18CS54 - Data Mining
Lecture 1: Course Overview

Dr.Vani Vasudevan

Professor –CSE, NMIT

COURSE COORDINATOR

Name: Dr. Vani Vasudevan

Room: 332, III Floor, Main Building

Email: Vani.V@nmit.ac.in

Phone: +918022167980

Name: Dr. Vani Vasudevan
Designation: Professor
Room #: 332, Level 3
Ext.: 980
E-mail: vani.v@nmit.ac.in

18CS54 Data Mining (Room No. C1: 138, C2: 132)
18CSE751 Introduction to Machine Learning (Room No. 307)

DAY	8:45 am	9:45 am	10:45-10:55 pm	10:55 am	11:55 am	12:55-1:30pm	1:30 pm	2:30 pm	3:30 pm
MON	CP		Break	18CSE751	OH	Lunch Break	PA		
TUE	CP			OH	18CS54 (C2)		CP	18CS54 (C1)	CP
WED	CP			18CSE751	OH		PA		
THU	CP	18CS54 (C2)		DA			PA		
FRI	18CSE751	CP		DA			PA		
SAT	CP	18CS54 (C2)		18CS54 (C1)			CP	18CS54 (C1)	CP

OH: Office Hours; **CP**: Course Preparation/Assessment; **DA**: Department Activities; **PA**: Project/Research Activities

Feel free to drop by my office during OH or send an email to make an appointment.

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 - Chapter1)

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.

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.

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:

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

COURSE ASSESSMENT PORTFOLIO

Type of Assessment	WEIGHT (Marks)	ASSESSMENT DATES
Continuous Internal Assessment(CIA)	50%(50)	
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