**AML – 64016 – Assignment 1**

[sgudise@kent.edu](mailto:sgudise@kent.edu)

GitHub link – <https://github.com/Samueljames647/sgudise_64061/tree/main/Assignment%201>

1. A screenshot of a computer program

   Description automatically generatedRead, Run and understand the python code in the Sample project uses the iris data

A screenshot of a computer program

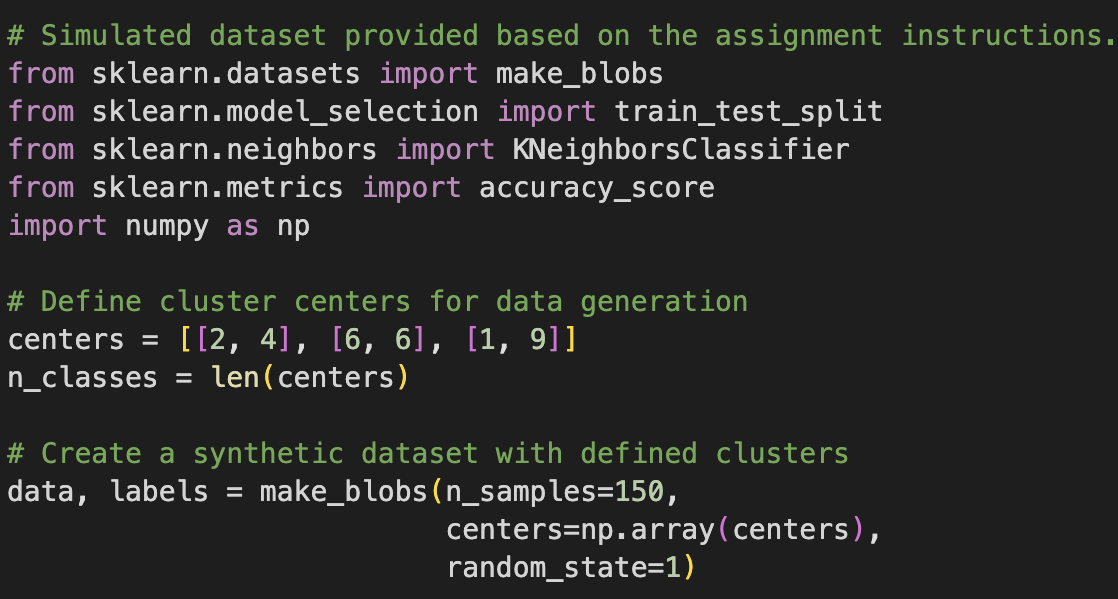
Description automatically generated

**NOTES:**

* Dataset Iris was transferred to Python, it was divided into training and testing datasets, with 80% for training and 20% testing.
* KNN (K-Nearest Neighbors) was trained by using the training dataset, after being obtained with default parameters.
* After that, the model was tested using the testing dataset, and the KNN algorithm's performance was assessed with various parameters.
* Each model's accuracy was computed and printed.

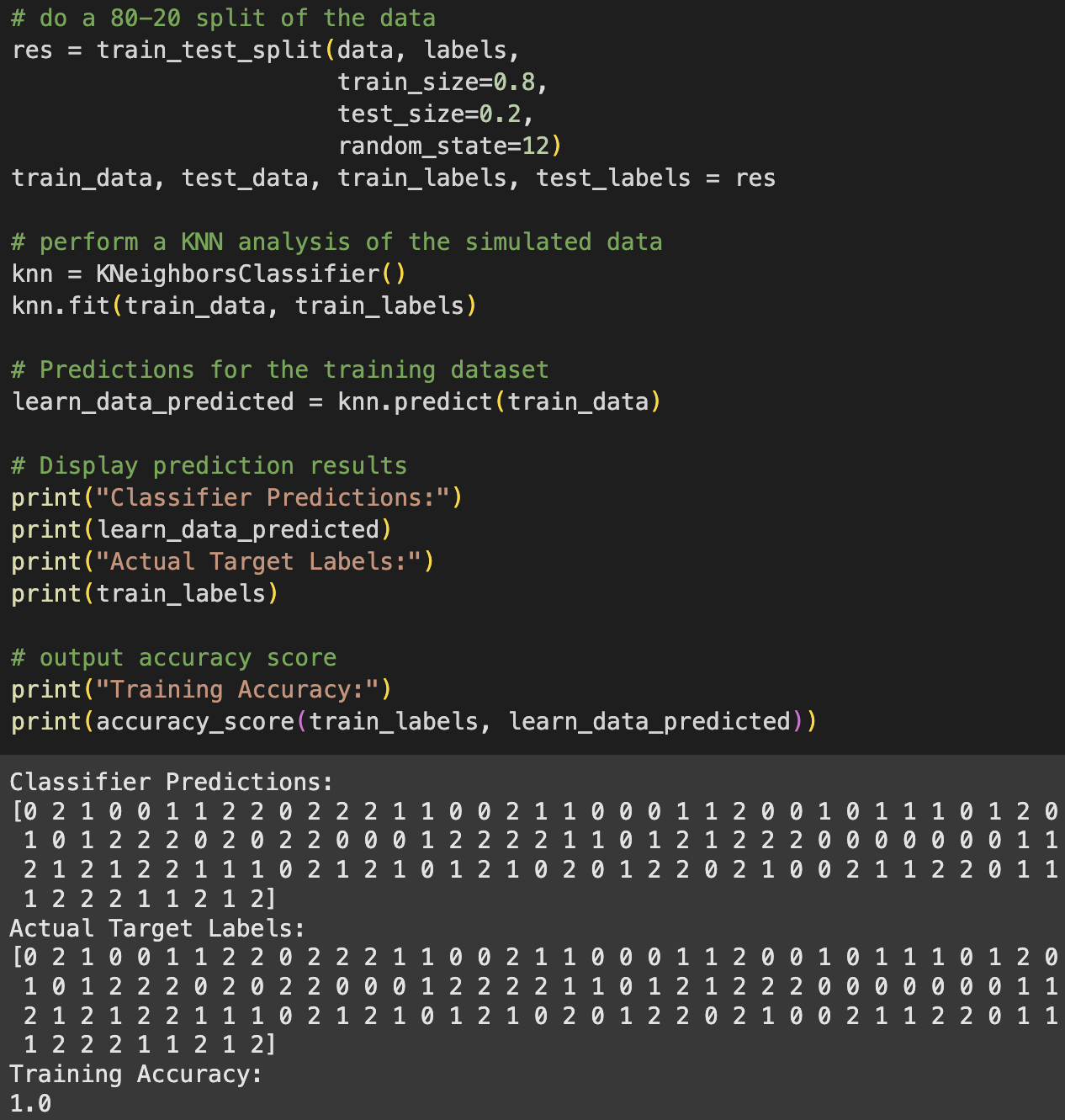
**Replicate the study by using a new simulated dataset.**

**Step-1**

****

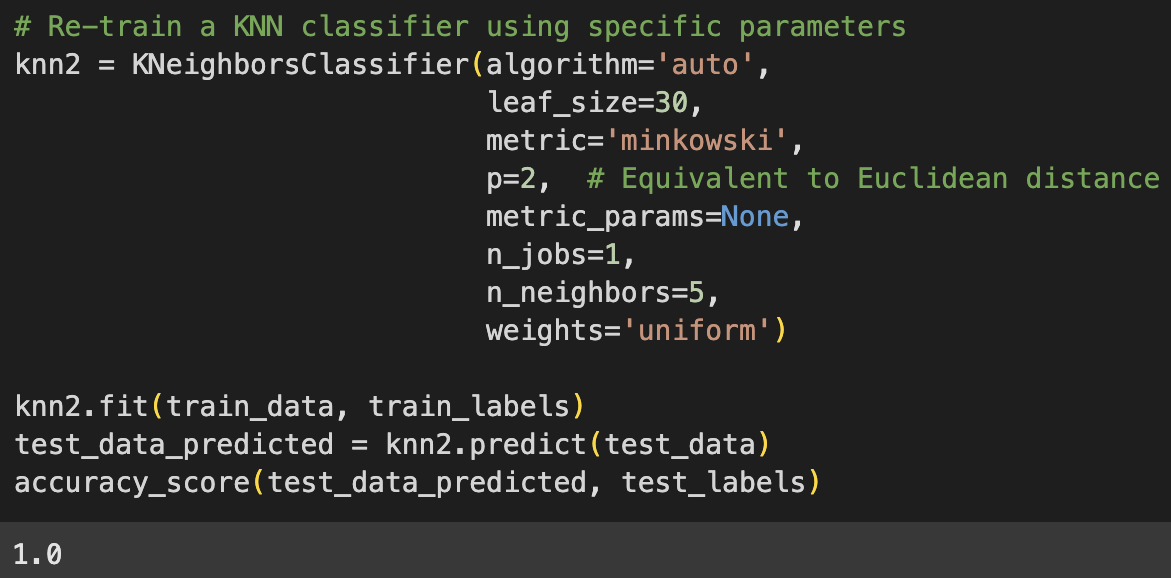
Created the set simulated dataset according to the instruction

**Step-2**

****

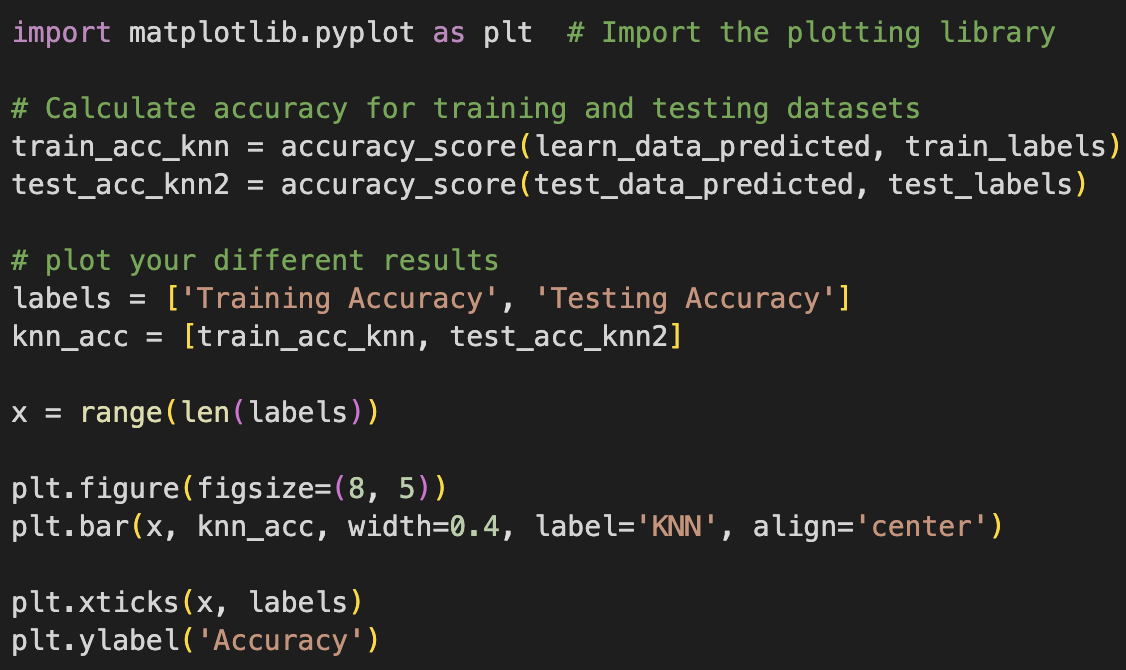
I divided the dataset into two parts: 80% for training the model and 20% for testing it. Then, I used KNN to train the model on the training data and tested it. The accuracy of the model on the training data was 100%

**Step-3**

****

I ran the KNN code with changes to the settings, using 2 for the Euclidean distance, and found the accuracy of the testing data to be 100%. Because the simulated dataset is evenly distributed, both the training and testing accuracy were 100%.

**Step-4**

****

I plotted the training and testing accuracy to create a visual graph for comparison. The graph below shows the results for graphical comparison.

