Customer Churn Prediction in the Telecom Industry Using ML

Team Members: Seokhoon Shin, Joshua Nahm, Jiwon Choi, Shinyeong Park, Jaejoong Kim

1. Project Topic & Problem Definition

According to Bain & Company, preventing 5% of customer churn can result in a 25-29% increase in operating profit. Acquiring new customers is generally more costly than retaining existing ones. Many of our teammates had experiences as marketing interns, and we wanted to see if we could apply churn prediction to marketing. We came across a sample dataset titled "Teleco Customer Churn" that included features such as customer tenure, monthly charges, contract type, and payment methods. This project aims to predict customer churn based on these key features. The primary objective is to explore several machine learning processes and build a model that will empower businesses to allocate resources efficiently, focusing on the most vulnerable customer segments while enhancing overall customer experience and loyalty.

2. Data Source

We are using the **Telco Customer Churn** dataset from Kaggle, which contains customer demographics, services, account information, and whether the customer has churned or not. The key features include customer tenure, monthly charges, contract type, and payment methods. The dataset has 7,043 records and 21 columns.

- **URL**: https://www.kaggle.com/blastchar/telco-customer-churn

Using the dataset above, we aim to build a customer churn prediction model to help telecom companies identify high-risk customers and take churn prevention strategies. By understanding which factors contribute most to churn, we can provide actionable insights for reducing customer loss and improving customer satisfaction.

3. Proposed Methodology

Our team will begin the data-cleaning process by addressing missing values and eliminating non-essential columns. Categorical features such as contract type and payment methods will be encoded using one-hot encoding to ensure seamless integration with ML models: Logistic Regression and Random Forest. Numerical features, including customer tenure and monthly charges, will be normalized to enhance model performance and accuracy.

Preliminary insights indicate that customers on month-to-month contracts tend to churn more than those with longer-term agreements. Moreover, using electronic checks as a payment method is strongly correlated with increased churn, suggesting underlying issues in the customer experience for this segment. Finally, customer tenure is a significant predictor of churn, with newer customers (0-6 months) particularly susceptible to leaving.