

Course Code: CSE 321

Credits: 2

Course Title: Data Mining and Machine learning

Total Marks: 100

Course Description:

An introduction to data mining and machine learning; Data preparation, model building, and data mining and machine learning techniques such as clustering, decisions trees and neural networks; Induction of predictive models from data: classification, regression, and probability estimation.

∅ **Course Learning Outcome:** (at the end of the course, student will be able to do:)

CLO1	Able to conceptualize basic applications, concepts, and techniques of data mining and machine learning
CLO2	Able to identify appropriate data mining and machine learning algorithms to solve real world problems
CLO3	Able to compare and evaluate different data mining and/or machine learning techniques like classification, prediction, clustering and association rule mining
CLO4	Able to apply knowledge of data mining and machine learning in developing research ideas

∅ **Mapping of Course Learning Outcomes to Program Learning Outcomes** [attainment level used for CLOs from 1(weak)-3(strong) correlation]

PLO's \ CLO's	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8	PLO9	PLO10	PLO11	PLO12
CLO1	3											
CLO2	3	3										
CLO3	2	3	2									
CLO4	3	3			2							1

∅ **Teaching and Learning Activities (TLA)**

TLA1	Interactive discussion using online/multimedia or whiteboard
TLA2	Interactive video and/or scenario-based presentation
TLA3	Case study and group discussion
TLA4	Real-life-project conceptualization in a team to apply data mining and machine learning knowledge

☒ Course Delivery Plan

Week/Lesson (hour)	Discussion Topic & Book Reference	Student Activities during Online and Onsite and TLA	Assessment and Mapping with CLO
<p>Wk. 1 Lesson 1 & 2</p>	<p>Lesson-1: Introduction of the Course Teacher to the students and vice versa; Sharing the course information, e.g. Course Teacher’s information, different information of BLC course, course contents, and course delivery plan and so on.</p> <p>Lesson-2: Introduction to data mining; Relationship to data warehousing; Why data mining is a discipline, Examples of research project using data mining. (Ref.: TB-I [Chapter 1])</p>	<p>Lesson-1 & 2: Online/Onsite discussion; Review Feedback online; Using Interactive content, e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.; <u>TLA1, TLA2</u></p>	<p>Background Preparation week</p>
<p>Wk. 2 Lesson 3 & 4</p>	<p>Lesson-3: Overview of data mining tasks: predictive tasks and descriptive task, e.g. clustering, classifications, rules learning etc. (Ref.: TB-I [Chapter 1])</p> <p>Lesson-4: Introduction to machine learning; Examples of research project using Data Mining; Differences between data mining and machine learning. (Ref.: TB-I [Chapter 1], TB-II [Chapter 1])</p>	<p>Lesson-3 & 4: Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.; <u>TLA1, TLA2</u></p>	<p>CLO1, CLO3</p> <p>Assignment 1 (will be due within Wk. 6)</p> <ul style="list-style-type: none"> - Data Mining Basics and Usage of data repository - Submission usingBLC
<p>Wk. 3 Lesson 5 & 6</p>	<p>Lesson-5: Discussion on data: types of data, quality of data, data preprocessing, and similarity and dissimilarity measures. (Ref.: TB-I [Chapter 2])</p> <p>Lesson-6: Continuation the lecture on Lesson-5, i.e. data preprocessing. (Ref.: TB-I [Chapter 2])</p>	<p>Lesson-5 & 6: Online/Onsite discussion; Review Feedbackonline; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.; <u>TLA1, TLA2, TLA4</u></p>	<p>CLO1, CLO2</p>

<p>Wk. 4 Lesson 7 & 8</p>	<p>Lesson-7: Discussion on data exploration: exploratory data analysis (EDA), summary statistics, data visualization, and OLAP and multidimensional data analysis. (Ref.: TB-I [Chapter 3])</p> <p>Lesson-8: Continuation the lecture on Lesson-7, i.e. data exploration. (Ref.: TB-I [Chapter 3])</p>	<p>Lesson-7 & 8: Online/Onsite discussion; Review Feedbackonline; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.; <u>TLA1, TLA2, TLA4</u></p>	<p><u>CLO1, CLO2</u></p>
<p>Wk. 5 Lesson 9 & 10</p>	<p>Lesson-9: Discussion on classification: a linear classifier, decision trees, model overfitting, evaluating the performance of a classifier etc. (Ref: TB-I [Chapter 4])</p> <p>Lesson-10: Continuation of Lesson-9, i.e. classification and prediction. (Ref: TB-I [Chapter 4])</p>	<p>Lesson-9 & 10: Online/Onsite discussion; Review Feedbackonline; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.; <u>TLA1, TLA2, TLA3</u></p>	<p><u>CLO1, CLO2, CLO3</u></p> <p><u>Class Test# 1</u> (Either online or onsite based on Wk.2 - Wk.4 discussion) based on CLO1 and CLO2</p>
<p>Wk. 6 Lesson 11 & 12</p>	<p>Lesson-11: Discussion on nearest-neighbor classifiers and Bayesian classifiers. (Ref: TB-I [Chapter 4])</p> <p>Lesson-12: Review discussion for Mid Exam; Exercises on classification and prediction. (Ref: TB-I [Chapter 4])</p>	<p>Lesson-11 & 12: Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.; <u>TLA1, TLA2, TLA3, TLA4</u></p> <p><u>Students Submit Assignment-1 through LMS or BLC (online)</u></p>	<p><u>CLO1, CLO2, CLO3</u></p>

Wk. 7	Midterm Exam Week Topics: Wk. 1 – Wk. 6		
Wk. 8	Presentation Week		<u>CLO1, CLO2, CLO3</u> <u>Project Concept Presentation by Team Lead (using Google meet)</u>
Wk. 9 Lesson 13 & 14	<u>Lesson-13:</u> Discussion on cluster analysis: basics, K-means, agglomerative hierarchical clustering, DBSCAN. (Ref: TB-I [Chapter 8]) <u>Lesson-14:</u> Continuation of Lesson-13, i.e. cluster analysis. (Ref: TB-I [Chapter 8])	<u>Lesson-13 & 14:</u> Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.;; <u>TLA1, TLA2, TLA3</u>	<u>CLO1, CLO2</u>
Wk. 10 Lesson 15 & 16	<u>Lesson-15:</u> Continuation of Lesson-13, i.e. cluster analysis. (Ref: TB-I [Chapter 8]) <u>Lesson-16:</u> Discussion on association rule mining: basics, frequent itemset generation, rule generation, etc. (Ref: TB-I [Chapter 6])	<u>Lesson-15 & 16:</u> Online/Onsite discussion; Review Feedback online; Using Interactive content, e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.;; <u>TLA1, TLA2, TLA3</u>	<u>CLO1, CLO2, CLO3</u>
Wk. 11 Lesson 17 & 18	<u>Lesson-17:</u> Continuation of Lesson-16, i.e. association rule mining. (Ref: TB-I [Chapter 6]) <u>Lesson-18:</u> Continuation of Lesson-16, i.e. association rule mining. (Ref: TB-I [Chapter 6])	<u>Lesson-17 & 18:</u> Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.;; <u>TLA1, TLA2, TLA3</u>	<u>CLO1, CLO2, CLO3</u>

<p>Wk. 12 Lesson 19 & 20 (1.5x2=3.0)</p>	<p>Lesson 19: Discussion on artificial neural networks: perceptron, multilayer artificial neural network, etc. (Ref.: TB-I [Chapter 5], TB-II [Chapter 4])</p> <p>Lesson 20: Discussion on artificial neural networks: perceptron, multilayer artificial neural network, etc. (Ref.: TB-I [Chapter 5], TB-II [Chapter 4])</p>	<p><u>Lesson-19 & 20:</u> Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over- PPT, Weekly Forum, PPT, Video, H5P etc; <u>TLA1, TLA2, TLA3</u></p>	<p><u>CLO1, CLO2</u> <u>Class Test# 3</u> (either online or onsite based on Wk. 8 and Wk. 9 discussion) based on CLO2 and CLO3</p>
<p>Wk. 13 Lesson 23 & 24</p>	<p>Lesson 23: Discussion on Research article writing, review and publishing</p> <p>Lesson 24: Review class on topics discussed of Wk. 9 – Wk. 12 for preparing for the final examination.</p>	<p><u>Lesson-21 & 22:</u> Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.; <u>TLA1, TLA2, TLA3</u></p>	<p><u>CLO1, CLO2, CLO3, CLO4</u></p>
<p>Wk. 14</p>	<p style="text-align: center;">Final Exam Week Topics: Wk. 9 - Wk. 12</p>		

Ø **Text Books:**

- I. Introduction to Data Mining: Concepts and Techniques, Pang-Ning Tan, Michael Steinbach, and Vipin Kumar, *Pearson Education, Inc.*, 2006.
- II. Machine Learning, Tom M. Mitchell, *McGraw-Hill*, March 1997.

Ø **Reference Books:**

- (1) Jiawei Han and Micheline Kamber, Data Mining: Concepts and Techniques, *Morgan Kaufmann Publishers*, August 2000.
- (2) S. Sumathi and S. N. Sivanandam, Introduction to Data Mining and its Applications, *Springer-Verlag*, Berlin Heidelberg 2006.
- (3) Robert Tibshirani and Jerome Friedman, The Elements of Statistical Learning, Springer, May 2001.
- (4) Ian H. Witten and Eibe Frank, Data Mining: Practical Machine Learning Tools and Techniques, Elsevier Inc., 2005.
- (5) Ian H. Witten and Eibe Frank, Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, Morgan Kaufmann Publishers, October 1999

Ø **Marking [100 Marks]**

Bloom's Criteria	Attendance (07)	Class Test (15)	Assignment (05)	Presentation (08)	Midterm Exam (25)	Midterm Exam (40)
Remember		05				05
Understand		05	02	02	05	05
Apply		05		03	05	15
Analyze			03		05	05
Evaluate					05	05
Create				03	05	05

Appendix-1: Program outcomes

POs	Category	Program Outcomes
PO1	Engineering Knowledge	Apply the knowledge of mathematics, science, engineering fundamentals and an engineering specialization to the solution of complex engineering problems.
PO2	Problem Analysis	Identify, formulate, research the literature and analyze complex engineering problems and reach substantiated conclusions using first principles of mathematics, the natural sciences and the engineering sciences.
PO3	Design/Development of Solutions	Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for public health and safety as well as cultural, societal and environmental concerns.
PO4	Investigations	Conduct investigations of complex problems, considering design of experiments, analysis and interpretation of data and synthesis of information to provide valid conclusions.
PO5	Modern tool usage	Create, select and apply appropriate techniques, resources and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.
PO6	The engineer and society	Apply reasoning informed by contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to professional engineering practice.
PO7	Environment and sustainability	Understand the impact of professional engineering solutions in societal and environmental contexts and demonstrate the knowledge of, and need for sustainable development.
PO8	Ethics	Apply ethical principles and commit to professional ethics, responsibilities and the norms of the engineering practice.
PO9	Individual work and teamwork	Function effectively as an individual and as a member or leader of diverse teams as well as in multidisciplinary settings.
PO10	Communication	Communicate effectively about complex engineering activities with the engineering community and with society at large. Be able to comprehend and write effective reports, design documentation, make effective presentations and give and receive clear instructions.

PO11	Project management and finance	Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work as a member or a leader of a team to manage projects in multidisciplinary environments.
PO12	Life Long Learning	Recognize the need for and have the preparation and ability to engage in independent, life-long learning in the broadest context of technological change.