DATAWARE HOUSING AND DATA MINING

VI Semester: CSE / IT

Course Code	Category	Hours / Week			Credits	Maximum Marks		
		L	T	P	C	CIA	SEE	Total
ACSB14	Core	2	1	-	3	30	70	100
Contact Classes: 45	Tutorial Classes: Nil	Practical Classes: Nil				Total Classes: 60		

OBJECTIVES:

Students will try to learn:

I The scope and essentiality of data warehousing and mining.

II The analyses of data, choosing relevant models and algorithms for respective applications.

III The process and mining of complex datatypes such as streams, spatial, web and multimedia.

IV The research perspectives towards advances in data mining.

COURSE OUTCOMES:

At the end of the course the students should be able to:

CO 1 Relate knowledge discovery in databases (KDD) process with the help of data warehouse fundamentals and data mining functionalities.

CO 2 Demonstrate analytical processing on voluminous data using OLAP operations.

CO 3 Interpret multi-dimensional modeling, for designing data warehouse / data mart / enterprise data warehouse specific to organization.

CO 4 Apply preprocessing techniques on real time data for usage of data analytics and mining algorithms

CO 5 Apply Apriori and FP growth methods on transaction data for frequent pattern mining.

CO 6 Make use of classification algorithms for construction of a data model.

CO 7 Extend data models via classification algorithms for providing predictions.

CO 8 Inspect data model accuracy with the help of respective measures

CO 9 Apply clustering methods for performing unsupervised classification.

CO 10 Infer complex data models with respect to multimedia, streams, spatial and web mining.

CO 11 Examine data mining algorithms for solving real world problems.

MODULE-I DATA WAREHOUSING

Introduction to Data warehouse, A Multi-dimensional data model- Star, Snow flake and Fact constellation schemas, Measures, Concept hierarchy, Data warehouse architecture- A three tier Data warehouse architecture, types of OLAP servers, Data warehouse Implementation, Data Marts, Differences between OLAT and OLTP.

MODULE-II DATA MINING

Introduction, What is Data Mining, Definition, Knowledge Discovery in Data (KDD), Kinds of data bases, Data mining functionalities, Classification of data mining systems, Data mining task primitives, Data Preprocessing: Data cleaning, Data integration and transformation, Data reduction, Data discretization and Concept hierarchy.

MODULE-III	ASSOCIATION RULE MINING
Association Rules: Problem Definition, Frequent item set generation, The APRIORI Principle, support and confidence measures, association rule generation; APRIORI algorithm. FP-Growth Algorithms, Compact Representation of Frequent item Set-Maximal Frequent item set, closed frequent item set.	
MODULE-IV	**CLASSIFICATION AND PRIDICTION**
Issues Regarding Classification and Prediction, Classification by Decision Tree Induction, Bayesian Classification, Classification by Back propagation, Classification Based on Concepts from Association Rule Mining, Other Classification Methods, Prediction, Classifier Accuracy.	
MODULE-V	**CLUSTERING**
Types of data, categorization of major clustering methods, K-means partitioning methods, hierarchical methods, density based methods, grid based methods, model based clustering methods, outlier analysis. Mining Complex Types of Data: Multidimensional Analysis and Descriptive Mining of Complex, Data Objects, Mining Spatial Databases, Mining Multimedia Databases, Mining Time-Series and Sequence Data, Mining Text Databases, Mining the World Wide Web.	

Text Books:

1. Jiawei Han, Michelin Kamber, "Data Mining-Concepts and techniques", Morgan Kaufmann Publishers, Elsevier, 2nd Edition, 2006
2. Alex Berson, Stephen J.Smith, "Data warehousing Data mining and OLAP", Tata McGraw- Hill, 2nd Edition, 2007

Reference Books:

1. Arum K Pujari, "Data Mining Techniques", 3rd Edition, Universities Press, 2005
2. Pualraj Ponnaiah, Wiley, "Data Warehousing Fundamentals", Student Edition, 2004.
3. Ralph Kimball, Wiley, "The Data Warehouse Life Cycle Toolkit", Student Edition, 2006.
4. Vikram Pudi, P Radha Krishna, ―Data Mining, Oxford University, 1st Edition, 2007.

Web References:

1. http://www.anderson.ucla.edu
2. https://www.smartzworld.com
3. http://iiscs.wssu.edu