

11B1WCI832: Information Retrieval and Data Mining

Course Credit: 3

Semester: VII

Introduction:

Data mining is an interdisciplinary subject, and more appropriately named “knowledge mining from data”. However, it is also known as knowledge discovery from data (KDD).

Course Objectives (Post-conditions) Knowledge objectives:

- Good knowledge to Mine Data from Large Amount of Data Sources
- You will be able to compare a two different data set and you can find out which one is better
- You will be able to apply many retrieval techniques on a given document
- You will be able apply some pattern matching algorithm and map the various expected pattern depend upon usage requirements

Application objectives:

Data Mining and Information Retrieval have tremendous advantages in term of their application areas which include; Financial Data Analysis, Retail and Telecommunication Industries, Intrusion Detection and Prevention, Bioinformatics and many more.

Expected Student Background (Preconditions)

Database Management Concepts, Probability and Basic Mathematical Concepts

Topics Outline:

Serial No.	Topics	Hrs
1	Introduction to Data Mining	1
2	Data Preprocessing	7
3	Data Generalization, Data Warehousing, and On-line Analytical Processing	8
4	Association Rule Mining and Frequent Patterns	4

5	Classification and Prediction	4
6	Cluster Analysis	5
7	Introduction to Information Retrieval	1
8	Boolean Retrieval	2
9	Probabilistic Information Retrieval	3
10	Vector Support Machines	3
11	Pattern Matching Algorithm	2
12	Link Analysis	2
Total Hours = 42		

Text Book(s):

1. Jiawei Han and Micheline Kamber, "Data Mining, Concepts and Techniques", Elsevier 2nd edition.
2. An introduction to Information Retrieval, 2008 Cambridge UP.
3. Thomas H. Corman, Charles E. Leiserson, and Ronald L. Rivest, Introduction to Algorithms,. The MIT Press,England, 1989.

Reference Book(s):

1. Pang-Ning Tan, Michael Steinbach, Vipin Kumar, "Introduction to Data Mining ", Pearson Education.
2. Richard O. Duda, Peter E. Hart, David G. Stork , "Pattern Classification", 2nd Edition, Wiley Publication, November 2000.
3. Rijsbergen C. J. ,"Information Retrieval" 2nd edition.

Evaluation Scheme:

S.No	Examination	Marks
1	T-1	15
2	T-2	25
3	T-3	35
4	INTERNAL	25
Total Marks: 100		

Internal Marks Breakdown:

Assignments: 9 marks (3x3)
 Quizzes: 12 marks (3x4)
 Regularity: 4 Marks