EECS 440: Machine Learning Fall 2015 Write-up for Project 3, assigned 10/13. Max points: 20.

Yingcheng Sun ID: yxs489

1. Answer the following questions:
   (a) What is the accuracy of naïve Bayes and logistic regression on the different learning problems? For each problem, perform a t-test to determine if either is superior with 95% confidence.

   Answer: Below is the accuracy from the experiments, run with self-adaptive parameter m and λ for naïve Bayes and logistic regression. The data sets were run with 5-fold cross validation and all results averaged.

Accuracy	Voting	Spam	Volcanoes
Naïve Bayes	0.979	0.760	0.780
logistic regression	0.980	0.698	0.769

   Before doing T-test, we need to examine the homogeneity of variance between two data sets. We chose F-test to check the homogeneity of variance.

Standard Variance	Voting	Spam	Volcanoes
Naïve Bayes	0.008	0.008	0.009
logistic regression	0.013	0.008	0.123

   From the table we can see that the standard variance for Spam is the same, then calculated P – value which is 0.085695 larger than 0.5 and we can say that two sets of data are not significantly different for the results of Spam problem.

   (b) Compare these two methods with your previously implemented methods (trees, perceptrons, ANNs). Discuss which, if any, seems to be better on the provided learning problems in terms of (i) accuracy, precision, recall and the area under ROC, (ii) runtime, and (iii) code complexity and difficulty of implementation. Note that this question is about general trends---it may not happen that any one algorithm will turn out to be superior across all problems (cf. the NFL theorem). So use your judgment to determine "betterness." For methods that appear about equal in terms of accuracy, look at the precision and recall and try to explain any patterns you see.

   Answer: for the voting problem

	Accuracy	Precision	Recall	AROC	Runtime(s)
Naïve Bayes	0.979	0.967	0.780	1.00	0.042
logistic regression	0.950	0.967	0.769	0.989	0.012
Perceptrons	0.980	0.967	0.769	0.989	0.242
ANNs	0.984	0.995	0.969	0.989	0.628
trees	0.985	0.745	0.769	0.889	0.528

for the Spam problem

	Accuracy	Precision	Recall	AROC	Runtime(s)
Na ïve Bayes	0.625	0.625	1.000	0.447	3.17
logistic regression	0.697	0.732	0.818	0.447	1.167
perceptrons	0.615	0.625	1.000	0.795	1.167
ANNs	0.525	0.625	0.810	0.795	9.167
trees	0.225	0.625	1.000	0.795	10.167

for the Volcanoes problem

	Accuracy	Precision	Recall	AROC	Runtime(s)
Na ïve Bayes	0.625	0.625	0.982	0.751	4.852
logistic regression	0.746	0.710	0.564	0.331	6.838
perceptrons	0.746	0.760	0.564	0.761	8.858
ANNs	0.633	0.730	0.680	0.931	5.957
trees	0.633	0.770	0.600	0.331	1.957

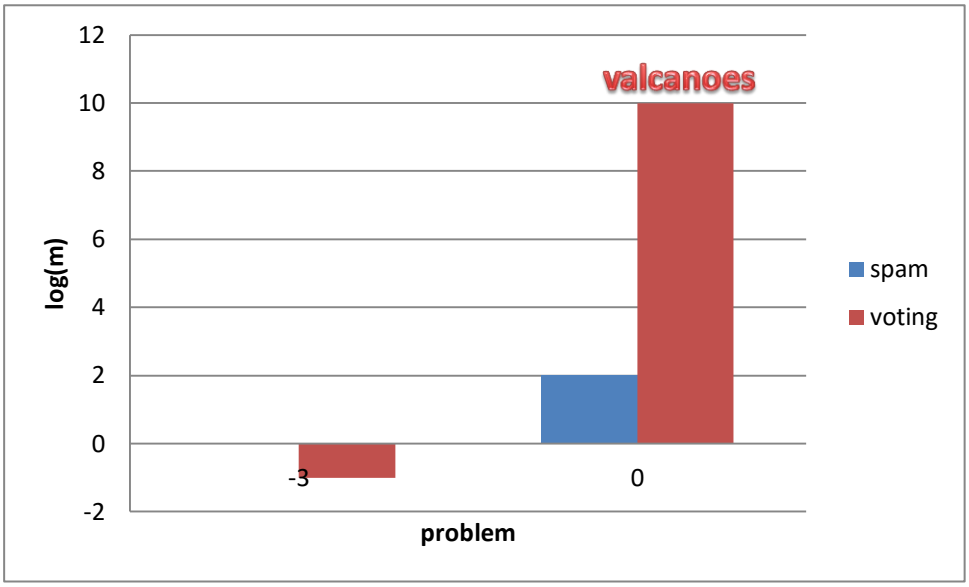
For the three problems: from the three tables we can see that all the machine learning algorithms have better performance for the voting problem, worse performance for the other two problems.

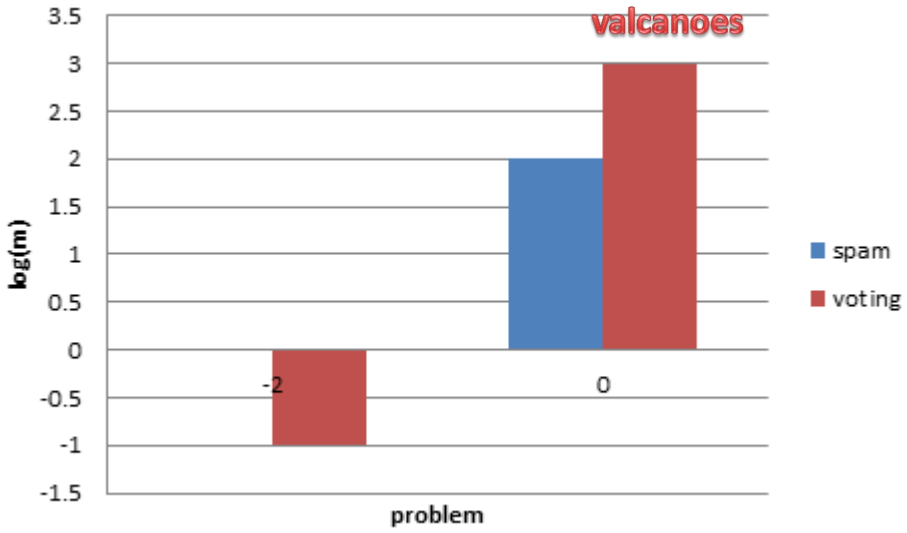
For the five methods: Na ïve Bayes, logistic regression and perceptrons have better runtime and Accuracy and Recall, and need less training time than the other two.
The performance of ANNs depends on its number of units in the hidden layer and times of iteration. For less units and iterations, ANN does not has good result. But for 5 units and more than 100 iterations, it has excellent testing results performance. Decision tree is similar to the Perceptrons with similar testing results.

The most difficult coding mothod is decision tree, and the easiest coding method is logistic regression and ANNs.

(c) Plot the choices of m and λ for the different folds and datasets on a graph with the datasets on the x-axis, the average choice on the y-axis with max and min values indicating the max and min choices on that dataset. Use a log scale for the y-axis. Discuss what these choices imply about the data and the algorithms.

Answer: for the m value





The data in the volcanoes dataset are more sparse.

(d) Compare the performance of the tuned classifiers with the MLE classifiers. Do you notice a benefit to tuning the parameters in this way?

Answer:

From the table we can see that it is more flexible for the tuned classifiers than MLE to choose the parameter and therefore it has better testing results and performance.