Machine Learning HW 2

In this assignment I implemented KNN algorithm using numpy, pandas, and collections. To run the program simply have, in the same directory, a folder called Data with relevant training and test files. Have the two files, HW2.py and KNN.py in the directory and run HW2.py which will process the data files given for HW2. The file HW2.py separates the data, creates an instance of KNN and then feeds each test data point to KNN then printing how many correct and false then printing a final accuracy. KNN itself has a constructor which takes in K and a distance metric. It has a fit function to actually save the training data, x and y. Then it has a predict method which takes in a data point, calculates the distance from that point to the rest in the training set and then sorts them by distance, smallest to largest, then grabs the first K of them, counts the label frequency using a dynamic dictionary then returns the labels with the largest frequency.

My KNN model with Ks between 2-6 had 100 percent accuracy for the given data set.