

# Information Retrieval

Assignment 3

Team 3:

Shayan Amani

Soheil Gharatappeh

1. Please find our codes for assignment 3 in the following GitHub url:

<https://github.com/SHi-ON/InfoRet/tree/master/src/main/java/edu/unh/cs/ir/pa3>

You can find the information regarding to install the project on your system in project's README file.

<https://github.com/SHi-ON/InfoRet>

You can see the results from consul and they are going to be printed into a file named run\_file. There are two options to change TFIDF. You can change the similarityBase(), or TFIDFSimilarity(). Because of some confusions, we had about which one to choose, and if choose one, which parameters to change, we couldn't manage to change the search engine. However, we understand where we need to change those class, and we got some ideas about how these classes/methods work.

We used TFIDFSimilarity() methods and override them to implement problem requirements. An input argument for IndexSearcher3() can specify which variant is desired for TFIDF computations. In order to code for Inc.ltn we used queryNorm()'s argument to keep the summation of squared weights as sumSW then by overriding tf() method we wrote the logarithmic variant of TF and also for IDF.

We have problems to figure out how we should change the similarities to implement ddd and qqq parts of SMART notion using TFIDFSimilarity or other similarities. In fact, we couldn't answer to this question that which similarity/method/argument is belong to the document vector and which one is belong to the query vector.