

Reg. No. :

--	--	--	--	--	--	--	--	--	--	--

**Question Paper Code: 39025**

B.E. / B.Tech. DEGREE EXAMINATION, NOV 2017

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. 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) Elaborate on various frequentist inference based statistical inference mechanisms used in big data. (16)

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 stochastic 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) Elaborate on handling large datasets in main memory. (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) With a neat diagram explain MapReduce programming. (16)
- Or
- (b) Write short note on HDFS Architecture. (16)
-