COURSE DETAILS

Name of the Programme	:	**Master of Science (Data Science)**
Semester	:	**I**
Course Name	:	**Data Management**
Course Code	:	**T3364**
No. of Credit	:	**4**

Course Outline:

Module No.	Topic
1	Introduction- Basic Concepts & Principles of Data Management
2	Types of Digital data – Structured, unstructured and semi-structured, real-time, right-time data
3	Data preprocessing – Data Cleaning, Data Transformation, Data reduction, Discretization and generating concept hierarchies – related experiments
4	Data Processing pipelines in science and business
5	Data Mining knowledge representation – task relevant data - Background knowledge- Interestingness measures - Representing input data and output knowledge – Visualization techniques
6	Attribute-oriented analysis - Attribute generalization - Attribute relevance - Class comparison - Statistical measures
7	Data mining algorithms: Association rules - Motivation and terminology Example: mining weather data - Basic idea: item sets - Generating item sets and rules efficiently - Correlation analysis

8	Data mining algorithms: Classification - Basic learning/mining tasks Inferring rudimentary rules: 1R algorithm - Decision trees - Covering rules
9	Sample data set analysis and Data Modeling
10	OLAP -Operational Data store, data dimensions, No SQL data store
11	Comparison of data models with examples /cases
12	Data Provenance and Linked Data

Books Recommended:

1. MASTER DATA MANAGEMENT AND DATA GOVERNANCE, 2/E 2nd Edition by Alex Berson
2. Data Mining: Concepts and Techniques, Third Edition Jiawei Han, Micheline Kamber, Jian Pei, Morgan Kaufmann, 2011.
3. Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations - Ian H. Witten, Eibe Frank, Morgan Kaufmann.
4. Data Science For Dummies By Lillian Pierson