Description of the file directory and steps to run the script

Directory Tree:-

* Fraud-Detection is the main folder which contains the following:
* main script ‘fraud\_detection.py’ (the script itself contains detailed analysis and explanation)
* script to tune hyperparameters ‘fine\_tune\_model.py’
* script which contains all the important helper functions ‘utils.py’ (all the helper functions are thoroughly explained in the script)
* ‘requirements.txt’ file which contains all the libraries with the exact versions used in this project
* the 2-slide presentation pitch ‘Pitch.pdf’
* a folder called **data** which contains the following:
  + ‘labels\_obf.csv’ 🡪 fraud labels
  + ‘transactions\_obf.csv’ 🡪 transaction data
* a folder called **figs** which contains the following figures:
  + ‘pipeline.png’
  + ‘feature\_engineering.png’
  + ‘feature\_importance.png’
  + ‘XGBoostBase.png’
  + ‘pr-best-model.png’
* a folder called **processed\_data** which contains:
  + ‘processed\_df.csv’ 🡪 data after processing and feature engineering
  + ‘precision-recall-best-model.csv’
* a folder called **model** which contains:
  + ‘best\_model.pkl’ 🡪 final trained model

Steps to run the script:-

1. Extract the zip file into any folder
2. Open a terminal or command prompt
3. Change the directory into this folder using the following command:
   1. “cd *location\_of\_the\_folder*” 🡪 this will take you insidse the folder
4. Now we need to install the required libraries to run our script which can be done using the command below:
   1. “pip install -r requirements.txt”
5. Now after successfully installing the required libraries, we can run our main script by the command below:
   1. “python fraud\_detection.py” 🡪 all windows users
   2. “python3 fraud\_detection.py” 🡪 some mac users who have multiple versions of Python installed