Course Code: **CSE450** **Credits**: 3

Course Title: **Data Mining** **CIE Marks**: 60

SEE Marks: 40

Course Description (from syllabus)/Rational:

An introduction to data mining; Data preparation, model building, and data mining techniques such as clustering, decisions trees and neural networks; Induction of predictive models from data: classification, regression, and probability estimation; Application case studies; Data-mining software tools review and comparison.

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
CLO2	Able to identify appropriate data mining algorithms to solve real world problems
CLO3	Able to compare and evaluate different data mining techniques like classification, prediction, clustering and association rule mining
CLO4	Able to apply knowledge of data mining 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 in a team to apply data structure knowledge

Course Delivery Plan (include Lab if any)

Week/Lesson (hour)	Discussion Topic & Book Reference	Student Activities during Online and Onsite and TLA	Assessment and Mapping with CLO
Wk 1 Lesson 1 & 2 (1.5x2 = 3.0)	**Lesson-1**: Introduction to data mining; Relationship to data warehousing; Why data mining is a discipline?, Examples of research project using Data Mining **Lesson-2**: Overview of data mining tasks: Clustering, Classifications, Rules learning etc; Working with Google Colab and Learning Basic of Python; Introduction to Weka;Text Ref: *Chapter 1*	Lesson-1 & 2: Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc; **TLA1, TLA2** Tools to use: Weka for Data Mining; Google Colab (https://colab.research.google.com/)	Background Preparation week
Wk 2 Lesson 3 & 4 (1.5x2 = 3.0)	**Lesson-3**: Review of data mining task and related application examples; Introduction to Data Repository (local and global) including UCI and others **Lesson-4**: Course Project Team and discussion on presentation and deliverables; Google site display of successful projects, Discussion on project plan; Working with basic python in Google Colab with NumPy and Panda	Lesson-3 & 4: Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc; **TLA1, TLA2** Hands-on: Working with Google Colab for basic python; Working with Numpy and Panda	**CLO1, CLO3** **Assignment 1** (will be due by Wk3) - Data Mining Basic and Usage of data repository -Submission using LMS (BLC)
Wk 3 Lesson 5 & 6 (1.5x2=3.0)	**Lesson-5**: Discussion on data mining process: Data preparation and cleansing and task identification (Ref: Chapter 3) **Lesson-6**: Review Discussion on Data preprocessing; Working with Weka and Google Colab for data preprocessing; Project Discussion and execution plan; *Discussion on*	Lesson-5 & 6: Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc; **TLA1, TLA2, TLA4** Hands-on: Working for data pre-processing using Google colab; Student form team for the project and fill the team	**CLO1, CLO2**

	Course Projects.	info using Google form. **Student Submit Assigment-1 in LMS or BLC (online)**	
Wk 4 Lesson 7 & 8 (1.5x2=3.0)	**Lesson-7:** Discussion on Classification and Prediction (Ref: Chapter 7); Classification: tree-based approaches, Neural Networks etc (Ref: Chapter 7) **Lesson-8:** Discussion on Classification; Working on classification using Weka and Google Colab	Lesson-7 & 8: Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc; **TLA1, TLA2** Hands-on: Working on classification using Weka and Google Colab *Course Project Allocation among teams.*	**CLO1**, **CLO2** **Class Test# 1** (Either online or onsite based on Wk2-Wk3 discussion) based on CLO1 and CLO2
Wk 5 Lesson 9 & 10 (1.5x2=3.0)	**Lesson-9**: Review discussion on Classification and Prediction; Working with Google colab for classification and prediction **Lesson-10**: Review Discussion on Classification and Prediction; Working with Google colab for classification and prediction;	Lesson-9 & 10: Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc; **TLA1, TLA2**, **TLA3** Hands-on: Working on classification using Weka and Google Colab	**CLO1**, **CLO2**
Wk 6 Lesson 11 & 12 (1.5x2=3.0)	**Lesson 11:** Discussion on Nearest Neighbour and Bayesian Classification; Example and exercise **Lesson 12:** Review discussion for Mid Exam; Exercises on classification and prediction	Lesson-11 & 12: Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc; **TLA1, TLA2**, **TLA3, TLA4** Hands-on: Working on classification using Weka and Google Colab	**CLO1**, **CLO2, CLO3** **PRN#1:Project Concept Presentation by Team Lead (using Google meet team record their presentation and upload the Google meet link in BLC)** **Class Test# 2** (either online or onsite based on

			Wk 4, Wk 5 discussion) based on CLO1, CLO2 **Assignment 2** (will be due by Wk 8) - Classification and Prediction
Wk 7	**Midterm Exam Week** **Topics: Wk 3 – Wk 6**		
Wk 8 Lesson 13 & 14 (1.5x2=3.0)	**Lesson 13:** Discussion on Clustering basic, Clustering – statistical approach an Neural network and other approaches (Ref: Chapter 8); Example and exercise **Lesson 14:** Review discussion on clustering; Case Study; Working with Weka and Google Colab for clustering	Lesson-13 & 14: Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc; **TLA1, TLA2, TLA3** Hands-on: Working on Clustering using Weka and Google Colab **Student Submit Assigment-2 in LMS or BLC (online)**	**CLO1**, **CLO2**
Wk 9 Lesson 15 & 16 (1.5x2=3.0)	**Lesson 15**: Discussion on Association rule mining (Ref: Chapter 6); Example and Exercise **Lesson 16**: Review discussion on Association Rule mining; Using Google Colab and Weka for accociation rule mining	Lesson-15 & 16: Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc; **TLA1, TLA2, TLA3** Hands-on: Working on association rule mining using Google Colab and Weka	**CLO1**, **CLO2**
Wk 10 Lesson 17 & 18 (1.5x2=3.0)	**Lesson 17:** Discussion on CRISP-DM; CRISP-DM Process model (Ref: Chapter 5, Chapter 10, Appendix A); Research Paper discussion on Classification and	Lesson-17 & 18: Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly	**CLO1**, **CLO2**

Lab Session 9 (3.0)	Association rule mining **Lesson 18:** Review discussion on Data Mining Model; Using Google Colab for Association Application	Forum, PPT, Video, H5P etc; **TLA1, TLA2, TLA3** Hands-on: Working on association rule mining using Google Colab and Weka	
Wk 11 Lesson 19 & 20 (1.5x2=3.0)	**Lesson 19:** Discussion on Time series mining; Mining data streams; Research reading **Lesson 20:** Discussion on Neural Network and applications; Using Octave or Matlab	Lesson-19 & 20: Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc; **TLA1, TLA2, TLA3** Hands-on: Working on neural network using Matlab or Octave	**CLO1**, **CLO2** **Class Test# 3** (either online or onsite based on Wk 8 and Wk 9 discussion) based on CLO2 and CLO3
Wk 12 Lesson 21 & 22 (1.5x2=3.0)	**Lesson 21:** Review discussion on Clustering and Association rule mining **Lesson 22:** Course Project presentation by the Team Lead	Lesson-21 & 22: Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc; **TLA1, TLA2, TLA3, TLA4** Hands-on: Course project presentation by team lead	**CLO1**, **CLO3** **PRN#2:Project Implementation Presentation by Team (using Google meet team record their presentation and upload the Google meet link in BLC)**
Wk 13 Lesson 23 & 24 (1.5x2=3.0)	**Lesson 23:** Discussion on Research article writing, review and publishing **Lesson 24:** Review class on topics discussed of Wk 8, Wk 9 and Wk 10 for preparing for the final exam	Lesson-21 & 22: Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc; **TLA1, TLA2, TLA3**	**CLO1**, **CLO2**, **CLO3**
Wk 14	**Final Exam Week** Topics: Wk 8, Wk 9, and Wk 10		

Text Book(s): Data Mining: Concepts and Techniques, Jiawei Han and Micheline Kamber, Morgan Kaufmaan Publishers, August 2000

Reference Material/Book(s):

(1) S. Sumathi, S.N. Sivanandam, Introduction to Data Mining and its Applications, Springer-Verlag Berlin Heidelberg 2006
(2) Ian H. Witten and Eibe Frank, Data Mining: Practical Machine Learning Tools and Techniques, Elsevier Inc., 2005
(3) Ian H. Witten and Eibe Frank, Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, Morgan Kaufmann Publishers, October 1999
(4) Schlar.google.com and Google search engine for research papers

CIE – Breakup (Theory) [60 marks]

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

CIE – Breakup (Lab) [100 marks]

Bloom's Criteria	**Attendance (10)**	**Lab Performance (25)**	**Lab Report (25)**	**Lab Final (40)**
Remember				
Understand		05	05	10
Apply		05	05	05
Analyze		05	05	10
Evaluate			10	05

Create		10		10

SEE – Semester End Examination [40 marks] {Theory}

Bloom Criteria	**Score for the Test**
Remember	05
Understand	05
Apply	15
Analyze	05
Evaluate	05
Create	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.