

At first we write a code to generate mallet input file.

This file is like this:

```
1 l1 f1 سلام
2 l2 f1 سلام
3 l1 f1 ايا ف2 خوبي
4 l2 f1 ف2 خوبيد سلام
```

l1 and l2 are labels of chats.

F is feature and our features are fires five words do each message.

If a message has less than 5 words, it doesn't matter.

Then we run following commands:

```
bin/mallet import-file --input sample-data/mallet.txt --output
labeled.mallet
```

this command generate labeled.mallet file.

```
bin/mallet train-classifier --input labeled.mallet --training-
portion 0.9 --trainer MaxEnt --trainer NaiveBayes
```

this command generate this output:

NaiveBayesTrainer

Summary. train accuracy mean = 0.5442661502282473 stddev = 0.0
stderr = 0.0

Summary. test accuracy mean = 0.5258245177349098 stddev = 0.0
stderr = 0.0

Summary. test precision(l1) mean = 0.5177065767284992 stddev = 0.0
stderr = 0.0

Summary. test precision(l2) mean = 0.5486935866983373 stddev = 0.0
stderr = 0.0

Summary. test recall(l1) mean = 0.763681592039801 stddev = 0.0
stderr = 0.0

Summary. test recall(l2) mean = 0.2876712328767123 stddev = 0.0
stderr = 0.0

Summary. test f1(l1) mean = 0.6170854271356784 stddev = 0.0 stderr
= 0.0

Summary. test f1(l2) mean = 0.37745098039215685 stddev = 0.0
stderr = 0.0

MaxEntTrainer, gaussianPriorVariance=1.0

Summary. train accuracy mean = 0.5740766357725826 stddev = 0.0
stderr = 0.0

Summary. test accuracy mean = 0.5637834474175483 stddev = 0.0
stderr = 0.0

Summary. test precision(l1) mean = 0.5482661668228679 stddev = 0.0
stderr = 0.0

Summary. test precision(l2) mean = 0.5944444444444444 stddev =
0.0 stderr = 0.0

Summary. test recall(l1) mean = 0.7276119402985075 stddev = 0.0
stderr = 0.0

Summary. test recall(l2) mean = 0.39975093399750933 stddev = 0.0
stderr = 0.0

Summary. test f1(l1) mean = 0.6253340459647247 stddev = 0.0 stderr
= 0.0

Summary. test f1(l2) mean = 0.47803425167535374 stddev = 0.0
stderr = 0.0

you see that precision and recall of maxent is bigger than naivebayes.(except recall of l1)
and as you see accuracy of maxent is better.