GUJARAT TECHNOLOGICAL UNIVERSITY

Syllabus for Master of Business Administration, 3rd Semester
Functional Area Specialization: Information Technology Management
Subject Name: Data Warehousing and Data Mining (DWDM)
Subject Code: 4539251

With effective from academic year 2018-19

1. Learning Outcomes:

Learning Outcome Component	Learning Outcome (Learner will be able to)
Business Environment and Domain Knowledge (BEDK)	• *Explain* and *discuss* the importance of Data Warehouses as a part of a firms' IT infrastructure. • *Evaluate* how Data Warehouses can be used for decision making • *Comment* on the opportunities and challenges of managing data analytics
Critical thinking, Business Analysis, Problem Solving and Innovative Solutions (CBPI)	• *Assess* role of data mining techniques in business decision making. • *Explain* why data and data science capability are strategic assets.
Global Exposure and Cross-Cultural Understanding (GECCU)	• *Evaluate* the role played by data-warehouses and data mining in providing business insights in different industries across the world. • *Explore* recent trends in data mining such as web mining, spatial-temporal mining.
Social Responsiveness and Ethics (SRE)	• *Assess* the role of data warehousing in customer relationship management systems
Effective Communication (EC)	• *Describe* different methodologies used in data mining and data ware housing.
Leadership and Teamwork (LT)	• *Evaluate* the different models of OLAP and data preprocessing for a given system. • *Design* a data mart or data warehouse for any organization.

LO – PO Mapping: Correlation Levels:
1 = Slight (Low); 2 = Moderate (Medium); 3 = Substantial (High), "-"= no correlation

Sub. Code: 4539251	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9
LO1: Explain and discuss the importance of Data Warehouses as a part of a firms' IT infrastructure.	3	3	1	3	1	1	-	1	2
LO2: Evaluate how Data Warehouses can be used for decision making	3	3	3	2	2	1	-	1	2
LO3:Comment on the opportunities and challenges of managing data analytics	2	3	3	3	2	1	-	-	2
LO4: Evaluate the role played by data-warehouses and data mining in providing business insights in different industries across the world.	2	3	3	2	3	2	1	1	2
LO5: Explore recent trends in data mining such as web mining, spatial-temporal mining	3	3	2	2	3	1	-	2	2

LO6: Assess the role of data warehousing in customer relationship management systems	3	3	3	2	2	1	-	1	2
LO7: Describe different methodologies used in data mining and data ware housing.	3	2	1	1	-	-	-	-	-
LO8: Evaluate the different models of OLAP and data preprocessing for a given system.	3	2	3	1	-	3	-	-	1
LO9: Design a data mart or data warehouse for any organization	3	3	2	1	-	3	-	2	1

2. Course Duration: The course duration is of **40 sessions of 60 minutes each.**

3. Course Contents:

Module No:	Contents	No. of Sessions	70 Marks (External Evaluation)
I	**RDBMS concepts:** • Introduction • Normalization(1NF to BCNF) • Structured Query Language (SQL) • Features of SQL • Data Definition Language (DDL) • Data Manipulation Language (DML) • Views, Functions in SQL • Group By and Having Clauses • Subqueries • Examples of SQL	10	18
II	**Data warehousing concepts:** • Difference between DWH and OLTP-based DBMS environments • Development Process, DW development life cycle • DW development Methodologies • DW Process framework • Data warehouse Design • Detailed Dimensional Modelling • Reporting and Query tools • Data Extraction • Transformation and Loading Process • Meta Data Management, Data Marts. **Data Mining concepts:** • Data Pre-processing ▪ Data types, attributes and properties	10	18

	▪ Data Quality ▪ Pre-processing ▪ Types of Data Mining, cleaning, integration and reduction		
III	**Association Rule Mining And Classification:** Mining Frequent Patterns **Associations And Correlations:** • Mining Methods • Association Rules – Correlation Analysis , Constraint Based Association Mining **Classification And Prediction:** • Basic Concepts • Decision Tree Induction • Bayesian Classification, Rule Based Classification • Classification by Back Propagation • Support Vector Machines • Associative Classification • Lazy Learners • Other Classification Methods – Prediction.	10	17
IV	**Clustering And Trends In Data Mining:** • Cluster Analysis: ▪ Types Of Data ▪ Categorization Of Major Clustering Methods ▪ K-Means – Partitioning Methods, Hierarchical Methods, Density-Based Methods, Grid Based Methods, Model-Based Clustering Methods, Clustering High Dimensional Data, Constraint Based Cluster Analysis • **Outlier Analysis.** • **Overview of Text Mining, Web mining & Multimedia. Data Mining.** • **Data Mining Applications.**	10	17
V	**Practical:** Hands on training on the concepts taught using tools such as XML Miner & WeKA. Students are required to make presentation on applications of Data mining in business areas like Risk management and targeted marketing, Customer profiles and feature construction, Medical applications, Scientific Applications etc.	---	(30 marks CEC)

4. Pedagogy:

- ICT enabled Classroom teaching
- Case study
- Practical / live assignment
- Interactive class room discussions

5. Evaluation:

Students shall be evaluated on the following components:

A	**Internal Evaluation**	**(Internal Assessment- 50 Marks)**
	• Continuous Evaluation Component	30 marks
	• Class Presence & Participation	10 marks
	• Quiz	10 marks
B	**Mid-Semester examination**	**(Internal Assessment-30 Marks)**
C	**End –Semester Examination**	**(External Assessment-70 Marks)**

6. Reference Books:

No.	Author	Name of the Book	Publisher	Year of Publication / Edition
1	Alex Berson, Stephen Smith	Data Warehousing, Data Mining and OLAP	McGraw Hill	2004 / 1st
2	Jaiwei Han, Jain Pei, Michelin Kamber	Data Mining: Concepts and Techniques	Elsevier	2011 / 3rd
3	George M. Marakas	Modern Data Warehousing, Mining and Visualization: Core Concepts	Pearson	2003 / 1st
4	SoumedraMohanty	Data Warehousing: Design, Development and Best Practices	McGraw Hill	2005
5	PaulrajPonnaiah	Data Warehousing Fundamentals for IT Professionals	Wiley – Blackwell	2010 / 2nd
6	Ralph Kimball	The Data Warehouse Toolkit	Wiley	2013 / 3rd
7	Alan R. Simon, Steven L. Shaffer	Data Warehousing and Business Intelligence for E-commerce	Morgan Kauffman	2001 / 1st
8	Jeffrey A. Hoffer, V. Ramesh, HeikkiTopi	Modern Database Management	Pearson	2016 / 12th
9	Pang-Ning Tan, Michael Steinbach, AnujKarpatne, Vipin Kumar	Introduction to Data Mining	Pearson	2018 / 2nd

Note: Wherever the standard books are not available for the topic appropriate print and online resources, journals and books published by different authors may be prescribed.

7. List of Journals / Periodicals / Magazines / Newspapers / Web resources, etc.

1. International Journal of Data Mining and Emerging Technologies
2. International Journal of Data Mining, Modeling and Management
3. International Journal of Data Warehousing and Mining
4. Analytics India (Magazine)
5. https://onlinecourses.nptel.ac.in/noc19_mg14/preview
6. https://onlinecourses.nptel.ac.in/noc19_cs15/preview