customers with high monthly charges.