

## CMSC 170: Introduction to Artificial Intelligence

### Laboratory Exercise Journal

Accomplish this journal while working on or when you are done with the laboratory exercise for the week.

1. Problems encountered. Explain the **specifics** of the problem/s. *You may include code snippets and/or screenshots.* (Minimum of 2 sentences)

One problem I encountered while working on the exercise was calculating the distance of the csv files. At first I was thinking to append the distance to the element after computing. However, it would be very inefficient when solving for the k lowest distances for all the data in the csv files. Another problem I encountered was finding an efficient way to get the most common classification in the k nearest neighbors. My first thinking was to have a counter for the ones and zeros. However, this will have a problem in the second data set (fruits) since the classifications are not ones and zeros only.

2. How the problems were resolved and what are the **specific fixes** done. *You may include code snippets and/or screenshots.* (Minimum of 3 sentences)

For the distance problem, I was able to solve it by appending the index of that element along with the distance. I also created another variable called 'distances' which is a list of tuples where the first element is the index, and the second is the calculated distance. With this, I solved the problem on how I could find the index of a certain element. I was also able to sort out the 'distances' list using a built-in python function.

As for the classification problem, I created another list called 'classifications' where all the classifications of the k nearest neighbors are stored. From here, I used a python module that gets the most common elements in the list.

3. Learnings from the exercise / lesson. Explain in your own words. *Avoid merely listing laboratory topics like "I learned how to use Inheritance. I learned about Encapsulation."* Explain and analyze. (Minimum of 5 sentences)

I learned how to determine the classification of a data using the classification of the nearest neighbors to it. Using the euclidean distance, we can determine which ones are the nearest in that certain plot. From here, we get the most common classification among the nearest neighbors. And this will be the classification of our data. K Nearest Neighbors is a an example of supervised learning algorithm.