National Computing Education Accreditation Council
NCEAC

COURSE DESCRIPTION FORM

INSTITUTION   FAST School of Computing, National University of Computer and Emerging Sciences, Islamabad

PROGRAM(S) TO BE EVALUATED   **BS-DS FALL 2024**

Course Description

(Fill out the following table for each course in your computer science curriculum. A filled-out form should not be more than 2-3 pages.)

Course Code	DS2001
Course Title	Introduction to Data science
Credit Hours	3
Prerequisites by Course(s) and Topics	
Grading Policy	Absolute Grading
Assessment Instruments with Weights (homework, quizzes, midterms, final, programming assignments, lab work, etc.)	Assessment items of Theory Part (see table below)

Assessment items of Theory Part

Assessment Item	Number	Weight (%)
Assignments	4	10
Quiz	4	10
Sessional-1	1	12.5
Sessional-2	1	12.5
Project	1	10
Final Exam	1	45

Policy about missed assessment items in the	Retake of missed assessment items (other than midterm/ final exam) will not be held. For a missed midterm/ final exam, an exam retake/ pretake application along

National Computing Education Accreditation Council
NCEAC

course	with necessary evidence are required to be submitted to the department secretary. The examination assessment and retake committee decides the exam retake/ pretake cases.
Course Plagiarism policy	Plagiarism in project or midterm/ final exam may result in F grade in the course. Plagiarism in an assignment will result in zero marks in the whole assignments category.
Course Instructors	Ms. Saira Qamar
Lab Instructors (if any)	Mr. Talha Tariq

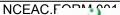| **Course Coordinator**		Ms. Saira Qamar
URL (if any)		Google Classroom Code: w4u66t3
Current Catalog Description		Data Science is the study of the generalizable extraction of knowledge from data. Being a data scientist requires an integrated skill set spanning mathematics, statistics, machine learning, databases and other branches of computer science along with a good understanding of the craft of problem formulation to engineer effective solutions. This course will introduce students to this rapidly growing field and equip them with some of its basic principles and tools as well as its general mindset. Students will learn concepts, techniques and tools they need to deal with various facets of data science practice, including data collection and integration, exploratory data analysis, predictive modeling, descriptive modeling, data product creation, evaluation, and effective communication. The focus in the treatment of these topics will be on breadth, rather than depth, and emphasis will be placed on integration and synthesis of concepts and their application to solving problems.
Textbook (or **Laboratory Manual** for Laboratory Courses)		1. Severance, C.R., 2016. "Python for everybody: Exploring data using Python 3." CreateSpace Independent Publ Platform. 2. Miller, B.N., Ranum, D.L. and Anderson, J., 2019. "Python programming in context." Jones & Bartlett Pub. 3. McKinney, W., 2012. "Python for data analysis: Data wrangling with Pandas, NumPy, and IPython." O'Reilly Media, Inc.
Reference Material		1. Joshi, P., 2017. "Artificial intelligence with python." Packt Publishing Ltd. 2. Janert, P.K., 2010. "Data analysis with open source tools: a hands-on guide for programmers and data scientists." O'Reilly Media, Inc.
Course Learning Outcomes		**A. Course Learning Outcomes (CLOs)**
		CLO-1 Comprehend the fundamental constructs of programming language for data analysis and representation. CLO-2 Understand and apply the Object-oriented concepts in the programming language CLO-3 Solve and analyze programming and data analysis problems using standard

libraries and/or toolboxes of the programming language.

PLO:

By the time of graduation, the students develop an ability to:

1) Apply knowledge of computing and mathematics that is appropriate for the program.
2) Analyze a problem and define computing requirements that are appropriate to its solution.
3) Design, implement, and evaluate a computer-based system, process, component or program to meet desired needs.
4) Work in a team to accomplish a common goal.
5) Understand professional, ethical, and social issues and responsibilities.
6) Communicate effectively with different audiences.
7) Learn programming for large-sized datasets
8) Identify useful and hidden patterns from data.
9) Improve decision making skills by mining data from various aspects.
10) Solve real world problems by applying mathematical and computational approaches.
11) Change the world for the better – in areas like healthcare, transportation, and education etc.

C. Mapping of CLOs on PLOs
(CLO: Course Learning Outcome, PLOs: Program Learning Outcomes)

		PLOs										
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
CLOs	1	✔	✔									
	2		✔	✔								
	3			✔	✔		✔					

	Topics to be covered:			
Topics Covered in the Course, with Number of Lectures on Each Topic (assume 15-week instruction and one hour lectures)	List of Topics	No. of Weeks	Contact Hours	CLO(s)
	Introduction: What is Data Science? • Why now? – Datafication • Current landscape of perspectives • Skill sets needed	1	3	1
	Introduction to Python, Object oriented concepts	2	6	2
	Assignment 1 (Python)			
	Numpy, Pandas, Matplotlib, Seaborn	2	6	2

Basic Exploratory Data Analysis and the Data Science Process
    • Basic tools (plots, graphs and summary statistics) of EDA
    • Philosophy of EDA - The Data Science Process
    • A Case Study
EDA on different types of datasets
	1	2	**2**

Assignment 2 (Pandas and EDA)			
Playing with time series data (Univariate, Multivariate, Filling missing values, Plotting etc) Text Processing Audio Processing Image Processing Basics	**2**	**3**	**7,8,9,11**

Introduction to Machine Learning
    -Machine learning Basics
    -Why we need it.
    -Supervised learning
-Unsupervised Learning
-Reinforcement Learning
-Classification vs Regression
-Pipeline for training machine learning model
-Data Preprocessing
-Detail about confusion matrix
-Scoring
    -Accuracy
    -F measure
    -Precision
    -Recall
	2	**6**	**2**

Assignment 3 (Machine Learning)			
Machine Learning Algorithms Natural Language Processing	4.5	**8**	**2**
Assignment 4 (Machine Learning)			

Feature Generation and Feature Selection (Extracting Meaning from Data) • Motivating application • Feature Generation Feature Selection algorithms	**2**	**4**	**2**	
Assignment 5				
Assignment Presentation	**1**	**2**	**3,4,5,6,9,10**	
Total	**15**	**45**		

Laboratory	No lab
Projects/Experiments Done in the Course	
Programming Assignments Done in the Course	Yes, there are five programming assignments.

Class Time Spent (in hours)	**Theory**	**Problem Analysis**	**Solution Design**	**Social and Ethical Issues**
	14.5	17	13	0.5

Oral and Written Communications	Every student is required to submit at least __1___ written reports of typically ___5__ pages and to make __1___ demonstration of typically ____10___ minutes duration.