

## Helper classes which could help analyze the data but not important

### CountTotalDoc:

count all documents in microblogTrec2011 7 folders.  
input files: microblogTrec2011

### CountRelevantDocs:

count the relevant documents in the given materials for each query.  
input files: rel/ folder which contains all relevant documents extracted from  
microblogTrec2011  
output file: CountRel

### GetTheSameRelevant:

compare the result\_file from Indri and the relevant documents feedback in qrels.  
input file:  
    result\_file from indri  
    relevantQrels (which is the end with "1" line in qrels, I did not put in the folder)  
output file: sameRelevant which contains only the matched results that indri could get  
and it is a true relevant document.

### GetDifferentNumber:

compare the the sameRelevant file from GetTheSameRelevant with the CountRel, we  
can know how many relevant queries don't be taken by Indri.  
input file: CountRel  
    sameRelevant  
output file: diffNum

### ExtractFromTrec: (just for Ideal Case, not for formal run)

Extract all the relevant feedback from microblogTrec2011, thus we can analyze which  
terms are in the relevant tweets. This is not for the formal run, it is only for analysis.  
input file: relevantQrels — a file containing only relevant line from "qrels"  
    microblogTrack2011  
output file: rel/ folder. (This folder contains all the true relevant tweets extracted from  
the 7 folders for each query.)