

DOCUMENTATION (Search the BIBLE)

Libraries imported:

nltk – natural language toolkit for tokenizing and stemming

math – to perform sqrt and logarithms

csv – to read the csv file

flask – for all the flask components like flash-messages, templates, etc.,

Core Functions:

get_all_shingles():

returns all the shingles in a list

create_matrix():

returns the document_shingle matrix

create_k_signature():

returns the signature matrix using the hash functions

query_signature_matrix():

returns the query shingle matrix and query signature matrix

LSH Class:

returns the 3-D bucket list and query_bucket_list

getdocuments():

returns the candidate documents by taking the 3-D bucket list and query_bucket_list

getValidatedDocuments()

returns the filtered documents list, by taking the candidate documents

findSimilarity() returns the similarity value between two vectors

HTML:

Layout.html:

Basic layout of the all pages such as the title and extends to other pages.

Document.html:

displays the selected document.

_navbar.html:

For the navigation bar.

searchResults.html:

shows the search results as a list of document names

Home.html:

home page containing the search bar and which shows the results upon searching.

_messages.html:

Helps display the flash messages