COURSE DESCRIPTION: MT2005 Probability and Statistics

INSTITUTION	National University of Computer and Emerging Sciences, Chiniot-Faisalabad Campus, Pakistan
PROGRAM TO BE EVALUATED	BS- Computer Science BS- Software Engineering BS- Artificial Intelligence

Course Description

Course Code	MT2005
Course Title	Probability and Statistics
Credit Hours	3 (3-0)
Prerequisites by Course(s) and Topics	NA
Grading Policy	Relative (Interquartile Range Method)
Policy about missed assessment items in the course	Retake of missed assessment items (other than midterm/ final exam) will not be held. A student who misses (or late submits) an assessment item (other than the sessional/final exam) is awarded zero marks in that assessment item. For a missed midterm/ final exam, an exam re-take/ pre-take application along with necessary documents must be submitted to the department secretary. The examination assessment and retake committee will decide the exam re-take/ pre-take cases.
Course Plagiarism Policy	Plagiarism in the sessional/final exam may result in an F grade in the course. Plagiarism in the project/assignment will result in zero marks in the **whole assignments** category.
Assessment Instruments with Weights (homework, quizzes, midterms, final, programming assignments, lab work, etc.)	

Assessment Item	Number	Weight (%)
Assignment	3-4	10
Quiz	3-4	10
Project	1	7
Sessional I	1	14
Sessional II	1	14
Final Exam	1	45

Course Coordinator	Dr. Haris Khurram
URL (if any)	

Course Catalog Description	This course introduces fundamental concepts of Statistical analysis and probability theory such as exploratory data analysis and data visualization. Moreover, using data to find the probability for intelligent systems using the probability distribution's function. Also, computational regression modeling uses R language with statistical testing and confidence interval estimation. Finally, equipping students with the tools to interpret and analyze real data.
Textbook (or **Laboratory Manual** for Laboratory Courses)	1. **Probability & Statistics for Engineers and Scientists** by R.E Walpole 2. **A First Course in Probability** by Sheldon Ross
Reference Material	1. **Probability & Statistics with R for Engineers and Scientists** by Michael Akritas

Course Goals

A. Course Learning Outcomes (CLOs)

Upon the successful completion of this course, students will have to know how to:

CLO	Course Learning Outcome (CLO)	Domain	Taxonomy level	GAs	Tools*
1	Use the ability to collect, Identify, and analyze the data using descriptive statistics and exploratory data analysis.	Cognitive	3	02	A1, Q1, S-I, F
2	Apply the probability theory and distribution on data and able to solve the real problems.	Cognitive	3	02	A2, A3, Q2, Q3, Q4, S-II, F
4	Analyze the data using statistical and computational models and draw inferences in the computing field.	Cognitive	4	02	A4, Q5, P, F

B. Graduate Attributes

For each attribute below, indicate whether this attribute is covered in this course or not. Leave the cell blank if the enablement is little or non- existent.

GA-2. Knowledge for Solving Computing Problems	Apply knowledge of computing fundamentals, knowledge of a computing specialization, and mathematics, science, and domain knowledge appropriate for the computing specialization to the abstraction and conceptualization of computing models from defined problems and requirements	√
GA-3. Problem Analysis	Identify, formulate, research literature, and solve complex computing problems reaching substantiated conclusions using fundamental principles of mathematics, computing sciences, and relevant domain disciplines	√

* A: Assignment; Q: Quiz; S: Seasonal; P: Project; F: Final

Weeks	Topics	CLO's	Assessments
1	Definition of Statistics, Data/Variable, Its Types (Quantitative and Qualitative Variables, Discrete and Continuous Variables). Big Data and Technical Data. Difference Between Sample and Population, Difference Between Parameters and Statistics, Type of Statistics (Descriptive and inferential), Types of data collection (Primary Data and Secondary Data).	1	A1: Application of Statistics in ———
2	Exploratory Data Analysis (EDA) as a Data Scientist, Presentation of data. Tabular representation of the data. Introduction to graphical Representation, Simple Bar Chart, Multiple Bar Chart, Component bar Chart, Pie Diagrams, Histogram (Shape of the data: Symmetric and skewed), Examples and Applications	1	Q1:
3	Measures of Central Tendency. Arithmetic Mean, Median, Mode, Quantiles, and Weighted Arithmetic Mean Examples and Application, Introduction to R language.	1	A2: Quartile, Percentile, Decile
4	Measures of Dispersion, Range, Variance and Standard Deviation, Interquartile range, Box Plot. Absolute and Relative Measures of Dispersion. Qualitative-wise Quantitative Analysis, Comments on the use of EDA. Examples and Application. Statistical Analysis and EDA of real data using R	1	
5	Introduction to probability, Importance in Artificial Intelligence, Objective and subjective approaches of probability. Random experiments, Sample space, Events, Simple and composite events. Mutually Exclusive Events, Independent and Dependent Events. Examples and Application.	2	Mid 1
6	Possibilities using Counting Techniques. Examples and Application, and its use in finding Probabilities.	2	A3: Bayes and Naïve Bayes
7	Conditional probability, Bayes' theorem, and Naïve Bayes Classifier, Example and Applications	2	Q2
8	Random variable, Discrete random variable, and its probability distribution, Continuous random variable and its probability density function.	2	
9	Binomial probability distribution and its Properties, Applications, and real-life Problems.	2	
10	Poisson Distribution, Properties, Applications, and real-life Problems. Hypergeometric distribution, Properties, Applications, and real-life Problems.	2	Q3
11	Uniform Distribution, Properties, Applications, and real-life problems. Normal distribution, Properties, and Estimation of Area under the normal curve using tables. Example and real-life Problems	2	Q4
12	Problems related to the Inverse use of the Standard Normal distribution using the table. Use of the R language for solving probability and Probability distribution problems	2	Mid 2 Project
13	Bivariate Data Analysis and its importance. Measuring Association among data. Covariance and Correlation, Application and real-life Examples. Introduction to Applied Data Modelling. Linear Regression Models using the OLS Method, Applications, and examples.	3	A4: Regression and OLS

Weeks	Topics	CLO's	Assessments
14	Prediction of real-life Problems. Matrix method for solving multiple linear regression models. Introduction to Generalized Linear Models.	3	Q5
15	Advances in Statistical Modelling, Regression, and Classification Problems in Machine Learning using R, Understanding the R output for modeling. Statistical Inference of the Models and Parameter Estimation	3	
16	Confidence Intervals and Testing of Hypotheses, including mean testing. Application of testing of hypotheses using R. Concept of P-values. The importance of data-driven decision-making.	3	Final

Programming Language for Course and Assignments (ifany)	R language			
Class Time Spent (in percentage)	**Theory**	**Problem Analysis**	**Solution Design**	**Social and Ethical Issues**
	35	25	35	5

Instructor(s) Name:

Dr. Haris Khurram

Instructor Signature:

_____

Date: 29/Jan/2025