

Global Initiative for Academic Network



Ministry of Human Resource Development, Government of India



A Short Course On

Data Mining: Concepts and Techniques

June 20th to July 1st, 2016

Overview

The quantity and variety of online data is increasing very rapidly. With the proliferation of data repositories in enterprises and organizations and the acceleration of computing power, there is an increasing demand for making sense of the archived data. Data mining has emerged as one of the most powerful techniques to achieve the above goal. This course aims to provide students with a broad background on data mining concepts and techniques, exposure to data mining tools, application of data mining techniques to improve marketing, sales, and customer relationship management. It covers data mining methodology and algorithms, including association rule learning, decision trees, artificial neural networks, clustering, and link analysis. This course will cover these issues and will illustrate the whole process by examples of practical applications from the life sciences, computer science, and commerce. Given the close connection of data mining to real world application and its potential in improving the performance of business operations, data mining concepts and techniques will be discussed in various business and organizational contexts. This course can contribute to the scope for research work. Course participants will learn these topics through lectures and hands-on experiments. Also case studies and assignments will be shared to stimulate research motivation of participants.

Objectives

Faculties and Students are expected to demonstrate the following competencies:

1. Evaluating and implementing the wide range of emerging and newly adopted methodologies and technologies to facilitate knowledge discovery.
2. Ability to preprocess data into suitable form to serve as input for a range of data mining algorithms.
3. Designing and implementing data mining applications using sample, real data sets and modern tools.
4. Assessing and operationalizing the interesting patterns mined from different kinds of databases.
5. Gaining experience by doing independent study and research.

Modules

Module A: Introductory Concepts : June 20 - June 24

Lecture 1: Introduction, Data Preprocessing, Measure of similarity and dissimilarity.

Lecture 2: Classification, Decision tree induction, Model Over fitting, Performance of a Classifier, Comparing Classifiers

Tutorial 1: Problem solving and Lab session with examples of Decision tree and Model Over fitting.

Lecture 3: Classification Techniques: Rule-Based Classifier, Nearest-Neighbor classifiers.

Lecture 4: Bayesian Classifiers, Artificial Neural Network (ANN) Approach.

Tutorial 2: Problem solving and Lab session with examples on different classifiers.

Lecture 5: Associations, Frequent Item set Generation, Rule Generation, Evaluation of Association Patterns, FP-Growth Algorithm

Lecture 6: Association pattern Analysis: Advanced Concepts

Tutorial 3: Problem solving and Lab session with examples of Association rule generation

Lecture 7 : Clustering , Cluster Analysis, Hierarchical clustering

Lecture 8: K-means clustering, DBSCAN, Cluster evaluation

Tutorial 4 : Problem solving and Lab session with examples of clustering and K-means algorithm

Lecture 9 : Anomaly Detection

Lecture 10: Clustering over categorical data

Tutorial 5: Problem solving and Lab session with examples of Anomaly Detection.

Module B: Advance Techniques : June 26 - June 30

Lecture11: Spatial Data Mining

Lecture12: Graph Mining

Tutorial 6: Problem solving examples of Graph and Spatial data mining

Lecture 13: Data Stream analysis

Lecture 14:Text Mining

Tutorial 7: Problem solving examples of Data stream analysis, and Text Mining

Lecture 15: Boolean Matrix Decomposition Basics

Lecture 16: Boolean Matrix Decomposition: Advanced Methods

Tutorial 8: Problem solving examples of Matrix Decomposition

| | |
|-------------------------------------|---|
| | <p>Lecture 17: Privacy-Preserving Data Mining</p> <p>Lecture 18: Privacy-Preserving Data Mining</p> <p>Tutorial 9: Problem solving examples of Privacy-Preserving Data Mining</p> <p>Lecture 19: Case studies-I (network intrusion)</p> <p>Lecture 20: Case studies-I I (earth science /social network analysis)</p> <p>Tutorial 10: Discussion on recent trends of research in data mining</p> |
| Examination and Certificates | <ul style="list-style-type: none"> ▪ Examination will be conducted on 1st July 2016 and Grade Certificate of 2 Credit will be awarded to the participants who passed the examination. ▪ Participation Certificate will be provided to each participant. |
| Seats | Number of participants for the course will be limited to Forty . |
| You Should Attend If... | <ul style="list-style-type: none"> ▪ You are a computer engineer or research scientist interested in data mining techniques. ▪ You are a student from academic institution in Mathematical and Computational Sciences. ▪ You are a faculty from academic institution in Mathematical and Computational Sciences. |
| Course Registration Fee | <ul style="list-style-type: none"> ▪ Participants from abroad : US \$500 ▪ Industry/ Research Organizations: Rs. 10000/- ▪ Academic Institutions: <ul style="list-style-type: none"> • Host Institution (Jabalpur Engineering college): Faculty and students :Nil • Other institution: <ul style="list-style-type: none"> ○ Faculty : Rs. 5000/- ○ Student: Rs. 1000/- (For SC/ST : Rs. 500/-) <p>The above fee includes all instructional materials, computer use for tutorials and free internet facility. Food, transport and accommodation of course participants will be borne by the individual course participants themselves. On the request the participants will be provided with single bedded accommodation on payment basis in the Guest House/Hostel.</p> |
| Registration Process | Interested participants will have to first register with the GIAN website http://www.gian.iitkgp.ac.in for one-time registration with registration fees of Rs. 500.00 which will enable them to enrol for any number of courses being offered. Participants then needs to select “Data Mining: Concepts and Techniques” course from the list at “Course Registration”. Subsequent registration for this course will have to be done with Jabalpur Engineering College, Jabalpur by submitting the registration form as attached with the brochure, to the Course Coordinator by online and offline mode. They need to enclose the demand draft of requisite fees in favour of “Principal Jabalpur Engineering College, Jabalpur”. |
| Important Dates | Registration in GIAN Portal and Receipt of Course registration form to course Coordinator : 12 June 2016 Intimation to the shortlisted participants : 15 June 2016 |

The Faculty



Dr. Jaideep Vaidya is a Professor of Computer Information Systems at Rutgers University. He has a Ph.D. and Masters degree in Computer Science from Purdue University and a

Bachelor's degree in Computer Engineering from the University of Mumbai. Dr. Vaidya's work is at the confluence of privacy, security, data mining, and data management. He is particularly interested in Security and Privacy issues raised by data mining - the area of Privacy preserving data mining; Data Mining techniques applied to interoperation of heterogeneous information sources; and the use of data mining for enhancing security. In general, he also has broad interests in applied cryptography and general Secure Multi-Party Computation techniques including RFID security, privacy preserving optimization and secure supply chain management. Other research interests include novel applications of data mining and databases such as role engineering, spatio-temporal data mining, mining of environmental data and environmental data warehousing.



Dr. Mamta Lambert is currently an Associate Professor in the Department of Computer Applications. Her subject of interest includes Computer Graphics, Data mining and

Decision Support System.



Dr. A. Hemlata is working as Assistant Professor in the Department of Computer Science & Engineering. Her subject of interest includes Computer Graphics, Image Processing and

Data mining.

Course Coordinators

Dr. Mamta Lambert
Associate Professor

Department of Computer Applications

Phone: 0761-2433553, 91-9406740301

E-mail: lambert_mamta@yahoo.co.in

Dr. A. Hemlata
Assistant Professor

Department of Computer Science & Engineering

Phone: 0761-2433455, 91-9424418804

E-mail: a.hemlata@rediffmail.com

Local Coordinator (GIAN)

Dr. S.S. Thakur
Professor & Head

Department of Applied Mathematics

Phone : 0761-2433522, 91-9827444784

E-mail: ssthakur@jecjabalpur.ac.in

<http://www.gian.itkgrp.ac.in/GREGN>

Venue



JABALPUR ENGINEERING COLLEGE

(Declared Autonomous by Govt. of M.P.)

Gokalpur, Jabalpur (M.P.) India- 482011

Phone : +91-761-2331953, +91-761-2431355, Web: jecjabalpur.ac.in

Global Initiative on Academic Network (GIAN)

Jabalpur Engineering College Jabalpur

A Short Course On

Data Mining: Concepts and Techniques

June 20th to July 1st, 2016

REGISTRATION CUM ACCOMODATION REQUEST FORM

(To be submitted by the candidates to the course coordinator after one time registration in the GIAN portal. This form should reach by June 12, 2016)

1. Name (Block Letters): M/F:

2. Registration ID generated by GIAN Portal:

3. Participant Type:

I. Student:

Course (B.E/ B. Tech. /M.C.A/ M. Tech. / M. Sc. /M.Phil. /Ph.D.):

Branch & Semester:

Student ID Number:

Institute:

II. Faculty :

Designation:

Department:

Organization:

III. Industry:

Designation/Professional Title:

Organization:

4. Address:

.....

.....

Tel.: Mobile:

E-mail:

5. Accommodation Required (Yes/ No):

The Registration fee of Rupeeshas been paid via Demand Draft No.....in favour of "Principal Jabalpur Engineering College, Jabalpur" Demand Draft has been enclosed herewith.

Date:

Signature

Forwarded:

Head of the Institution

(With Seal)