Name : Abdulrahman Almyman ID:

IDE: Jupyter Notebook because easy to use and you can divide you work into cells and review your mistakes fast and solving problems. (Note: sometime I use Colab when I am far from my PC )

Lib: pandas ( To Preparation dataset easily ) , scikit-learn (train\_test\_split, accuracy\_score, RandomForestClassifier, etc. )

1.1.4 :

**Chosen Algorithm**: Decision Tree Classifier because it is the highest Accuracy

**Hyperparameters**: I made it default ( RandomForestClassifier() )

1.1.5.:

-read xAPI-Edu-Data.csv by pandas.read\_csv

1- Encoded categorical variables using LabelEncoder

2- Separate features (X) and target (y)

3- Split the dataset into training and testing sets (e.g., 80% train, 20% test) with a random state set to 92 .

4- - using accuracy\_score and classification\_report to measure its performance on the test dataset also use ConfusionMatrixDisplay.

1.1.6.:

**1- IMPORT DATA FROM KAGGEL**

**2- Data Preparation**

3- Encode categorical features

4- # Separate features (X) and target (y)

5- # Split the data into training set (80%) and test set (20%) & RS =91 as the report says

6- # Initialize the Algorithm (chosen: the Random Forest Classifier )

7- # Train the model

8- # Make predictions

9 - # Calculate the accuracy,F1 score, and confusion matrix of the model

1.1.7:

Random Forest Classifier : Model accuracy: 83.33% , F1-score: 0.84

Decision Tree Classifier: Model accuracy: 77.08% , F1-score: 0.78

KNN Classifier : Model accuracy: 56.25% , F1-score: 0.57

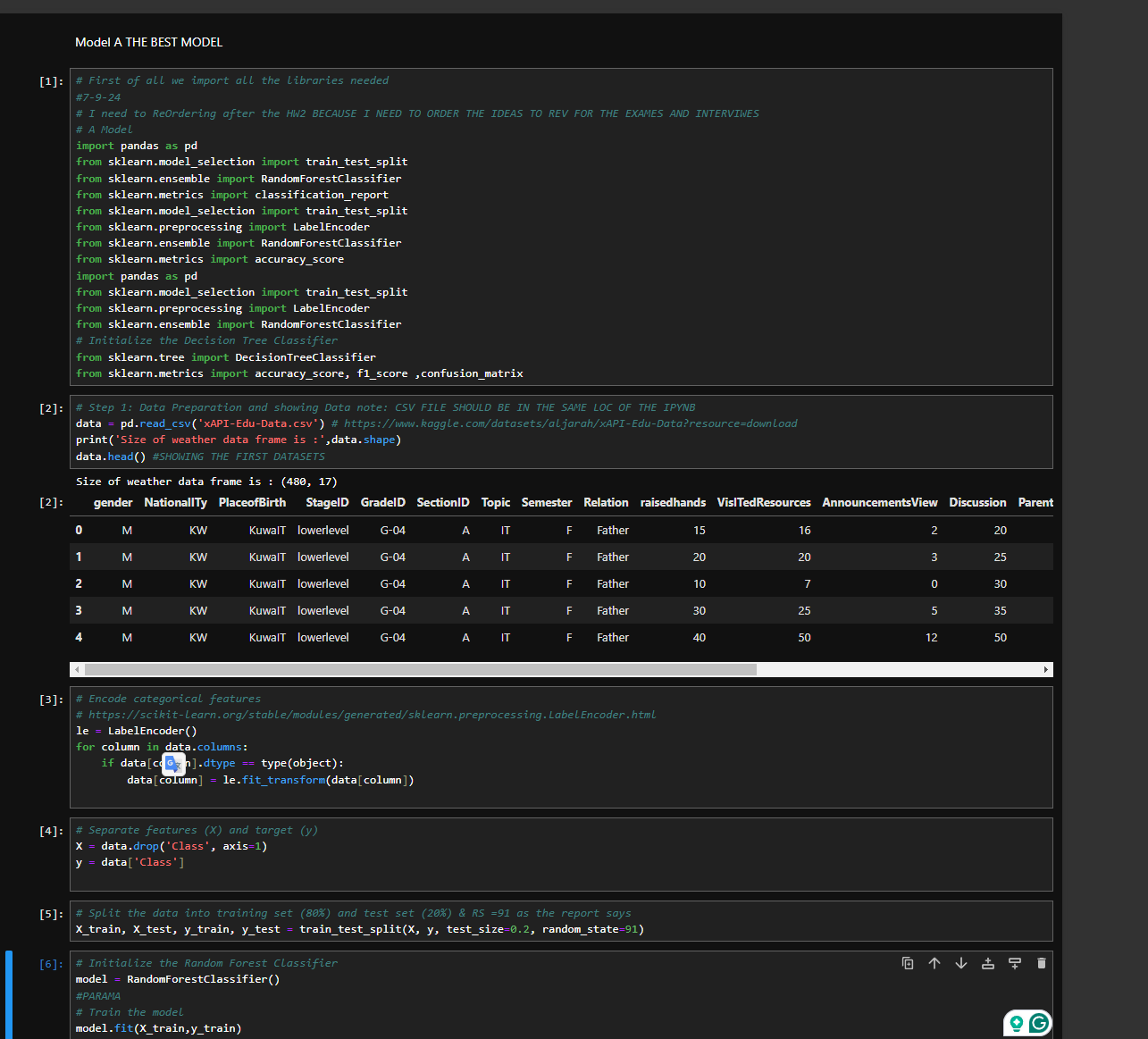
SVC: Model accuracy: 62.50%, F1-score: 0.63

Logistic : 76.04%, F1-score: 0.77

Last update today HW2\_1.1

​

Photo: ​



صورة تحتوي على نص, لقطة شاشة, برمجيات, برامج الوسائط المتعددة

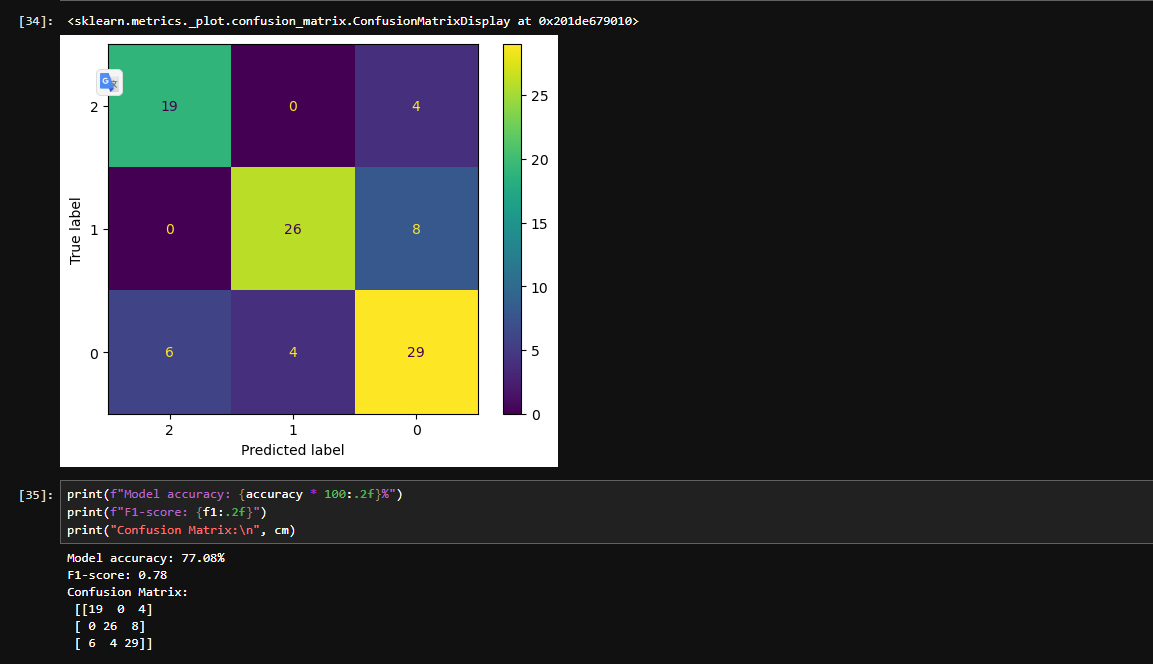
تم إنشاء الوصف تلقائياً

صورة تحتوي على نص, لقطة شاشة, برمجيات, برامج الوسائط المتعددة

تم إنشاء الوصف تلقائياً

The other results

DecisionTreeClassifier



KNeighborsClassifier

صورة تحتوي على لقطة شاشة, نص, برمجيات, برامج الوسائط المتعددة

تم إنشاء الوصف تلقائياً

SVC

صورة تحتوي على نص, لقطة شاشة, برمجيات, برامج الوسائط المتعددة

تم إنشاء الوصف تلقائياً

LogisticRegression

صورة تحتوي على نص, لقطة شاشة, عرض, برمجيات

تم إنشاء الوصف تلقائياً

Recources:

<https://aws.amazon.com/what-is/hyperparameter-tuning/?nc1=h_ls>

<https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.LabelEncoder.html>