		WAREHOUSING		
	-	tem (CBCS) scheme		
(Effective froi	m tne academic SEMESTER –	year 2016 -2017) VI		
Subject Code	15CS651	IA Marks	20	
Number of Lecture Hours/Week	3	Exam Marks	80	
Total Number of Lecture Hours	40	Exam Hours	03	
100001100010120000010110000	CREDITS - 0		1 00	
Course objectives: This course will				
Define multi-dimensional data				
• Explain rules related to associ		tion and clustering ana	ılysis.	
 Compare and contrast betwee 		_	•	hms
Module – 1				Teaching
				Hours
Data Warehousing & modeling		-		8 Hours
Operational Database systems and	l Data warehoi	use, Data Warehousi	ng: A	
multitier Architecture, Data warehou	ise models: Ente	erorise warehouse Dat	a mart	
		rprise		
	Transformati	on and loading Ma		
•		on and loading, Me	etadata	
Repository, Data warehouse design	and usage: Bu	siness Analysis frame	etadata ework,	
Repository, Data warehouse design	and usage: Bu	siness Analysis frame	etadata ework,	
Repository, Data warehouse design Data warehouse design process and	and usage: But usage for info	siness Analysis frame formation processing, (etadata ework, Online	
Repository, Data warehouse design Data warehouse design process and analytical processing to multidin	and usage: Bud usage for informational data	ormation processing, omining. Data Cub	etadata ework, Online be: A	
Repository, Data warehouse design Data warehouse design process and analytical processing to multidin multidimensional data model, Sta	and usage: Bud usage for informensional data	siness Analysis frame ormation processing, of mining. Data Cub s and Fact constells	etadata ework, Online be: A ations:	
Repository, Data warehouse design Data warehouse design process and analytical processing to multidin multidimensional data model, Sta Schemas for multidimensional Data	and usage: Bud usage for informensional data ars, Snowflakes a models, Dime	ormation processing, ormation processing, ormation Data Cubs and Fact constellarsions: The role of consions:	etadata ework, Online be: A ations: oncept	
Repository, Data warehouse design Data warehouse design process and analytical processing to multidir multidimensional data model, Sta Schemas for multidimensional Data Hierarchies, Measures: Their Categ	and usage: Bud usage for informensional data ars, Snowflakes a models, Dime	ormation processing, ormation processing, ormation Data Cubs and Fact constellarsions: The role of consions:	etadata ework, Online be: A ations: oncept	
Repository, Data warehouse design Data warehouse design process and analytical processing to multidim multidimensional data model, Sta Schemas for multidimensional Data Hierarchies, Measures: Their Categ Operations.	and usage: Bud usage for informensional data ars, Snowflakes a models, Dime	ormation processing, ormation processing, ormation Data Cubs and Fact constellarsions: The role of consions:	etadata ework, Online be: A ations: oncept	
Repository, Data warehouse design Data warehouse design process and analytical processing to multidin multidimensional data model, Sta Schemas for multidimensional Data Hierarchies, Measures: Their Categ Operations. Module – 2	and usage: Bud usage for informensional data ars, Snowflakes models, Dime gorization and control of the state	usiness Analysis frame ormation processing, of mining. Data Cub is and Fact constellar insions: The role of computation, Typical	etadata ework, Online be: A ations: oncept OLAP	9 Hayye
Repository, Data warehouse design Data warehouse design process and analytical processing to multidimultidimensional data model, Sta Schemas for multidimensional Data Hierarchies, Measures: Their Categ Operations. Module – 2 Data warehouse implementation	and usage: Bud usage for informensional data ars, Snowflakes models, Dime gorization and control of the control	risiness Analysis frame ormation processing, ormation processing, ormation. Data Cubs and Fact constellar sions: The role of computation, Typical ormation, Typical ormation.	etadata ework, Online be: A ations: oncept OLAP	8 Hours
Repository, Data warehouse design Data warehouse design process and analytical processing to multidimultidimensional data model, Sta Schemas for multidimensional Data Hierarchies, Measures: Their Categ Operations. Module – 2 Data warehouse implementation computation: An overview, Indexing	and usage: Bud usage for informensional data ars, Snowflakes models, Dime gorization and contact are properties.	ring: Efficient Data Bitmap index and join	etadata ework, Online be: A ations: oncept OLAP Cube index,	8 Hours
Repository, Data warehouse design Data warehouse design process and analytical processing to multidin multidimensional data model, Sta Schemas for multidimensional Data Hierarchies, Measures: Their Categ Operations. Module – 2 Data warehouse implementation computation: An overview, Indexing Efficient processing of OLAP Queries	and usage: Bud usage for informensional data ars, Snowflakes models, Dime gorization and control of the Data ming OLAP Data: Eas, OLAP server	ring: Efficient Data Bitmap index and join Architecture ROLAP	etadata ework, Online be: A ations: oncept OLAP Cube index, versus	8 Hours
Repository, Data warehouse design Data warehouse design process and analytical processing to multidim multidimensional data model, Sta Schemas for multidimensional Data Hierarchies, Measures: Their Categ Operations. Module – 2 Data warehouse implementation computation: An overview, Indexing Efficient processing of OLAP Queries MOLAP Versus HOLAP.: Introduc	and usage: Bud usage for informensional data ars, Snowflakes models, Dime gorization and compared to the compared of the compa	siness Analysis frame ormation processing, ormation processing, ormation, Data Cubs and Fact constells and Fact constells are role of computation, Typical sing: Efficient Data Bitmap index and join Architecture ROLAP at a mining, Challenges	etadata ework, Online be: A ations: oncept OLAP Cube index, versus s, Data	8 Hours
Repository, Data warehouse design Data warehouse design process and analytical processing to multidir multidimensional data model, State Schemas for multidimensional Data Hierarchies, Measures: Their Catego Operations. Module – 2 Data warehouse implementation computation: An overview, Indexing Efficient processing of OLAP Querie MOLAP Versus HOLAP.: Introduc Mining Tasks, Data: Types of Data,	and usage: Bud usage for informensional data ars, Snowflakes models, Dime gorization and compared to the compared of the compa	siness Analysis frame ormation processing, ormation processing, ormation, Data Cubs and Fact constells and Fact constells are role of computation, Typical sing: Efficient Data Bitmap index and join Architecture ROLAP at a mining, Challenges	etadata ework, Online be: A ations: oncept OLAP Cube index, versus s, Data	8 Hours
Repository, Data warehouse design Data warehouse design process and analytical processing to multidin multidimensional data model, Sta Schemas for multidimensional Data Hierarchies, Measures: Their Categ Operations. Module – 2 Data warehouse implementation computation: An overview, Indexing Efficient processing of OLAP Querie MOLAP Versus HOLAP.: Introduc Mining Tasks, Data: Types of Data, of Similarity and Dissimilarity,	and usage: Bud usage for informensional data ars, Snowflakes models, Dime gorization and compared to the compared of the compa	siness Analysis frame ormation processing, ormation processing, ormation, Data Cubs and Fact constells and Fact constells are role of computation, Typical sing: Efficient Data Bitmap index and join Architecture ROLAP at a mining, Challenges	etadata ework, Online be: A ations: oncept OLAP Cube index, versus s, Data	8 Hours
Repository, Data warehouse design Data warehouse design process and analytical processing to multidir multidimensional data model, Sta Schemas for multidimensional Data Hierarchies, Measures: Their Categ Operations. Module – 2 Data warehouse implementation computation: An overview, Indexing Efficient processing of OLAP Querie MOLAP Versus HOLAP.: Introduc Mining Tasks, Data: Types of Data, of Similarity and Dissimilarity, Module – 3	and usage: Bud usage for informensional data ars, Snowflakes models, Dime corization and compared to the Data ming OLAP Data: Eas, OLAP servertion: What is data Quality, Data Quality, Data Quality, Data data data data data data data data	ring: Efficient Data Sitmap index and join Architecture ROLAP ata Preprocessing, Medical States of the control of the computation of the computati	ctadata ework, Online be: A ations: oncept OLAP Cube index, versus s, Data easures	
Repository, Data warehouse design Data warehouse design process and analytical processing to multidin multidimensional data model, State Schemas for multidimensional Data Hierarchies, Measures: Their Categorations. Module – 2 Data warehouse implementation computation: An overview, Indexing Efficient processing of OLAP Querie MOLAP Versus HOLAP.: Introduc Mining Tasks, Data: Types of Data, of Similarity and Dissimilarity, Module – 3 Association Analysis: Association Association Analysis: Association	and usage: Bud usage for informensional data ars, Snowflakes models, Dime corization and compared to the control of the contro	ring: Efficient Data Cub and Fact constellar computation, Typical comput	ctadata ework, Online oe: A ations: oncept OLAP Cube index, versus s, Data easures	8 Hours
Repository, Data warehouse design Data warehouse design process and analytical processing to multidin multidimensional data model, State Schemas for multidimensional Data Hierarchies, Measures: Their Categoperations. Module – 2 Data warehouse implementation computation: An overview, Indexing Efficient processing of OLAP Querie MOLAP Versus HOLAP.: Introduc Mining Tasks, Data: Types of Data, of Similarity and Dissimilarity, Module – 3 Association Analysis: Association Association, Rule generation. Alto	and usage: Bud usage for informensional data ars, Snowflakes models, Dime corization and compared to the control of the contro	siness Analysis frame ormation processing, ormation processing, ormation, Data Cubs and Fact constellansions: The role of computation, Typical situation, Typical situation, Typical situation, Efficient Data Situation, Challenges at a Preprocessing, Memory and Definition, Frequences for Generating Frequences.	ctadata ework, Online oe: A ations: oncept OLAP Cube index, versus s, Data easures	
Repository, Data warehouse design Data warehouse design process and analytical processing to multidir multidimensional data model, State Schemas for multidimensional Data Hierarchies, Measures: Their Categoperations. Module – 2 Data warehouse implementation computation: An overview, Indexing Efficient processing of OLAP Queries MOLAP Versus HOLAP.: Introduc Mining Tasks, Data: Types of Data,	and usage: Bud usage for informensional data ars, Snowflakes models, Dime corization and compared to the control of the contro	siness Analysis frame ormation processing, ormation processing, ormation, Data Cubs and Fact constellansions: The role of computation, Typical situation, Typical situation, Typical situation, Efficient Data Situation, Challenges at a Preprocessing, Memory and Definition, Frequences for Generating Frequences.	ctadata ework, Online oe: A ations: oncept OLAP Cube index, versus s, Data easures	
Repository, Data warehouse design Data warehouse design process and analytical processing to multidim multidimensional data model, Sta Schemas for multidimensional Data Hierarchies, Measures: Their Categ Operations. Module – 2 Data warehouse implementation computation: An overview, Indexing Efficient processing of OLAP Querie MOLAP Versus HOLAP.: Introduc Mining Tasks, Data: Types of Data, of Similarity and Dissimilarity, Module – 3 Association Analysis: Association Association, Rule generation. Alto	and usage: Bud usage for informensional data ars, Snowflakes models, Dime corization and compared to the control of the contro	siness Analysis frame ormation processing, ormation processing, ormation, Data Cubs and Fact constellansions: The role of computation, Typical situation, Typical situation, Typical situation, Efficient Data Situation, Challenges at a Preprocessing, Memory and Definition, Frequences for Generating Frequences.	ctadata ework, Online oe: A ations: oncept OLAP Cube index, versus s, Data easures	

Module – 3 Association Analysis: Association Analysis: Problem Definition, Frequent Item set Generation, Rule generation. Alternative Methods for Generating Frequent Item sets, FP-Growth Algorithm, Evaluation of Association Patterns. Module – 4 Classification: Basics: General approach to solve classification problem, Decision Trees Induction, Model Over fitting, Evaluating the performance of a classifier, Method for Comparing Classifiers, Rule Based Classifiers, Nearest Neighbor Classifiers, Bayesian Classifiers. Module – 5 Clustering Analysis: Overview, K-Means, Agglomerative Hierarchical Clustering, DBSCAN, Cluster Evaluation, Density-Based Clustering, Graph-Based Clustering, Scalable Clustering Algorithms. Course outcomes: The students should be able to:

Identify data mining problems and implement the data warehouse

- Write association rules for a given data pattern.
- Choose between classification and clustering solution.

Question paper pattern:

The question paper will have TEN questions.

There will be TWO questions from each module.

Each question will have questions covering all the topics under a module.

The students will have to answer FIVE full questions, selecting ONE full question from each module.

Text Books:

- 1. Pang-Ning Tan, Michael Steinbach, Vipin Kumar: Introduction to Data Mining, Pearson, First impression, 2014.
- 2. Jiawei Han, Micheline Kamber, Jian Pei: Data Mining -Concepts and Techniques, 3rd Edition, Morgan Kaufmann Publisher, 2012.

Reference Books:

- 1. Sam Anahory, Dennis Murray: Data Warehousing in the Real World, Pearson, Tenth Impression, 2012.
- 2. Michael.J.Berry,Gordon.S.Linoff: Mastering Data Mining, Wiley Edition, second edition, 2012.