❖ Project Proposal

This repo is one of the T5 Bootcamp requirements.

❖ Abstract

This project aims to help a foreign telecom company determine the likelihood of a customer leaving the company and not activating the subscription, as well as identify the factors that cause customers to abandon the subscription activation. The data for this study came from Kaggle. The dataset will be built to predict whether or not a particular customer will stay with the company.

❖ Design

This project is intended for the T5 Data Science BootCamp requirements. Data provided by Kaggle has been used in this project. It consists of company records of different parameters for each employee (such as Customers who left within the last month, Customer account information, Demographic info about customers, etc.). Customers at danger of leaving can be classified using machine learning algorithms that are trained from a data set, which can then be used to pinpoint the fundamental cause.

❖ Data

The dataset is provided in .csv format. It contains 7043 row (customers), and 21 columns (features). The "Churn" column, which indicates if customers have left during the last month, is project's focus.

❖ Algorithms

a. What is the percentage for each gender which left the company? Does gender has an effect on churn?
b. What is the percentage of elderly people in the company?
c. What is the effect of tenure length on churn?

❖ Tools

- NumPy and Pandas for data manipulation
- Scikit-learn for modeling
- Matplotlib and Seaborn for Visualization