EECS 440 Machine Learning Programming Problem 3 Writeup

Sixiao Zhang sxz603, Jianzhe Zhang jxz851, Junyi Zheng jxz990

For all questions, we use naïve bayes. We choose number of bins to be 5, and m, which is for m-estimate, to be 10.

(a) For all problems and any learning algorithm, compare the accuracy of the ensemble versions (10 iterations) to the base learners. Produce a table with the accuracies of the base learner, the bagged learner, and the boosted learner. Perform paired *t*-tests to determine if any of the ensemble methods are significantly better than the base learner with 95% confidence.

Table 1 shows the results for different datasets.

Table. 1 Accuracies of bagging, boosting and base on different datasets

	Voting	Volcanoes	Spam
Bagging	0.973	0.703	0.617
Boosting	0.975	0.747	
Base	0.975	0.724	0.678

For voting, the result of t-test with 95% confidence is shown below.

```
For bagging and the base, the interval is [-0.004, 0.008], the null hypothesis can't be rejected
For boosting and the base, the interval is [-0.025, 0.025], the null hypothesis can't be rejected
```

For volcanoes, the result of t-test with 95% confidence is shown below.

```
For bagging and the base, the interval is [0.007, 0.035], the null hypothesis can be rejected
For boosting and the base, the interval is [-0.093, 0.046], the null hypothesis can't be rejected
```

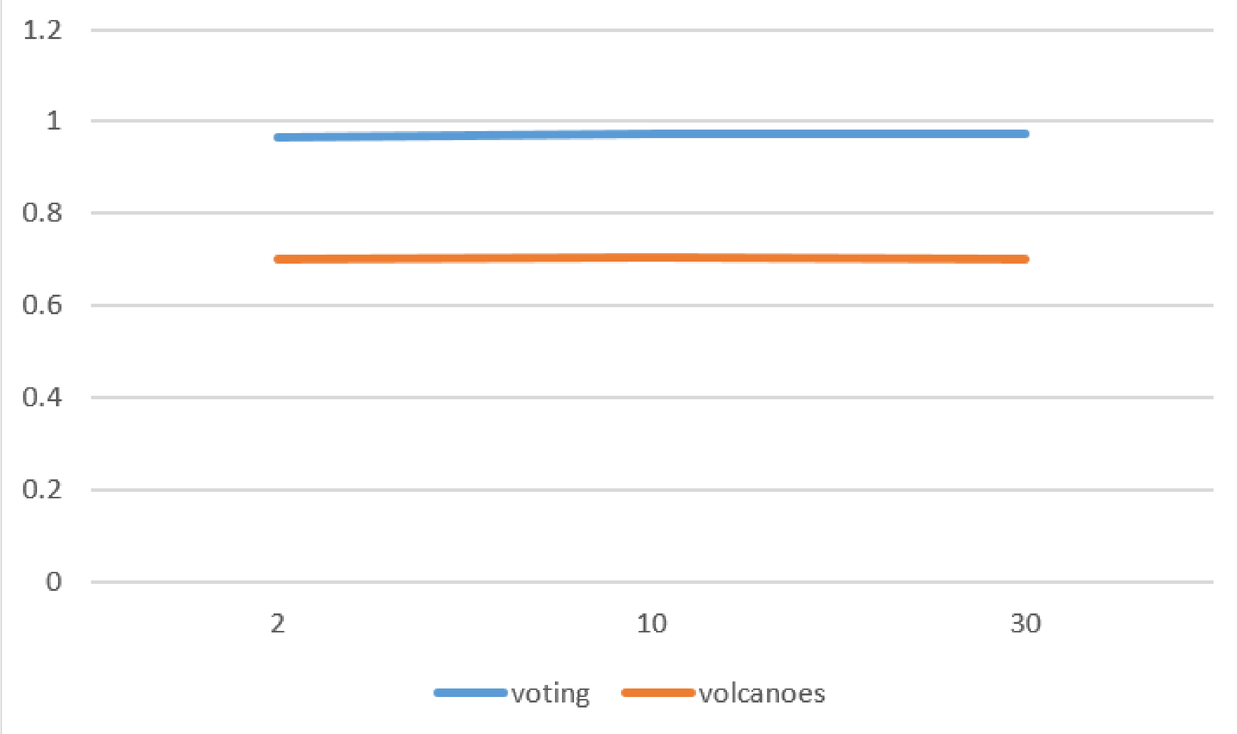
For spam, the result of t-test with 95% confidence is shown below.

(b) For any two problems and any learning algorithm, evaluate how the accuracy of bagging changes with the number of iterations. Pick at least three iteration values between 2 and 50, and plot the accuracy on a graph. Do you see any difference by problem?

Table 2 shows the results of bagging with different iterations.

Table. 2 Accuracies of bagging using voting and volcanoes with different iterations

	ITER=2	ITER=10	ITER=30
Voting	0.966	0.973	0.975
Volcanoes	0.701	0.706	0.702



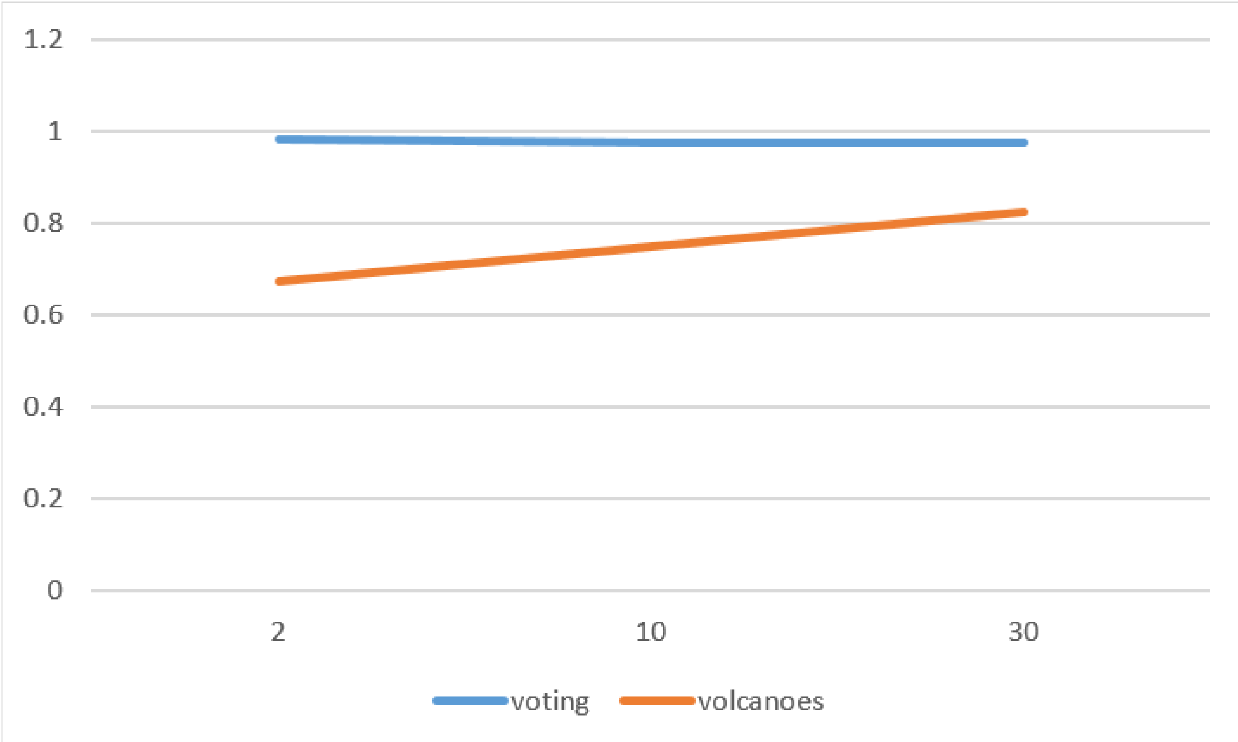
It seems that the accuracy doesn't change on either problem.

(c) Repeat (b) for boosting.

Table 3 shows the results of boosting with different iterations.

Table. 3 Accuracies of boosting using voting and volcanoes with different iterations

	ITER=2	ITER=10	ITER=30
Voting	0.980	0.975	0.973
Volcanoes	0.672	0.747	0.825



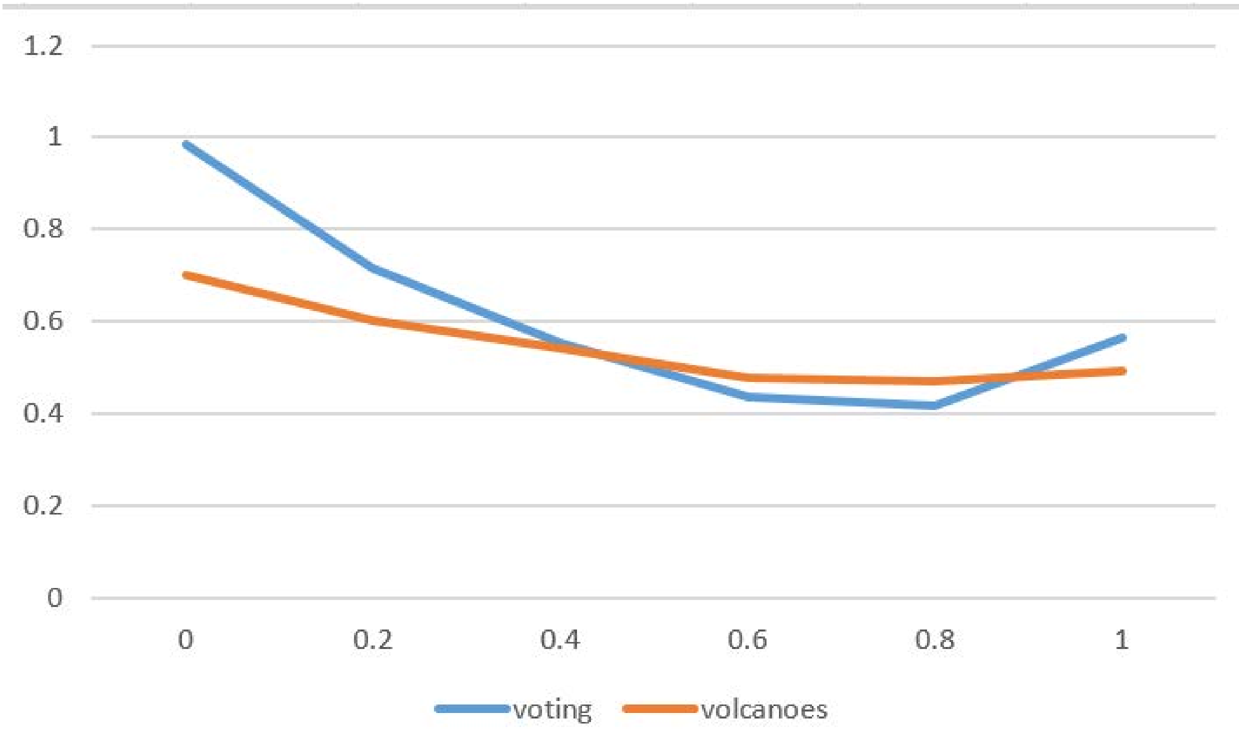
For voting, the accuracy doesn't change when iteration increases. But for volcanoes, the accuracy increases significantly when iteration increases. This might be because that voting is small and accurate so there is not much improvement. But volcanoes is large and not very accurate, so it improves a lot when iteration increases.

(d) Evaluate the sensitivity of bagging to noise as follows. When training, after constructing the training sample, flip an example's label with probability *p*. Then use this noisy sample in your bagging algorithm and evaluate the resulting classifier on the usual (noise free) test set. For any two problems and any learning algorithm, plot a graph with *p* on the x-axis and the test-set accuracy of bagging (30 iterations) on the y-axis. You can use results from the previous questions for a *p=0* point. Discuss how resilient bagging is to noise based on your observations.

Table 4 shows the results.

Table. 4 Sensitivity of bagging to noise

	P=0	P=0.2	P=0.4	P=0.6	P=0.8	P=1
Voting	0.984	0.717	0.553	0.438	0.419	0.566
Volcanoes	0.702	0.603	0.543	0.478	0.470	0.495



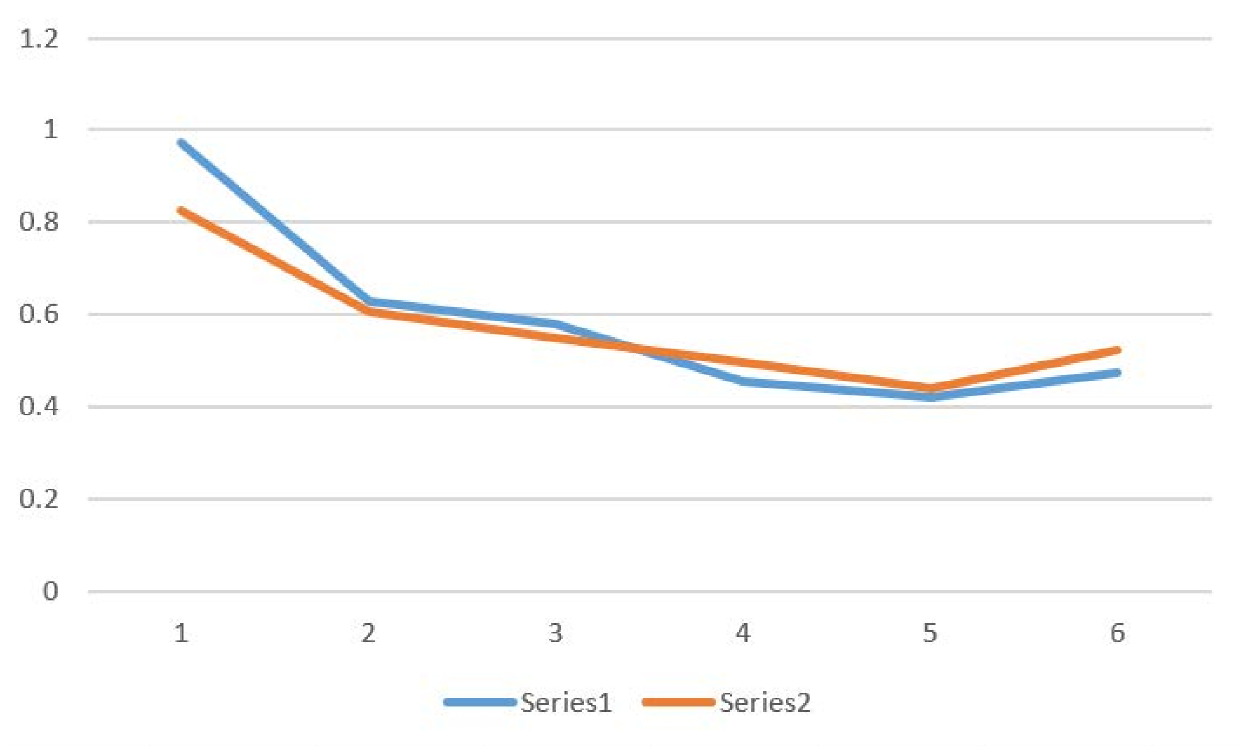
From the plot, we can see that when p is small, about 0.3 and smaller, the accuracy decreases significantly when p increases. When p is large, about 0.3 to 0.8, the accuracy decreases slowly when p increases. But when p is above 0.8, the accuracy slightly increases.

(e) Repeat (d) for boosting.

Table 5 shows the results.

Table. 5 Sensitivity of boosting to noise

	P=0	P=0.2	P=0.4	P=0.6	P=0.8	P=1
Voting	0.973	0.629	0.580	0.456	0.421	0.477
Volcanoes	0.825	0.606	0.550	0.496	0.443	0.526



From the plot, we can see that when p is small, the accuracy changes significantly when p increases. When p is large, the accuracy changes slowly when p increases.