

## Tentative Teaching Plan

Name of the Faculty: Nidhi Passi  
 Name of the Course: B.Sc. (H) Computer Science  
 Semester : VI Sec (if any): - A and B  
 Title of the Paper : Data Mining

### Course Learning Outcomes

On successful completion of the course, students will be able to do following:

1. Pre-process the data, and perform cleaning and transformation.
2. Apply suitable classification algorithm to train the classifier and evaluate its performance.
3. Apply appropriate clustering algorithm to cluster data and evaluate clustering quality
4. Use association rule mining algorithms and generate frequent item-sets and association rules

Month	Topics Covered	References
January	<p>Chapter 1 : Introduction : 1.1 - What Is Data Mining? 1.2 Challenges 1.3 Data Mining Origins 1.4 Data Mining Tasks</p> <p>Chapter 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, 2.3.5 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</p> <p><b>Practical:</b> Introduction to R, Que 1</p>	<p>[1] <b>Course Book:</b>                      1. Introduction to Data Mining, Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Pearson Education.</p>
February	<p>Chapter 4: 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</p> <p>Chapter 5: Classification: Alternative Techniques :                      5.1 – Rule Based Classifier (upto page 212), 5.2 – Nearest Neighbor Classifiers</p> <p><b>+TEST 1 ( 28<sup>th</sup> February )</b>  <b>+ Assignment ( 11<sup>th</sup> March; Submission 24<sup>th</sup> March)</b>  <b>Practical:</b> Que 2, 3, 4</p>	
March	<p>Chapter 5- 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</p> <p>Chapter 6: Association Rules: 6.1-Problem definition, 6.2-Frequent itemset generation, 6.3-Rule generation till Pg 351</p> <p><b>+Test 2 ( 12<sup>th</sup> April )</b>  <b>Practical:</b> Que 5 + Project</p>	
April	<p>Chapter 8: 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</p> <p>+ Revision  <b>+Test 3 (if feasible/required) ( 18<sup>th</sup> April)</b>  <b>Practical:</b> Que 6 + Project Presentation</p>	

### Course Books:

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

### References:

2. Data Mining: Concepts and Techniques, 3rd 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