USER GUIDE

This is a guide to help you run all experiments in the artefact with the three jupyer notebooks supplied to reproduce results.

PC requirements:

The experiments for the artefact are a bit intensive and are run on two platforms;

- The Notebooks **Data_Cleaning_Preprocessing.ipynb** and **EDA_Feature_Engineering_Collab.ipynb** were run on my local PC with Visual Studio Code. They do not need much processing power and the results can be reproduced on a local PC with specifications similar to mine (Lenovo Legion Slim 5, AMD Ryzen 7, Nvidia 4060 GPU).

- The Notebook **LSTM_Implementation.ipynb** requires more proceesing power and was run on a paid google collab account for extra GPU and High RAM. The results can be reproduced if the experiments are run on a collab notebook with the specifications (100 Google Collab Pro units (T4 and L4 GPU)).
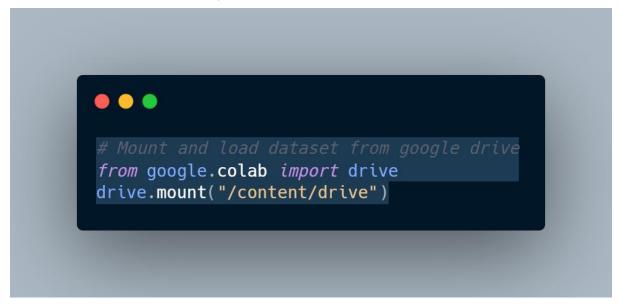
Directory structure and File Path;

The Notebooks are run locally and online and have some differences in the file paths to load the datasets.

If the notebooks are executed locally, follow the steps below;

- Ensure the "**Data**" folder containing the datasets and the "**requirements.txt**" file are in the same file directory as the Jupyter notebooks.

- Create a virtual environment and run the code dependencies and libraries "**requirements.txt**" file in your virtual environment to install all project dependencies.

- If the notebook is executed locally, then load the dataset with the file path **'data/cleaned_medical_appointments.csv'**

If the notebook is executed on Google collab then follow the steps below;

- Upload the Data file to the Google Drive account.
- Run the code snippet at the start of the notebook to mount Google Drive at a location on the collab file directory

```
# Mount and load dataset from google drive
from google.colab import drive
drive.mount("/content/drive")
```

- Then load the dataset with the file path
  "/content/drive/MyDrive/data/cleaned_medical_appointments.csv"