

**Department of Computer Science and Engineering**  
**National Institute of Technology, Calicut**  
**Tentative Course Details - Monsoon 2016-17**  
**CS4038: Data Mining**

### **Weekly Course Schedule**

Tue	10:15am - 11:15am.
Thurs	11.15am-12.15pm.
Fri	8:00am - 09:00am.

### **Instructor**

Name:	<b>Ibrahim bin Abdul Majeed</b>
Office:	<b>CSE 103-C</b>
Mobile:	<b>8277397407</b>
Email:	<b>ibrahim@nitc.ac.in</b>

### **Course Objective**

Objective of the course is a first level study of the area of Data mining.

### **Course Outcomes**

- Differentiate the kinds of data mining tasks such as classification, clustering, decision trees, association analysis.
- Relate fundamental computer science and mathematical principles of algorithms, probability theory, linear algebra and optimization to solve data mining tasks.
- Describe what makes data and how preprocessing of data helps in data mining tasks.
- Use of data mining tools and libraries such as weka, iPython notebook etc., for solving real world data mining problems.
- Model various socially important problems related to domains such as health care, weather forecasting, object recognition, anomaly detection etc., as data mining problems, by extraction of proper features, selection of right algorithms, and evaluation of results.

### **References**

1. Pang-Ning Tan, Micheal Steinbach and Vipin Kumar, *Introduction to Data Mining*, Pearson Education 2006.
2. Ian H. Witten, Eibe Frank, Mark A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques* (3/e), Morgan Kaufmann, 2011.
3. Han and Kamber, *Data Mining: Concepts and Techniques* (2/e), Morgan kaufmann, 2005.

### **Evaluation and Grading**

- $15\%(T1) + 20\%(T2) + 50\%(Final) + 15\%(Assignment)$
- Attendance is mandatory. Late submissions will not be done without prior written permission.
- All issues regarding evaluations must be resolved within one week after the marks are announced.
- Grading will be relative.

## Tentative Course Schedule

WEEK 1,2	Data preprocessing, exploration
WEEK 3,4	Classification: Basic Concepts, Decision Trees
WEEK 5,6	Classifiers: Rule based, Naïve Bayes, Support Vector Machines,
WEEK 7,8	Association analysis
WEEK 8,9	Cluster Analysis
WEEK 10	Anomaly detection

## Standard of Conduct

Each student is expected to adhere to high standards of ethical conduct , especially those related to cheating and plagiarism. Any submitted work must be an individual effort. Any academic dishonesty will result in zero marks in the corresponding evaluation and will be reported to the department council for record keeping and for permission to assign F grade in the course. The department policy on academic integrity can be found at: <http://cse.nit.ac.in/sites/default/files/Academic-Integrity.pdf>