# Information Retrieval and Web Search Individual Assignment 1 Instruction to run the application

**Paritosh Chauhan**

chauhapa@tcd.ie

## 1 Introduction

Assignment 1 is to create a searcher for the information on the Cranfield Collection with the set of 225 queries from the same collection. The implementation is expected with different analyzers and scoring methods to understand the working of information retrieval using Apache Lucene ( Version 8.4.1) .

| Analyzer | | Relevance |
|---|---|---|
| English Analyzer | | BM25 |
| Standard Analyzer | | LM Dirichlet Similarity |
| Classic Analyzer | | Classic Similarity |

## 2 How to Run

### 2.1 Accessing Ubuntu instance in AWS

1. Windows - open command prompt in Administrator mode or Unix machine open terminal and navigate to the path of the folder where the .PEM file is stored

2. Run the command on the terminal to get the instance

- *ssh -i "IRWS_Paritosh.pem" ubuntu@ec2-34-229-251-75.compute-1.amazonaws.com*

3. If prompted for password enter - *Pari0@aws*

4. If the login is successful the user displayed must be *ubuntu@ip-172-31-42-230: $*

5. ```sudo su```

   To access machine as a root user of the system and the current user will change *root@ip-172-31-42-230:/home/ubuntu*

### 2.2 Project folder to run the application

6. ```cd /root/InformationRetrivalWebSearcher/```

   Entering will land on the folder were the files are stored and the user path changes to *root@ip-172-31-42-230: /InformationRetrivalWebSearcher*.

### 2.3 Running the application and getting the scores from the table

7. ```mvn clean package install```

   Running this will clear the previous jar from the target folder and progresses to build followed by packaging the project. .

8. ```mvn exec:java -Dexec.mainClass="com.myproject.to```

   Run this command to execute the main file that will generate the results from the permutation and combination of all the analyzers used along with the path to the output file and filename where the results are stored .

### 2.4 Running Trec Eval

9. Travel to the folder were Trec Eval is situated

   ```cd trec_eval/```

10. To make sure the trec eval has the update make to evaluate

    ```make```

11. To get the result select the result file you want to evaluate.

    ```./trec_eval test/qrel.new ../output/results_Engl```