**CS 519 Applied Machine Learning I Project – Stage 2**

**Team Members: 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. However, predicting Customer Churn is always a difficult task to do specially for Telecom service providers. Developing a predictive model that accurately identifies customers who are likely to churn from telecom services can help a Telecom company to retain their customers and improve their profit.

**Motivations:** On the one hand,Ishtiaq worked in a Telecom company for around 8 years. During his tenure, he always saw how they were 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. On the other hand, Israel had one-year academic experience in Chile in 2013, where he was part of a Business Intelligence Diploma at University of Chile, where this problem was a core business case they studied as a very practical type of need where BI and ML solutions can contribute. 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:

1. *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.
2. *Classification.* It is to predict which customers are going to be churned. By applying classification techniques, we will try to classify the customers in two classes that is churn or not churn.
3. *Clustering.* We 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 customers.

**Related works:** Predicting Customer Churns in Telecom industry has always been difficult due to the complex behavior of customers and their changing preferences. There were lots of research work has been done, and many are on-going in this area.

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

Another research project was *“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. Conclusions of this study show that this mentioned algorithm gives better results than the classification algorithms like Naïve Bayes (NB) and KNN.