**Information Retrieval - Assignment I**

1. Build an uncompressed index from the given dataset. You can do the coding either using Java or Python.
2. Execute the code on the given data set and generate the Index for it.
3. You must submit both the code and the executable version (accordingly for Java/Python).

Submit the exe and src folder in an archive

|-- Archive

|--bin

|--src

|--meta

|--docs

|--Instructions or README

|--Generated index file

|-- Write up about over all algo and design of application

|--Write up on “number of Dictionary terms generated” and “Sample some Conjunctive Boolean Queries and results”

1. Submit the generated output of the Index.
2. Give a report on how your algorithm works, whether stop words removed/not, stemming/Lemmatization used or not etc.
3. Also based on the output generated summarize the number of Dictionary terms generated.
4. Give some Conjunctive Boolean Queries and test your Index generated. Provide the generated output for this.

Hello,

Enter your search query and hit Enter key. Wrap your query in "" for searching a phrase.

If you want to exit, just hit Enter key!

Enter your search query: person

Showing 4 results for "person"

- Doc1.docx

- Doc2.docx

- Doc3.docx

- Doc5.docx

Enter your search query: need

Showing 2 results for "need"

- Doc2.docx

- Doc6.docx

Enter your search query: friend

Showing 1 results for "friend"

- Doc2.docx

Enter your search query: family

Showing 1 results for "family"

- Doc5.docx

Enter your search query: person & family

Showing 1 results for "person & family"

- Doc5.docx

Enter your search query: person & friend

Showing 1 results for "person & friend"

- Doc2.docx

Enter your search query: !person

Showing 2 results for "!person"

- Doc6.docx

- Doc4.docx

Enter your search query: !person

Showing 2 results for "!person"

- Doc6.docx

- Doc4.docx

Enter your search query: !need

Showing 4 results for "!need"

- Doc3.docx

- Doc5.docx

- Doc1.docx

- Doc4.docx

Enter your search query: Person & need & friend & family

No matching results found for "person & need & friend & family"

**Guidelines for submission of Assignment**

**DATE OF SUBMISSION: MAY 6TH (SUNDAY)**

Submit as ZIP file which comprises of the following files.

1. Entire code for your Solution a) To build the Index file b) To evaluate the Boolean Queries on the Index file generated
2. Run the coding and the output generated for your code both 1a and 1b must be given.
3. Executable version, preferably with a command line argument as input. Also give a description about how to provide inputs to your code. (Give the sample Conjunctive Boolean Queries you have used for testing your code, as the same can be compared with the output you have provided).
4. A detailed description of your algorithm, how it works, what are the methodologies used, assumptions made for generating the Index file must be given. (To get a snapshot about your coding).

**ALL THE FOUR FILES SHOULD BE INCLUDED IN THE ZIP FILE TO BE CONSIDERED FOR FURTHER EVALUTION.**

**PLEASE STICK ON TO THE SUBMISSION DATE. NO EXTENSION WILL BE PROVIDED AS THE MARKS MUST BE SUBMITTED IMMEDIATELY.**