Kaixin Zhang

kz1021

CS980

Assignment 1 – Search and Indexing

The implementations for this assignment:

1. Indexing function using Trec-Car-Tool to index the paragraphCorpus Data.
2. Lucene Search function with BM25 retrieval mode
3. Java File input function to create a run file for TREC eval tool based on the results from Lucene search.

GitHub Link: https://github.com/zkx0804/SearchAndIndexing

Evaluation Results:

Because indexing the data corpus is unable to finished in time, so currently I don't have results. However, the code for output run file are ready and was tested on other data set.

I will update the evaluation results once the index is finished.