

**(FOR THE CANDIDATES ADMITTED
DURING THE ACADEMIC YEAR 2020 ONLY)**

20 UBC 6E5

REG.NO. :

N.G.M.COLLEGE (AUTONOMOUS) : POLLACHI

END-OF-SEMESTER EXAMINATIONS : MAY - 2023

COURSE NAME: B.C.A

MAXIMUM MARKS: 70

SEMESTER: VI

TIME : 3 HOURS

PART - III
DATA MINING

SECTION - A (10 X 1 = 10 MARKS)

ANSWER THE FOLLOWING QUESTIONS.

MULTIPLE CHOICE QUESTIONS.

(K1)

1. Removing duplicate records is a data mining process called _____.
a) Data isolation b) Recovery c) Data pruning d) Data cleaning
2. Incorrect or invalid data is known as _____.
a) Outlier b) Missing data c) Changing data d) Noisy data
3. Decision trees are built using _____.
a) Heuristics b) Greedy algorithms
c) Dynamic programming d) Divide and conquer strategy
4. Enrichment is _____.
a) A stage of the KDD process in which new data is added to the existing selection
b) The process of finding a solution for a problem simply by enumerating all possible solutions according to some pre-defined order and then testing them
c) The distance between two points as calculated using the Pythagoras theorem.
d) The process of finding the right formal representation of a certain body of knowledge
5. Which of the following is a data warehouse?
a) Can be updated by end users.
b) Contains numerous naming conventions and formats.
c) Organized around important subject areas.
d) Contains only current data.

ANSWER THE FOLLOWING IN ONE (OR) TWO SENTENCES.

(K2)

6. What is Data Warehousing?
7. Define data Reduction.
8. Write the parts of Association rule.
9. Define Coding.
10. Expand OLTP.

SECTION - B (5 X 4 = 20 MARKS)

ANSWER EITHER (a) OR (b) IN EACH OF THE FOLLOWING QUESTIONS. (K3)

11. a) List the major issues in Data mining.

(OR)

- b) Summarize the concept of Pixel oriented data visualization techniques.

(CONTD 2)

12.a) Explain in detail data cleaning as a process.

(OR)

b) Describe the concept of redundancy and correlation analysis.

13. a) Write short notes on various types and sources of data.

(OR)

b) Narrate the concept KNN algorithm.

14. a) Discuss the need of data enrichment.

(OR)

b) List the 10 golden rules in data selection and explain it.

15. a) Elaborate the concept of top down design approach in data Warehousing.

(OR)

b) Give notes on attribute oriented induction.

SECTION - C

(4 X 10 = 40 MARKS)

ANSWER ANY FOUR OUT OF SIX QUESTIONS

**(16th QUESTION IS COMPULSORY AND ANSWER ANY THREE QUESTIONS
(FROM Qn. No : 17 to 21) (K4 (Or) K5)**

16. Classify the qualitative and quantitative attributes in datamining.

17. Discuss the various data reduction techniques.

18. List the OLAP tools and explain it.

19. Analyze genetic algorithm in detail.

20. Write down the steps involved in a KDD Process.

21. Narrate the concept of Data Warehousing life cycle model with neat diagram.

ETHICAL PAPER