

## NLP Assignment 2

In this assignment, you are asked to write a program to evaluate the predicted result.

The documents will be about five different domains, namely: *book*, *dvd*, *health*, *music*, and *toys\_games*. Following two files are the gold standard and predicted result of this document classification problem:

*gold\_standard.xml* : The gold standard of the documents.

*test\_outcome.xml* : The predict result by our system.

You need to calculate:

1. Precision, recall, f1-measure of *book* category
2. Precision, recall, f1-measure of *dvd* category
3. Precision, recall, f1-measure of *health* category
4. Precision, recall, f1-measure of *music* category
5. Precision, recall, f1-measure of *toys\_games* category
6. Micro-average precision, recall, f1-measure
7. Macro-average precision, recall, f1-measure

The sample output is shown below.

```
book
  p :0.933127572016461
  r :0.907
  f :0.919878296146045
dvd
  p :0.913179916317992
  r :0.873
  f :0.892638036809816
health
  p :0.747393744987971
  r :0.932
  f :0.829550511793502
music
  p :0.932112068965517
  r :0.865
  f :0.897302904564316
toys_games
  p :0.880713489409142
  r :0.79
  f :0.832894043226147
Micro-average
  p :0.8734
  r :0.8734
  f :0.8734
Macro-average
  p :0.881305358339416
  r :0.8734
  f :0.877334871424899
```

References:

1. <http://rushdishams.blogspot.tw/2011/08/micro-and-macro-average-of-precision.html>
2. [http://en.wikipedia.org/wiki/F1\\_score](http://en.wikipedia.org/wiki/F1_score)

Please upload:

1. Your program (include source code)
2. The report: briefly describe the method (include the snapshot)