README FILE

Running Instructions

1. Place *classification_main.ipynb*, *clustering_main.ipynb* and *requirements.txt* in a single folder.
2. Download the required data using the link - *https://www.kaggle.com/c/2551/download-all* and unzip the files *cs-test.csv* and *cs-training.csv* into the folder.
3. Navigate to the folder in Terminal/PowerShell/Cmd and run the following command to install dependencies and modules required to execute the Jupyter Notebooks.
   pip install -r requirements.txt
4. Run *classification_main.ipynb* which contains the code for the classification task.
5. Run *clustering_main.ipynb* which contains the code for clustering customers and performing cluster analysis.

Team

Name	NetID	UIN
Arianna Brown	abrown85	654697231
Warisa Siewsrichol	wsiews3	674193989
Vignesh Narayanaswamy	Vnaray21	660472930

Contributions

Below is the contribution of each person in the team.

TASK	SUB-TASK	DONE BY
CLASSIFICATION	Exploratory Data Analysis	Vignesh, Arianna, Warisa
	Data Cleaning	Vignesh
	Hyper-parameter Tuning	Vignesh
	Training and Evaluation	Vignesh
	Predicting test data	Arianna, Warisa, Vignesh
CLUSTERING	Data Preparation	Arianna, Warisa
	Exploring clustering techniques	Warisa, Arianna
	Hyper-parameter tuning	Warisa, Arianna
	Plots and Visualization	Vignesh
	Results and Conclusions	Vignesh, Warisa, Arianna
REPORT	Structure	Arianna, Warisa, Vignesh
	Content	Arianna, Warisa, Vignesh
	Editing and Review	Vignesh, Arianna, Warisa
README FILE	Content	Vignesh
PRESENTATION SLIDES	Structure	Warisa, Arianna
	Content	Warisa, Arianna
	Editing and Review	Warisa, Arianna, Vignesh