

Roll No.

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

B.E / B.Tech (FT) END SEMESTER EXAMINATIONS – APRIL / MAY 2019

INFORMATION TECHNOLOGY

Seventh Semester

IT 8701 DATA ANALYTICS

(Regulation 2012)

Time: 3 Hours

Answer ALL Questions

Max. Marks 100

PART-A (10 x 2 = 20 Marks)

1. What are the characteristics of Big Data?
2. List out the differences between analysis and reporting?
3. What is a regression analysis? How is it different from classification?
4. What is a neural network? List out the applications of Neural networks.
5. Draw the block diagram of a data stream architecture?
6. What is meant by filtering and bloom filter?
7. Write an algorithm for hierarchical clustering?
8. How clustering is done for large data sets?
9. What is Hive? What are the advantages of Hive?
10. List out the difference between SQL and NoSQL databases?



Part – B (5 x 16 = 80 marks)
(Question No.11 is Compulsory)

11. 1. What is Big Data? List out the data sources and its classification? (8 Marks)
2. What is Data analytics? List out the types of data analytics with an example (8 marks)
12. a) What is a principal component analysis and explain how is it carried out?
(OR)
b) Write an algorithm for fuzzy decision trees? Illustrate it with a simple example.
13. a) Write an algorithm for implementing bloom filter. Illustrate it with a simple example.
(OR)
b) Write a procedure for counting distinct elements in a data stream. Illustrate it with an example.

14. a) Write apriori algorithm for association analysis. Illustrate it with a simple example.

(OR)

b) Write an algorithm for K-means algorithm? Illustrate it with an example.

15. a) What is a NoSQL data base? Illustrate the basics of NoSQL database such as MongoDB with neat illustrations.

(OR)

b) Explain Hadoop architecture with a neat illustration. Explain Hadoop ecosystem components with its associated usages.

