

**Krantiguru Shyamji Krishna Verma Kachhh University**  
**Master of Science (Information Technology)**  
**Semester: I**

<b>Paper Code:</b> CCCS103		<b>Total Credit :</b> 4
<b>Title of Paper:</b> Data Warehousing and Data Mining		<b>Total Marks :</b> 70
		<b>Time :</b> 3 Hrs
<b>Unit</b>	<b>Description</b>	<b>Weighting</b>
<b>I</b>	<p><b>Introduction</b></p> <p>An overview and definition along with clear understanding of the four appearing in the definition.</p> <p>Differences between Operational Database Systems and Data Warehouses</p> <p>Overview of Multi-dimensional Data Model, and the basic differentiation between "Fact"and "Dimension"; Multi-dimensional Cube Concept</p> <p>Hierarchies of "Dimensions" Parameters: Examples and the advantages.</p> <p>Star, Snowflakes, and Fact Constellations Schemas for Multi-dimensional Databases</p> <p>Measures: Their Categorization and Computation, Pre-computation of Cubes, Constraint on Storage Space, Possible Solutions</p> <p>OLAP Operations in Multi-dimensional Data Model: Roll-up, Drill-down, Slice &amp; Dice, Pivot (Rotate). Indexing OLAP Data; Efficient Processing of OLAP Queries. Type of OLAP Servers: ROLAP versus MOLAP versus HOLAP, Metadata Repository</p>	<b>20%</b>
<b>II</b>	<p><b>Data warehouse Architecture</b></p> <p>The Design of A Data Warehouse: A Business Analysis Framework;</p> <p>The Process of Data Warehouse Design, A 3-Tier Data Warehouse Architecture; Enterprise Warehouse, Data mart, Virtual Warehouse, Discovery-Driven Exploration of Data Cubes; Complex Aggregation at Multiple Granularity: Multi-feature Cubes, Constrained Gradient Analysis of Data Cubes</p>	<b>20%</b>
<b>III</b>	<p><b>Pre-Processing</b></p> <p>The need for Pre-processing, Descriptive Data Summarization</p> <p>Data Cleaning: Missing Values, Noisy Data, Data Cleaning as a Process</p> <p>Data Integration &amp; Transformation, Data Cube Aggregation; Attribute Subset Selection, Dimensionality Reduction:(Basic Concepts only).</p> <p>Numerosity Reduction: Regression &amp; Log-linear Models, Histograms, Clustering, Sampling. Data Discretization &amp; Concept Hierarchy Generation</p> <p>For Numerical Data: Binning, Histogram Analysis, Entropy-based Discretization, Interval Merging by <math>\chi</math> Analysis, Cluster Analysis, Discretization by Intuitive Partitioning For Categorical Data</p>	<b>20%</b>
<b>IV</b>	<p><b>Data Mining- An Introduction</b></p> <p>An Overview; What is Data Mining; Data Mining - on What Kind of Data</p> <p>Data Mining Functionalities - What Kind of Patterns Can be Mined;</p> <p>Concept/Class Description: Characterization &amp; Discrimination; Mining Frequent Patterns, Associations, and Correlations; Classification &amp; Prediction; Cluster Analysis; Outlier Analysis, Classification of Data Mining Systems</p> <p>Data Mining Task Primitives, Integration of a Data Mining System with a Database or Data Warehouse System, Major Issues in Data Mining</p>	<b>20%</b>
<b>V</b>	<p><b>Mining Frequent Pattern, Association and correlations</b></p> <p>Basic Concepts: Market Basket Analysis; Frequent Itemsets, Closed Itemsets, and Association Rules; Frequent Pattern Mining: A Roadmap</p> <p>Apriori Algorithm: Finding Frequent Itemsets Using Candidate Generation;</p> <p>Generating Association Rules from Frequent Itemsets; Improving the Efficiency of Apriori. From Association Mining to Correlation Analysis;</p> <p>Interesting: An Example; From Association Analysis to Correlation Analysis</p> <p>Introduction to Classification and Prediction, Supervised learning, Unsupervised learning, Classification by decision tree induction</p>	<b>20%</b>
<b>Basic Text &amp; Reference Books :-</b>		
<b>1.</b>	Jiawei Han & Micheline Kamber, "Data Mining: Concepts & Techniques", Morgan Kaufmann Publishers (2002)	

**Krantiguru Shyamji Krishna Verma Kachchh University**  
**Master of Science (Information Technology)**  
**Semester: I**

<b>Paper Code:</b> CCCS103			<b>Total Credit : 4</b> <b>Total Marks : 70</b> <b>Time : 3 Hrs</b>
<b>Title of Paper:</b> Data Warehousing and Data Mining			
<b>Unit</b>	<b>Description</b>	<b>Total Marks</b>	
I	Q.1 (A) Answer the Following. (Definitions, Blanks, Full Forms, True/False, Match the Following)	06	14
	Q.1 (B) Medium / Long Questions. (With Internal Option)	08	
II	Q.2 (A) Answer the Following. (Definitions, Blanks, Full Forms, True/False, Match the Following)	06	14
	Q.2 (B) Medium / Long Questions. (With Internal Option)	08	
III	Q.3 (A) Short / Medium Questions (With Internal Option)	06	14
	Q.3 (B) Medium / Long Questions. (With Internal Option)	08	
IV	Q.4 (A) Short / Medium Questions (With Internal Option)	06	14
	Q.4 (B) Medium / Long Questions. (With Internal Option)	08	
V	Q.5 (A) Short / Medium Questions (With Internal Option)	06	14
	Q.5 (B) Medium / Long Questions. (With Internal Option)	08	