```
In [1]:
          import pandas as pd
          import numpy as np
          import matplotlib.pyplot as plt
          from sklearn.model_selection import train_test_split
          from sklearn.tree import DecisionTreeClassifier
          from sklearn import metrics
          from sklearn.exceptions import NotFittedError
          from sklearn.preprocessing import StandardScaler, LabelEncoder
          from sklearn.datasets import make_classification
          from sklearn.model_selection import KFold
          from sklearn.model_selection import cross_val_score
          from sklearn.metrics import accuracy_score
          from sklearn.model_selection import StratifiedKFold
          from sklearn.linear_model import LinearRegression
 In [2]:
          Data = pd.read_csv("P2_Data.csv") # Upload the data
 In [3]:
          Data
                   F1
                            F2
                                   F3
                                                   F5
                                                         F6
                                                                F7
                                                                      F8
                                                                             F9
                                                                                    F10 ...
                                                                                              F28
                                                                                                       F29 F30
                                                                                                                  F31
                                                                                                                         F32
                                                                                                                               F33
                                                                                                                                       F34
                                                                                                                                             F35
                                            F4
                                                                                                                                                     F36
                                                                                                                                                          Target
 Out[3]:
                      -15267.84
                                                                           -38.02
                                                                                   8.35 ... -100.43 10497.32
                854.04
                                193.04 12132.20 464.22 -19.81
                                                             920.42 11.84
                                                                                                             8 -46.22
                                                                                                                        91.89
                                                                                                                             -41.70
                                                                                                                                     -849.78
                                                                                                                                            -5.12 4261.68
                                                                                                                                                          609.09
            1 1748.76
                       -4299.45 439.72 18046.14 540.36
                                                       -8.55 1913.78 11.28
                                                                          122.02
                                                                                   -8.83 ... -97.41 16792.88
                                                                                                             4 51.24 -450.34 -75.30 -1801.59
                                                                                                                                            -0.16 2737.58
                                                                                                                                                          232.66
                834.44 -18927.93
                                -85.86 18533.56 611.34 -15.88 1956.90 10.00
                                                                           -24.26
                                                                                  -54.33 ... -135.59
                                                                                                  -5647.94
                                                                                                                40.93
                                                                                                                      -471.02 -55.50
                                                                                                                                   -2529.75
                                                                                                                                             8.50 3696.64
                                                                                                                                                          480.01
                                                                                                                                             9.76 2919.54 1685.20
            3 1429.36
                        3653.73
                                -59.69
                                      31314.52 408.45
                                                                           90.88
                                                                                  -38.87 ...
                                                                                                   6869.46
                                                                                                                      -290.45
                                                       -9.67 1628.18 11.59
                                                                                            -85.77
                                                                                                                 -9.19
                                                                                                                             -67.86
                                                                                                                                   -2915.34
                        -3932.07
               188.80
                                128.70 22837.34 530.22 -14.93 1218.16
                                                                     9.55
                                                                           -93.78
                                                                                  -61.95 ... -75.76 16974.08
                                                                                                             2 100.86
                                                                                                                      -850.10 -64.95 -2154.12 13.42 2693.48
                                                                                                                                                          617.96
                987.12
                        4369.68 -223.67 25595.70 395.58 -14.05 2503.50 13.57
                                                                           10.14
                                                                                  -36.93 ...
                                                                                            -95.59
                                                                                                   9304.98
                                                                                                                 -7.54
                                                                                                                      -845.24 -71.10 -1960.50
                                                                                                                                             9.02 3062.52 1057.77
          1495
                                                                                                            10
          1496
                966.92 -12845.01 -735.65 25065.72 613.14 -15.62
                                                                                  -69.32 ... -129.22 16906.02
                                                                                                            10
                                                                                                                16.42 -360.75 -60.21
                                                                                                                                    -753.06
                                                                                                                                            -1.60 4167.22 1232.01
                                                             859.76 11.04
                                                                           43.12
                860.70
                                       31577.66 577.62
                                                                                                   4687.12
          1497
                         -641.70
                               -393.30
                                                       -9.49
                                                             617.64
                                                                     5.80
                                                                           -63.68
                                                                                  -89.18 ... -201.85
                                                                                                            10
                                                                                                                -60.35
                                                                                                                      -477.90
                                                                                                                             -60.36 -1749.51 20.26 4466.74 2037.64
                                                                          -172.42
                                                                                  -27.69 ... -133.37
                                                                                                             8 -16.85
                                                                                                                                             8.90 3599.06
          1498
                428.82 -16169.43
                                -84.49 26820.44 555.30 -11.96 1631.22
                                                                     9.80
                                                                                                   9937.10
                                                                                                                      -458.68 -42.75 -1714.83
                                                                                                                                                           -27.86
                753.78 10536.00 -238.10 34668.58 557.04 -17.71 2344.28 25.30
                                                                           97.58 -101.30 ... -65.15 21663.04
                                                                                                                 0.39 -620.86 -59.76 -3509.94 15.88 4241.28 1012.31
         1500 rows × 37 columns
 In [4]:
          columns = ['F1', 'F2', 'F3', 'F4', 'F5', 'F6','F7','F8','F9','F10','F11','F12','F13','F14','F15','F16','F17','F18','F19','F20','F21','F22','F23','F
 In [5]:
          df = pd.DataFrame(data = Data, columns = columns)
 In [6]:
          # Map strings to integers
          Data["F20"] = Data["F20"].map({"Low": 99, "Very low": 100, "High": 101, "Very high": 102, "Medium": 103})
          Data["F27"] = Data["F27"].map({"USA":5,"UK": 6,"Europe":7,"Rest":8})
 In [7]:
          X = Data.drop('Target', axis = 1) # Excluding target variable
          y = Data['Target'] # target variable
 In [9]:
          # Split the dataset into train and test subsets
          X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.3, random_state = 0)
          # To train the model, we use X_train as the features and y_train as true labels.
          # To test the model, we use X_test as the features and y_test is used to validate the predicted labels
          # Test size is the percentage of data that should be used for testing/validation.
          print("Number transactions X_train dataset: ", X_train.shape)
          print("Number transactions y_train dataset: ", y_train.shape)
          print("Number transactions X_test dataset: ", X_test.shape)
          print("Number transactions y_test dataset: ", y_test.shape)
          Number transactions X_train dataset: (1050, 36)
          Number transactions y_train dataset: (1050,)
          Number transactions X_test dataset: (450, 36)
          Number transactions y_test dataset: (450,)
In [11]:
          # create a regressor: ExtraTrees Regressor
          from sklearn.ensemble import ExtraTreesRegressor
          X_train, X_test, y_train, y_test = train_test_split(X, y, random_state=0)
          reg = ExtraTreesRegressor(n_estimators=100, random_state=0).fit(X_train, y_train)
          reg.score(X_test, y_test) # Calculate the accuracy
          0.6925554721731707
Out[11]:
In [12]:
          # create a regressor: Decision Tree Regressor
          from sklearn.tree import DecisionTreeRegressor
           regre = DecisionTreeRegressor(random_state=0)
          regre.fit(X_train, y_train) # fit the data into model
          y_predict = regre.predict(X_test) # Predict the value of dataset on the test subset
          regre.score(X_test, y_test) # Calculate the accuracy
          0.2689172595321214
Out[12]:
In [13]:
          # create a regressor: Linear Regressor
          from sklearn.linear_model import LinearRegression
          lin = LinearRegression()
          lin.fit(X_train, y_train) # fit the data into model
          y_predict = lin.predict(X_test) # Predict the value of dataset on the test subset
          lin.score(X_test, y_test) # Calculate the accuracy
          0.5818213125282481
Out[13]:
```

ExtraTrees Regressor imparts the highest accuracy, i.e. 69%, whereas Decision Tree Regressor exhibits the least accuracy, i.e. 27%.

In []: