Help on module tf\_idf:

NAME

tf\_idf

CLASSES

builtins.object

tf\_idf

class tf\_idf(builtins.object)

| This class contains methods which generate the tf-idf scores

| for query string and the documents of the dataset.

|

| Methods defined here:

|

| tf\_doc\_specific(self, myList, global\_word)

| This method is used to find the tf-idf score of the specific document which has been pa

ssed as an argument

| :param myList: The list of words obtained from an individual document

| :param global\_word: The list of all the words in the entire corpus

| :return: a dictionary which maps the tf-idf scores of the words in the specific documen

t

|

| tf\_idf\_query(self, Query, global\_word, doc\_freq, doc\_count)

| This method is responsible for computing the tf-idf score of the user query

| :param Query: list of strings present in the user query

| :param global\_word: list of all the words present in the corpus

| :param doc\_freq: dictionary which maps each word of the corpus to the number of documen

ts it appears in

| :param doc\_count: integer denoting total number of documents in the corpus

| :return: dictionary of user query words mapped on to the corresponding tf-idf scores

|

| ----------------------------------------------------------------------

| Data descriptors defined here:

|

| \_\_dict\_\_

| dictionary for instance variables (if defined)

|

| \_\_weakref\_\_

| list of weak references to the object (if defined)

DATA

docFreq = {}

FILE

c:\users\smit\pycharmprojects\searchengine\tf\_idf.py