RuiMing Li
DS3000
March 27th, 2022
Assignment 9 Part 2

2.

Accuracy is how much data you label correctly in machine learning. It is the ratio of the number of items labeled correctly over the total number of items. An example would be if you had three dogs and you labeled one of them incorrectly as a cat through machine learning then the accuracy would only be 67%.

Precision is a measure of the correctly predicted positive observations to the total predicted positive observations. An example would be if you had three dogs and machine learning labeled all three as dogs, then you would have 100% precision.

Recall is the measure of the correctly predicted positive observations to the observations in actual class. An example would you had an ad blocker that is designed to block specifically ads about clothing, you would have high recall if the program was able to correctly locate the clothing ads and prevent you from seeing it.

An alternative to the train split method, we could use a k-fold mechanism which use all the data points in the data set as the training data and test data. K-fold separates the "dataset into multiple sets of indices with zero overlaps to pull random sets of data from the dataset."

https://towardsdatascience.com/how-to-select-a-data-splitting-method-4cf6bc6991da#:~:text=As%20an%20alternative%20to%20train,data%20from%20your%20data%20set.