B.Sc. (Hons.) Computer Science (CBCS)
Semester VI
Data Mining Guidelines

Sr. No.	Units	Topics	Chapter	No. of Lectures
1	Introduction	1.1 - What Is Data Mining? 1.2 Challenges 1.3 Data Mining Origins 1.4 Data Mining Tasks	1	5L
2	Data mining techniques	2.1- Types of data, 2.2 – Data Quality, 2.3.1 Aggregation, 2.3.2 Sampling, 2.3.3 Dimensionality reduction – upto pg 51, 2.3.4 Feature subset selection upto pg 52, 2.4.5 Feature creation upto pg 55, 2.3.6 Discretization upto pg 59, 2.3.7 variable transformations 2.4.3 Dissimilarity among data objects 2.4.4 similarity among data objects	2	10L
3	Classification	4.1 – Preliminaries, 4.2 – General Approach to Solving a Classification Problem, 4.3 Decision Tree Induction (Till Pg. 165), 4.5 – Evaluating the Performance of a Classifier	4	7L
4		5.1 – Rule Based Classifier (upto page 212),5.2 – Nearest Neighbor Classifiers, 5.3– Bayesian Classifiers (Complete for discrete data and only introduction of Bayes classifier for continuous attributes) till pg. 233, 5.7.1 – Alternative Metrics	5	8L
5	Association Rules	6.1-Problem definition, 6.2-Frequent itemset generation, 6.3-Rule generation till Pg 351	6	10L
6	Clustering	8.1 Basic concepts of clustering analysis, 8.2 K-Means (8.2.1-8.2.5 except 8.2.3), 8.3 Agglomerative Hierarchical Clustering (except pg 522-524), 8.4 DBSCAN	8	12L

Course Books:
1. Introduction to Data Mining, Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Pearson Education.

References:
2. Data Mining: Concepts and Techniques, 3nd edition,Jiawei Han and Micheline Kamber
3. Data Mining: A Tutorial Based Primer, Richard Roiger, Michael Geatz, Pearson Education 2003.
4. Introduction to Data Mining with Case Studies, G.K. Gupta, PHI 2006

5. Insight into Data mining: Theory and Practice, Soman K. P., DiwakarShyam, Ajay V., PHI 2006