COURSEPACK

SCHEME

<table>
<tr><td>Course Title</td><td colspan="3">Data Mining and Web Algorithms</td><td colspan="2">Course Type</td><td colspan="6">Theory</td></tr>
<tr><td>Course Code</td><td colspan="3">R1UC617T</td><td colspan="2">Class</td><td colspan="6">B.Tech</td></tr>
<tr><td rowspan="6">Instruction delivery</td><td>Activity</td><td>Credits</td><td>Credit Hours</td><td colspan="4" rowspan="2">Total Number of Classes per Semester</td><td colspan="2" rowspan="2">Assessment in Weightage</td></tr>
<tr><td>Lecture</td><td>3</td><td>3</td></tr>
<tr><td>Tutorial</td><td>0</td><td>0</td><td rowspan="4">Theory</td><td rowspan="4">Tutorial</td><td rowspan="4">Practical</td><td rowspan="4">Self-study</td><td rowspan="4">CIE</td><td rowspan="4">SEE</td></tr>
<tr><td>Practical</td><td>0</td><td>0</td></tr>
<tr><td>Self-study</td><td>0</td><td>0</td></tr>
<tr><td>Total</td><td>3</td><td>3</td></tr>
<tr><td>Course Lead</td><td colspan="2">Mr.V.Janakiraman</td><td>Course Coordinator</td><td colspan="6">Mr. Soumalya Ghosh</td></tr>
<tr><td>Names Course Instructors</td><td colspan="4">Theory</td><td colspan="6">Practical</td></tr>
<tr><td colspan="4">1. Mr.V.Janakiraman
2. Mr. Soumalya Ghosh</td><td colspan="6"></td></tr>
</table>

COURSE OVERVIEW

This course is intended to provide a thorough and sound understanding of the essential theoretical base for data mining algorithms along with the web analysis algorithms. The background provided by the course should ensure the real-world applicability of the data mining algorithms and how to address the problems.

PREREQUISITE COURSE

PREREQUISITE COURSE REQUIRED	No	
If, yes please fill in the Details	**Prerequisite course code**	**Prerequisite course name**
	NA	**NA**

COURSE OBJECTIVE

1. To build up the basic theoretical background on data mining.
2. To apply different data mining algorithms to address the problem.
3. To analyze strategies to develop solutions for the real-world situation.
4. To apply the web algorithms.

COURSE OUTCOMES (COs)

After the completion of the course, the student will be able to:

CO No.	Course Outcomes
R1UC617T.1	Get Basic introductory knowledge on data mining
R1UC617T.2	Get the idea about Data Mining Techniques
R1UC617T.3	Introduction to classifications algorithms
R1UC617T.4	Student able to get extensive knowledge on Clustering approaches
R1UC617T.5	Student apply the Web mining algorithms.

BLOOM'S LEVEL OF THE COURSE OUTCOMES

INTEGRATED

CO No.	Remember **BTL1**	Understand **BTL2**	Apply **BTL3**	Analyse **BTL4**	Evaluate **BTL5**	Create **BTL6**
R1UC617T.1	√	√				
R1UC617T.2			√			
R1UC617T.3				√	√	
R1UC617T.4				√	√	
R1UC617T.5						√

PROGRAM OUTCOMES (POs): AS DEFINED BY CONCERNED THE APEX BODIES

COURSE ARTICULATIONMATRIX

COs#/POs	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**	**PO9**	**PO10**	**PO11**	**PO12**	**PSO1**	**PSO2**
R1UC602C.1	2	2			3									
R1UC602C.2	1	2			3				1					
R1UC602C.3					3						1			
R1UC602C.4					3	3							2	2
R1UC617T.5						3				2			3	3

Note: 1-Low, 2-Medium, 3-High

Assessment Pattern for Theory Course:

Type of Course (T)	**CIE**			**Total Marks**		**Final Marks CIE*0.5+SEE*0.5**
	IA1#	MTE	IA2#	CIE	SEE	
THEORY	25	50	25	100	100	100

COURSE CONTENT

Content
THEORY: Basic Data Mining Tasks – Data Mining Versus Knowledge Discovery in Data Bases – Data Mining Issues – Data Mining Matrices – Social Implications of Data Mining – Data Mining from Data Base Perspective. Data Mining Techniques – a Statistical Perspective on data mining – Similarity Measures – Decision Trees – Neural Networks – Genetic Algorithms. Classification: Introduction – Statistical – Based Algorithms – Distance Based Algorithms – Decision - Tree Construction Principle- Pruning Techniques - Integration of Pruning and Construction - Ideal Algorithm - Issues and Challenges in DM . Clustering Tree – Based Algorithms – Neural Network Based Algorithms – Rule Based Algorithms – Combining Techniques: Introduction – Similarity and Distance Measures – Outliers – Hierarchical Algorithms. Partitioned Algorithms. Web mining algorithms - Algorithms for mining streaming data - streaming data types - Online mining - various streaming algorithms - application and comparison of web mining algorithms - Web Structure Mining - Web Usage Mining - Text Mining - Unstructured Text.

LESSON PLAN FOR COMPREHENSIVE COURSES

FOR THEORY 15 weeks * 3 Hours = 45 Classes) (1credit = 1Lecture Hour)

L-No	Topic for Delivery	Skill	Competency
1	Basic Data Mining Tasks	Student will get introductory knowledge on data mining	Student able to build the foundation on data mining
2	Data Mining Versus Knowledge Discovery in Data Bases		
3	Data Mining Issues		
4	Data Mining Matrices		
5	Social Implications of Data Mining		
6	Data Mining from Data Base Perspective.		
7	Data Mining Techniques	Get the idea about Data Mining Techniques	Student should able to build the strong concept on data mining
8	a Statistical Perspective on data mining		
9	Similarity Measures		
10	Decision Trees		
11	Neural Networks		
12	Genetic Algorithms		
13	Classification	Getting knowledge on classifications algorithms	
14	Introduction		
15	Statistical		
16	Based Algorithms		
17	Distance Based Algorithms		
18	Decision.		
19	Clustering Tree	Understand the Clustering approaches	
20	Based Algorithms		
21	Neural Network Based Algorithms		
22	Rule Based Algorithms		
23	Combining Techniques: Introduction		
24	Similarity and Distance Measures		
25	Outliers		
26	Tree Construction Principle		
27	Hierarchical Algorithms		
28	Partitioned Algorithms.		
29	Web mining algorithms	Get concept on web mining	Able to apply the Web mining algorithms
30	Algorithms for mining streaming data		
31	streaming data types		
32	Online mining		
33	various streaming algorithms		
34	Revision		
35	application of web mining algorithms		
36	comparison of web mining algorithms		
37	Pruning Techniques		
38	Integration of Pruning and Construction		
39	Ideal Algorithm		
40	Issues and Challenges in DM		

41	Web Structure Mining		
42	Web Usage Mining		
43	Text Mining		
44	Unstructured Text.		
45	Revision		

Text Book

- Jiawei Han, Jian Pei, Hanghang Tong, “Data Mining Concepts and Techniques”, Elsevier Science

Reference Books

- Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar - Introduction to Data Mining, Pearson
- Patel and Barik, ”Introduction to Web Technology & Internet”, Acme Learning.
- Mohammed J. Zaki and Wagner Meira, Jr, Data Mining and Machine Learning: Fundamental, Cambridge University Press

Journals/Magazines/Govt. Reports/Gazatte/Industry Trends

- Data Mining and Knowledge Discovery: https://link.springer.com/journal/10618

Webliography

- Tutorial point https://www.tutorialspoint.com/data_mining/index.htm

SWAYAM/NPTEL/MOOC Certification

- NPTEL: Data Mining, Prof. Prabitra Mitra,IIT Kharagpur (https://onlinecourses.nptel.ac.in/noc21_cs06/preview)

PRACTICE PROBLEMS

SNo	Problem	BTL

STUDENT-CENTEREDLEARNING (SELF-LEARNING TOWARDS LIFE-LONG-LEARNING)

A) COURSE-BASEDPROJECT (Psychomotor skills) (Min 45 Projects*)

SNo	Suggested Projects	BTL

B) SELF-LEARNINGTHROUGHMOOCs(CognitiveSkills):