DATA MINING LABORATORY (Professional Elective – III)

Course Code: KG23ACM315	L	T	P	C
	0	0	3	1.5

B.Tech. III Year II Sem.

Prerequisites

- A course on "Database Management System

Course Objectives: The objectives of this course for the student are to:

1. The course is intended to obtain hands-on experience using data mining software.
2. Intended to provide practical exposure of the concepts in data mining algorithms
3. Understand the concept of classification and its approaches
4. Understand the construction of data warehouse
5. Learn how to perform OLAP operations

Course Outcomes: After completion of this course, the students will be able to:

CO1. **Apply** preprocessing statistical methods for any given raw data. **(K3)**
CO2. **Construct** a data warehouse. **(K6)**
CO3. **Implement** various algorithms for data mining in order to discover interesting patterns from large amounts of data. **(K3)**
CO4. **Apply** OLAP operations on data cube construction **(K3)**
CO5. **Evaluate** the classification of data using different approaches **(K5)**

LIST OF EXPERIMENTS: Experiments using Weka/ Pentaho/Python

1. Data Processing Techniques:

(i) Data cleaning (ii) Data transformation – Normalization (iii) Data integration

2. Partitioning - Horizontal, Vertical, Round Robin, Hash based

3. Data Warehouse schemas – star, snowflake, fact constellation

4. Data cube construction – OLAP operations

5. Data Extraction, Transformations & Loading operations

6. Implementation of Attribute oriented induction algorithm

7. Implementation of apriori algorithm

8. Implementation of FP – Growth algorithm

9. Implementation of Decision Tree Induction

10. Calculating Information gain measures

11. Classification of data using Bayesian approach

12. Classification of data using K – nearest neighbour approach

13. Implementation of K – means algorithm

14. Implementation of BIRCH algorithm

15. Implementation of PA2 algorithm

16. Implementation of DBSCAN algorithm

TEXT BOOKS:

1. Data Mining – Concepts and Techniques - JIAWEI HAN & MICHELINE KAMBER, Elsevier.

2. Data Warehousing, Data Mining & OLAP- Alex Berson and Stephen J. Smith- Tata McGraw-Hill Edition, Tenth reprint 2007

REFERENCE BOOK:

1. Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Anuj Karpatne, Introduction to Data Mining,

Pearson Education

CO-PO Mapping

COs	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12	PSO 1	PSO 2	PSO 3
CO1	**1**	2	3	2				2	2			3	H	2	H
CO2	**3**	3	**2**	**1**				3	**2**			**3**	H	2	2
CO3	**2**	**3**	**2**	**3**				2	**3**			2	H	2	H
CO4	**1**	3	**1**	2				**1**	3			2	H	2	2
CO5	**2**	2	**2**	2				2	2			L	L	2	H

HIGH=3, **MEDIUM**=2, **LOW**=1