AMIRAJ COLLEGE OF ENGINEERING AND TECHNOLO

DEPARTMENT OF COMPUTER ENGINEERING

TERM DATE: 20-JUN-2016 TO 23-0CT-2016

HOURS	DETAILS OF TOPIC TO BE COVERED	PLANNED D
I	**Introduction to data mining (DM)**	
1	Motivation for Data Mining - Integration of a Data Mining system with a Database or a Data Warehouse	
2	Data Mining-Definition and Functionalities	1/7/2
3	Classification of DM Systems	TC
4	DM task primitives Languages and System Architectures: Task relevant data - Kind of Knowledge to be mined	11/7/
5	Integration of a Data Mining system with a Database or a Data Warehouse	
6	Issues in DM	
7	KDD Process	
II	**Data Pre-processing**	
8	Why to pre-process data?	12/7/
9	Data cleaning: Missing Values, Noisy Data - Data Integration and transformation	TC
10	Data Reduction: Data cube aggregation	26/7/
11	Dimensionality reduction	
12	Data Compression - Numerosity Reduction	
13	Discretization and Concept Hierarchy	
III	**Concept Description and Association Rule Mining**	
10	What is concept description? - Data Generalization and summarization	
11	based characterization - Attribute relevance	
12	class comparisons	28/7/
13	Association Rule Mining: Market basket analysis	TC
14	basic concepts - Finding frequent item sets	11/8/
15	Apriori algorithm - generating rules	
16	Improved Apriori algorithm	
17	Incremental ARM	
18	Associative Classification	
19	Associative Classification – Rule Mining	
IV	**Overview and concepts Data Warehousing and Business Intelligence**	

20	Why reporting and Analysing data, Raw data to valuable information-Lifecycle of data	
21	What is Business Intelligence - BI and DW in today's perspective	12/8/
		TC
22	What is data warehousing - The building Blocks: Defining Features	30/8/
23	Data warehouses and data marts	
24	Overview of the components - Metadata in the data warehouse - Need for data warehousing	
25	Basic elements of data warehousing - trends in data warehousing.	
V	**The Architecture of BI and DW**	
26	BI and DW architectures and its types - Relation between BI and DW	1/9/2
27	OLAP (Online analytical processing) definitions	TC
28	Difference between OLAP and OLTP	15/9/
29	Dimensional analysis - What are cubes?	
30	Drill-down and roll-up - slice and dice or rotation	
31	OLAP models - ROLAP versus MOLAP	
32	defining schemas: Stars, snowflakes and fact constellations	
VI	**Classification and Prediction**	
33	What is classification and prediction? – Issues regarding Classification and prediction	16/9/
34	Classification methods: Decision tree,	TC
35	Bayesian Classification,	30/9/
36	Rule based, CART	
37	Neural Network	
38	Prediction methods: Linear and nonlinear regression	
39	Logistic Regression	
VII	**Data Mining for Business Intelligence Applications**	
40	Clickstream Mining, Market Segmentation, retail industry, telecommunications industry, banking & finance and CRM etc.,	1/10/
41	Data Analytics Life Cycle: Introduction to Big data Business Analytics - State of the practice in analytics role of data scientists	TC
		12/10/
42	Key roles for successful analytic project - Main phases of life cycle - Developing core deliverables for stakeholders.	

VII	**Advance topics**	
43	Introduction and basic concepts of following topics.	
44	Big Data: Introduction to big data: distributed file system – Big Data and its importance, Four Vs, Drivers for Big data, Big data analytics, Big data applications.	13/10/ TO
45	Algorithms using map reduce, Matrix-Vector Multiplication by Map Reduce	20/10/
46	Hadoop Architecture, Hadoop Storage: HDFS, Common Hadoop Shell commands , Anatom	
47	NameNode, Secondary NameNode, and DataNode, Hadoop MapReduce paradigm, Map	
48	Hadoop Configuration – HDFS Administering –Monitoring & Maintenance.	

GY

ATES	ACTUAL DATES
:016) 2016	
2016) 2016	
2016) 2016	

2016) 2016		
:016) 2016		
2016) 2016		
2016) /2016		

/2016 D /2016	