**CS 519 Applied Machine Learning I Project**

**Team Member: Md Ishtiaq Ahmed, Israel Gonzalez**

**Problem Statement:** Predicting customer churn is an important part of customer relationship management and can help companies improve their customer retention, reduce costs, and improve their financial performance. But predicting customer churn is always a difficult task to do specially for the telecom service providers. Developing a predictive model that accurately identifies customers who are likely to churn from telecom services can help a telecom a company to retain their customers and improve their profit.

**Motivations:** I (Ishtiaq) worked in a telecommunication company for around 8 years. During my tenure, I always saw that how they are struggling to retain their customers. It is also difficult to find out a solid reason behind the churn. If they can predict accurately which customers are about churn then they can take predictive measures to avoid that. So, this problem has real-world implications for business performance and profitability which worked as a motivation for us. Also, predicting customer churn is a well-studied problem in machine learning, and there are many techniques and algorithms that can be applied to this problem. By working on this problem, we can apply different techniques learned from the course.

**Analysis Task:** The analysis can be divided into three parts. First part is the exploratory data analysis. This is important to understand the characteristic of data. It will help to identify potential issues with the data and can provide insights about the structure of the data. Which is important to select appropriate machine learning models.

Second part is to predict which customers are going to be churned. This is mainly a classification problem. By applying classification techniques, we will try classifying the customers in two class that is churn or not churn.

Third part of analysis will be to divide the customers in different groups based on their behavior. We will try to analyze if there is any behavioral pattern that can segregate a group of customer. Our idea is to apply clustering technique to do this analysis.

**Related works:** Predicting customer churns in telecom industries is always been difficult due to the complex behavior of the customers and their changing preferences. There were lot of research work was done and many are on-going in this area.

One of the research projects we found with title " Behavior-Based Telecommunication Churn Prediction with Neural Network Approach”. This study used neural networks to predict customer churn in a telecom company. Customer service usage information are used as the features. Customer churn was predicted using clustering algorithms.

Another research project is “ Intelligent Decision Forest Models for Customer Churn Prediction ” . In this paper several techniques were used to predict churn including random forest algorithm, functional tree algorithm and logistic model tree (LMT) algorithm. Result of this study shows that this mentioned algorithm gives better result than the classification algorithms like Naïve Bayes(NB) and KNN.