

In general, the directory `index_dir` is use for store the out put of index.

I remove all those words with single character like one number or one symbol and one letter. An also put url in a single token. It will be able to recognize all kinds of normal url.

In search part I implement Rocchio Algorithm.

When doing the search after the document print on terminal, one request will also be print ask user to type the relevant document.(I don't do any validation on that step, so the input must be follow the form given below. sorry about that). The form of input should be : document name white space document name white space document name etc....  
example:

`doc1.txt doc2.txt doc3.txt`

When finish typer enter. And the new document and new cosine score will be print out.