


Name:			
Enrolment No:			
UPES End Semester Examination, May 2024			
Programme Name: BTech CSE Big Data		Semester : VI	
Course Name : Big Data Search		Time : 03 hrs	
Course Code : CSBD3008		Max. Marks: 100	
Instructions: Attempt All questions			
SECTION A (5Qx4M=20Marks)			
S. No.		Marks	CO
Q 1	Define the terms document, index and segments used in Apache Lucene	4	CO2
Q 2	Identify some common challenges in maintaining big data search systems.	4	CO1
Q 3	Justify the importance of ranking/relevance scoring in search results	4	CO1
Q 4	Discuss the features of Elasticsearch.	4	CO3
Q 5	Explain how big data search and analytics improves the e-commerce industry.	4	CO1
SECTION B (4Qx10M= 40 Marks)			
Q 6	Discuss the features and benefits of Lucene. Also, explain its limitations.	10	CO2
Q 7	How do you define and measure the relevance of search results in a big data context? Or Summarize the role of text transformation in improving search quality. List and explain the steps involved in text transformation.	10	CO2
Q 8	What are the main parameters over which the performance and efficiency of various big data search tools can be compared?	10	CO1
Q 9	Discuss the different backend components of Elasticsearch.	10	CO3
SECTION-C (2Qx20M=40 Marks)			
Q 10	Elaborate the concept of Index, shards, and replicas in Elastic search. Also, explain the best practices considered regarding shards in ES.	20	CO3
Q 11	i. Explain the Boyers-Moore algorithm of data search. ii. How does inserting a new string into a trie work?	20	CO1