Artificial Intelligence Projects

- Registration ends: 14/4/2022.

- Registration link: https://docs.google.com/forms/d/e/1FAIpQLScvHa7JlnUalWdyH5tEp4oQ_KAa-TpWTbeYVSrgRQ14Z2Hd4g/viewform

- Minimum number of members in team is 5 and maximum is 7

- You must deliver a detailed report for the project contains all your work (Preprocessing, algorithms used in the module and the achieved accuracy).

Note: **Report will be graded**

Project (1): Service cancellation predictor

Description:

Service cancellation is simply when customers leave doing business with an entity. It involves determining the possibility of customers stopping doing business with an entity. In other words, if a consumer has purchased a subscription to a particular service, we must determine the likelihood that the customer would leave or cancel the membership. It is a critical prediction for many businesses because acquiring new clients often costs more than retaining existing ones. For many businesses, the ability to predict that a particular customer is at a high risk of canceling service, while there is still time to do something about it. Whereas the company will try to offer some extra functionalities for not leaving the service.

Dataset:

Dataset link: Service Cancellation DataSet

Dataset description:

The dataset consists of 7043 rows and 21 columns, where rows represent the number of customers in the dataset and the columns represent each customer's attribute. The attributes are used to predict the service cancellation of a particular customer.

Each row represents a customer, each column contains customer's attributes described on the column Metadata.

There are 21 columns so we will divide them into independent and dependent columns:-

1- Independent variables: ['customerID', 'gender', 'SeniorCitizen', 'Partner', 'Dependents', 'tenure', 'PhoneService', 'MultipleLines', 'InternetService', 'OnlineSecurity', 'OnlineBackup', 'DeviceProtection', 'TechSupport', 'StreamingTV', 'StreamingMovies', 'Contract', 'PaperlessBilling', 'PaymentMethod', 'MonthlyCharges', 'TotalCharges']
   a. Services that each customer has signed up for – phone, multiple lines, internet, online security, online backup, device protection, tech support, and streaming TV and movies
   b. Customer account information – how long they've been a customer, contract, payment method, paperless billing, monthly charges, and total charges

c. Demographic info about customers – gender, age range, and if they have partners and dependents

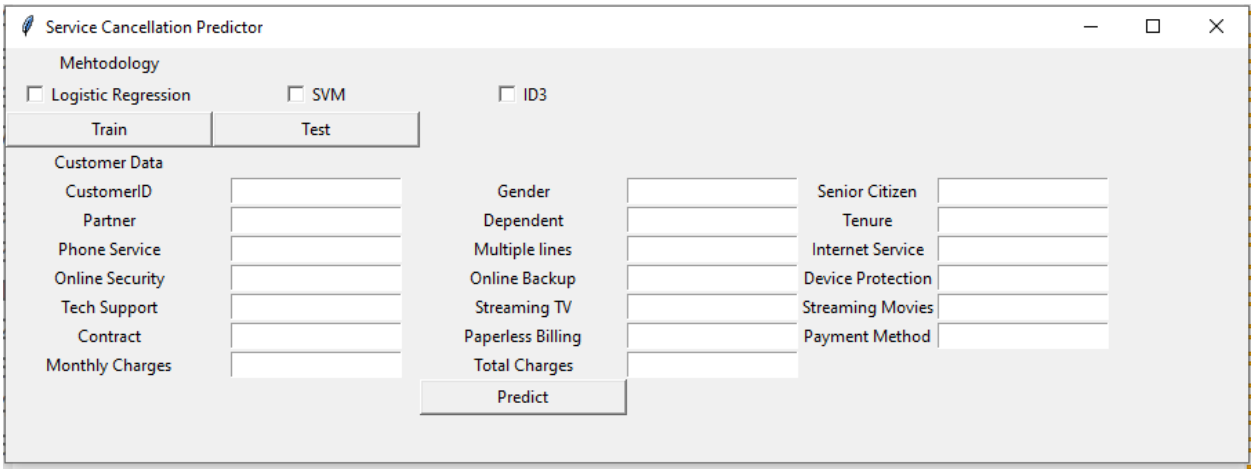
2- Dependent variables: ['Churn']
   a. Customers who left within the last month – the column is called Churn

Main Steps:

1. You need first to apply some preprocessing on the data to make sure that it is ready to use it. Preprocessing phase includes handling unwanted features, checking if data types of columns are correct, null values, categorical values, and data scaling
2. Choose the best technique to predict the service cancellation with a high accuracy. You must prove by code that the chosen algorithm produces a higher accuracy than some of the other algorithms. try to use logistic regression, SVM, and Decision Tree ID3.

Deliverables:

An application with a simple GUI that accepts input classification method and the accuracy of the used classification method. Also, the gui can accept data of a customer and predict if he\she may cancel the subscription.



Hint:

For GUI, you can use the **tkinter** library.

Mentor: T.A. Aya Saad

Email: aya.saad@cis.asu.edu.eg