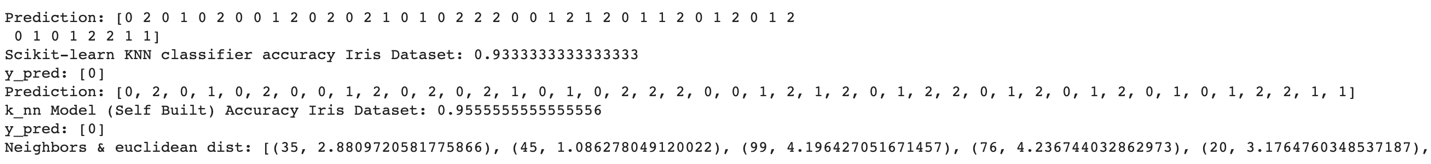
K-Nearest Neighbor – Implement Class Model from Scratch & Compare with Sklearn

What is K-Nearest Neighbors?

K-Nearest Neighbors is an algorithm used in classification and regression problems; however, this application uses the algorithm as a classifier. This is a form supervised learning where classification outputs are class membership however this specific algorithm is considered lazy learning. This object is assigned a class based on the classes of its neighbors based on the number of neighbors provided (K). In regression the outputs are the property values for the specified objects. The value is the average of the values of k nearest neighbors. The goal here was to create a model that takes in row values, computes the Euclidean distance between one another and output a prediction of the proper class. Then the results were compared to the sklearn model that can be imported with the Scikit-Learn library.

The class was implemented to fit the data, predict on the data, and display the data. There is a class definition that calculates the Euclidean distance used for the algorithm. Next the input data is fit into an X\_train and y\_train (target). Then doing a bunch of fancy math, calculating distances between points, class predictions, and more; the model is able to predict a class. Lastly, I provided a quick function to be able to display the data similar to the Sklearn model such that comparison can be easy for the three datasets tested on. These datasets are Iris, Breast Cancer, and Wine datasets from sklearn.

The results and comparison between the imported model and my model are shown below for the listed datasets:



Calendar

Description automatically generated

Chart, scatter chart

Description automatically generated

Results indicate our hand built model performs as well if not better in most rudimentary basics. Please feel free to take this code to the next level and see what you can make of it! Code can be found for the model, testing, and blog are on [github](https://github.com/blakelobato/CS-Data-Science-Build-Week-1). The medium article is [here](https://blakelobato1.medium.com/).