|  |  |
| --- | --- |
|  | **FAST National University of Computer and Emerging Sciences** |
| **Department of Computer Science, Peshawar Campus**  **Course Content, Objectives and Outline**  **(Spring 2023)** |

**PROGRAM:** BS-Computer Science / BS-Software Engineering

**SEMESTER:** Fall (2023)

**COURSE:** Data Science (CS4048)

**PRE-REQUISITE:**

**CREDIT HRS:** 3

**INSTRUCTOR:** Dr. Musadaq Mansoor

**Email:** musadaq.mansoor@nu.edu.pk

**Course Content:**

Introduction to Data Science, Data Science Life Cycle, Data, Types of Data, Data Acquisition, Extract Transform Load, Data Wrangling, Exploratory data analysis, Data Visualization, Data Analysis, Text Analysis, Machine Learning, Supervised Learning, Unsupervised Learning, Clustering, Data Science success stories.

**Objectives:**

Understanding basic concepts of Data Science (1)

Understanding Data Science Life Cycle (4)

Understanding data and its types (1)

Implementation of Data Science concepts (5)

Understanding Machine learning concepts for Data Science (3)

**Reference Books:**

1. Python for Data Analysis, 1st Edition, William McKinney

2. Doing Data Science, 1st Edition, Cathy O'Neil and Rachel Schutt

3. Introduction to Data Science. A Python Approach to Concepts, Techniques and Applications, 1st Edition, Laura Igual.

**Course Evaluation Criteria:**

|  |  |  |
| --- | --- | --- |
| **#** | **Evaluation** | **Weightage %** |
| 1 | Project | 4 |
| 2 | Quiz | 10 |
| 3 | Presentation | 4 |
| 4 | Class Tasks | 2 |
| 5 | Sessional-I | 15 |
| 6 | Sessional-II | 15 |
| 7 | Finals | 50 |
| **Total** | | **100** |

**Weekly Distribution:**

|  |  |
| --- | --- |
| **Week No.** | **Topic Description** |
| 1 | * Introduction * Data Science Life cycle & process |
| 2 | * Introduction to Data: * Types of data and datasets |
| 3 | * Python –I * Python - II |
| 4 | * Data Acquisition * BeautifulSoup |
| 5 | * LXML * PyPDF |
| 6 | * ETL * PETL |
| 7 | * Data Wrangling |
| 8 | * Data Analysis |
| 9 | * Pandas – I * Pandas - II |
| 10 | * Exploratory Data Analysis |
| 11 | * Matplotlib * Seaborn |
| 12 | * Introduction to Machine Learning |
| 13 | * Supervised Learning * Unsupervised Learning |
| 14 | * Model Evaluation and Performance Metrics * Introduction to Scikit – Learn |
| 15 | * Data Science case studies |
| 16 | * Presentations |