

<b>Branch Name:</b>	IMCA
<b>Program Code:</b>	CS301
<b>Course Name:</b>	Data Warehouse and Data Mining
<b>Course Code:</b>	1CS3010603T
<b>Pre-requisite Course:</b>	Knowledge of RDBMS, OLTP and OLAP

**Course Objective:**

1. To understand the need of Data Warehouses over Databases, and the difference between usage of operational and historical data repositories
2. To understand the concept of Analytical Processing (OLAP) and its similarities & differences with respect to Transaction Processing (OLTP).
3. To understand the need for pre-processing
4. To learn the algorithms used for various type of Data Mining problems

**Teaching and Examination Scheme:**

Teaching Scheme (Hours per week)				Evaluation Scheme (Marks)				Total
Lecture	Tutorial	Practical	Credit	Theory		Practical		
				University Assessment	Continuous Assessment	University Assessment	Continuous Assessment	
4	1		5	60	40	-	-	100

**Course Contents:**

Unit No	Topics	Total Hours	Weightage (%)
1	<p><b>Data Warehouse and OLAP Technology:</b> Introduction to Data Warehouse, Multidimensional Data Model, Data Warehouse Architecture, OLAP Servers.</p> <p><b>Data Mining Introduction:</b> What is Data Mining – on What Kind of Data, Data Mining Functionalities – What Kind of Patterns Can be Mined, Are All of the Patterns Interesting, Classification of Data Mining Systems, Major Issues in Data Mining.</p>	10	25
2	<p><b>Data Preprocessing:</b> Need for Pre-processing, Data Cleaning, Data Integration &amp; Transformation, Data Reduction, Concept Hierarchy Generation. Thrashing, Locality.</p> <p><b>Mining Frequent Patterns, Associations, and Correlations:</b> Basic Concepts: Market Basket Analysis, Efficient and Scalable frequent Itemset Mining Methods, From Association Mining to Correlation Analysis</p>	10	25
3	<p><b>Classification:</b> Introduction to Classification and Prediction, Issues Regarding Classification and Prediction, Classification by Decision Tree Induction, Bayesian Classification.</p> <p><b>Cluster Analysis:</b> Introduction to Cluster Analysis, Types of Data in Cluster Analysis, A categorization of Major Clustering Methods, Partitioning Methods.</p>	10	25

<b>4</b>	<p><b>Application and Trends in Data Mining:</b> Data Mining Applications, Data Mining System Products and Research Prototypes.</p> <p>Data Extraction, Transformation and Loading: Objectives, ETL Overview, ETL Requirements and Steps, Data Extraction- Source Identification, Data Extraction Techniques, Evaluation of Techniques, Data Transformation</p>	10	25
----------	---	----	----

**Text Books:**

1. Vikram Pudi P. Radha Krishna, "Data Mining", Oxford University Press, 1st Edition 2009.

**References Books:**

1. Jiawei Han & Micheline Kamber, "Data Mining: Concepts & Techniques", Morgan Kaufmann Publishers (2002)
2. W. H. Inmon, "Building the Data Warehouse", Wiley Dreamtech India Pvt. Ltd.
3. Mohanty, Soumendra, "Data Warehousing: Design, Development and Best Practices", Tata McGraw Hill (2006)

**Course Learning Outcomes (CLO): On completion of this course, the students will be able to:**

CLO	Description	Bloom's Taxonomy Level
CLO1	Ability to apply pre-processing on existing operational & historical data for creation of Data Warehouse	1 Remembering 2 Understanding
CLO2	Ability to apply Apriori algorithm for Association Mining	2 Understanding, 3 Applying
CLO3	Ability to apply Decision Tree and Bayesian algorithms for Classification	3 Applying, 4 Analyze, 5 Evaluate, 6 Create

**Mapping of CLOs with POs & PSOs**

Course Learning Outcomes	Program Out comes(POs)												Program Specific Outcomes(PSOs)	
	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12	PSO1	PSO2
CLO1	H	M	H	M	L	M	M	L		L	H	H	H	H
CLO2	M	H	H	H	H	M	M	L		L	M	L	H	H
CLO3	M	M	M	M	M	M		L		M	M	L	H	H

**H:High, M:Medium, L:Low**