

<b>Devi Ahilya University, Indore, India</b> <b>Institute of Engineering &amp; Technology</b>				<b>I Year M.E. (Computer Engineering Sp. in Software Engineering )</b> <b>(Full Time)</b>			
<b>Subject Code &amp; Name</b>	<b>Instructions Hours per Week</b>			<b>Credits</b>			
<b>SER2G1</b>	<b>L</b>	<b>T</b>	<b>P</b>	<b>L</b>	<b>T</b>	<b>P</b>	<b>Total</b>
<b>Data Mining &amp; Warehousing</b>	<b>3</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>1</b>	<b>1</b>	<b>5</b>
<b>Duration of Theory Paper: 3 Hours</b>							

### Objectives:

- To develop the abilities of critical analysis to data mining systems and applications.
- To implement practical and theoretical understanding of the technologies for data mining
- To understand the strengths and limitations of various data mining models.

### Unit-I

**Data Mining** -Introduction: Data Mining Primitives, Languages, and System Architectures: Data mining primitives, Query language, Designing GUI based on a data mining query language, Knowledge Discovery in Databases (KDD), KDD Process, Data Preprocessing, Data Cleaning, Data Transformation, Data Compression and Dimension Reduction, Principal Component Analysis, Binning Methods.

### Unit-II

**Data Warehousing** –Introduction and Design: Overview and Concepts: Need for data warehousing, basic elements of data warehousing, Architecture and Infrastructure: Architectural components, Infrastructure and metadata. Data Design And Data Representation: Principles of dimensional modeling, Dimensional modeling, data extraction, transformation and loading, data quality. OLAP in data warehouse –ROLAP, MOLAP, HOLAP. OLTP Vs OLAP, Various Data Warehouse Schemas.

### UNIT-III

**Association & Classification Techniques:** Introduction, Frequent itemset mining methods – Apriori, FP-Growth, Pattern evaluation methods, Basic concepts of classification, Decision tree induction, Bayes classification, Rule-based classification.

## UNIT-IV

**Clustering Techniques:** Introduction, Clustering paradigms; Partitioning algorithms – K-Means, K-Medoid, CLARA; Partition based clustering – BIRCH; Density based clustering - DBSCAN; Categorical clustering algorithms, Evaluation of Clustering.

## UNIT-V

**Other DM techniques & Web Mining:** Spatial Mining, Spatial Mining tasks, Spatial clustering, Spatial Trends. Web Mining : Introduction Web content mining, Web structure Mining, Web Usage Mining.

**Temporal and spatial DM:** Temporal association rules, Sequence Mining, GSP, SPADE, SPIRIT, and WUM algorithms, Episode Discovery, Event prediction, Time series analysis.

### Reference Books:

1. Data Mining Techniques; ArunK.Pujari ; University Press.
2. Data Mining Concepts and Techniques, Jiawei Han Micheline Kamber,Jianpei, Morgan Kaufmann.
3. Data Mining; Adriaans&Zantinge; Pearson education.
4. Mastering Data Mining; Berry Linoff; Wiley.
5. PaulrajPonniah, “Data Warehousing Fundamentals”, John Wiley.
6. Text Mining Applications, Konchandy, Cengage