

Adaptive Filtering Software: Future Work

Wish list for the software

This section lists some nice-to-add features.

Real-time adaptive filtering

It would be desirable to modify the adaptive filtering to operate with real-time feedback: that is, after a document has been judged by the classifier as potentially promising for some topics, the human expert(s) would have a chance to look at them and provide the system with the feedback that it can use to improve the classifier. Moreover, it would be nice to be able to augmented the test set during the run, as new documents keep coming in from the outside world. In the current implementation (Jan 2004), none of these features are available: the entire test set and all the "feedback" have to be supplied to the classifier beforehand. This is fine for experiments and for competitions such as TREC, but is still makes the "adaptiveness" of our adaptive filtering more of a toy than a practical feature.

Input file checks

The program has an unpleasant property of dying with a segmentation fault when some of the files it expects to find do not exist. Proper checks should be provided instead. (I am not sure if it is my fault or Lemur's).