Homework 3 for LING 572

Haotian He

January 25 2014

Collaborators: Haotian He and Jonggun Park

1   Q1

The training accuracy is 0.9444444444444444.
The test accuracy is 0.8966666666666666.

2   Q3

Table 1: Results of my Bernoulli NB model

cond_prob_delta	Training accuracy	Test accuracy
0.1	0.9592139414163886	0.9124579124579124
0.5	0.9555061179087876	0.9023569023569024
1.0	0.9514275120504264	0.8956228956228957
2.0	0.9421579532814238	0.8787878787878788

3   Q4

Table 2: Results of my multinomial NB model

cond_prob_delta	Training accuracy	Test accuracy
0.1	0.9577308120133482	0.9090909090909091
0.5	0.9525398591027067	0.9023569023569024
1.0	0.9480904708935854	0.898989898989899
2.0	0.9406748238783834	0.8922558922558923

4   Q5

Table 3: Results of my multinomial NB model with binary features

cond_prob_delta	Training accuracy	Test accuracy
0.1	0.9595847237671487	0.9124579124579124
0.5	0.9562476826103078	0.9057239057239057
1.0	0.9529106414534668	0.898989898989899
2.0	0.9477196885428254	0.898989898989899

5   Q6

(a):
1. All these three tables show, Bernoulli NB model, multinomial NB model, and multinomial NB model with binary features share the same trend that both the training and test accuracies go down, corresponding to the increase of the conditional probability delta value.
2. The accuracies of Bernoulli NB model go down (0.0336700336700336) a little more than multinomial NB model (0.0168350168350168), and even more than multinomial NB model with binary features (0.0134680134680134).
3. When the conditional probability delta equals to 0.1, the test accuracies of Bernoulli NB model and multinomial NB model with binary features are the same and higher (0.9124579124579124) than that of multinomial one (0.9090909090909091).
4. When the conditional probability delta equals to 2.0, the test accuracies of multinomial NB model with binary features stay the highest (0.898989898989899).
5. The testing accuracies of my multinomial NB model with binary features stay the same for both the conditional probability delta 1.0 and 2.0.

(b):