Reg. No. : | | | | | | | | | | |

Question Paper Code: 39217

B.E. / B.Tech. DEGREE EXAMINATION, MAY 2022

Elective

Computer Science and Engineering

01UCS917 - MASSIVE DATASET ANALYTICS

(Regulation 2013)

Duration: Three hours Maximum: 100 Marks

Answer ALL Questions

PART A - (10 x 2 = 20 Marks)

1. List the characteristics of big data and challenges in handling big data.
2. Write any two possible web data from which effective analysis can be carried out.
3. Highlight the uses of regression modeling.
4. Define principal component analysis.
5. List out any four NoSQL databases.
6. What is visual analytics?
7. Differentiate Fuzzy logic and Neural Networks.
8. What is SVM?
9. Define Resampling.
10. What are the four V's of Big Data?

PART - B (5 x 16 = 80 Marks)

11. (a) (i) Discuss the evolution of big data analytics. (8)

(ii) Explain in detail about the major resampling techniques. (8)

Or

(b) Discuss in detail the evolution of analytic scalability. (16)

12. (a) Explain the implication of PCA in Data visualization. (16)

Or

(b) Describe various stochasic search methods in detail. (16)

13. (a) With a neat sketch explain Stream Data Model. (16)

Or

(b) With an example explain the counting of distinct elements in a stream. (16)

14. (a) Explain Apriori algorithm and with an example show how association rules are generated from frequent item sets. (16)

Or

(b) Discuss in detail about the algorithm that handles non-main-memory data, but does not require a Euclidean space. (16)

15. (a) Discuss about Hadoop Distributed File System architecture with a neat diagram. (16)

Or

(b) Write short note on HDFS Architecture. (16)