Project 3 Report

Shunyu Wang

40043915

COMP 479 Fall 2019

1. Refer to the submission package.

2. Using the index you compiled, design and test queries to try to satisfy the following information needs:

(Compare and report on at least 3 queries for each information need.)

1. Which departments have AI research?

Please refer to these files in the submission

Which\_departments\_have\_AI\_research\_tf\_idf.txt

Which\_departments\_have\_AI\_research\_bm25.txt

1. Which researchers are working on AI research?

Please refer to these files in the submission

Which\_researchers\_are\_working\_on\_AI\_research\_tf\_idf.txt

Which\_researchers\_are\_working\_on\_AI\_research\_bm25.txt

1. What AI research is being conducted at Concordia?

Please refer to these files in the submission

Which\_researchers\_are\_working\_on\_AI\_research\_tf\_idf.txt

Which\_researchers\_are\_working\_on\_AI\_research\_bm25.txt

3. Test and compare two ranking schemes for your queries: tf-idf and BMI as

implemented in Project 2:

|  |  |  |
| --- | --- | --- |
| **Ranking schemes** | **Tf-idf** | **BM25** |
| **@10** | Which\_departments\_have\_AI\_research\_compare10.txt | Which\_departments\_have\_AI\_research\_compare10.txt |
| **@50** | Which\_researchers\_are\_working\_on\_AI\_research\_compare50.txt | Which\_researchers\_are\_working\_on\_AI\_research\_compare50.txt |
| **@100** | Which\_researchers\_are\_working\_on\_AI\_research\_compare100.txt | Which\_researchers\_are\_working\_on\_AI\_research\_compare100.txt |

4. Compute df for all final terms in AIindex and report your vocabulary with df values.

Please refer to these files in the submission

AIindex\_term\_df.txt

5. Use the df weights of AIindex in your ranking for your queries.

Summary (trouble encountered, issues resolved, and findings.)

Teamwork