Build Script (Team3)

Amith, Pooja, Vaughan, Sepideh

December 7, 2018

1 Dependency

Java version: Java8 or Higher
Build Tool: Maven

2 Data Path

Since all the team members used different kind of index on the paragraph corpus, we hard-coded the directory path to the Indexed paragraph corpus in the main function and reduced the number of inputs the program takes in to make it easier.

All the paragraph and word embedding files are in the **"/home/team3"**

3 How to Run

We noticed that once we cloned the repository on the server, the permission of the directory needs to be changed to rwx for all user,group and owner in order to compile and build.

Please perform the below steps

- **sudo git clone https://github.com/poojahoza/CS853FinalProject.git**

- **sudo chmod 777 CS853FinalProject**

- **cd CS853FinalProject**

- **./install.sh**

This will run all the methods using the Benchmark-Test-y1 Outlines as the Query and we have used Benchmark-Test-y1 released Ground truth for Evaluation.
install.sh can be run by passing 4 arguments in this format

./install.sh TEST-OUTLINE TEST-QREL TRAIN-QREL TRAIN-OUTLINE

The easiest way is to run ./install.sh with no arguments.

With the LambdaBM25 training we have used the Benchmark Y1 train data set to create our initial model. We have also pre-processed the test data set for use with RankLib.

Once the process has run we will need to generate our ranklib model
Please perform the below steps

- **java -jar RankLib.jar -train {Running directory}/rankingDoc.txt -ranker 5 -metric2t NDCG@10 -save mymodel.txt -qrel {Running directory}/lambdaRankOutput.qrels**

- **java -jar RankLib.jar -load mymodel.txt -metric2t MAP -rank {Running directory}/testDoc.txt -indri testIndri.txt**

- **Migrate testIndri.txt to /home/team3 and rerun the program"**

You may now use your normal evaluation tool to process /home/team3/testIndri2.txt