Course Code: CSE 322 **Credits: 3**

Course Title: Data Mining and Machine learning Lab **Total Marks: 100**

Course Description:

An introduction to Weka and Python; 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 possess the basic knowledge of Weka and Python concerning data mining and machine learning
CLO2	Able to implement different data mining and machine learning algorithms like classification, prediction, clustering and association rule mining to solve real-world problems using Weka and Python
CLO3	Able to compare and evaluate different data mining and machine learning algorithms like classification, prediction, clustering and association rule mining using Weka and/or Python
CLO4	Able to apply implementation 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
Wk. 1	**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. **Lesson-2**: Introduction to Weka: installation, accessing a data set, introduction to different file formats, running a simple classifier.	Online/Onsite discussion; Using Interactive content, e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.; **TLA1, TLA2**	Background Preparation week
Wk. 2	Data Visualization Using Weka	Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.; **TLA1, TLA2**	**CLO1, CLO3**
Wk. 3	Feature/Attribute Selection Using Weka	Online/Onsite discussion; Review Feedbackonline; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.; **TLA1, TLA2, TLA4**	**CLO1, CLO2**

Wk. 4	Classification Using Weka: ☼ Algorithm type: decision tree ☼ Method of classifier evaluation ☼ Parameter tuning ☼ Performance evaluation	Online/Onsite discussion; Review Feedbackonline; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.; **TLA1, TLA2, TLA4**	**CLO1, CLO2**
Wk. 5 Lesson 9 & 10	Classification Using Weka: ☼ Algorithm type: Bayesian, instance-based ☼ Method of classifier evaluation ☼ Parameter tuning ☼ Performance evaluation	Lesson-9 & 10: Online/Onsite discussion; Review Feedbackonline; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.; **TLA1, TLA2, TLA3**	**CLO1**, **CLO2, CLO3**
Wk. 6	Cluster Analysis Using Weka: ☼ Algorithm type: partitional (*K*-means), hierarchical, density-based ☼ Method of classifier evaluation ☼ Parameter tuning ☼ Performance evaluation	Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc.; **TLA1, TLA2, TLA3, TLA4**	**CLO1**, **CLO2, CLO3**

Wk. 7	**Midterm Exam Week**		
Wk. 8	**Presentation of Project # 1 (Using Weka) by Students' Groups**		**CLO1, CLO2, CLO3, CLO4** **Project Presentation by Groups (Using Google Meet)**
Wk. 9 Lesson 13 & 14	Introduction to Python: installation, accessing a data set, introduction to different file formats, running a simple classifier.	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**
Wk. 10	Classification Using Python: ☼ Algorithm type: Decision tree ☼ Method of classifier evaluation ☼ Parameter tuning ☼ Performance evaluation	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. 11	Classification Using Python (Continued): ☼ Algorithm type: Bayesian, instance-based ☼ Method of classifier evaluation ☼ Parameter tuning ☼ Performance evaluation	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. 12	Cluster Analysis Using Python: ☼ Algorithm type: partitional (*K*-means), hierarchical, density-based ☼ Method of classifier evaluation ☼ Parameter tuning ☼ Performance evaluation	Online/Onsite discussion; Review Feedback online; Using Interactive content e.g. Voice over PPT, Weekly Forum, PPT, Video, H5P etc; **TLA1, TLA2, TLA3, TLA4**	**CLO1**, **CLO2**
Wk. 13	**Presentation of Project # 2 (Using Python) by Students' Groups**		**CLO1**, **CLO2**, **CLO3, CLO4**
Wk. 14	**Final Exam Week** Topics: Wk. 1 - Wk. 12		

∅ Recommended Books:

I. Ian H. Witten, Eibe Frank, Mark A. Hall, Christopher J. Pal, Data Mining: Practical Machine Learning Tools and Techniques, 4th Edition, Morgan Kaufmann, 2016.

II. Jason Brownlee, Machine Learning Mastery with Weka, 2016.

III. Trevor Hastie, Robert Tibshirani and Jerome Friedman, The Elements of Statistical Learning, Springer, 2008.

IV. Andreas C. Müller and Sarah Guido, Introduction to Machine Learning with Python, O'Reilly Media, Inc., October 2016.

V. Ian H. Witten and Eibe Frank, Data Mining: Practical Machine Learning Tools and Techniques, Elsevier Inc., 2005.

VI. Ian H. Witten and Eibe Frank, Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, Morgan Kaufmann Publishers, October 1999.

∅ Marking [100 Marks]

Bloom's Criteria	Attendance (10)	Class Performance (25)	Project / Lab Report (25)	Final Exam (40)
Remember		05		05
Understand		05	05	05
Apply		05	05	15
Analyze		05	05	05
Evaluate		05	05	05
Create			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.