

LING 572 Homework 5

David McHugh and Chris Curtis

14 Feb 2013

Question 1

The commands we used to import the scripts into MALLET binary format are:

```
% mallet import-svmlight --input /opt/dropbox/12-13/572/hw5/examples/train2.vectors.txt \
  --output q1/train2.vectors
```

```
% mallet import-svmlight --input /opt/dropbox/12-13/572/hw5/examples/test2.vectors.txt \
  --output q1/test2.vectors --use-pipe-from q1/train2.vectors
```

After building the binary vector files, we trained and tested the classifier using the following command:

```
% vectors2classify --training-file q1/train2.vectors --testing-file q1/test2.vectors \
  --trainer MaxEnt --output-classifier q1/m1
```

Finally, we created the text representation of the model as follows:

```
% classifier2info --classifier q1/m1 > q1/m1.txt
```

Using these commands, we achieved the following accuracies:

	accuracy
Training	0.968519
Test	0.826667

Question 2

When running our MaxEnt classifier implementation via

```
% maxent_classify.sh test2.vectors.txt q1/m1.txt q2/res > q2/acc
```

We achieved a test accuracy of 0.826667, which is identical to the result from MALLET. The results are identical because the model is exactly the model MALLET trained, and so applying the MaxEnt decoding formula results in the same class probabilities for each instance; thus, since all class label decisions are the same the accuracy is identical.

Question 3

Nothing to report for this question.

Question 4

Nothing to report for this question.