**Academic Year: 2023-24**

**TermTest 2 Syllabus**

**Course Name: Data Warehousing and Mining Course Code:** **DJ19CEC501**

**Class**: TY BTech **Sem:** V

|  |  |
| --- | --- |
| **Unit No.** | **Topics** |
| 2 | **ETL Process and OLAP:** Major steps in ETL process, Data extraction: Techniques, Data transformation: Basic tasks, Major transformation types, Data Loading: Applying Data  Type 1 and Type 2 changes |
| 4 | **Classification:**  Basic Concepts of classification, Decision Tree Induction, Attribute Selection Measures using Information Gain, Tree pruning  Bayes Classification Methods: Bayes’ Theorem, Naïve Bayesian Classification  Rule - Based Classification: Using IFTHEN Rules for classification, Rule Extraction from a Decision Tree, Rule Quality Measures, Rule Pruning  Model Evaluation & Selection: Metrics for Evaluating Classifier Performance, Holdout Method and Random Subsampling, Cross Validation, Bootstrap, |
| 5 | **Clustering:**  Cluster Analysis and Requirements of Cluster Analysis  Partitioning Methods: k-Means, k-Medoids  Hierarchical Methods: Agglomerative, Divisive  Density Based Methods: DBScan |
| 6 | **Mining Frequent Patterns and Association Rules:**  Apriori and FP tree algorithm |

**Subject Incharge:**

Dr. Meera Narvekar

Dr. Kiran Bhowmick

Dr. Kranti Ghag

Prof. Lynette D’mello