|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| D:\UAAR\UIIT\courseOutlineCommittee\CourseContents_Final_V02\New folder\logo4.png | **PMAS Arid Agriculture University Rawalpindi**  **University Institute of Information Technology** | | | |
| AI-403 Data Mining | | | | | |
| **Credit Hours:** | | **3(2-3)** | **Prerequisites:** | **None** | |
| **Teacher:** | | Dr. Ghulam Mustafa |  |  | |

|  |
| --- |
| **Course Description:** |
| This course serves as a guide to techniques and applications of data mining including Machine Learning and Data Mining, Classification and its methods, Knowledge representation, Data Preparation for Knowledge Discovery, decision trees etc. |
| **Course Outcomes:** |
| On completion of this module, students should be able to: Apply the specific concepts in real practice, Demonstrate specific methods in Data Mining; Demonstrate an understanding of the fundamental ideas of Knowledge Representation and Reasoning; Demonstrate an understanding of the principles of a number of different approaches in Machine Learning; Demonstrate the ability to apply AI and Computational Intelligence techniques to a variety of research and application projects. |
| **Teaching Methodology:** |
| Lectures, Assignments, Presentations, Course Project etc. Major component of the course should be covered using practical implementation of AI techniques. |
| **Courses Assessment:** |
| Exams, Assignments, Quizzes. Course will be assessed using a combination of written examinations. |
| **Books &Reference Materials:** |
| * Jiawei Han, Jian Pei, Hanghang Tong, Data Mining Concepts and Techniques 4th Edition - July 2, 2022. * Dr. David L. Olson, Dr. Dursun Delen, Advanced Data Mining Techniques Latest Edition * Witten and Eibe, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, Latest Edition |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Theory | Practical | | |
| Week 1 | Introduction: Data Mining  * Why Data Mining? * What is Data Mining? * Knowledge Discovery | * Introduction of Weka * Versions and Installation | | |
| * (KDD) Process * What Technologies Are Used? | * Introduction of Datasets formats used in Weka * Load Datasets in Weka | | |
|  | * Introduction of ML algorithms used in Weka. | | |
| Week 2 | * Data Mining in Business Intelligence * KDD Process: A Typical View from ML   and Statistics   * Multi-Dimensional View of Data Mining | Introduction of Data Preprocessing Algorithms on Weka | | |
| * Data Mining: On What Kinds of Data * What Kinds of Patterns Can Be Mined? * Data Mining Functions | Demonstration of data-preprocessing on ARFF datasets | | |
|  | Implementation of Data Preprocessing algorithms using Python | | |
| Week 3 | Getting to Know Your Data   * + Data Objects and Attribute Types | Feature Selection: Introduction | | |
| * Important Characteristics of Structured Data * Basic Statistical Descriptions of Data | Demonstration of Feature Selection: process and visual representation | | |
|  | Implementation of Feature Selection using Python | | |
| Week 4 | * Data Preprocessing: An Overview   + Major Tasks in Data Preprocessing   + Data Cleaning | Introduction of:  Training set  Supplied test set | | |
| * Data Integration * Data Reduction | Cross-validation  Percentage split | | |
|  | Demonstrate in Weka | | |
| Week 5 | * Mining Frequent Patterns | Introduction of Association Rule Mining in Weka | | |
| * + Mining Various Kinds of Association   Rules | Demonstration of Association rule process on ARFF datasets | | |
|  | | | Implementation of Data Association rule mining using Python |
| Week 6 | * Classification: * Decision Tree * Decision Tree Induction | | | Introduction of Navie Bayes in Weka |
| * + Attribute Selection Measures   + Tree Pruning | | | Demonstration of Navie Bayes on ARFF datasets |
|  | | | Implementation of Navie Bayes algorithm using Python |
| Week 7 | * Bayes Classification Methods   + Bayes’ Theorem | | | Introduction of Logistic Regression in Weka |
| * + Na¨ıve Bayesian Classification | | | Demonstration of Logistic Regression on ARFF datasets |
| * Rule-Based Classification | | | Implementation of Logistic Regression algorithm using Python |
| Week 10 | MIDTERMS | | |  |
| Week 9 | * Genetic Algorithms | | | Introduction of Support Vector Machines in Weka |
| * Genetic Algorithm Operators | | | Demonstration Support Vector Machines on ARFF datasets |
|  | | | Implementation of Support Vector Machines using Python |
| Week 10 | * Cluster Analysis: Concepts and   Methods | | | Hierarchal Clustering and Filtered Clustering |
| * + Cluster Analysis and Partitioning Methods | | | Canopy and cobweb clustering |
|  | | | Implement using Weka |
| Week 11 | * K means Clustering | | | Introduction of K means clustering in Weka |
| * + Support Vector Machine | | | Demonstration K means clustering on ARFF datasets |
|  | | | Implementation of K means clustering on using Python |
| Week 12 | Neural Network  Single Layer | | Introduction of Multilayer Perceptron in Weka | |
| * + Multi-Layer | | Demonstration of Multilayer Perceptron | |
|  | | Implementation of Multilayer Perceptron using Python | |
| Week 13 | * Apriori algorithm: Introduction and Working | | Introduction of Apriori algorithm in Weka | |
| * Numerical examples | | Demonstration of Apriori algorithm | |
|  | | Implementation of Apriori algorithm using Python | |
| Week 14 | * Expert Systems | | Machine Learning Platform: Tensor Flow | |
| * Expert Systems Working | | Introduction and Overview | |
|  | | Installation | |
| Week 15 | * Decision Support Systems | | Convolutional Neural Networks | |
| * DSS: Working and Model | | Demonstration of Convolutional Neural Networks on Tensor Flow. 1 | |
|  | | Demonstration of Convolutional Neural Networks on Tensor Flow. 2 | |
| Week 16 | Presentations | | | |
| Presentations | | | |
| Presentations | | | |

|  |  |  |
| --- | --- | --- |
| **Course Learning Outcomes (CLOs):** | | |
| At the end of the course the students will be able to: | **Domain** | **BT Level\*** |
| 1. Students should be able to describe the underlying concepts of data mining. Understand the Data Mining Process and implement data mining process based solutions. | C | 1 |
| 1. Application of data mining algorithms. | P | 3 |
| 1. Analysis of valid patterns in test data using data mining experiments with test data. | C | 4 |
| \* BT= Bloom’s Taxonomy, C=Cognitive domain, P=Psychomotor domain, A= Affective domain | | |