**Student number**------------23025664

**Module name**--------------- MACHINE LEARNING AND PREDICTIVE ANALYTICS

**Module number**------------ UFCFMJ-15-M

**Programme**------------------Msc Financial Technology

**Step-by-Step Guide to Run the Code**

1. **Set Up the Environment**

Ensure you have Python installed on your system. You can download it from the official python website.

1. **Install Necessary Libraries**

For Data Handling

import pandas as pd

import numpy as np

For Data Visualization

import matplotlib.pyplot as plt

import seaborn as sns

For Data Preprocessing

from sklearn.model\_selection import train\_test\_split, GridSearchCV

from sklearn.preprocessing import LabelEncoder, StandardScaler

from sklearn.compose import ColumnTransformer

from sklearn.preprocessing import OneHotEncoder

from typing import Dict

For Modelling

from sklearn.linear\_model import LogisticRegression

from sklearn.ensemble import RandomForestClassifier

from sklearn.cluster import KMeans

from sklearn.metrics import accuracy\_score, classification\_report

from sklearn.pipeline import Pipeline

For Statistic

from statsmodels.graphics.gofplots import qqplot

from sklearn.metrics import classification\_report

from scipy.stats import Shapiro

import time.

Metrics

from sklearn.metrics import confusion\_matrix, classification\_report

from sklearn.metrics import accuracy\_score, f1\_score, precision\_score, recall\_score, roc\_auc\_score, log\_loss

from sklearn.metrics import cohen\_kappa\_score, average\_precision\_score, precision\_recall\_curve, roc\_curve

import shap

1. **Load the Data**

Ensure the dataset is in a suitable format, typically a CSV file.

1. **Write the Code**

**As seen in the file**

1. **Run the Code**

Run the code using python.

1. **Interpret the Results**

Analyze the plots to understand the relationships between the variables in your dataset. Also evaluate the performance metrics of the models to check their accuracy and precision.

**DATASET**

A loan approval prediction dataset was downloaded from Kaggle through this link <https://www.kaggle.com/datasets/architsharma01/loan-approval-prediction-dataset> and saved as CSV file to the same working directory as my jupyter notebook, and imported using panda library

The dataset and my code are in the folder saved as Machine learning assessment and can be accessed via my onedrive link <https://uweacuk-my.sharepoint.com/:f:/g/personal/ronke2_igabor_live_uwe_ac_uk/Elm_wUBSkXNPo1pCUjwKduwBcsvmeMKLwQe3pO5O7VOn6A?e=D6eS2K> or my Github link <https://github.com/rv2-igabor/MACHINE--LEARNING> for the code and dataset.

Thank you