Adaboosting Using Decision Stump

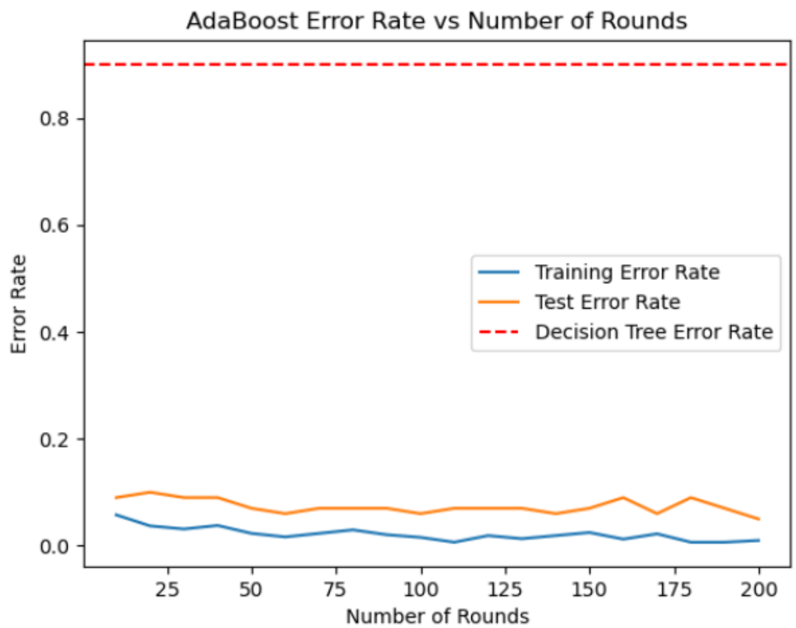
Findings

Number of Rounds	Training Error Rate	Training Accuracy	Test Error Rate	Test Accuracy
1	0.06	0.942339373970346	0.09	0.91
50	0.03	0.9769357495881383	0.07	0.93
100	0.02	0.984349258649094	0.06	0.94
150	0.03	0.9752883031301482	0.07	0.93
200	0.01	0.9901153212520593	0.05	0.95

Main Takeaways

Our model's generalization performance exhibits some overfitting, as indicated by the lower test accuracy compared to the training accuracy. The model's performance significantly improves with lower training errors, which is expected since the model tends to overfit to the training data. Specifically, our single decision tree achieved a training accuracy of 1.0, while the accuracy increased to 0.99 when used in conjunction with adaboost for 200 iterations. However, after 170 iterations, the error rate exceeded 0.5 in some cases, triggering a weight reset and resulting in a decrease in accuracy by 0.1. Our train error is notably lower than our test error. Increasing the number of iterations would likely result in a training accuracy of 1.0 and an error of 0. However, our test accuracy increased from 0.91 to 0.95. As the test accuracy is lower than the train accuracy, our model is overfitting. Nonetheless, adaboost with decision stump achieves an accuracy of 1.0. Although decision stump is a weak learner, adaboost optimization can enable it to behave like a strong learner, such as a decision tree. In summary, our single decision tree model is overfitting since its test accuracy is 0.91, and its train accuracy is 1.0.
The decision tree's training error rate is significantly lower than its test error rate, which suggests that it is probably overfitting the training data and not generalizing effectively to new data. This implies that while the variance of the decision tree is high (because it does not generalize well to new data), the bias of the decision tree is low (because it fits the training data well). The AdaBoost model's test error rate with decision stumps is invariably lower than the test

error rate with the decision tree, indicating that it is more generalizable to fresh data.This may suggest that while the variance is smaller, the AdaBoost model's bias is larger than the decision tree (because it may not suit the training data as well) (since it generalizes better to new data). The AdaBoost model's test error rate has been trending downward with increasing rounds, which shows that when more decision stumps are added, the model is becoming less overfit and generalizing to new data more effectively. This is in line with the notion that a model's complexity can both reduce bias and increase variation. However, ensemble techniques like AdaBoost can help to decrease variance by mixing simpler models.



In this project, we applied AdaBoost using decision stumps learned using the Gini index as the weak learners to classify handwritten digits 3 and 5 from the NIST dataset. We also estimated test error using a single decision tree learned using any library without any pruning. In this report, we present the results obtained from these experiments and analyze the behavior of both the training set error and the test set error as we increase the number of rounds.