COURSE OUTLINE

INSTITUTION **University of Management & Technology, Lahore**

Course Description:

|  |  |
| --- | --- |
| Course Code | CS458 |
| Course Title | Data Mining |
| Credit Hours | 3 |
| Course Instructor | Dr. Noshad Jamil |
| Contact | Email: noshad.jamil@umt.edu.pk; Office: CB1-507 |
| Lectures | Section A1 : Tuesday (17:00 – 18:15) , Wednesday (15:30 – 14:45)  Section D1 : Monday (12:30 – 13:45), Wednesday (11:00 – 12:15) |
| Office Hours |  |
| Prerequisites by  Course(s) and Topics | Advance Statistics, Introduction to Data Science |
| Assessment Instruments with Weights (homework, quizzes, midterms, final, programming assignments, lab  work, etc.) | * Quizzes (Approx. 6) **15%** * Assignments (Approx. 4) **20%** * Mid Term **25%** * Final Exam **40%** |
| URL (if any) |  |
| Course Moderator |  |
| Textbook (or Laboratory Manual for Laboratory Courses) | Jiawei Han & Micheline Kamber, Jian Pei (2011). Data Mining: Concepts and Techniques, 3rd Edition. |
| Reference Material | **RefA:** Pang-Ning Tan, Michael Steinbach, and Vipin Kumar (2005). Introduction to Data Mining.  **RefB:** Charu C. Aggarwal (2015). Data Mining: The Textbook  **RefC:** D. Hand, H. Mannila, P. Smyth (2001). Principles of Data Mining. MIT Press.  **RefD:** Ian H. Witten, E Frank, M A Han. Data Mining Practical Learning Tools & Techniques |
| Course Goals/Objectives | Data Mining has emerged at the confluence of artificial intelligence, statistics, and databases as a technique for automatically discovering hidden patterns in large datasets. The main purpose of this course is the ability to analyze and construct knowledge from data. The aims of this course are to:   * Expand on the student’s understanding and awareness of the concepts of data mining basics, techniques, and application. * Introduce the concepts of Data Pre-processing and Summary Statistics. * Introduce the concepts of Frequent Item Set Generation, Associations and Correlations measures. * Introduce the concepts of Classification, Prediction, and Clustering algorithms.   Build on the programming and problem-solving skills developed in previous subjects studied by the student, to achieve an understanding of the development of Classification, Prediction, and Clustering applications. |

Course Learning Outcomes (CLOs):

|  |  |  |  |
| --- | --- | --- | --- |
|  | **CLOs** | **Description** | **Domain & BT Level \*** |
| CLO 1 | **Apply** preprocessing techniques on any given raw data. | C3 (Apply) |
| CLO 2 | **Apply** proper data mining algorithm to discover interesting patterns | C3 (Apply) |
| CLO 3 | **Analyze** and extract patterns to solve problems and point out how to deploy solution | C4 (Analyze) |
| CLO 4 | **Analyze** systematically supervised, semi-supervised and unsupervised models and algorithms for their accuracy. | C4 (Analyze) |
| \* BT= Bloom’s Taxonomy, C=Cognitive domain, P=Psychomotor domain, A= Affective domain | | | |

Mapping of CLOs to Program Learning Outcomes (PLOs):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| CLOs/PLOs | CLO 1 | CLO 2 | CLO 3 | CLO 4 |
| PLO 1: Academic Education |  |  |  |  |
| PLO 2: Knowledge for Solving Computing Problems |  |  |  |  |
| **PLO 3: Problem Analysis** | 🗸 | 🗸 |  |  |
| **PLO 4: Design and Development of Solutions** |  | 🗸 | 🗸 | 🗸 |
| **PLO 5: Modern Tool Usage** |  |  | 🗸 | 🗸 |
| PLO 6: Individual and Teamwork |  |  |  |  |
| PLO 7: Communication |  |  |  |  |
| PLO 8: Computing Professionalism and Society |  |  |  |  |
| PLO 9: Ethics |  |  |  |  |
| PLO 10: Life Long Learning |  |  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tentative Course Plan** | | | | |
| **Week** | **Topics** | **Chapter** | **Assessment** | **CLOs** |
| 1 | Introduction to data mining and basic concepts | Ch 1 |  | 1 |
| 2 | Revision of Statistics and Database Management Systems | Ch 1 & Ch 2 |  | 1 |
| 3 | Data Preprocessing | Ch 3 | Assignment 1 | 1 |
| 4 | Quiz 1 | 1,2,3 |
| 5 | Data Warehousing & Online Analytical Processing | Ch 3 |  | 1,2 |
| 6 | Data Cube Technology | Ch 5 | Assignment 2 | 1 |
| 7 | Frequent Pattern Analysis using Association Rule Mining | Ch 6 | Quiz 2 | 2,3,4 |
| 8 | Mid-Term and Review |  | Midterm Exam | 1,2,3,4 |
| 9 | Classification | Ch 8 & Ch 9 |  | 2,3,4 |
| 10 | Assignment 3 | 2,3,4 |
| 11 | Cluster Analysis | Ch 10 & Ch 11 | Quiz 3 | 2,3,4 |
| 12 |  | 2,3,4 |
| 13 | Outlier Detection | Ch 12 | Quiz 4 | 2,3,4 |
| 14 | Introduction to Weka Data Mining Workbench | Ch 10 (RefD) | Assignment 4 | 2,3 |
| 15 | Preprocessing and Classification in Weka | Ch 17 (RefD) |  | 2,3,4 |
| 16 | Final |  | Final Exam | 1,2,3,4 |