

(6 Pages)

Reg. No. :

Code No. : 20448 E Sub. Code : ASCS 31

B.Sc. (CBCS) DEGREE EXAMINATION,
NOVEMBER 2021.

Third Semester

Computer Science

Skill Based Subject — INTRODUCTION TO BIG DATA
ANALYTICS

(For those who joined in July 2020 onwards)

Time : Three hours

Maximum : 75 marks

PART A — ($10 \times 1 = 10$ marks)

Answer ALL questions.

Choose the correct answers :

1. Facebook feed is an ———— data.
(a) Structured (b) Semi structured
(c) Unstructured (d) Well structured
2. The heterogeneity of big data is known as ————
(a) Volume (b) Velocity
(c) Variety (d) All of the above

3. Using data to predict the future event is known as _____ analysis.
- (a) Descriptive (b) Predictive
(c) Futuristic (d) Prescriptive
4. KDD stands for
- (a) Knowledge Discovery in Database
(b) Knowledge Discovery in Datamining
(c) Known Discovery in Database
(d) Knowledge Dissection in Database
5. Platform used for big data analysis _____
- (a) Hadoop (b) C
(c) C++ (d) BASIC
6. Which of the following is not an analytics methods?
- (a) Decision tree (b) K-means
(c) Random forest (d) Binary tree
7. A supervised learning algorithms uses _____ for discrete values.
- (a) Classification (b) Regression
(c) Clustering (d) None of the above

8. ————— learning algorithm can be used with labeled data.
- (a) Supervised (b) Unsupervised
- (c) Reinforcement (d) Semi supervised
9. ————— is used for diagnosis.
- (a) Alexa (b) Siri
- (c) IBM Watson (d) Diagoniser
10. Selection of data analysis algorithm depends on
- (a) Data quality (b) Parameters used
- (c) Data source (d) All of the above

PART B — ($5 \times 5 = 25$ marks)

Answer ALL questions, choosing either (a) or (b)

Each answer should not exceed 250 words.

11. (a) Discuss about the challenges of big data.

Or

- (b) Narrate what are the additional dimensions of big data in addition to 3V.

12. (a) Describe the traditional analytics used before bigdata analysis.

Or

- (b) What are the challenges faced by statisticians in bigdata analysis?

13. (a) Discuss how machine learning is associated with bigdata analysis.

Or

- (b) Explain about text mining in bigdata analysis.

14. (a) Explain the following :

- (i) Euclidean distance
- (ii) Manhattan distance.

Or

- (b) Describe random forest algorithm.

15. (a) Explain how e-Bay uses bigdata.

Or

- (b) Describe how AI is affecting our lives.

PART C — ($5 \times 8 = 40$ marks)

Answer ALL questions, choosing either (a) or (b)

Each answer should not exceed 600 words.

16. (a) Describe in detail about types of data.

Or

- (b) Discuss how big data analytics helps in the decision making process.

17. (a) Write about the bigdata analysis pipeline.

Or

- (b) Narrate in detail about various applications of bigdata analysis.

18. (a) Illustrate the different phases of bigdata analysis.

Or

- (b) Enumerate various skills needed for bigdata analysis.

19. (a) Discuss in detail about supervised learning.

Or

- (b) Explain about clustering in bigdata analysis.

20. (a) Discuss how IBM Watson is used for pattern recognition.

Or

- (b) Explain how to decide whether to use supervised algorithm or unsupervised algorithm.
