



L. D. College of Engineering, Ahmedabad – 15

LESSON PLAN

Over all Term Planning	
Branch:	Information Technology
Semester:	B.E 7 th SEM
Subject Name:	Data Mining & Business Intelligence
Subject Code:	21710715
Affiliating University:	Gujarat Technological University
Starting date of the term:	18-06-2018
Ending date of the term:	17-10-2018
Course Teacher:	Prof. S.A.Solanki

University Structure of the subject:

Teaching Scheme			Credits C	Examination Marks						Total Marks
L	T	P		Theory Marks			Practical Marks			
				ESE (E)	PA (M)		ESE (V)		PA (I)	
		PA	ALA		ESE	OEP				
3	0	2	5	70	20	10	20	10	20	150

L- Lectures; T- Tutorial/Teacher Guided Student Activity; P- Practical; C- Credit; ESE- End Semester Examination; PA- Progressive Assessment;

Syllabus:

Sr. No.	Content	Total Hrs	% Weightage
1	<p>Overview and concepts Data Warehousing and Business Intelligence</p> <p>Why reporting and Analysing data, Raw data to valuable information- Lifecycle of Data - What is Business Intelligence - BI and DW in today's perspective - What is data warehousing - The building Blocks: Defining Features - Data warehouses and data 1marts - Overview of the components - Metadata in the data warehouse - Need for data warehousing - Basic elements of data warehousing - trends in data warehousing.</p>	5	12%

2	The Architecture of BI and DW BI and DW architectures and its types - Relation between BI and DW - OLAP (Online analytical processing) definitions - Difference between OLAP and OLTP - Dimensional analysis - What are cubes? Drill-down and roll-up - slice and dice or rotation - OLAP models - ROLAP versus MOLAP - defining schemas: Stars, snowflakes and fact constellations	7	16%
3	Introduction to data mining (DM) Motivation for Data Mining - Data Mining-Definition and Functionalities – Classification of DM Systems - DM task primitives - Integration of a Data Mining system with a Database or a Data Warehouse - Issues in DM – KDD Process	4	08%
4	Data Pre-processing Why to pre-process data? - Data cleaning: Missing Values, Noisy Data - Data Integration and transformation - Data Reduction: Data cube aggregation, Dimensionality reduction - Data Compression - Numerosity Reduction - Data Mining Primitives - Languages and System Architectures: Task relevant data - Kind of Knowledge to be mined - Discretization and Concept Hierarchy.	7	16%
5	Concept Description and Association Rule Mining What is concept description? - Data Generalization and summarization-based characterization - Attribute relevance - class comparisons Association Rule Mining: Market basket analysis - basic concepts - Finding frequent item sets: Apriori algorithm - generating rules – Improved Apriori algorithm – Incremental	7	16%
6	Classification and Prediction What is classification and prediction? – Issues regarding Classification and prediction: Classification methods: Decision tree, Bayesian Classification, Rule based, CART, Neural Network Prediction methods: Linear and nonlinear regression, Logistic Regression Introduction of tools such as DB Miner /WEKA/DTREG DM Tools	7	16%
7	Data Mining for Business Intelligence Applications Data mining for business Applications like Balanced Scorecard, Fraud Detection, Clickstream Mining, Market Segmentation, retail industry, telecommunications industry, banking & finance and CRM etc., Data Analytics Life Cycle: Introduction to Big data Business Analytics - State of the practice in analytics role of data scientists Key roles for successful analytic project - Main phases of life cycle - Developing core deliverables for stakeholders.	4	08%
8	Advance topics Introduction and basic concepts of following topics. Clustering, Spatial mining, web mining, text mining, 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. Algorithms using map reduce, Matrix-Vector Multiplication by Map Reduce. Introduction to Hadoop architecture: Hadoop Architecture, Hadoop Storage: HDFS, Common Hadoop Shell commands , Anatomy of File Write and Read., NameNode, Secondary NameNode, and DataNode, Hadoop MapReduce paradigm, Map and Reduce tasks, Job, Task trackers - Cluster Setup – SSH & Hadoop Configuration – HDFS Administering –Monitoring & Maintenance.	4	08%

Reference Books:

1. J. Han, M. Kamber, “Data Mining Concepts and Techniques”, Morgan Kaufmann
2. M. Kantardzic, “Data mining: Concepts, models, methods and algorithms, John Wiley & Sons Inc.
3. Paulraj Ponnian, “Data Warehousing Fundamentals”, John Willey.

4. M. Dunham, “Data Mining: Introductory and Advanced Topics”, Pearson Education.

5. G. Shmueli, N.R. Patel, P.C. Bruce, “Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner”, Wiley India.

Lesson Plan

Sr. No.	Topic	Planned Date	Actual Date	Planned Date	Actual Date	Mode of Delivery	Resources required
		(Div A)	(Div A)	(Div B)	(Div B)		
1	Motivation for Data Mining - Data Mining-Definition and Functionalities	18/06/18 19/06/18		20/06/18		Chalk Board/ppt	Hand Outs/ppt
2	Classification of DM Systems - DM task primitives	25/06/18 26/06/18		27/06/18		Chalk Board/ppt	Hand Outs/ppt
3	Integration of a Data Mining system with a Database or a Data Warehouse	02/07/18 03/07/18		4/07/18		Chalk Board/ppt	Hand Outs/ppt
4	Issues in DM – KDD Process, End of Chapter Quiz						
5	Why to pre-process data? - Data cleaning: Missing Values, Noisy Data - Data Integration and transformation, NPTL vedio lecture	09/07/18 10/07/18		11/07/18		Chalk Board/ppt	Hand Outs/ppt
6	Data Reduction: Data cube aggregation, Dimensionality reduction	16/07/18 17/07/18		18/07/18		Chalk Board/ppt	Hand Outs/ppt
7	Data Compression - Numerosity Reduction	23/07/18 24/07/18		25/07/18		Chalk Board/ppt	Hand Outs/ppt
8	Data Mining Primitives - Languages and System Architectures: Task relevant data	30/07/18 31/07/18		01/08/18		Chalk Board/ppt	Hand Outs/ppt
9	Kind of Knowledge to be mined - Discretization and Concept Hierarchy End of Chapter Quiz	06/08/18 07/08/18		08/08/18		Chalk Board/ppt	Hand Outs/ppt
10	What is classification and prediction? – Issues regarding Classification and prediction:, NPTL vedio Lecture	13/08/18 14/08/18		29/08/18		Chalk Board/ppt	Hand Outs/ppt
11	Classification methods: Decision tree, Bayesian	20/08/18		5/09/181		Chalk	Hand

	Classification, Rule based, CART, Neural Network,	21/08/18				Board/ppt	Outs/ppt
12	Prediction methods: Linear and nonlinear regression, Logistic Regression End of Chapter Quiz	27/08/18 28/08/18		5/09/181		Chalk Board/ppt	Hand Outs/ppt
13	Introduction of tools such as DB Miner /WEKA/DTREG DM Tools, Practical Demo video Presentation by students on relevant topic	4/09/180 10/09/18		12/09/181		Chalk Board/ppt	Hand Outs/ppt
14	Data mining for business Applications like Balanced Scorecard, Fraud Detection, Clickstream Mining, Market Segmentation, retail industry,telecommunications industry, banking & finance and CRM, case study	11/09/18 17/09/18		19/09/181			
15	Data Analytics Life Cycle: Introduction to Big data Business Analytics - State of the practice in analytics role of data scientists	18/09/18 24/09/18		26/9/18		Chalk Board/ppt	Hand Outs/ppt
16	Key roles for successful analytic project - Main phases of life cycle - Developing core deliverables for stakeholders., End of Chapter Quiz	25/09/18 01/10/18		3/10/18		Chalk Board/ppt	Hand Outs/ppt
17	Introduction to R programming with basic commands	08/10/18 09/10/18		10/10/18		Chalk Board/ppt	Hand Outs/ppt

Prof. S.A.Solanki

Department of Information Technology

L. D. C. E. – Ahmedabad-15