



# INSTITUTE OF AERONAUTICAL ENGINEERING (Autonomous)

B.Tech VI Semester End Examinations (Regular), November – 2020

Regulation: IARE–R16

## DATA WAREHOUSING AND DATA MINING

**Time: 2 Hours**

**(CSE | IT)**

**Max Marks: 70**

---

**Answer any Four Questions from Part A**

**Answer any Five Questions from Part B**

---

### PART – A

1. Explain the schemas for multi-dimensional databases. [5M]
2. What are data mining primitives? Explain briefly. [5M]
3. Explain the association rule mining? [5M]
4. What are the advantages of FP-growth algorithm? [5M]
5. Write short notes on multimedia data mining. [5M]
6. Enumerate three categories of measures, based on the kind of aggregate functions used in computing a data cube. [5M]
7. Summarize the various issues that have to be addressed during data integration. [5M]
8. Explain any two techniques for assessing classifier accuracy. [5M]

### PART – B

9. What are the differences between three main types of data warehouse usage: information processing, analytical processing and data mining. [10M]
10. Suppose that a data warehouse consists of the four dimensions, date, spectator, location and game, and the two measures, count and charge, where charge is the fare that a spectator pays when watching a game on a given date. Spectators may be students, adults or seniors with each category having its own charge rate. Draw a Star schema diagram for the data warehouse. [10M]
11. Why preprocess of the data is done? Discuss issues to consider during data cleaning. [10M]
12. List the various stages of knowledge discovery process with a diagram. Explain how data mining is a part of knowledge discovery process. [10M]
13. What is market basket analysis? Explain association rules with confidence & support. [10M]
14. Illustrate about frequent item set? Write the Apriori algorithm for frequent item set generation? Describe with an example [10M]
15. What is decision tree? Explain how classification is done using decision tree induction. [10M]
16. Develop an algorithm for classification using Bayesian classification and illustrate the algorithm with relevant example. [10M]
17. What is an outlier? Explain different techniques for mining distance based outlier detection. [10M]
18. How agglomerative hierarchical clustering works? Explain with an example. [10M]