Program Name	B. Tech. (Computer Engineering)	Semester – VII
Course Code	R4CO4001T	
Course Title	Data Mining and Data Warehousing	
Prerequisite	Statistics, Database, Data Structures and Algorithms	

COURSE OUTCOMES: Students will be able to	
1.	Perform the preprocessing of data and apply mining and data warehousing techniques on it.
2.	Identify and Implement association rules, classification, and clustering algorithms
3.	Solve real world problems in business and scientific information using data mining
4.	Use data analysis tools for scientific applications

COURSE CONTENTS	Hrs	CO
1. **Introduction to Data Mining:** Introduction, Data Mining Techniques, Knowledge Discovery, KDD Process.	2	1
2. **Data Warehouse and OLAP Technology for Data Mining**: Introduction to Data Warehouse, Data Cube and OLAP, Data Warehouse Design and Usage, Data Warehouse Implementation.	4	1
3. **Data Processing**: Types of Data, Data Quality, Data Pre-processing, Data Cleaning, Outlier Detection, Missing Value Detection, Data Integration, Data Reduction, Data Transformation and Data Discretization, Similarity and Dissimilarity, Summary Statistics, Visualization,	7	1,2
4. **Classification:** Decision Tree, Model Overfitting, Evaluating Performance of Classifier, Rule-Based Classifier, Nearest-Neighbor Classifier, Bayesian Classifier, ANN, SVM, Ensemble Methods, Class Imbalance Problems.	7	2,3
5. **Association:** Frequent Item Generation, Rule Generation, Compact Representation of Frequent Itemsets, FP-Growth Algorithm	4	2,3
6. **Clustering**: K-Means, Agglomerative Hierarchical Clustering, DBSCAN, Cluster Evaluation, Density-Based Clustering.	4	2,3
7. **Web Mining:** Web Content Mining, Web Structure Mining, Web Usage mining.	3	2,3
8. **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.	4	4
9. **Applications:** Anomaly Detection, Time Series Prediction, Visual and Audio Data Mining, Recommendation Systems.	2	3,4
10. **Advances in Data Mining.**	2	3,4

TEXTBOOKS

1. Pang-Ning Tan, Michael Steinbach and Vipin Kumar, "Introduction to Data Mining", Pearson Education, ISBN: 978-93-3257-140-2

2. Jiawei Han, Micheline Kamber, and Jian Pei, "Data Mining Concepts and Techniques", 3rd Edition, Morgan Kaufmann, ISBN: 978-93-80931-91-3

RECOMMENDED READING

1. M. Berry and G. Linoff, "Mastering Data Mining", John Wiley and Sons, 2nd Edition.

2. I.H. Witten and E. Frantk, "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, 4th Edition.

Program Name	B. Tech. (Computer Engineering)	Semester – VII
Course Code	R4CO4001P	
Course Title	Data Mining and Data Warehousing Lab	
Prerequisite	Statistics, Data Structures and Algorithms	

COURSE OUTCOMES: Students will be able to	
1.	Investigate different data mining and data warehouse tasks and evaluate the algorithms with respect to their accuracy.
2.	Compare the results of a data mining exercise and analyze the results.
3.	Design a data mining solution to a practical problem.

LIST OF EXPERIMENTS:

		Hrs	CO
1.	To perform a multidimensional data model using SQL queries. e.g., snowflake, star and fact constellation schema.	2	1
2.	To perform various OLAP operations such as: slice, dice, roll up, drill up etc.	2	1
3.	To perform data cleaning and preparing for operations	2	1,2
4.	Study of Decision Trees and other classification Algorithms.	4	2,3
5.	To perform association rule mining	4	2, 3
6.	Study of predictive algorithms.	2	2, 3
7.	Study of clustering and its different techniques.	4	2, 3
8.	To perform text mining on the given data warehouse and perform correlation analysis between for the given data sets	2	2, 3

TEXTBOOKS

1   Pang-Ning Tan, Michael Steinbach and Vipin Kumar, "Introduction to Data Mining", Pearson Education, ISBN: 978-93-3257-140-2
2   Jiawei Han, Micheline Kamber, and Jian Pei, "Data Mining Concepts and Techniques", 3rd Edition, Morgan Kaufmann, ISBN: 978-93-80931-91-3

RECOMMENDED READING

1   M. Berry and G. Linoff, "Mastering Data Mining", John Wiley and Sons, 2nd Edition.

2   I.H. Witten and E. Frantk, "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, 4th Edition.