

# CMPE493 - ASSIGNMENT 4

## Bekir Yıldırım - 2014400054

All wanted info are in screenshots below:

```
Assignment4 git:(master) ✗ python3 main.py
*****
NO MUTUAL INFORMATION
*****
Vocabulary size: 14669
-----
tp: 230, fp: 4, fn: 10, tn: 236
Performance values for class legitimate:
-----
PRECISION: 0.9829059829059829
RECALL: 0.9583333333333334
F-measure: 0.970464135021097
-----
tp: 236, fp: 10, fn: 4, tn: 230
Performance values for class spam:
-----
PRECISION: 0.959349593495935
RECALL: 0.9833333333333333
F-measure: 0.97119341563786
-----
MACRO AVERAGED PRECISION: 0.9711277882009589
MACRO AVERAGED RECALL: 0.9708333333333333
MACRO AVERAGED F-MEASURE: 0.9708287753294784
-----
```

```
*****
WITH MUTUAL INFORMATION
*****
Most 100 discriminating words:
dict_keys(['language', '!', '5', 'free', 'remove', 'linguistic', 'http', 'com', 'money', 'linguist', 'check', 'mail', 'linguistics', 'university', 'market', 'best', 'site', 'cost', 'click', 'business', 'our', '0',
'100', 'product', 'internet', 'service', 'company', 'day', '8', 'english', 'today', 'advertise', 'million', 'www', 'sell', 'cash', 'hour', 'win', 'dollar', 'pay', 'home', 'bulk', 'x', 'web', 'call', 'card', 'query',
'save', 'income', 'credit', 'mailing', 'success', 'offer', 'guarantee', 'thousand', 'purchase', 'hundred', 'yours', 'earn', 'over', 'us', 'department', 'customer', 'instruction', 'easy', 'edu', 'yourself', 'profit',
t', 'speaker', 'reference', 'anywhere', 'online', 'grammar', 'name', 'want', '/', 'visit', 'address', '24', 'order', 'every', 'receive', 'theory', 'zip', 'phone', 'need', 'buy', 'personal', 'price', '20', 'syntax',
'10', 'here', 'return', 'off', '99', 'step', '1998', 'science', 'list'])
-----
Vocabulary size: 100
-----
tp: 228, fp: 1, fn: 12, tn: 239
Performance values for class legitimate:
-----
PRECISION: 0.9956331877729258
RECALL: 0.95
F-measure: 0.9722814498933903
-----
tp: 239, fp: 12, fn: 1, tn: 228
Performance values for class spam:
-----
PRECISION: 0.952191235059761
RECALL: 0.9983333333333333
F-measure: 0.973524215883947
-----
MACRO AVERAGED PRECISION: 0.9739122114163434
MACRO AVERAGED RECALL: 0.9729166666666667
MACRO AVERAGED F-MEASURE: 0.9729024357409926
-----
1500
1500
Approximate Randomization Test Result: 1.0
```