

INF5830, 2017, Obligatory assignment 2 - clarifications

I have received several questions regarding the test and training sets; in particular regarding the final test set, which is never used.

Test set

- We take away 400 items for a final testing. We are not performing the final testing as part of this exercise set, but we can do it at a later stage when we think we have found the best classifier. One point is that the same 400 items should be kept away during all the experiments.
- All the experiments in the assignment are performed with the same 1600 items as development set.
- In exercise 1, these items are further divided into 1400 for training and 200 for development test set. We do not touch the 400 final test set.
- In exercise 2, we use the whole development set of 1600 items for 10-fold cross-validation.

NLTK typo

As pointed out in lecture 2, slide 27, there is a mistake in the web version of the NLTK-book. It is not updated to the latest version of NLTK

The following does not yield the right result:

```
wrong_features = list(all_words)[:2000] #Wrong (earlier version)
```

You should instead use:

```
word_features = [w for (w,_) in all_words.most_common(2000)]
```

Precision and recall

The following request in exercise 1 was not very precise "Calculate accuracy, recall, precision and F-score." I have been asked whether you should calculate precision and recall for both classes. Please do!

(One word of caution. The terms 'pos(itive)' and 'neg(ative)' are used in two different ways in this exercise. As names for the two labels on the movie reviews. And when we talk about the true positives etc. of the classifier. Do not confuse them.)