EECS 440: Project 4 Writeup

Austin Feydt (apf31) and Mathanki Singaravelan (mxo174)

November 18, 2017

a) **Learning Algorithm: Decision Tree:**

Voting:

Base Learner: 0.988636363636%

Bagged Learner: 0.984090909091%

Boosted Learner: 0.988636363636%

Volcanoes:

Base Learner: 0.729717615419%

Bagged Learner: 0.729269385926%

Boosted Learner: 0.729717615419%

Spam:

Base Learner: 0.663870156283%

Bagged Learner: 0.691230212%

Boosted Learner: 0.663870156283%

As we can see, there doesn't seem to be a discernible difference between the decision stump and the ensemble methods implemented for this project, thus the t-tests aren't really necessary.

b) As we can see in Figure 1, there is an interesting shape in these graphs. As the number of iterations begins to increase, we see that the accuracy in both graphs actually begins to decrease (by a very small amount). Eventually, the accuracy begins to increase again, showing the overfitting of the data.

c) We can see in Figure 2 that the number of iterations of boosting has a small impact on both of the problems that we tested it on.

d) In Figure 3, we see how introducing noise into our dataset effects the power of bagging. For a small amount of noise, it seems that bagging is able to do still a pretty good job at classification. However, as we keep increasing our p value, the accuracy begins decreasing quickly, showing that perhaps bagging isn't the most resilient to noise.

e) In Figure 4, we see how introducing noise into our dataset effects the power of boosting. Similarly for bagging, small amounts of noise seem to not effect the ensemble too much. And as p continues to increasing, accuracy also decreases, but not quite as drastically as with bagging. This suggests that boosting can better deal with extraneous noise than bagging.

Across all 4 projects, one very common trend we've noticed is how much harder it can be to properly train learning algorithms on continuous attributes than just nominal values. For each project, we found ourselves having to take the continuous case into account every step of developing our algorithm, sometimes having to preprocess them and convert them into nominal attributes, thus losing precision. The voting dataset was a great dataset to have, because everything was nominal and it was a really small dataset. However, it gave us false hope that our algorithms would also run well on these larger, more complex data sets (like volcanoes and spam). In the real world, data isn't as nice and neat as it was in voting, which was a really important thing to acknowledge when our accuracies weren't as high as we were hoping.

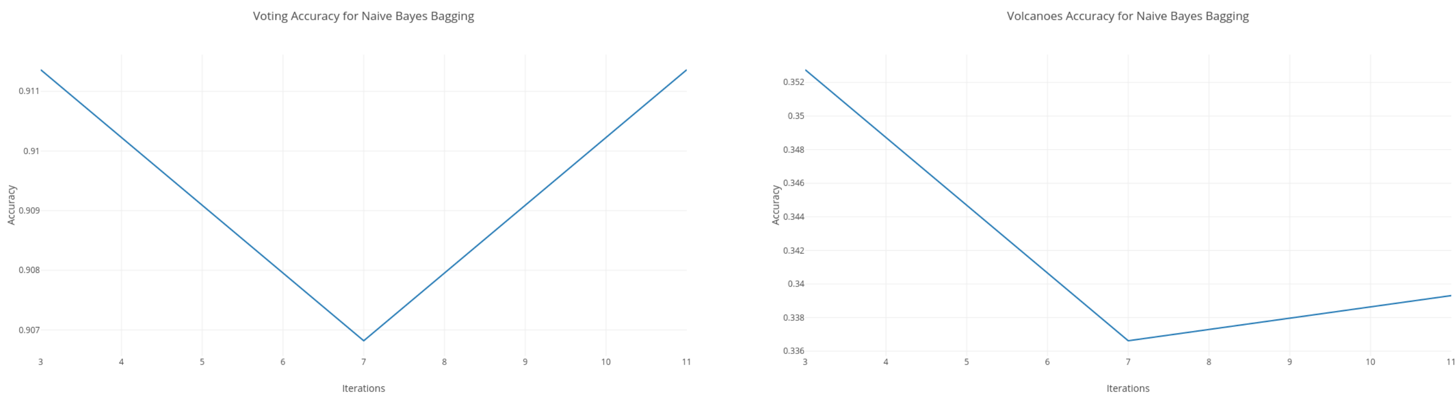

Figure 1: Part b): Accuracy of Bagging by varying number of weak classifiers in ensemble

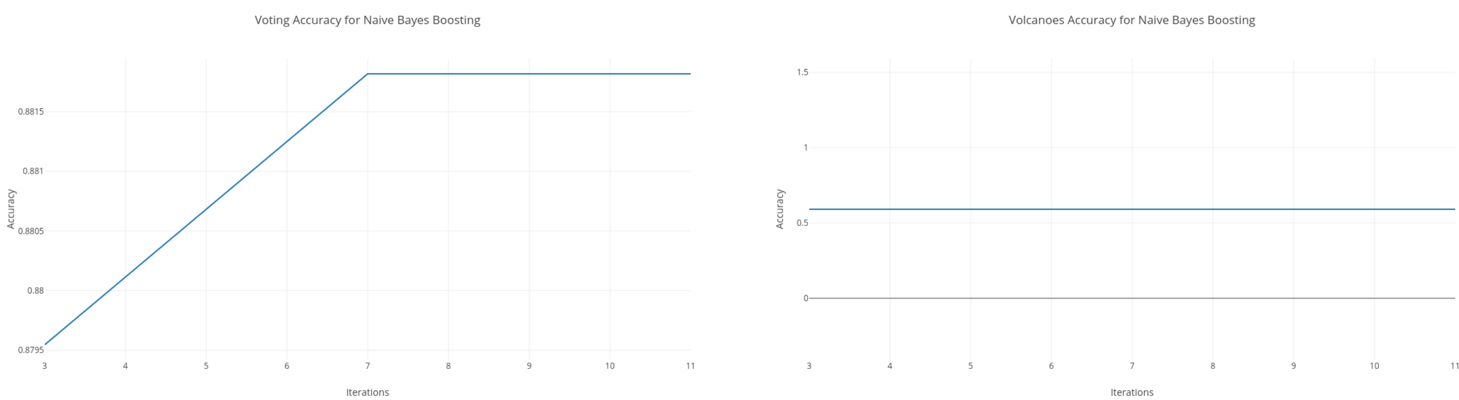

Figure 2: Part c): Accuracy of Boosting by varying number of weak classifiers in ensemble

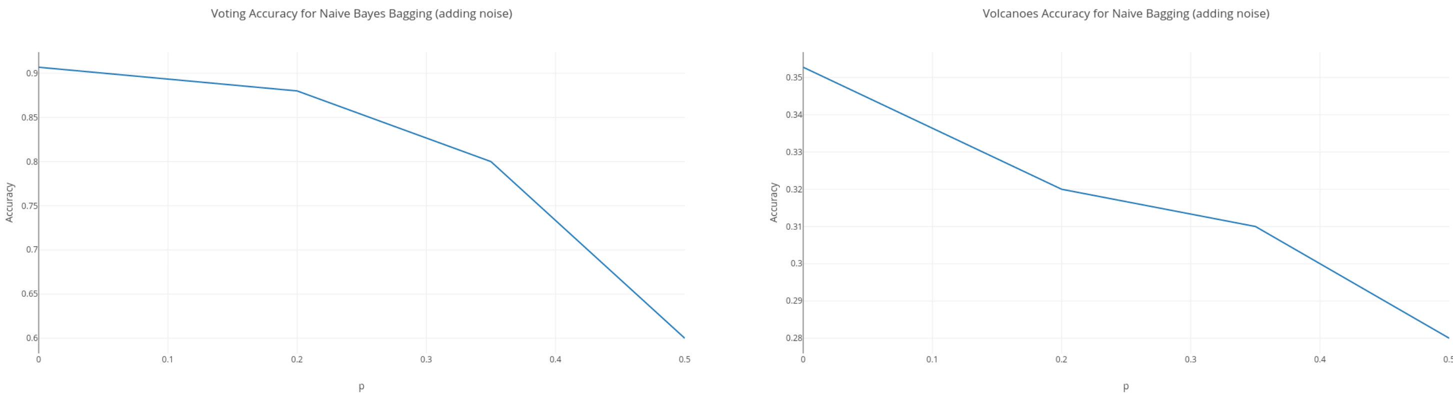

Figure 3: Part d): Accuracy of Bagging by introducing noise

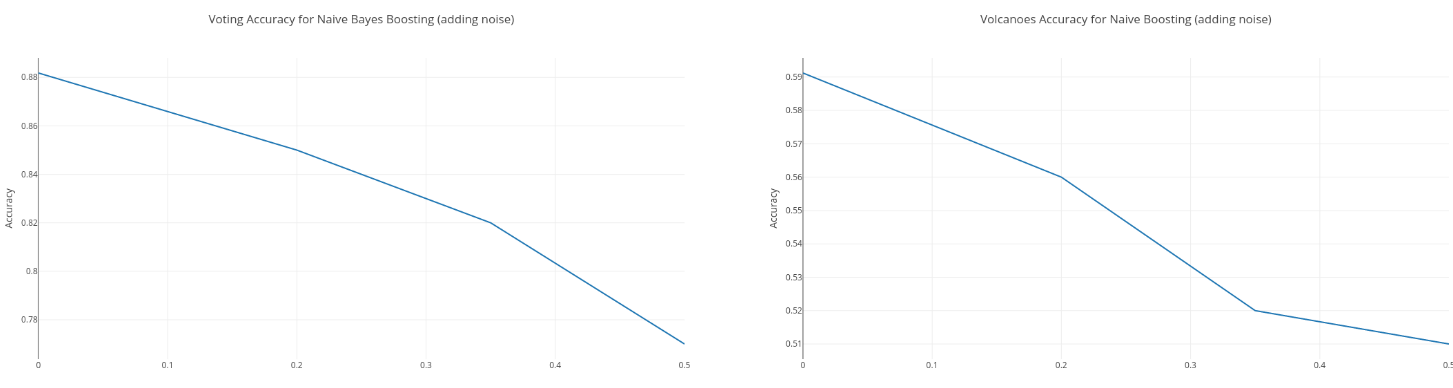

Figure 4: Part e): Accuracy of Boosting by introducing noise