

National University of Computer and Emerging Sciences, Lahore Campus

Course: Information Retrieval
Program: BS(Computer Science)
Duration: 60 Minutes
Paper Date: 23-Sept-19
Section: ALL
Exam: Midterm-1

Course Code: CS317
Semester: Fall 2019
Total Marks: 17
Weight: 13%
Page(s): 4
Roll No:

Instruction/Notes: *Attempt the examination on this question paper.. You can use extra sheets for rough work but do not attach extra sheets with this paper. Do not fill the table titled Question/marks*

Question	1	2	3	4	Total
Marks	/ 2	/ 4	/ 7	/ 4	/17

Q1) What is advantage of using inverted index over forward index ? [2 Points]

Q2) Let V = Vocabulary size, n = Total number of words, d = total documents, $AveD$ = Average Document Length, $|q|$ = query length, $|posting|$ = length of posting list of a word. [4 Marks]

- a) Briefly describe index creation method using hashmaps. What is time complexity of building inverted index using HashMaps?

- b) Briefly describe index creation method using sorting (without hashmap). What is time complexity of building inverted index using this method?

Q3) a) What is the advantage of using IDF weight in Tf.IDF weighting scheme. Illustrate with an example. [2 Points]

Q3)b) Given the three-document corpus and a stop word list below, answer the following question AFTER removing stopwords. Total Documents in corpus = 3

Document 1	Shares in apple company stock prices rise
Document 2	An apple a day keeps the doctor away
Document 3	Apple health benefits apple fruits mango
Query	Apple Company
Stopwords	the, in, a,

Calculate Tf.IDF score of **document 1** for given query.

[3 Points]

Q3)c) If document vectors are length normalized, the ranking according to Euclidean distance and ranking according to cosine of the angle between the document vectors gives the same ranking or different ranking? Justify your answer. [2 Points]

Q4) Use maximum likelihood estimate to compute query likelihood from the following three documents AFTER removing stopwords.

Document 1	Shares in apple company stock prices rise
Document 2	An apple a day keeps the doctor away
Document 3	Apple health benefits apple fruits mango
Query	apple company
Stopwords	the, in, a,

[4 Points]

a) Use Unigram Model

b) Use Bigram model