Write up
Project 3

1. Training Adaboost

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?

Ensemble classifiers consist of multiple classifiers that do not make the same mistake. Ensembles reduce the variance of the classifiers. Variance refers to the rate of change in the performance of the model when slight changes are made to the training data set.
An ensemble of decision tree stumps would ask just one question for each feature and therefore, they are more stable with low variance. The predictions will only change slightly with changes in the training dataset. With variance being lower, decision tree stumps can generalize to regions not covered by the training dataset. Using deeper decision trees in Adaboost is possible and may yield higher accuracy, however, it would be expensive to execute due to the high dimensional space. Therefore, it is better to use decision stumps since they are shorter trees and straightforward.

Random forest ensemble method uses decision trees as well but they are deeper and are all independent of each other which means the results do not depend on the outcome of the previously trained tree. Whereas, in Adaboost, each decision tree stump depends on the previous stump as it gives more importance to the previously misclassified samples with the help of higher weights.
Random forests consist of equal weights for each decision tree while voting for the final classification as opposed to adaboost decision tree stumps that have higher weights (adaptive parameter) for stumps with lower error and higher accuracy.

2. Testing Adaboost

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

To run the adaboost using perceptrons, we will need to add perceptrons as weak learners with a fixed number of epochs. The samples will also need to be sampled using the new weights in order to bias the training process towards the weak cases (misclassified samples).