Capstone 2 Project Proposal: Telco Customer Churn

Problem statement:

The objective of this project is to predict customer behavior in order to retain them. The goal is to analyze all relevant customer data and develop an algorithm that can model churn (customer retention index) with at least 80% accuracy. Also, by understanding the reasons behind customers leaving, focused customer retention programs can be developed to increase retention chances.

Context:

Telco is a fictional phone and internet services company providing services to its customers in California. Telco has collected data from a total of 7043 customers which includes multiple important customer demographics, services that each customer pays for, a satisfaction score and a 'churn' score. The churn score shows how long the customer has been using the services of Telco. The executives at Telco want to understand how accurately can the Churn feature can be predicted based on other features. The executives would like to see at least 80% accuracy in prediction. This will help understand which customers are more likely to leave and what specific actions can be taken to increase the chances of retention.

Data

There is one source dataset which contains information from 7043 customers (rows) and 21 features (columns). First the data will be analyzed to understand how many features are usable for modelling. In order to understand the relationship between different features and 'churn' correlation map will be generated. The dataset has to be divided into a training set ('train') and testing set ('test'). Different models will be derived using 'train' and the performance will be tested on the 'test'.

Dataset link:
https://www.kaggle.com/datasets/blastchar/telco-customer-churn

Final Deliverables

The final deliverables will include Jupyter notebook with code and comments, project report (pdf document) and a slides pack to summarize the detains for the executive team.