

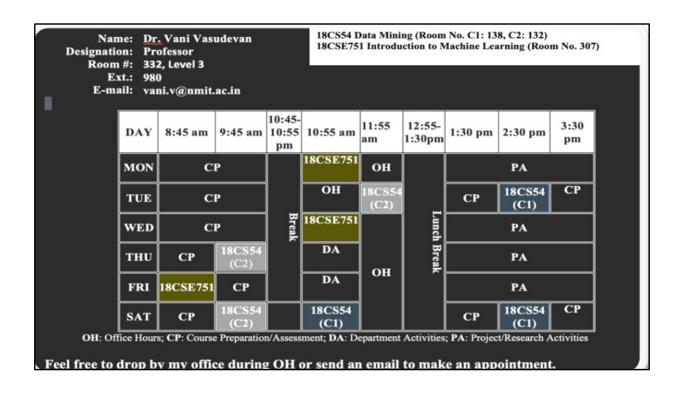
COURSE COORDINATOR

Name: Dr. Vani Vasudevan

Room: 332, III Floor, Main Building

Email: Vani.V@nmit.ac.in

Phone: +918022167980



COURSE DESCRIPTION

This course is an introductory course on data mining. It introduces students to the basic concepts, principles, methods, implementation techniques, and applications of data mining.

COURSE OBJECTIVES

- To understand the concepts of data mining, tasks of data mining and pre-processing techniques.
- To describe the data mining tasks classification, association analysis and cluster analysis.
- To solve problems based on the different data mining tasks.

COURSE CONTENTS...

• Unit -I (T1 - Chapter 1)

Data Mining: Introduction, KDD Process, Challenges, Data Mining Tasks, Data Mining Trends and Applications.

• Unit -II (T1 - Chapter2)

Data, Types of Data, Data Pre-processing, Measures of Similarity And Dissimilarity

• Unit -III (T1 Chapter 4 (4.1-4.3), Chapter 5(5.1, 5.2))

Classification: Basics, General Approach to Solve Classification Problem, Decision Tree Induction, Rule Based Classifiers, Nearest Neighbor Classifiers.

Dr. Vani Vasudevan

10/7/21

6

COURSE CONTENTS

• Unit –IV (T1-Chapter 6(6.1-6.3, 6.6 &6.7))

Association Analysis: Basic Concepts, Frequent Item set Mining Methods – Apriori Algorithm- Frequent item set generation, Rule Generation, FP Growth Algorithm, Evaluation of Association Patterns

• Unit -V (T2-Chapter 10(10.1,10.2,10.3-1-4,10.6))

Cluster Analysis, Partitioning Methods: k-Means, K-medoids, PAM, CLARA and CLARANS

Hierarchical Methods: AGNES, DIANA, BIRCH,

Density Based Methods: DBSCAN, Cluster Evaluation.

Dr. Vani Vasudevan

10/7/21

′

BOOKS

TEXTBOOKS:

- 1. PANG-NING TAN, VIPIN KUMAR, MICHAEL STEINBACH: INTRODUCTION TO DATA MINING, PEARSON, 2012.
- 2. JIAWEI HAN AND MICHELINE KAMBER: **DATA MINING CONCEPTS AND TECHNIQUES**, 3RD EDITION, MORGANKAUFMANN PUBLISHER, 2014

REFERENCE BOOK:

 G. K. GUPTA: INTRODUCTION TO DATA MINING WITH CASE STUDIES, 3RD EDITION, PHI, NEW DELHI, 2009.

Type of Assessment Continuous Internal Assessment(CIA)	WEIGHT (Marks) 50%(50)	ASSESSMENT DATES	
Mid-Semester Exam(MSE) 1	15	Nov 15-17, 2021(Mon-Wed)	
MSE 2	15	Dec 22-24, 2021 (Wed-Fri)	
MSE 3	15	Jan 24-27, 2022 (Mon,Tue, Thurs)	
Best of two MSEs	30		
Evaluation of Problem-Solving Exercises (LA I)	10	Dec 8-10,2021(Wed-Fri)	
Implementation of Laboratory based exercises (LA II)	10	Jan 17-19,2022 (Mon-Wed)	
Semester End Examination (SEE)	50% (100)	Feb 15,2022 onwards	

COURSE OUTCOMES

- CO 1: Outline the data mining tasks, challenges and data mining trends & applications
- CO 2: Demonstrate the various data pre-processing techniques, measures of similarity and dissimilarity on categorical and numeric data.
- CO 3: Illustrate the working of decision tree classifier, rule-based classifier and nearest neighbor classifier using datasets
- CO 4: Categorize frequent item sets from transactional data using FP growth algorithm and apriori algorithm
- CO 5: Illustrate partitioning methods and hierarchical methods of clustering and evaluate the clusters.

WEB RESOURCES

- https://cognitiveclass.ai/ Data Science and Cognitive Computing Courses
- https://www.kdnuggets.com/ Site on AI, Analytics, Big Data, Data Mining, Data Science, and Machine Learning
- https://www.kaggle.com/ ML & DS community
- https://www.cs.waikato.ac.nz/ml/weka/courses.html Waikato University Weka MOOC
- https://ocw.mit.edu/search/ocwsearch.htm?q=data%20mining -MIT open courseware
- https://nptel.ac.in/courses/106/105/106105174/ NPTEL