

VITH SEM, PE-I_BDA[CSE4059] _END SEM MAKE-UP_QUESTIONS_MAY2023

Type DES

Q1. Discuss issue related with terminology of unstructured data. (3)

Q2. Discuss on various classified data models of NoSQL database. (4)

Q3. Differentiate NoSQL database characteristics with traditional database. (3)

Q4. Illustrate MongoDB query for the following: (5)

(i) To create a collection by the name “food” and then insert documents into the “food” collection (5 id’s). Each document should have a “fruits” array.

(ii) To find those documents from the “food” collection where grapes is present in the 2nd index position of the “fruits” array.

(iii) To find the document with (_id:1) from the “food” collection and display two elements from the array “fruits”, starting with the element at 1st index position.

(iv) To find all documents from the “food” collection which have elements “orange” and “grapes” in the array “fruits”.

Q5. Discuss on features of Cassandra. (3)

Q6. Discuss functional component of reducer class. (2)

Q7. Discuss on Recommendation systems. (2)

Q8. Discuss on requirements of Scala. (3)

Q9. Discuss on solutions for following problems related to distributed hardware performance woes. (5)

Q10. Discuss on Hadoop Limitations. (2)

Q11. Discuss on high level of Hadoop architecture. (4)

Q12. Discuss Gossip Protocol with an example. (4)

Q13. Discuss on Spark SQL Architecture. (3)

Q14. Discuss the advantages and drawback of anomaly detection. (2)

Q15. Illustrate a program to demonstrate how to load a LIBSVM data file, parse it as an RDD of LabeledPoint and then perform regression using a decision tree with variance as an impurity measure and a maximum tree depth of 5. The Mean Squared Error (MSE) is computed at the end to evaluate goodness of fit. (5)