

selective - 2

Total number of printed pages:2

PG/3rd/PCSE316

2021

INFORMATION RETRIEVAL

Full Marks: 100

Time: Three hours

The figures in the margin indicate full marks for the questions.

Answer any five questions.

1. a) What is information retrieval (IR)? Is an IR system same as that of a web search? Justify your answer. 2+3=5
- b) What are the major components of an IR system? Draw the architecture of an IR system. 2+3=5
- c) Briefly explain the function of the major components of the indexing process of an IR system with a suitable diagram. 10
2. a) What is a vector space model? Explain the vector space model for an IR system with a suitable example. 2+10=12
- b) What is text mining? Briefly explain the functions of text mining techniques. 3+5=8
3. a) What do you mean by Boolean model in terms of IR? Explain the Boolean model for IR system with a suitable example. 3+7=10
- b) Explain the terms Term Frequency (TF) and Inverse Term Frequency (IDF). 3+3=6
- c) Mention the advantages and disadvantages of Boolean model. 2+2=4
4. a) What is MapReduce? How does MapReduce work? Explain MapReduce with a suitable example. 2+4+6=12
- b) What is an organic list and a sponsored list of web search process? 3+2=5
- c) What do you mean by probabilistic ranking? 3
5. a) What are the major components of query process of an IR system? Explain the function of each component with a suitable diagram. 3+9=12
- b) What is an inverted list? Explain the inverted list with a suitable 2+6=8

example.

6. Write short notes on (any four)

5*4=20

- a) Binary independence model
- b) Crawling
- c) Hadoop
- d) Latent semantic
- e) Link analysis

