Executive Summary

The U.S. based telecommunication company is trying to see if it can identify factors that cause customers to cancel their internet and cellular services. Also, if the company can predict whether a customer will cancel their service in the future. During the past couple of years, the company has experienced record levels of customers canceling their service. This led to huge financial losses to the telecom company. It is crucial for the company to determine the factors responsible for customer attrition and predict if customers will cancel their service in the future, as it minimizes future financial losses, build satisfactory customer relationships, and offer better delivery methods to customers.

The main aim of the analysis is to perform an exploratory data analysis on the given telecom company dataset and discover the variables that drive the differences between customers who do and do not cancel their service. Then build a machine learning classification algorithm to forecast if customers will cancel company service. The company can utilize analysis results in optimizing their product, reduce customer attrition and company financial losses.

The objective of the exploratory data analysis of the telecom data is to find out if there is a relationship between customers canceling the service and other various factors such as senior citizen, average calls in minutes, internet service, and many more. One of the interesting findings represents that there is a correlation between customers canceling service and average monthly international calls. The customers who cancel the telecommunication service tend to have a lesser average monthly duration of international calls when compared to those who do not. Similarly, there is a strong relationship between service cancellation and the total number of months with the company. The customers who cancel the service tend to remain fewer months with the company than those who do not.

The exploratory data analysis reveals certain potential patterns for the telecom company to lessen the customer attrition rate. The company can significantly benefit from this to optimize its business productivity, service quality, enable in influencing customer loyalty and customer retention.

Among Logistic regression, Random forest, Decision tree, the best classification model identified based on the analysis of the performance of the model is Random Forest. False-positive error rate refers to the model predicting customers as ‘yes’ under canceled service among those who did not cancel the service. Whereas false-negative error rate refers to customers as ‘no’ under canceled service among those who did cancel the service. The false-positive rate and false-negative rate of the random forest model on the testing dataset are 11.22% and 38.67% respectively, which are comparatively less than other models.

The area under the ROC curve shows how much the model can distinguish between different classes. The higher the value of area under ROC value, the better the model in predicting the classes as did or did not cancel the service. The random forest model can predict with the highest accuracy of 88.46%. Hence, the random forest is the best model in predicting whether a customer did cancel the service or not.

The telecom company should start offering customers a higher international call duration. The summary results and histogram plot 1 indicate that the customers who cancel the telecommunication service tend to have a lesser average monthly duration of international calls when compared to those who do not. Among 427 customers who cancel the service, the mean average monthly international call duration was

93.57 minutes. This is around 20 minutes less than the mean of average international call minutes of customers who do not cancel the service.

Likewise, the company must grant benefits to customers like high-speed internet, an increase in monthly national and international call duration, online security, and tech support. This will help the company to retain customers indefinitely. The summary results and box plot 3 specifies that the customers who cancel the service tend to remain fewer months with the company than those who do not. Among 427 customers who cancel the service, the average total number of months remained with the company is around 17. This is around 21 months less than the average number of months who remained with the company and did not cancel the service.

In conclusion, it is recommended for the company to extend in monthly duration of international calls and other perks to customers, as they remain satisfied. This in turn benefits company in building satisfactory customer relationships, customer retention, offer better delivery methods to customers, and minimize future financial losses.