Krantiguru Shyamji Krishna Verma Kachchh University, Bhuj
Master of Science (Computer Applications & Information Technology)
Semester: VII

Paper Code: CCCS728	Total Credit : 4
Title of Paper: Data Warehousing and Data Mining	**Total Marks :** 70 **Time :** 3 Hrs

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

Paper Code: CCCS728		**Total Credit :** 4 **Total Marks :** 70
Title of Paper: Data Warehousing and Data Mining		**Time :** 3 Hrs

Unit	Description		Total Marks
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	