**INTRODUCTION:**

The objective of this analysis is to assess the model's accuracy using simulated data and a KNN (k-nearest neighbors) examination. Using the given dataset makeblob’s was imported and divided it into 3 clusters where the centers are at (2,2), (6,6) and (1,9).

Text

Description automatically generated

**METHOD:**

**TESTING AND TRAINING:**

From the sklearn model selection model, the training and testing split function was used to split the data into a training set which is 80% of the data and testing set which is 20% of the data.



**K-NEIGHBORS CLASSIFIER:**

From the sklearn.neighbors model, the KNeighbors classifier was used with K=5 and was setup on the training set of the data. By using the accuracy score method from the sklearn.metrics package was used to predict all labels of the test data using the classifier, the precision of the model was determined.

Text

Description automatically generated

To plot the findings matplotlib.pylot was used. The KNN classifiers decision function was well imposed on the simulated data’s plot with different colors for each class.

**Summary:**

The KNN classifier performed effectively, determining the highly accurate score of 0.966, which is an indication that the model classified the data appropriately. As the dataset got divided into 3 clusters it was generated well, which made it easy to categorize the real-world data with more intricate and overlap boundaries. The plot demonstrates that the classifier differentiated 3 groups successfully, but there were certain spots along the boundaries in between the classes where as the classifier is less sure but it would be negotiable.

Chart, scatter chart

Description automatically generated

The KNN classifier decision boundary shows the 3 different spaces defining that it is dependent on the distances to a K Nearest Neighbors in training set, which is usually caused by the density of training points in those areas.

The whole outcome is that the analysis represents how well the K Nearest Neighbors classification worked on the simulated data and given a proper presentation of how well it actually functions.

**REFERNCES:**

[https://colab.research.google.com/drive/1EmBEipQnpJxwuDtE7FAuaO4u1Awjg-iY?usp=sharingLinks to an external site.](https://colab.research.google.com/drive/1EmBEipQnpJxwuDtE7FAuaO4u1Awjg-iY?usp=sharing)

[**https://kent.instructure.com/courses/57954/pages/python-numpy?module\_item\_id=2944498**](https://kent.instructure.com/courses/57954/pages/python-numpy?module_item_id=2944498)

[**https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning**](https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning)