

Reg. No. :

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

Question Paper Code: 49217

B.E./B.Tech. DEGREE EXAMINATION, SEP 2020

Elective

Computer Science and Engineering

14UCS917 - MASSIVE DATASET ANALYTICS

(Regulation 2014)

Duration: One hour

Maximum: 30 Marks

PART A - (6 x 1 = 6 Marks)

(Answer any six of the following questions)

1. What was Hadoop named after? CO1- R
(a) Model Planning (b) Model Building (c) Data Preparation (d) Operationalize
2. Near real time processing deals with _____ characteristics of data. CO1- R
(a) Creator Doug Cutting's favorite circus act (b) Cutting's high school rock band
(c) The toy elephant of Cutting's son (d) A sound Cutting's laptop
3. _____ is an index plot of the principal component variances. CO2- R
(a) Vector diagram (b) Decision Tree (c) Binary Tree (d) Scree diagram
4. Fuzzy logic is a form of CO2- R
(a) Two-valued logic (b) Crisp set logic (c) Many-valued logic (d) Binary set logic
5. Using of the main memory as a bit array is called CO3- R
(a) Bloom filter (b) Window filter (c) Blur filter (d) Drop filter
6. Bloom filter consists of _____. CO3- R
(a) Array (b) Vector (c) Key values (d) Both A & C
7. _____ model of data is used to describe a common form of many to many relationship between two kinds of objects. CO4- R
(a) Hierarchical (b) Market-basket (c) Structured (d) Vertical

8. The best known family of clustering algorithms is CO4- R
 (a) A-priori (b) Limited pass (c) K-means (d) Multihash
9. _____ was the first to publicize MapReduce – a system they CO5- R
 had used to scale their data processing needs.
 (a) Yahoo (b) Google (c) Microsoft (d) Linux
10. “Sharding” a database across many server instances can be achieved with CO5- R
 (a) LAN (b) SAN (c) MAN (d) All of the above

PART – B (3 x 8= 24 Marks)

(Answer any three of the following questions)

11. Discuss the challenges with Big Data. CO1- U (8)
12. Write short notes on Bayesian modeling. CO2- U (8)
13. Explain in detail about Alon-Matias-Szegedy algorithm for second CO2- U (8)
 moments
14. How does the Multistage algorithm take more than two passes to find CO4-U (8)
 the frequent pairs? Explain.
15. Describe the various visualization techniques that can be used for CO5- U (8)
 visualizing data.