Rikita Patangay
CS4375.003

Narrative Document

a. how SVM works, and how SVM kernels work, your impression of the strengths and weaknesses of SVM

SVM (Support Vector Machine) is a popular AI algorithm. SVM works by mapping data to a high-dimensional feature space so that the data points can be categorized. Kernels provide a window to manipulate data. Kernels work by transforming the training data so that a non-linear decision surface is able to transform into a linear equation in a higher number of dimension spaces. Some strengths of SVM are that it works really well with a straightforward margin of separation, it is effective in high dimensional spaces, it is also effective in cases where the number of dimensions is greater than the number of samples, and it is memory efficient. SVM is not very efficient with a larger dataset, the required training time is higher.

b. how Random Forest works, how the other 2 algorithms you used work compared to the simple decision tree, and what is your impression of the strengths and weaknesses of these ensemble techniques

Random Forest is known as a classification algorithm consisting of many sets of decision trees. It uses bagging and random selection when building each individual tree to try to create an uncorrelated forest of trees. Random Forest is more useful for situations with large data sets vs. a simple decision tree. It also avoids overfitting by utilizing multiple trees. The advantage is that this method uses training data to make predictions. Single data trees' disadvantage is that they lead to overfitting, and tend to learn the data too well.