Reg. No. :

Question Paper Code: 39217

B.E. / B.Tech. DEGREE EXAMINATION, NOV 2018

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. Write down the specification for Bloom's filter.

2. What are realtime analytics platform?

3. Define subspace clustering.

4. What is "market-basket" model?

5. Write any two examples for stream data.

6. What is visual analytics?

7. Differentiate Fuzzy logic and Neural Networks.

8. Define K-Means clustering algorithm.

9. Define Resampling.

10. What are the components of Hadoop framework?

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) (i) With a neat diagram explain MapReduce programming. (8)

(ii) Highlight the features of NoSQL. (8)

Or

(b) (i) Discuss about Hadoop Distributed File System architecture with a neat diagram. (10)

(ii) Write short notes on Visualization for Big Data. (6)

39217