

Cmpe 493 Introduction to Information Retrieval

Assignment 2, Spring 2020

Egemen Göl

2016400009

In this project, I implemented three different classifiers to use in the sentiment analysis task of some reviews. All algorithms use laplace smoothing with $\alpha=1$.

```
~/code/sentiment_nlp
▶ python3 measure.py
=== Training all.
=== Classic


|       | Precision | Recall | F-measure |
|-------|-----------|--------|-----------|
| Micro | 0.82      | 0.82   | 0.82      |
| Macro | 0.82      | 0.82   | 0.82      |



=== Binary


|       | Precision | Recall | F-measure |
|-------|-----------|--------|-----------|
| Micro | 0.82      | 0.82   | 0.82      |
| Macro | 0.82      | 0.83   | 0.82      |



=== Bernoulli


|       | Precision | Recall | F-measure |
|-------|-----------|--------|-----------|
| Micro | 0.80      | 0.80   | 0.80      |
| Macro | 0.80      | 0.82   | 0.80      |



=== Randomization Tests

Classic & Binary      0.89

Classic & Bernoulli   0.15

Binary & Bernoulli    0.06
```

All precision, recall and F-measure values are similar, no significant difference there.

If we look at the randomization outputs, we see a similarity between Classic & Binary methods, Bernoulli methods differentiates itself.

This means Bernoulli classified our test input with somewhat the same correctness but the answers were different for the same questions. It classified differently, with a similar level of success.