Nilkamal School of Mathematics, Applied Statistics & Analytics

B.Sc (Data Science)

- **Program Educational Objectives (PEOs)**
- **Program Outcomes (POs)**
- **Course Outcomes (COs)**

Program Educational Objectives (PEOs):

1. **Professional Skills**
2. **Career Growth**
3. **Higher Studies**

Program Outcomes (POs):

PO-1: Apply the knowledge of Mathematics, Statistics and Computer Science fundamentals and specialization to solution of complex scientific problems (Scientific Knowledge)

PO-2: Identify, formulate, research literature, and analyze complex scientific problems reaching substantiated conclusions using principles of Data Sciences. (Problem analysis)

PO-3: Design of solutions for complex Data Analytical problems and design of system components or processes that meet the specified needs with appropriate considerations of public health and safety, and cultural, societal, and environmental considerations (Design/development of solutions)

PO-4: Use research-based methods including design of experiments, analysis and interpretation of data and Mathematical/Statistical knowledge leading to logical conclusions (Conduct investigations of complex problems)

PO-5: Create, select, and apply appropriate techniques, resources, and modern scientific and IT tools including analysis, prediction and modeling complex scientific activities with an understanding of limitations (Modern tool usage)

PO-6: Apply reasoning within the contextual knowledge to access societal, health, safety, legal, and cultural issues and the con-sequent responsibilities relevant to the professional Mathematical, Statistical and Computer Sciences practice (The Mathematician and society)

PO-7: Apply ethical principles and commit to professional ethics and responsibilities and norms of scientific practice (Ethics)

PO-8: Function effectively as an individual independently and as a member or leader in diverse teams, and in multidisciplinary settings (Individual and team work)

PO-9: Communicate effectively on complex scientific activities with the science community and with society at large such give and receive clear instructions (Communication)

PO-10: Demonstrate knowledge and understanding of scientific management principles and apply those to one's own work as a member and leader of a team to manage projects in multidisciplinary environments (Project management and finance)

PO-11: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change (Life-long Learning).

Courses and Course Outcomes (COs):

Descriptive Statistics – I

- **CO-1**: To prepare the data and select appropriate methods to represent data graphically and derive the basic descriptive statistics of the data.

Introduction to Probability Theory

- **CO-1**: To understand basic elements of probability theory and apply them to solve real life problems.

Univariate Calculus

- **CO-1**: Understanding of mathematical calculus through visualizations. Will have strong foundations for theory of probability and Statistics

Elementary Number Theory

- **CO-1**: Students will be exposed to numbers and will be introduced with its application in a modern field like Cryptography which is essential for any data science programs.

Discrete Mathematics

- **CO-1**: To study or describe objects or problems in computer algorithms and programming languages

Foundations of Computer Science

- **CO-1**: Understand the basic concepts of computer systems, types of hardware and software.
- **CO-2**: Understand basic concepts of networks, operating systems and data representation.
- **CO-3**: Use MS-Office application to create documents, ppt. presentations and data management.

Introduction to R

- **CO-1**: After completion of the course, students would be able to understand and write functions in R. Perform basic data analysis in R.

Environmental Studies

- **CO-1**: Understanding of the working of ecosystem.
- **CO-2**: Exploring the impacts of natural, manmade and legislative events on the natural world and its inhabitants

Descriptive Statistics – II

- **CO-1**: Identify the nature of the given variable.
- **CO-2**: Understand and apply appropriate measure of association between two variables using a relevant software.
- **CO-3**: Fit and interpret the simple linear regression model given the dataset.
- **CO-4**: Understand the role of various Index Numbers in real life.

Probability Models for Discrete Data

- **CO-1**: Understand the nature of various discrete distributions
- **CO-2**: Identify and fit the appropriate discrete distribution given the discrete variable

Probability Models for Continuous Data

- **CO-1**: Understand the nature of various continuous distributions
- **CO-2**: Identify and fit the appropriate continuous distribution given the continuous variable

Linear Algebra

- **CO-1**: Understanding of the various fundamental concepts of Linear algebra.
- **CO-2**: Understanding its wide applications in Statistics and Analytics.
- **CO-3**: Solving theoretical and application-based problems

Numerical Methods

- **CO-1**: To solve transcendental equations, polynomial approximations and integrations by different techniques.
- **CO-2**: Solving the problems by writing the programs on the relevant software

Introduction to Programming

- **CO-1**: After completion of the course, students would be able to implement basic programs using python programming language

Effective Communication

- **CO-1**: Understand various ways of communication
- **CO-2**: Illustrate professional requirements through communication proficiency

Sampling Distributions & Applications

- **CO-1**: Carryout test of significance of various hypotheses
- **CO-2**: Understand the need for sampling distributions various statistics
- **CO-3**: Apply statistical tests of hypotheses based on sampling distributions

Statistical Inference for Data Science – I

- **CO-1**: Identify appropriate population probability model and its parameter
- **CO-2**: Chose an estimator according to the criteria of good estimators
- **CO-3**: Understand different methods of estimation, both frequentist and Bayesian and their properties
- **CO-4:** Use the confidence interval approach to infer about population parameters

Multivariate Calculus

- **CO-1**: Understanding of the mathematical concepts of limit, continuity in higher dimensions with emphasis on applications.

Data Management

- **CO-1**: Design and draw ER and EER diagram for the real-life problem.
- **CO-2**: Create database through Relational Database concepts and retrieving the data.
- **CO-3**: Use different Information retrieval models to describe a problem and select the modelling approach that is best suited to a problem

Research Writing

- **CO-1**: Be familiar with and fully comply with art of research writing
- **CO-2**: Articulate essentials of a good research writing

Research Initiative in Data Science – I

- **CO-1**: Analyse the data by applying concepts and techniques learnt in various courses in this semester
- **CO-2**: Evince interest to read journal articles and pose research problems related to the courses studied.

Statistics Lab – I

- **CO-1**: Obtain the point estimates of parameters of standard distributions given the random samples
- **CO-2**: Use statistical software to run the data analysis related to tests of significance and tests of hypotheses
- **CO-3:** Interpret the results of the tests using p-value

Technology Lab – I

- **CO-1**: To design and draw ER and EER diagram for the real life problem.
- **CO-2**: To create database through Relational Database concepts and retrieving the data. The SQL queries are implemented using Oracle.

Mathematics Lab – I

- **CO-1**: Understanding of the mathematical concepts of limit, continuity in higher dimensions with emphasis on applications.

Data Analysis using Python

- **CO-1**: Use NumPy library for creating an array and then reshaping, slicing and indexing a NumPy array
- **CO-2**: Use NumPy for carrying arithmetic and Linear algebra operations.
- **CO-3**: Use appropriate scaling method for the given variables

Theory of Optimization & Graph Theory

- **CO-1**: Understanding of the mathematical concepts of Linear Programming and Graph Theory with emphasis on applications.

Statistical Inference for Data Science – II

- **CO-1**: Understand concepts of Statistical hypothesis, developing tests to test the hypotheses.
- **CO-2**: Formulate Statistical hypothesis in real life situations. Apply appropriate test to validate the hypothesis

Regression Analysis

- **CO-1**: Formulate simple and multiple regression models;
- **CO-2**: Carry out tests of linear hypothesis;
- **CO-3**: Perform validation of a regression model;
- **CO-4**: Select the important explanatory variables

Designs of Experiments

- **CO-1**: Understand planning and conducting the experiment and analyzing the data collected through the experiment

Machine Learning - I

- **CO-1**: Analyze and identify significant characteristics of data sets.
- **CO-2**: Develop an understanding of training a learning algorithm including over-fitting, noise, convergence and stopping criteria.
- **CO-3**: Match a data set with the most promising inductive learning algorithms.
- **CO-4**: Understand and implement the training, testing, and validation phases of learning algorithms development and deployment.

- **CO-5:** Determine the computational complexity associated with development and execution of learning algorithms for a given data set.

Research Ethics

- **CO-1**: Understand meaning and scope of Research and Publication Ethics.

Research Initiative in Data Science - II

- **CO-1**: Analyse the data by applying concepts and techniques learnt in various courses in this semester
- **CO-2**: Evince interest to read journal articles and pose research problems related to the courses studied.

Statistics Lab - II

- **CO-1**: Understand and analyse the design of experiments models
- **CO-2**: Use appropriate parametric and non-parametric tests for the analysis of a given hypotheses and data
- **CO-3**: Build a regression model with several variables and check the adequacy of the model.

Technology Lab - II

- **CO-1**: Implement basic models of supervised machine leaning using python programming language

Data Wrangling with Python

- **CO-1**: Identify the data types of various variables present in the dataset
- **CO-2**: Efficiently handle missing data
- **CO-3**: Detect and treat outliers
- **CO-4**: Use appropriate scaling method for the given variables
- **CO-5**: Identify and handle Imbalanced datasets

Mathematics Lab - II

- **CO-1**: Understanding of the mathematical concepts of Linear Programming and Graph Theory with emphasis on applications

Operations Research

- **CO-1**: Identify an appropriate model to be used in the real-life situation.
- **CO-2**: Analyse the data and interpret the results based on the model.

Multivariate Analysis

- **CO-1**: prepare the data for analysis.
- **CO-2**: select appropriate methods of multivariate data analysis for given study objectives.
- **CO-3**: interpret the results of multivariate analysis.

Differential Equations

- **CO-1**: Solve first order differential equations utilizing the standard techniques.
- **CO-2**: Analyse any mathematical model which is reduced to ordinary differential equations.

Machine Learning - II

- **CO-1**: Recognize the characteristics of machine learning that make it useful to real-world problems.
- **CO-2**: Characterize machine learning algorithms as supervised and unsupervised.
- **CO-3**: Identify frequent Item-sets and derive association rules.
- **CO-4**: Build recommendation systems.
- **CO-5**: Understand text mining.

Research Initiative in Data Science - III

- **CO-1**: Use multivariate techniques to analyse real life data
- **CO-2**: Formulate and solve optimization problems using OR techniques
- **CO-3**: Handle big data and use ML and Big data analytics techniques

Professional Skills

- **CO-1**: Identify and use appropriate words for communication, choose proper tools to communicate, use positive body language while communicating, maintain proper eye contact to build trust and confidence.

Statistics Lab - III

- **CO-1**: Prepare the data for multivariate analysis/Collect data for operations research techniques.
- **CO-2**: Choose appropriate technique/model for given dataset and perform data analysis using software.
- **CO-3**: Interpret the results to achieve the objective.

Technology Lab - III

- **CO-1**: After completion of the course, students would be able to implement basic models of unsupervised, reinforcement machine

Mathematics Lab - III

- **CO-1**: Solve first order differential equations using various standard methods.
- **CO-2**: Interpret the results of a mathematical model in first order differential equations.

Big Data Analytics

- **CO-1**: Demonstrate capability to use Big Data Frameworks like Hadoop
- **CO-2**: Using tools like pig, NO SQL and MongoDB for Big data Applications
- **CO-3**: Construct scalable algorithms for large Datasets using Map Reduce techniques
- **CO-4**: Develop a BDA application for real life problem

Capstone Project

- **CO-1**: Demonstrate ability to carry out a data science project from end to end.
- **CO-2**: Demonstrate proficiency in preparation and walk through of a presentation.
- **CO-3**: Demonstrate ability to carry out a literature search and summarize the state of the art.
- **CO-4**: Demonstrate ability to translate the project objects into a realistic work plan that draws on multiple people.

Markov Chains

- **CO-1**: Model any given situation as a Markov chain.
- **CO-2**: Obtain transition probability matrix and classify the states.
- **CO-3**: Make predictions using Markov chain approach.

Time Series & Forecasting

- **CO-1**: Decompose a time series and estimate its component.
- **CO-2**: Fit an appropriate forecasting model on the given data.

Statistical Process Control

- **CO-1**: Use the Lean – Six – Sigma tools conveniently and effectively including statistical techniques.
- **CO-2**: Handle data in the organization more efficiently and with disciplined culture in executing the responsibilities.
- **CO-3**: Blend Lean Six-sigma with Data Science to create a culture of continuous process improvement.

Deep Learning Techniques

- **CO-1**: Learn to design neural network architectures and training procedures
- **CO-2**: Understand the data needs of deep learning
- **CO-3**: Explain strengths and challenges of deep learning.

Entrepreneurship Skills

- **CO-1**: Approach Entrepreneurship with clarity and focus.
- **CO-2**: Apply the various tools and critical-thinking skills needed for business success.

Statistics Lab – IV

- **CO-1**: Analyse real life situations using Markov Chains.

Data Visualization and Modelling

- **CO-1**: Understand and describe the main concepts of data visualization

Technology Lab – IV

- **CO-1**: After completion of the course, students would be able to implement basic models of deep