Sri Krishna Institute of Technology,
Bangalore

COURSE PLAN

Academic Year 2019-2020

Program:	B E – Computer Science &Engineering
Semester :	6
Course Code:	17CS651
Course Title:	Data Warehousing and Data Mining
Credit / L-T-P:	3/ 3-0-0
Total Contact Hours:	40
Course Plan Author:	Shweta S Bagali

Academic Evaluation and Monitoring Cell

Sri Krishna Institute of Technology
#29,Chimney hills,Hesaraghata Main road, Chikkabanavara Post
Bangalore – 560090, Karnataka, INDIA
Phone / Fax :08023721477/28392221/23721315

Table of Contents

A. COURSE INFORMATION

1. Course Overview

Degree:	BE	Program:	CSE
Semester:	6	Academic Year:	2019-2020
Course Title:	Data Mining and Data Warehousing	Course Code:	17CS651
Credit / L-T-P:	3-0-0	SEE Duration:	180minutes
Total Contact Hours:	40	SEE Marks:	60
CIA Marks:	40	Assignment	1/Module
Course Plan Author:	Shweta S Bagali	Sign ..	
Checked By:		Sign ..	
CO Targets	CIA Target :	SEE Target:	

Note: Define CIA and SEE % targets based on previous performance.

2. Course Content

Content / Syllabus of the course as prescribed by University or designed by institute.

Module	Content	Teaching Hours	Blooms Learning Levels
1	Basic Concepts: Data Warehousing: A multitier Architecture, Data warehouse models: Enterprise warehouse,Data mart and virtual warehouse, Extraction, Transformation and loading, Data Cube: A multidimensional data model, Stars, Snowflakes and Fact constellations: Schemas for multidimensional Data models, Dimensions: The role of concept Hierarchies, Measures: Their Categorization and computation, Typical OLAP Operations.	8	L1-L2
2	Efficient Data Cube computation: An overview, Indexing OLAP Data: Bitmap index and join index, Efficient processing of OLAP Queries, OLAP server Architecture ROLAP versus MOLAP Versus HOLAP.: Introduction: What is data mining, Challenges, Data Mining Tasks, Data: Types of Data, Data Quality, Data Preprocessing, Measures of Similarity and Dissimilarity	8	L1-L2
3	Association Analysis: Problem Definition, Frequent Item set Generation, Rule generation. Alternative Methods for Generating Frequent Item sets, FP-Growth Algorithm, Evaluation of Association Patterns	8	L1-L3
4	Decision Trees Induction,Method for Comparing	8	L1-L2

	Classifiers, Rule Based Classifiers, Nearest Neighbor Classifiers,Bayesian Classifiers			
5	Overview, K-Means, Agglomerative Hierarchical Clustering, DBSCAN, Cluster Evaluation, Density- Based Clustering, Graph- Based Clustering, Scalable Clustering Algorithms	8		L1-L3
-	Total			

3. Course Material

Books & other material as recommended by university (A, B) and additional resources used by course teacher (C).

1. Understanding: Concept simulation / video ; one per concept ; to understand the concepts ; 15 – 30 minutes
2. Design: Simulation and design tools used – software tools used ; Free / open source
3. Research: Recent developments on the concepts – publications in journals; conferences etc.

Modules	Details	Chapters in book	Availability
A	**Text books (Title, Authors, Edition, Publisher, Year.)**	-	-
3,4,5	1.Pang-Ning Tan, Michael Steinbach, Vipin Kumar: Introduction to Data Mining, Pearson, First impression,2014.		In Lib
1,2	2. Jiawei Han, MichelineKamber, Jian Pei: Data Mining - Concepts and Techniques, 3rd Edition,Morgan Kaufmann Publisher, 2012		In Lib
B	**Reference books (Title, Authors, Edition, Publisher, Year.)**	-	-
	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 edtion,2012		In Lib
C	**Concept Videos or Simulation for Understanding**	-	-
C1	https://www.youtube.com/channel/UC4EX8zLiBUalk704IX_zu1Q		You tube
C2	https://www.youtube.com/channel/UCkw4JCwteGrDHIsyIIK04tQ		
C3	https://youtu.be/OGjrtkrOzWl		
C4	https://www.youtube.com/watch?v=e_uoTQUJJTQ		
C5			
D	**Software Tools for Design**	-	-
E	**Recent Developments for Research**	-	-
	Machine Learning		
	Artificial intelligence		
	Data Mining for biological Data Analysis		
F	**Others (Web, Video, Simulation, Notes etc.)**	-	-
1	https://www.slideshare.net/NithinGowda3/vtu-data-mining15cs651-notes-by-nithin-vvcemysuru		
	https://lecturenotes.in/subject/32/data-mining-and-data-warehousing-dmdw		

4. Course Prerequisites

Refer to GL01. If prerequisites are not taught earlier, GAP in curriculum needs to be addressed. Include in Remarks and implement in B.5.

Students must have learnt the following Courses / Topics with described Content . . .

Modules	Course Code	Course Name	Topic / Description	Sem	Remarks	Blooms Level
	18CS53	Database Management System	Relational Database Design using ER-to-Relational mapping, SQL data definition and data types,**Normalization**	5		l3

5. Content for Placement, Profession, HE and GATE

The content is not included in this course, but required to meet industry & profession requirements and help students for Placement, GATE, Higher Education, Entrepreneurship, etc. Identifying Area / Content requires experts consultation in the area.

Topics included are like, a. Advanced Topics, b. Recent Developments, c. Certificate Courses, d. Course Projects, e. New Software Tools, f. GATE Topics, g. NPTEL Videos, h. Swayam videos etc.

Modules	Topic / Description	Area	Remarks	Blooms Level
1	Data warehouse,Data mining	Entrepreneurship		L4
2	Knowledge Discovery	HE		L2
3	Clustering,Classification	HE		L2

B. OBE PARAMETERS

1. Course Outcomes

Expected learning outcomes of the course, which will be mapped to POs.

Modules	Course Code.#	Course Outcome **At the end of the course, student should be able to . . .**	Teach. Hours	Instr Method	Assessment Method	Blooms' Level
1	17CS651	Understand the functionality of the various data mining and data warehousing component	8	Lecture, discussion, PPT	Viva, Assignment	L2
2	17CS651	Design Multidimensional data model for datawarehouse and analyze the market needs by applying suitable OLAP operations.	8	Lecture / PPT, problem solving	Assignment, seminar	L3
3	17CS651	Explain the concept of Data mining system and apply the various preprocessing techniques on large dataset.	8	Lecture / PPT,	Assignment, seminar	L2
4	17CS651	Apply Association rules, classification and clustering techniques to discover various mining techniques.	8	Lecture, discussion/PPT	Question and answer, test	L3
5	17CS651	Explore recent trends in data mining such as web mining, spatial-temporal mining	8	Lecture / PPT, problem solving	Presentation, assignment	L4
-	-	**Total**	**40**	-	-	

2. Course Applications

Write 1 or 2 applications per CO.

Students should be able to employ / apply the course learnings to . . .

Mod	Application Area	CO	Level

ules	Compiled from Module Applications.		
1	Used in Banking, Education, Health,Insurance, Finance, Government,Retail,Services.	1	L3
2	**Used in marketing, industry,** Banking, Education, Health, Insurance	2	L3
3	**Used in marketing, industry,** Banking, Education, Health, Insurance	3	L4
4	Used in market Basket analysis, Medical diagnosis, CRM of credit card business	4	L4
5	Helps in industry, research area.	5	L3

3. Articulation Matrix

CO – PO Mapping with mapping level for each CO-PO pair, with course average attainment.

Modules	CO.#	Course Outcomes — At the end of the course student should be able to ...	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3	Level
1	17CS651	Understand the functionality of the various data mining and data warehousing component	3	1	2	3										2		
2	17CS651	Design Multidimensional data model for datawarehouse and analyze the market needs by applying suitable OLAP operations.	3	3	2	3		2		2						2		
3	17CS651	Explain the concept of Data mining system and apply the various preprocessing techniques on large dataset.	3	3	3	2									2	2		
4	17CS651	Apply Association rules, classification and clustering techniques to discover various mining techniques.	3	2	3	2		2		2					3	2		
5	17CS651	Explore recent trends in data mining such as web mining, spatial-temporal mining	3	3	3	1	2								3	3		
-		Average																-
-	PO, PSO	1.Engineering Knowledge; 2.Problem Analysis; 3.Design / Development of Solutions; 4.Conduct Investigations of Complex Problems; 5.Modern Tool Usage; 6.The Engineer and Society; 7.Environment and Sustainability; 8.Ethics; 9.Individual and Teamwork; 10.Communication; 11.Project Management and Finance; 12.Life-long Learning; S1.Software Engineering; S2.Data Base Management; S3.Web Design																

4. Curricular Gap and Content

Topics & contents not covered (from A.4), but essential for the course to address POs and PSOs.

Modules	Gap Topic	Actions Planned	Schedule Planned	Resources Person	PO Mapping
1					
2					

C. COURSE ASSESSMENT

1. Course Coverage

Assessment of learning outcomes for Internal and end semester evaluation.

Modules	Title	Teach. Hours	No. of question in Exam						CO	Levels
			CIA-1	CIA-2	CIA-3	Asg	Extra Asg	SEE		
1	Data warehouse	8	2			1	1	2	CO1	L2
2	Data mining	8	2			1	1	2	CO2	L3
3	Association Analysis	8		2		1	1	2	CO3	L2
4	Classification	8		2	2	1	1	2	CO4	L3
5	Clustering	8			2	1	1	2	CO5	L3
-	**Total**	**40**	**4**	**4**	**4**	**5**	**5**	**10**	**-**	**-**

2. Continuous Internal Assessment (CIA)

Assessment of learning outcomes for Internal exams. Blooms Level in last column shall match with A.2.

Modules	Evaluation	Weightage in Marks	CO	Levels
1, 2	CIA Exam – 1	30	CO1,CO2	L2,L3
3, 4	CIA Exam – 2	30	CO3,C04	L3
5	CIA Exam – 3	30	CO5	L4
1, 2	Assignment - 1	10		
3, 4	Assignment - 2	10		
5	Assignment - 3	10		
1, 2	Seminar - 1		-	-
3, 4	Seminar - 2		-	-
5	Seminar - 3		-	-
1, 2	Quiz - 1		-	-
3, 4	Quiz - 2		-	-
5	Quiz - 3		-	-
1 - 5	Other Activities – Mini Project	-		
	Final CIA Marks	40	**-**	**-**

D1. TEACHING PLAN - 1

Module - 1

Title:	Basic Concepts: Data Warehousing		Appr Time:	10 Hrs
a	**Course Outcomes**		**CO**	**Blooms**
	The student should be able to:			
	Understand the functionality of the various data mining and data warehousing component			
b	**Course Schedule**		-	-
Class No	**Portion covered per hour**		**-**	**-**
1	Data Warehousing & modeling: Basic Concepts: Data Warehousing:		1	L2
2	A multitier Architecture		1	L2
3	Data warehouse models: Enterprise warehouse, Data mart and virtual		1	L2

	warehouse			
4	Extraction, Transformation and loading,	1	L2
5	Data Cube: A multidimensional data model, Stars, Snowflakes and Fact constellations	1	L2
6	Schemas for multidimensional Data models	1	L2
7	Dimensions: The role of concept Hierarchies,	1	L2
8	Measures: Their Categorization and computation, Typical OLAP Operations	1	L2
c	**Application Areas**		
-	Students should be able employ / apply the Module learnings to . . .		
1	Banking,Education,Finanace,Indusrty		
2	Retailer,Health		
d	**Review Questions**		
-			
1	Define online analytical processing?	1	1
2	List the key features of data warehouse?	1	2
3	Differentiate operational database systems and data warehousing?	1	4
4	Discuss briefly about the multidimensional data models?	1	2
5	Describe the three-tier data warehousing architecture?	1	2
6	Compare Enterprise warehouse, data mart and virtual warehouse?	1	2
7	Name the OLAP operations? Express what is slice and dice operation?	1	2
8	List the various multidimensional models? Explain about the star schema, snowflake schema and fact constellation schema?	1	2
9	Define the use of concept hierarchy?	1	1
e	**Experiences**	-	-
1			
2			

Module – 2

Title:		Appr Time:	10 Hrs
a	*Course Outcomes*	CO	Blooms
-		-	Level
	The student should be able to:		
	Design Multidimensional data model for datawarehouse and analyze the market needs by applying suitable OLAP operations.	CO2	L3
b	*Course Schedule*	-	-
Class No	Portion covered per hour	-	-
1	Data warehouse implementation& Data mining: Efficient Data Cube computation	CO2	
2	Data warehouse implementation& Data mining: Efficient Data Cube computation	CO2	
3	An overview, Indexing OLAP Data: Bitmap index and join index	CO2	
4	An overview, Indexing OLAP Data: Bitmap index and join index	CO2	
5	Efficient processing of OLAP Queries, OLAP server Architecture ROLAP versus	CO2	

6	MOLAP Versus HOLAP. : Introduction: What is data mining, Challenges,		CO2	
7	Data Mining Tasks, Data: Types of Data		CO2	
8	Data Quality, Data Preprocessing, Measures of Similarity and Dissimilarity		CO2	
c	**Application Areas**		-	-
-	Students should be able employ / apply the Module learnings to . . .		-	-
1	Banking,Education,Finanace,Indusrty			
2	Retailer,Health			
d	**Review Questions**		-	-
-				
1	Discuss the efficient processing of OLAP queries?			
2	Describe the common techniques are used in ROLAP and MOLAP?			
3	Discuss about types of OLAP Servers?			
4	Explain Indexing OLAP Data?			
5	Define data mining?Discuss motivating challenges in data mining?			
6	Explain the steps involved in data preprocessing?			
7	Explain the types of the data sets?			
8	Define data mining. Explain the process of knowledge discovery in databases			
9	Define DW by Inmon and Imhoff ?Explain in detail with example			
10	What is data cube mesures?How it is categorized? Explain			
e	**Experiences**		-	-
1			CO3	L2
2				

E1. CIA EXAM – 1

a. Model Question Paper - 1

Crs Code:	17cs651	Sem:6	I	Marks:30		Time:		
Course:	Data mining and Data warehousing							
-	-	**Note: Answer all questions, each carry equal marks. Module : 1, 2**				**Marks**	**CO**	**Level**
		PART –A						
1	a	Define DW by Inmon and Imhoff ?Explain in detail with example				5	CO1	L4
	b	Explain ETL process and explain three data warehouse models.				5	CO1	L2
	c	Define 1) Dimension 2)Measures 3) Fact table 4)Foreign key 5) Primary key				5	CO1	L2
		OR						
2	a	Give the definition of Data warehousing. With a schematic diagram explain the working of general DW architecture				8	CO1	L1,L2
	b	Describe the operations of data cubes(OLAP operations)?				7	CO1	L2
		Part – B						
3	a	Describe the servers involved in implementatiton of a warehouse server				7	CO2	L4
	b	Define data mining. Explain the process of knowledge discovery in databases				4	CO2	L1,L2
	c	Explain indexing OLAP data.Bit map index with example				4	CO2	L2
		OR						
4	a	What are the properties necessary to describe attributes? Explain different types of attributes?				6	CO2	L2
	b	Explain various tasks of data mining with example for each				9	CO2	L2

b. Assignment -1

Model Assignment Questions						
Crs Code:	17CS651	Sem:6	Marks:10	Time:		
Course:	Data Mining and Data warehousing					

SNo	Assignment Description	Marks	CO	Level
1	Define online analytical processing?	2	1	L1
2	List the key features of data warehouse?	3	1	L1
3	Differentiate operational database systems and data warehousing?	5	1	L4
4	Discuss briefly about the multidimensional data models?	8	1	L2
5	Describe the three-tier data warehousing architecture?	8	1	L2
6	Compare Enterprise warehouse, data mart and virtual warehouse?	6	1	L2
7	Name the OLAP operations? Express what is slice and dice operation?	8	1	L1
8	List the various multidimensional models? Explain about the star schema, snowflake schema and fact constellation schema?	10	1	L1
9	Define the use of concept hierarchy?	5	1	L1
10	Discuss the efficient processing of OLAP queries?	5	2	L2
11	Describe the common techniques are used in ROLAP and MOLAP?	5	2	L2
12	Discuss about types of OLAP Servers?	8	2	L2
13	Explain Indexing OLAP Data?	6	2	L1
14	Define data mining?Discuss motivating challenges in data mining?	6	2	L1
15	Explain the steps involved in data preprocessing?	9	2	L1
16	Explain the types of the data sets?	8	2	L1
17	Explain various tasks of data-mining with example for each	8	2	L1
18	Define data mining. Explain the process of knowledge discovery in databases	6	2	L1
19	What are the properties necessary to describe attributes? Explain different types of attributes	6	2	L1
20	Explain features subset selection process with a flow chart?	5	2	L2
21	Explain the types of data?	8	2	L2
22	Explain different sampling approaches with example?	6	2	L2
23	Consider the following vectors x and y. Find cosine, correlation , Euclidean, jaccard x=(0, 1, 0,1) y=(1,0,1,0)	6	2	L3

D2. TEACHING PLAN - 2

Module – 3

Title:	Association Analysis		Appr Time:	10 Hrs
a	*Course Outcomes*		CO	Blooms
-	At the end of the topic the student should be able to . . .		-	Level
	Explain the concept of Data mining system and apply the various preprocessing techniques on large dataset.		CO3	L3

b	*Course Schedule*		
Class No	**Portion covered per hour**	**-**	**-**
1	Association Analysis: Problem Definition,		
2	Frequent Item set Generation,		
3	Frequent Item set Generation		
4	Rule generation. Alternative Methods for Generating Frequent Item sets		
5	Rule generation. Alternative Methods for Generating Frequent Item sets		
6	FPGrowth Algorithm		
7	Evaluation of Association Patterns		
8	Evaluation of Association Patterns		
c	**Application Areas**	**-**	**-**
-	Students should be able employ / apply the Module learnings to . . .	-	-
	Market Basket Analysis		
d	**Review Questions**	-	-
-	The attainment of the module learning assessed through following questions	CO3-	L2
1	Explain frequent item set generation of the Aprior algorithm	CO3-	L2
2	Explain rule generation in Aprior algorithm	CO3-	L2
3	Explain alternative methods for generating frequent itemsets	CO3-	L2
4	Explain briefly FP-growth algorithm	CO3-	L2
5	What is Association Analysis?Define Support & confidence. With an example	CO3-	L1
6	Explain the various measures of evaluating association patterns	CO3-	L2
7	Consider the following transaction dataset.Describe the construction of FP-tree in FP-Growth algorithm 	Tid	Items	 	1	{a,b}	 	2	{b,c,d}	 	3	{a,c,d,e}		CO3-	L3
e	**Experiences**	-	-
1			
2			

Module – 4

	Title:	Classification	Appr Time:	10 Hrs
a	*Course Outcomes*		**CO**	**Blooms Level**
-	At the end of the topic the student should be able to . . .		-	
	Apply Association rules, classification and clustering techniques to discover various mining techniques.		CO4	L3
b	*Course Schedule*			
Class No	**Portion covered per hour**		**-**	**-**
1	Classification: Decision Trees Induction,			
2	Classification: Decision Trees Induction,			
3	Method for Comparing Classifiers			
4	Method for Comparing Classifiers			
5	Rule Based Classifiers			
6	Rule Based Classifiers, Nearest Neighbor Classifiers			
7	Nearest Neighbor Classifiers			
8	Bayesian Classifiers.			
c	**Application Areas**		**-**	**-**
-	Students should be able employ / apply the Module learnings to . . .		-	-
d	**Review Questions**		-	-
-	The attainment of the module learning assessed through following questions		CO4	-
1	With neat block diagram, explain general approach to solve classification problem			
2	Explain how to build a decision tree using Hunt's algorithm			L3
3	Explain rule based classifiers with illustration			L2
4	Explain K-nearest neighbour classification algorithm with example			L2
5	Explain characteristics of nearest neighbour classifies			L2
6	Explain different methods for expressing attribute test condition			L2
7	Explain the characteristics of decision tree induction			L2
8	Explain Naives bayes classifier and its characteristics			L2
9	How Bayes theorem can be used for solving a classification problem? Explain			L3

e	Experiences		-	-
1				
2				

E2. CIA EXAM – 2

a. Model Question Paper - 2

Crs Code:	17CS651	Sem:6		Marks:30		Time		
Course:	DMDW							
-	-	Note: Answer all questions, each carry equal marks. Module : 3, 4				Marks	CO	Level
		PART A						
	1a	Explain frequent item set generation of the Aprior algorithm				5	CO3	2
	b.	What is Association Analysis?Define Support & confidence. With an example				5	CO3	1
	c.	Explain briefly FP-growth algorithm				5	CO3	2
		OR						
	2.a	Explain rule generation in Aprior algorithm				6	CO3	2
	b	Explain alternative methods for generating frequent itemsets				9	CO3	2
		PART B						
	3 a	With neat block diagram, explain general approach to solve classification problem				6	CO4	3
	b	How Bayes theorem can be used for solving a classification problem? Explain				9	CO4	4
		OR						
	4a	Explain K-nearest neighbour classification algorithm with example				7	CO4	3
	b	Explain Naives bayes classifier and its characteristics				8	CO4	3

b. Assignment – 2

	Model Assignment Questions								
Crs Code:	17CS651	Sem:	6th	Marks:	10	Time:			
Course:	DMDW								
SNo	Assignment Description						Marks	CO	Level
1	Explain frequent item set generation of the Aprior algorithm						5	3	2
2	Explain rule generation in Aprior algorithm						6	3	2
3	Explain alternative methods for generating frequent itemsets						6	3	2
4	Explain briefly FP-growth algorithm						5	3	2
5	What is Association Analysis?Define Support & confidence. With an example						5	3	2
6	Explain the various measures of evaluating association patterns						7	3	1
7	Consider the following transaction dataset.Describe the construction of FP-tree in FP-Growth algorithm						7	3	3
		Tid	Items						

	1	{a,b}				
	2	{b,c,d}				
	3	{a,c,d,e}				
8	With neat block diagram, explain general approach to solve classification problem			7	4	2
9	Explain how to build a decision tree using Hunt's algorithm			6	4	2
10	Explain rule based classifiers with illustration			6	4	2
11	Explain K-nearest neighbour classification algorithm with example			7	4	2
12	Explain characteristics of nearest neighbour classifies			5	4	2
13	Explain different methods for expressing attribute test condition			6	4	2
14	Explain the characteristics of decision tree induction			6	4	2
15	Explain Naives bayes classifier and its characteristics			8	4	2
16	How Bayes theorem can be used for solving a classification problem? Explain			9	4	3

D3. TEACHING PLAN - 3

Module – 5

Title:	Clustering Analysis	Appr Time:	10 Hrs
a	***Course Outcomes***	**CO**	**Blooms**
-	At the end of the topic the student should be able to . . .	-	**Level**
	Explore recent trends in data mining such as web mining, spatial-temporal mining	CO5	L4
b	***Course Schedule***	CO5-	-
Class No	**Portion covered per hour**	**-**	**-**
1	Clustering Analysis: Overview		
2	K-Means,		
3	Agglomerative Hierarchical Clustering		
4	DBSCAN, Cluster Evaluation		
5	Density-Based Clustering		
6	Graph-Based Clustering		
7	Graph-Based Clustering		
8	Scalable Clustering Algorithms		
c	**Application Areas**	**-**	**-**
-	Students should be able employ / apply the Module learnings to . . .	-	-
	Medical field		
		CO5	

d	Review Questions	-	-
-	The attainment of the module learning assessed through following questions	-	-
1	Describe different types of clustering mechanisms		L1
2	Explain DBSCAN algorithm. How the parameters are selected		L2
3	List out important issues for cluster validation		L1
4	Illustrate Grid-based clustering algorithms.How clusters are formed from dense-grid cells.		L2
5	Develop DENCLUE algorithm for kernel density estimation		L3
6	What is cluster analysis?		L1
7	Explain briefly agglomerative hierarchical clustering with example		L2
8	Briefly explain BIRCH scalable clustering algorithm		L2
9	Explain the different types of cluster analysis methods and discuss their features		L2
10	Write and Explain basic K-means algorithm		L4
11	Explain different methods for computing distance between clusters		L2
e	Experiences	-	-
1		CO10	L2
2		CO9	

E3. CIA EXAM – 3

a. Model Question Paper - 3

Crs Code	17CS651	Sem:6		Marks:30		Time:		
Course:	DMDW							
-	-	Note: Answer all questions, each carry equal marks. Module : 5				Marks	CO	Level
		PART A						
	1 a	Describe different types of clustering mechanisms				6	5	L2
	b	Explain DBSCAN algorithm. How the parameters are selected				7	5	L2
	c	What is cluster analysis? with example				2	5	L1
		OR						
	2a	List out important issues for cluster validation				6	5	L1
	b	Illustrate Grid-based clustering algorithms.How clusters are formed from dense-grid cells.				9	5	L3
		PART B						
	3 a	Develop DENCLUE algorithm for kernel density estimation				7	5	L2
	b	Explain briefly agglomerative hierarchical clustering with example				8	5	L2
		OR						
	4 a	Briefly explain BIRCH scalable clustering algorithm				8	5	L2
	b	Write and Explain basic K-means algorithm				7	5	L4

b. Assignment – 3

Crs Code:17CS651	Sem:6	Marks:10	Time:
Course:DMDW			

SNo	Assignment Description	Marks	CO	Level
1	Describe different types of clustering mechanisms	6	5	
2	Explain DBSCAN algorithm. How the parameters are selected	7	5	
3	List out important issues for cluster validation	6	5	
4	Illustrate Grid-based clustering algorithms.How clusters are formed from dense-grid cells.	9	5	
5	Develop DENCLUE algorithm for kernel density estimation	7	5	
6	What is cluster analysis?	2	5	
7	Explain briefly agglomerative hierarchical clustering with example	8	5	
8	Briefly explain BIRCH scalable clustering algorithm	8	5	
9	Explain the different types of cluster analysis methods and discuss their features	8	5	
10	Write and Explain basic K-means algorithm	7	5	
11	Explain different methods for computing distance between clusters	6	5	

F. EXAM PREPARATION

1. University Model Question Paper

Course:		Data mining and Data warehousing				Month / Year	June 2020		
Crs Code:		17CS651	Sem:	6	Marks:	60	Time:	180 minutes	
Module		Answer all FIVE full questions. All questions carry equal marks.					Marks	CO	Level
1	1.a	What is datawarehouse? Differentiate between ODS and data warehouse					5	CO1	L1
	B	Explain With a neat diagram, a three-tier data warehouse architecture					7	CO1	L2
		OR							
	2.a	Explain OLAP operations with examples					8	CO1	L2
	b	What is metadata in data warehouse?What it contains?					4	CO1	L1
2	3 a	Explain indexing OLAP data:Bitmap index and join index with example					8	CO2	L2
	b	Explain ROLAP versus MOLAP					4	CO2	L2
		OR							
	4 a	What is data mining?Briefly explain the motivating challenges					5	CO2	L1
	b	Explain data pre-processing steps					7	CO2	L2
3	5 a	Explain frequent itemset generation of the Aprior algorithm					6	CO3	L4
	b	Explain rule generation in Apriori algorithm					6	CO3	L2
		OR							
	6 a	Explain alternative methods for generating frequent itemsets					6	CO3	L2
	b	Explain briefly FP-growth algorithm					6	CO3	L2
4	7 a	With neat block diagram explain general approach to solve classification					6	CO4	L3

		problem			
	b	Explain how to build a decision tree using Hunt's algorithm	6	CO4	L2
		OR			
	8 a	Explain rule based classifiers with illustration	6	CO4	L2
	b	Explain K-nearest neighbour classification algorithm with example	6	CO4	L2
5	9 a	What is cluster analysis?Explain different types of clusterings	6	CO5	L1
	b	Explain briefly agglomerative hierarchical clustering with example	6	CO5	L2
		OR			
	10 a	Explain DBSCAN algorithm with example	6	CO5	L2
	b	Briefly explain BIRCH scalable clustering algorithm	6	CO5	L2

2. SEE Important Questions

Course:		Data mining and Data warehousing				Month / Year		June 2020	
Crs Code:		17CS651	Sem:	6	Marks:	60	Time:	180 minutes	
Module		Answer all FIVE full questions. All questions carry equal marks.					**Marks**	**CO**	**Level**
1	1.a	Describe the operations of data cubes(OLAP operations)?					6	CO1	L1
	B	Explain With a neat diagram, a three-tier data warehouse architecture					6	CO1	L2
		OR							
	2.a	Explain ETL process and explain three data warehouse models					6	CO1	L2
	b	What is metadata in data warehouse?What it contains?					4	CO1	L1
	c	Define DW by Inmon and Imhoff?					2	CO1	L1
2	3 a	Explain indexing OLAP data:Bitmap index and join index with example					8	CO2	L2
	b	Explain ROLAP versus MOLAP					4	CO2	L2
		OR							
	4 a	What is data mining?Briefly explain the motivating challenges					5	CO2	L1
	b	Explain data pre-processing steps					7	CO2	L2
3	5 a	Explain frequent itemset generation of the Aprior algorithm					6	CO3	L4
	b	Explain rule generation in Apriori algorithm					6	CO3	L2
		OR							
	6 a	Explain alternative methods for generating frequent itemsets					6	CO3	L2
	b	Explain briefly FP-growth algorithm					6	CO3	L2
4	7 a	With neat block diagram explain general approach to solve classification problem					6	CO4	L3
	b	Explain how to build a decision tree using Hunt's algorithm					6	CO4	L2
		OR							

	8 a	Explain rule based classifiers with illustration	6	CO4	L2
	b	Explain K-nearest neighbour classification algorithm with example	6	CO4	L2
5	9 a	What is cluster analysis?Explain different types of clusterings	6	CO5	L1
	b	Explain briefly agglomerative hierarchical clustering with example	6	CO5	L2
		OR			
	10 a	Explain DBSCAN algorithm with example	6	CO5	L2
	b	Briefly explain BIRCH scalable clustering algorithm	6	CO5	L2

Course Outcome Computation

Academic Year:

Odd / Even semester

LV Threshold : 3:>60%, 2:>=50% and <=60%, 1: <=49%
CO1 Computation :(2+2+2+3)/4 = 10/4=2.5

PO Computation