

Reg No.: _____

Name: _____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
EIGHTH SEMESTER B.TECH DEGREE EXAMINATION(S), OCTOBER 2019

Course Code: IT404

Course Name: DATA ANALYTICS

Max. Marks: 100

Duration: 3 Hours

PART A

Answer any two full questions, each carries 15 marks.

Marks

- | | | | |
|---|----|---|-----|
| 1 | a) | Tabulate the difference between analysis and reporting in data analytics. | (7) |
| | b) | What is re-sampling? Explain about different re-sampling techniques. | (8) |
| 2 | a) | Explain multilayer perceptron network with a neat diagram. | (8) |
| | b) | How splitting decision is made in decision tree induction? | (7) |
| 3 | a) | Explain bootstrapping in sample distribution. | (7) |
| | b) | How principal component analysis is used for feature selection? | (8) |

PART B

Answer any two full questions, each carries 15 marks.

- 4 a) Consider the following transactional data. Given frequent itemset {A,B,E}, (8)
min_sup=2 and min_conf=50%, what are the strong association rules.

TID	List of Items
1	A,B,E
2	B,D
3	B,C
4	A,B,D
5	A,C
6	B,C
7	A,C
8	A,B,C,E
9	A,B,C

- b) Analyze the advantages and limitations of hierarchical clustering over other clustering approaches (7)
- 5 a) Give Gartner's definitions of 3Vs in big data. (9)
- b) What is big data acquisition? (6)
- 6 a) List and explain a few applications of the market based model. (7)
- b) Explain any two technologies related to big data (8)

PART C

Answer any two full questions, each carries 20 marks.

- 7 a) Explain the following functions in R with suitable examples (10)
- i. summary()
 - ii. read.csv()
 - iii. head()
 - iv. rbind()
 - v. data()
- b) What is meant by exploratory data analysis? Mention some of the basic visualization techniques (6)
- c) What is dirty data? (4)
- 8 a) Discuss the following big data applications: (10)
- (i) Recommender systems
 - (ii) Social Media Analytics
- b) Illustrate and explain HDFS architecture with its features (10)
- 9 a) Explain t-distribution? Mention any two applications of t-distribution? (10)
- b) Give an example of how fraud detection is done using social network analysis (10)
