Telecom Customer Churn Rate Predictor

Group Members
Alba Mustafaj 21500009, Alp Ege Baştürk 21501267, Berat Biçer 21503050, Bora Ecer 21501757, H. Buğra Aydın 21501555

Dataset Description
The primary dataset for the project is the customer data of a telecommunication company, TelCo. The dataset contains customer IDs, their age group (senior or not), gender, marital status, dependents, tenure, information about used services and payment [1]. The dataset contains 21 columns. The last column stores information about churn rate, indicating the decision of the customer. This is also the target of this project. Dataset includes text fields, which will be converted to numerical form for the use in the project. Fields like seniority and dependents are Booleans and fields about payment amounts are float. Some of the columns are given below:

customerID Customer ID	**gender** Customer gender (female, male)	**SeniorCitizen** Whether the customer is a senior citizen or not (1, 0)	**Partner** Whether the customer has a partner or not (Yes, No)	**Dependents** Whether the customer has dependents or not (Yes, No)
7590-VHVEG	Female	0	Yes	No

Project Description
The objective of this project is predicting user's churn rate, which occurs when a customer stops using the service. We will be using the features mentioned previously to predict if a user is planning to stop using the service. The dataset will be divided into train and test sets. We are planning to create several models with the training dataset using several machine learning methods like SVM, kNN and logistic regression, decision tree classification, random forest classification and deep learning. After the training phase, models should be able to predict whether a customer is leaving or not by looking at the customer's data according to the features determined at the time of training. Then the comparison will be done between the real values and predictions.
Since we will be using different machine learning algorithms with the data, we will be able to compare our results to determine which algorithm generates a better model. Additionally, we are planning to tune the features to increase the success of predictions of the algorithms.

Expectations and Results
We expect to see the relation of data to churning. The algorithms will give binary outputs indicating a user with given data will leave or not. It's expected that customers with long tenure and seniority are less likely to leave the operator however we expect to uncover hidden relations of features.

Works Cited
[1] Blast Char "Telco Customer Churn". *Kaggle.* [Online]. Available: https://www.kaggle.com/blastchar/telco-customer-churn. [Accessed: Oct. 2, 2018]