

III B.C.A- V SEMESTER

17UCAAED2

17UCLSAED2

DATA MINING AND DATA WAREHOUSING

UNIT I

Data Mining – Introduction – Functionalities – Classification of Data Mining System – Issues – Data Preprocessing – Data Cleaning – Data Integration and Transformation – Data Reduction.

UNIT II

Data Mining Primitives – Query Language Architecture of Data Mining System – Data Generalization and Summarization Based Characterization – Analytical Characterization – Mining Class Comparisons.

UNIT III

Association Rule Mining – Mining Single Dimensions Boolean Association Rules from Transactional Databases – Multilevel Association Rule – Classification and Prediction – Classification by Decision Tree Induction – Bayesian Classification – Predictive.

UNIT IV

Web Mining: Introduction - Web Content Mining - Web Structure Mining - **Advanced Topics:** Spatial Mining - Temporal Mining- **Visualization:** Data Generalization and Summarization- Based characterization - Analytical Characterization- Analysis of Attribute Relevance - Mining Class Comparisons - Discriminating between different Classes - Mining Descriptive Statistical Measures in Large Databases

UNIT V

Data Warehouse: Overview - Architecture – Back Room Technical Architecture – Architecture for Front Room Meta Data and Meta Data Catalog. **Security:** Vulnerabilities- Solutions- Managing Security for Data Warehouse Environment Physical Design – Data Staging.

Reference Books:

1. Data Mining Concepts and Techniques - Jiawei Han, Micheline Kamber - Elsevier Publishers.
2. The Data Warehouse Life Cycle Toolkit - Ralph Kimball - Wiley Publishers.
3. M.H. Dunham, “Data Mining Introductory and Advanced Topics”, Pearson Education - Wiley Publishers.

RAMAKRISHNA MISSION VIVEKANANDA COLLEGE EVENING
COLLEGE(AUTONOMOUS)
CHENNAI 600 004.
MODEL QUESTION PAPER

CLASS B.Sc(CS)/BCA/M.Sc IT/MSW

MAX :75 MARKS
TIME:3 HOURS

SECTION A

ANSWER ANY **TEN** QUESTION

(10x2=20)

- Answer any ten out of twelve question (each carry 2 marks)

SECTION -B

ANSWER ANY **FIVE** QUESTION

(5x5=25)

- Answer any five out of seven question (each carry 5 marks)

SECTION -C

ANSWER ANY **THREE** QUESTION

(3x10=30)

- Answer any three out of five question (Each carry 10 marks)