P3 Writeup

Bill Tong

November 14, 2022

1 Training Adaboost

1.1 Question: What is the reason for using a decision tree stump rather than a decision tree with a greater depth? How does this differentiate adaboost from a random forest ensemble method?

The probability of the tree overfitting increases as the depth of a decision tree grows. Adaboost is different from a random forest ensemble method which builds a full tree, with each individual tree being fitted independently from other trees. In adaboost, it can be observed that there is a sequential structure for building the trees, with the weights from the previous tree affecting the current one. If overfitting occurs, it will also continue to occur as each stump is generated.

2 Testing Adaboost

2.1 Question: What would need to change to run an adaboost algorithm with a perceptron rather than a decision tree?

If we use a perceptron rather than a decision tree, then we wouldn't have to change the data and create duplicates to reflect the weights of each sample. This is because the perceptron algorithm and adaboost both deal with weights in a similar way. The perceptron functions by using epochs, stepping through all the data to update weights, and the adaboost steps through all the data to create classifiers. With the perceptron however, recall that the order in which the samples are put in will effect the output.