10601-HW5 REPORT

Name: Yi Song
ID: yis1

- Did you receive any help whatsoever from anyone in solving this assignment? Yes / No. If you answered 'yes', give full details: No.

- Did you give any help whatsoever to anyone in solving this assignment? Yes / No. If you answered 'yes', give full details: No.

- **1.1**

- **Specify what classifier you use, and report their 95% and 99% confidence interval on number of partition 2 and 10.What do you observe in the result?**

- In this part of hw, I used Naïve Bayes as my classifier. And 5-1.mat as data.

	Accuracy	0.95 confidence	0.99 confidence
Partition 2	0.52	0.3815~0.6585	0.3377~0.7023
Partition 10	0.60	0.2964~0.9036	0.2003~0.9997

As I randomly choose the test data set, so the accuracy value fluctuates a lot. In this figure we can find that 0.99 confidence interval is larger than 0.95 confidence interval.

- **1.2**

- **Specify what classifier you use, and report their 95% and 99% confidence interval for 2 and 10 cross sets.What difference do you see from 1.1?**

- In this part of hw, I used Naïve Bayes as my classifier. And 5-1.mat as data.

	Accuracy	0.95 confidence	0.99 confidence
Cross Sets 2	0.52	0.4221~0.6179	0.3911~0.6489
Cross Sets 10	0.50	0.4020~0.5920	0.3710~0.6290
Cross Sets 10(2)	0.56	0.4627~0.6573	0.4319~0.6881

The confidence interval is smaller than 1.1. With the test data number increasing, the confidence of Accuracy increase.

- **2.1**

- **Specify what two classifiers you use, report their accuracy and p-value under one-tailed test and two-tailed test for drawing 1/5 of instances as testing set and choose k to be 10.**

  In this part of hw, I choose Naïve bayes and perceptron as my classifiers. And 5-2.mat as data.

	Accuracy1	Accuracy2	P-value
One Tailed	0.76	0.54	0.0034
Two Tailed	0.76	0.54	0.0067

  As P-value<0.05 so the null hypothesis was rejected. Classifier1 is better than classifier 2.

- **2.2**

- **Specify what two classifiers you use, report their accuracy and p-value under one-tailed test and two-tailed test with k=10. Compare with the result you got from 2.1.**

- In this part of hw, I choose Naïve bayes and perceptron as my classifiers. And 5-2.mat as data.

	Accuracy1	Accuracy2	P-value
One Tailed	0.6960	0.6560	0.0977
Two Tailed	0.6960	0.6560	0.1954

As the gap of two accuracy value is not large. So the P-value turns out to be quite larger than 2.1. P-value in this test is larger than 0.05, so the hull hypothesis was not rejected.