

COURSE TITLE: DATA ANALYTICS AND MINING**COURSE OBJECTIVES:**

- To acquire the knowledge of various concepts and tools behind mining data for business intelligence
- To Study data mining algorithms, methods and tools
- To Identify business applications of data mining

LEARNING OUTCOMES:

The students will be able to:

- Apply data mining concepts for data analysis and report generation
- Develop industry level data mining skills using software tools
- Make use of relevant theories, concepts and techniques to solve real-world business problems

| Code | Course | Teaching Period / Week | | Credit | | | Duration of Theory Exam (in Hrs.) |
|--------|---------------------------|------------------------|---------|--------|------|-------|-----------------------------------|
| | | L | Pr./ Tu | Int. | Ext. | Total | |
| MCS202 | Data Analytics and Mining | 4 | - | 2 | 2 | 4 | 2 |

| Module No. | Objective | Content | Evaluation |
|------------|--|--|--|
| 1 | This module introduces students to the concept of data analytics | Data Analytics Introduction, Data Summarization and visualization, Linear, Non-linear regression, model selection | Online Test (Marks 5) |
| 2 | This module provides background on data objects and statistical concepts. It introduces techniques for preprocessing data before mining. | Data Mining and Data Preprocessing What is data mining?, Knowledge discovery- KDD process, related technologies - Machine Learning, DBMS, OLAP, Statistics, Data Mining Goals, stages of the Data Mining Process, Data Mining Techniques, Knowledge Representation Methods. Data cleaning, Data transformation, Data reduction, Discretization and generating concept hierarchies. introduction to data warehousing, OLAP, and data generalization. Data Cube Computation and Multidimensional Data Analysis | Written Unit Test – I (Marks 25) |
| 3 | This unit covers supervised learning method as classification and Prediction | Classification and Prediction Decision tree, Bayesian classification, rule-based classification, neural networks, support vector machines, associative classification, k-nearest-neighbor classifier, case-based reasoning. | Assignments will be given for the above topics. (Marks 10) |

| | | | |
|---|---|--|---|
| 4 | This unit covers unsupervised learning method as clustering and association rule mining To gain detailed insights of outlier detection | Clustering and Association Rule Mining Partitioning, hierarchical, density-based, grid-based, and model-based methods data clustering. Mining Frequent Patterns, Associations, and Correlations Outlier Detection: Detection of anomalies, such as the statistical, proximity-based, clustering-based, and classification-based methods. | Assignments will be given for the above topics. (Marks 10) |
|---|---|--|---|

EVALUATION:

| Evaluation | Details | Marks |
|---|---|------------------|
| (* please give details of assessment in terms of Unit test/ Project/ quiz /or other assignments and marks allotted for it) | | |
| Internal | <ul style="list-style-type: none"> • Unit test • Online Test • Assignments | 50 Marks |
| External | Final Examination | 50 Marks |
| Total marks | | 100 Marks |

TEXT BOOKS:

1. Shashi Shekhar and Sanjay Chawla, (2003), *Spatial Databases: A Tour*, Prentice Hall (ISBN 013-017480-7)
2. Avi Silberschatz, Henry F. Korth, S. Sudarshan, *Database System Concepts*, 5th edition, (2010), McGraw-Hill

REFERENCE BOOKS:

1. Stefano Ceri and Giuseppe Pelagatti, (1984), *Distributed Database; Principles & Systems*, McGraw-Hill International Editions
2. Raghuram Ramakrishnan and Johannes Gehrke, (2002), *Database Management Systems*, 3rd edition, McGraw-Hill.
3. Elmasri and Navathe, (2003), *Fundamentals of Database Systems*, 6th Edition, Addison. Wesley.
4. Shio Kumar Singh, (2011), *Database Systems: Concepts, Design and Applications*, 2nd edition, Pearson Publishing
5. Multi-dimensional aggregation for temporal data. M. Böhlen, J. Gamper, and C.S. Jensen. In Proc. of EDBT-2006, pp. 257-275, (2006).
6. R.H. Güting and M. Schneider (2005), *Moving objects databases*, Morgan Kaufmann Publishers, Inc.
7. Paulraj Ponniah, (2010), *Data Warehousing fundamentals*, JohnWiley.