	IG AND DATA	WAREHOUSING		
[As per Choice B	Based Credit Sys	stem (CBCS) scheme		
(Effective fro	m the academic	year 2016 -2017)		
Subject Code	SEMESTER -			
Number of Lecture Hours/Week	15CS651	IA Marks	20	
Total Number of Lecture Hours	3	Exam Marks	80	
Total Number of Lecture Hours	40	Exam Hours	03	
Course objectives: This course will	CREDITS -	03		
Define multi-dimensional dat	enable students	to		
Explain rules related to accomp	a models.			
Explain rules related to assoc Compare and contrast between	nation, classifica	tion and clustering ana	lysis.	
Compare and contrast between Module – 1	en different class	ification and clustering	algori	thms
				Teaching
Data Warehousing & modeling	Davis G			Hours
Data Warehousing & modeling:	Basic Conce	pts: Data Warehousii	ng: A	8 Hours
multitier Architecture, Data warehouse and virtual warehouse Extraction	re models: Ente	rprise warehouse, Data	mart	((/ //)
and virtual warehouse, Extraction, multidimensional data model. Str	Transformation	and loading, Data Cu	be: A	
auta model Ali	TE Shouttlakes	and F		7
nutrumiensional Data	models Dimor	Scioner TL - I C		
Hierarchies, Measures: Their Categ Operations.	orization and co	omputation, Typical (DLAP	
Module – 2		4 \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \		
Data warehouse implementation computation: An overview Indonesia	& Data mini	ng: Efficient Data	Cube	8 Hours
Paradicin. The overview. Indexing	() AP Data: D	trans in des 1 1 1 1		
processing of OLAP (Ineria	C III A P CAPTION	Analis DOLLE		
TOLAY .: Introduct	ion: What is dat	a mining CL II	**	
Tusks, Data. Types of Data, 1	Data Quality, Da	ta Preprocessing, Mea	sures	
of Similarity and Dissimilarity, Module – 3)			
Set Concretion Pul	nalysis: Problen	n Definition, Frequent	Item	8 Hours
Association Analysis: Association A set Generation, Rule generation. Alt	ernative Method	e for Consumi	Item	8 Hours
Item sets, FP-Growth Algorithm, Eval	ernative Method	e for Consumi	Item quent	8 Hours
Item sets, FP-Growth Algorithm, Eval	uation of Associ	ls for Generating Fred ation Patterns.	quent	8 Hours
Item sets, FP-Growth Algorithm, Eval Module – 4 Classification: Decision Trees Indu	uation of Associ	ls for Generating Free lation Patterns.	quent	
Item sets, FP-Growth Algorithm, Eval Module – 4 Classification: Decision Trees Indu Rule Based Classifiers, Nearest Neigh	uation of Associ	ls for Generating Free lation Patterns.	quent	8 Hours
Item sets, FP-Growth Algorithm, Eval Module – 4 Classification: Decision Trees Indu Rule Based Classifiers, Nearest Neigh Module – 5	ernative Method luation of Associ action, Method bor Classifiers, I	ls for Generating Free lation Patterns. for Comparing Classi Bayesian Classifiers.	quent	
Item sets, FP-Growth Algorithm, Eval Module – 4 Classification: Decision Trees Indu Rule Based Classifiers, Nearest Neigh Module – 5 Clustering Analysis: Overview	uation of Association, Method bor Classifiers, I	ls for Generating Free lation Patterns. for Comparing Classi Bayesian Classifiers.	fiers,	8 Hours
Item sets, FP-Growth Algorithm, Eval Module – 4 Classification: Decision Trees Indu Rule Based Classifiers, Nearest Neigh Module – 5 Clustering Analysis: Overview, Clustering, DBSCAN, Cluster Evalu	uation of Association, Method bor Classifiers, I	ls for Generating Free lation Patterns. for Comparing Classi Bayesian Classifiers.	fiers,	
Item sets, FP-Growth Algorithm, Evaluate Module – 4 Classification: Decision Trees Induction Based Classifiers, Nearest Neight Module – 5 Clustering Analysis: Overview, Clustering, DBSCAN, Cluster Evaluated Clustering, Scalable Clustering	ernative Method luation of Association, Method bor Classifiers, I K-Means, Ag luation, Density- Algorithms	ls for Generating Free lation Patterns. for Comparing Classi Bayesian Classifiers.	fiers,	8 Hours
Item sets, FP-Growth Algorithm, Evaluation of Module – 4 Classification: Decision Trees Induction of Classification: Decision Trees Induction of Classifiers, Nearest Neight Module – 5 Clustering Analysis: Overview, Clustering, DBSCAN, Cluster Evaluation of Clustering, Scalable Clustering of Course outcomes: The students should be course outcomes:	ernative Method luation of Association, Method is bor Classifiers, I K-Means, Ag lation, Density- Algorithms. d be able to:	ls for Generating Free lation Patterns. for Comparing Classi Bayesian Classifiers. glomerative Hierarc Based Clustering, Gr	fiers,	8 Hours
Item sets, FP-Growth Algorithm, Evaluated Module – 4 Classification: Decision Trees Induction Based Classifiers, Nearest Neight Module – 5 Clustering Analysis: Overview, Clustering, DBSCAN, Cluster Evaluated Clustering, Scalable Clustering Course outcomes: The students should be Identify data mining problems	ernative Method luation of Association, Method bor Classifiers, I K-Means, Ag luation, Density- Algorithms. d be able to: and implement	ls for Generating Free lation Patterns. for Comparing Classi Bayesian Classifiers. glomerative Hierarc Based Clustering, Grand Classifiers.	fiers,	8 Hours
Item sets, FP-Growth Algorithm, Evaluated Module – 4 Classification: Decision Trees Induction Based Classifiers, Nearest Neigh Module – 5 Clustering Analysis: Overview, Clustering, DBSCAN, Cluster Evaluated Clustering, Scalable Clustering Course outcomes: The students should lidentify data mining problems • Write association rules for a given	ernative Method luation of Association, Method bor Classifiers, I K-Means, Ag lation, Density- Algorithms. d be able to: and implement to yen data nattern	ls for Generating Free lation Patterns. for Comparing Classi Bayesian Classifiers. glomerative Hierarc Based Clustering, Gr	fiers,	8 Hours
Item sets, FP-Growth Algorithm, Evaluated Module – 4 Classification: Decision Trees Induction Based Classifiers, Nearest Neigh Module – 5 Clustering Analysis: Overview, Clustering, DBSCAN, Cluster Evaluated Clustering, Scalable Clustering Course outcomes: The students should lidentify data mining problems • Write association rules for a given	ernative Method luation of Association, Method bor Classifiers, I K-Means, Ag lation, Density- Algorithms. d be able to: and implement to yen data nattern	ls for Generating Free lation Patterns. for Comparing Classi Bayesian Classifiers. glomerative Hierarc Based Clustering, Gr	fiers,	8 Hours
Item sets, FP-Growth Algorithm, Evaluated Module – 4 Classification: Decision Trees Induction Based Classifiers, Nearest Neight Module – 5 Clustering Analysis: Overview, Clustering, DBSCAN, Cluster Evaluated Clustering, Scalable Clustering Course outcomes: The students should be used Clustering and a mining problems • Write association rules for a given Choose between classification and course outcomes: The students of a given Choose between classification and course outcomes are classification and course outcomes.	ernative Method luation of Association, Method bor Classifiers, I K-Means, Agnation, Density-Algorithms. d be able to: and implement the land clustering so	ls for Generating Free lation Patterns. for Comparing Classi Bayesian Classifiers. glomerative Hierarc Based Clustering, Gr	fiers,	8 Hours
Item sets, FP-Growth Algorithm, Eval Module – 4 Classification: Decision Trees Indu Rule Based Classifiers, Nearest Neigh Module – 5 Clustering Analysis: Overview, Clustering, DBSCAN, Cluster Evalu Based Clustering, Scalable Clustering Ourse outcomes: The students should Identify data mining problems Write association rules for a giv Choose between classification a Puestion paper pattern: the question paper will have TEN questions.	ernative Method luation of Association, Method bor Classifiers, I K-Means, Agation, Density-Algorithms. d be able to: and implement the density of the data pattern, and clustering so	ls for Generating Free lation Patterns. for Comparing Classi Bayesian Classifiers. glomerative Hierarc Based Clustering, Gr	fiers,	8 Hours
Item sets, FP-Growth Algorithm, Eval Module – 4 Classification: Decision Trees Indu Rule Based Classifiers, Nearest Neigh Module – 5 Clustering Analysis: Overview, Clustering, DBSCAN, Cluster Evalue Based Clustering, Scalable Clustering Course outcomes: The students should Identify data mining problems Write association rules for a giv Choose between classification a	iction, Method luation of Association, Method lost Classifiers, I K-Means, Agation, Density-Algorithms. In the able to: and implement the and clustering so lestions.	ls for Generating Free lation Patterns. for Comparing Classi Bayesian Classifiers. glomerative Hierarc Based Clustering, Grant the data warehouse	fiers,	8 Hours

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.
- 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.
- Michael.J.Berry, Gordon.S.Linoff: Mastering Data Mining, Wiley Edition, second edtion,2012.

Dent. Of Information Science & Engineering Ava's Institute of Engg. & Technology Mijar, MOODBIDRI - 574 225