

N.B.

- 1) All questions are **COMPULSORY**.
- 2) Figures to the right indicate **FULL** marks.
- 3) Use of non - programmable **CALCULATOR** is allowed.
- 4) Assume suitable data if necessary.

Q.1 What are different Data Mining techniques / functionalities? Explain in detail. (10)

OR

Q.1 Differentiate between (10)

- i) OLTP and OLAP
- ii) Clustering and classification

Q.2 Explain data discretization and concept hierarchy generation. (10)

OR

Q.2 State and explain the methods of data cleaning in detail. (10)

Q.3 Discuss the characteristics of Decision Tree Algorithm with suitable example. (10)

OR

Q.3 Discuss the concept of Frequent Pattern Growth algorithm and generate FP tree for the following transaction data set. (10)

(Minimum Support = 30%)

Tid	Items
1	E,A,D,B
2	D,A,C,E,B
3	C,A,B,E
4	B,A,D
5	D
6	D,B
7	A,D,E
8	B,C

P.T.O.

Q.4 Discuss Heuristic function and its types with the help of diagrams. (10)

OR

Q.4 Explain Logistic Regression along with Logit Transform. (10)

Q.5 Discuss about Parametric and Nonparametric regression methods in detail (10)

OR

Q.5 Describe the Prediction model types (in the generalized linear model) to fit the following use cases: (10)

- i) Number of Website Visitors per hour
- ii) Temperature Vs Pressure readings for an electric furnace
- iii) Number of spam emails received by a person.

Q.6 What are the types of components of time series plot? Explain in detail (10)

OR

Q.6 Describe autocorrelation and its properties in detail. (10)

* * *

010623-e-col-mumbai