INTRODUCTION

This investigation's goal is to assess the model's precision using simulated data and a KNN k-nearest neighbors analysis. The sklearn's make blobs method was used to create the data. Using two and four centers, six and six centers, and one and nine centers, the datasets module was divided into three clusters.

METHODS

The data was divided into 80 and 20 percent where 80% is a training set and 20% is a test set. Utilizing sklearn train test split function. Model Selection Module. A KNN classifier with k = 5. was built for the training dataset using the KNeighbors Classifier classes from the sklearn.neighbors package. Sklearn's accuracy evaluation method. A metrics package was used to evaluate the accuracy of the model that was used with the classifier to forecast all the labels in the test data. The results were then plotted using the matplotlib.pyplot package. In beginning a plot of simulated data for every class was overlaid with a KNN classifier decision function.

RESULTS

The classification accuracy of KNN in the test data is known to be 0.9667, indicating that the model could correctly classify most of the test data. Plot of KNN classifier decision boundary and simulation data. The plot shows that the classifier was able to separate the three groups quite successfully although there are regions at the class limitations where the classifier is not so accurate.

DISCUSSION

The KNN classifier accomplished well on the simulated data, achieving a very accuracy of 0.9667. Data might have been easier to classify than actual data with more complex and overlapping class boundaries because they were generated with very well-defined clusters.