

FYUGP Data Science

Course Outline for Semester III

S. No.	Category	Course Code	Course Title	Credits	L	T	P	S	Hours per week
1	Major	DS200MJ	Probability Theory	4	4	0	9	9	4
		DS201MJ	Concepts of Programming using C and C++	4	3	0	2	0	5
2	Minor 3	MTHS203MN	Applied Linear Algebra	4	3	0	2	0	5
4	Multidisciplinary		Student to choose	3	3	0	0	0	3
5	Ability Enhancement			3	3	0	0	0	3
6	Skill Enhancement		Student to choose	2	3	0	0	0	3

FYUGP Data Science

Course Title: Introduction to Probability Theory	L	T	P	S	Semester: 3 rd
Course Code: DS200MJ	4	x	x	x	Max Marks: 100
Credits: 4					

Course Objectives: This is a theoretically rooted introductory course on fundamental concepts in probability. The objective is to lay a foundation for future courses in probability, statistics, and data science.

Course Outcomes: After successful completion of this course, student will be able to:

1. Apply foundational concepts of probability including classical, relative frequency, and axiomatic definitions; compute probabilities using laws of total probability, Bayes' theorem, and assess independence of events.
2. Differentiate between discrete and continuous random variables, interpret their distribution functions, and analyze transformations and functions of random variables.
3. Compute and interpret expectations, variances, higher moments, moment generating functions (MGFs), and characteristic functions for random variables.
4. Analyze joint, marginal, and conditional distributions of multiple random variables/vectors; calculate conditional expectations and covariance, in multivariate settings.
5. Identify and apply appropriate discrete and continuous probability distributions in modeling and solving real-world problems.
6. Explain and apply probability inequalities, and concepts of convergence; apply major theorems (Law of Large Numbers, Central Limit Theorem) and tools such as the Delta Method and Continuous Mapping Theorem in probabilistic analysis.

Unit I: Basic definitions and ideas such as random experiment, sample space and event, Classical definition and relative frequency definition of probability, Axiomatic definition of probability, Elementary properties of probability function, Probability inequalities such as Boole's inequality and Bonferroni inequality.

Conditional probability and its basic properties, Examples of conditional probability and multiplication law, Theorem of total probability and related examples, Bayes theorem and related examples, Independent events.

Unit II: Random variables and their distribution function, Induced probability space, Discrete and continuous random variables, Function of random variables (Discrete and Continuous), Expectation and moments of random variables, MGF of random variables and its application, Characteristics function and its application.

Random vectors and their joint distribution functions, Marginal distribution, independent random variables, Conditional distribution of random vectors/variables, Expectation and moments of random vectors, Conditional Expectation, variance and covariance and their applications

Unit III: Standard Discrete Distributions: Bernoulli, Binomial, Poisson, Geometric, Hypergeometric, Negative Binomial, Multinomial distributions- definition, properties and applications.

FYUGP Data Science

Standard Continuous Distributions: Uniform, Normal, Exponential, Gamma, Beta, Weibull, Cauchy and Laplace distributions, lognormal, -definition, properties and applications.

Unit IV: Markov, Chebyshev and Jensen's inequality. Idea of limiting distribution, Convergence in distribution and probability, and related results, Convergence of moments and almost sure convergence, Various examples and counter examples. Weak law of large numbers, Central limit theorem, Applications, e.g., continuous mapping theorem and delta method.

Text Books /References:

1. Hogg, R. V., McKean, J., Craig, A. T. Introduction to Mathematical Statistics. Pearson Education.
2. Hoel, P. G., Port, S. C., Stone, C. J. Introduction to Probability Theory, Houghton Mifflin, Boston, MA.
3. Mood A. M. , Graybill R. A. and Boes D. C., Introduction to the theory of Statistics, Tata McGraw Hill
4. Ross Sheldon M., A first Course in Probability, Pearson.
5. Rohtagi, V.K. and Md. Ehsanes Saleh A. K., An Introduction to Probability and Statistics, John Wiley & Sons.

FYUGP Data Science

Course Title: Concepts of Programming using C and C++	L	T	P	S	Semester: 3 rd
Course Code: DS201MJ	3	x	2	x	Max Marks: 100
Credits: 4					

Course Objective: This course introduces the fundamentals of C and C++ programming, emphasizing structured and object-oriented programming concepts, control structures, functions, arrays, pointers, and data handling through classes and objects.

Course Outcomes: After successful completion of this course, student will be able to:

1. Understand the fundamental concepts of C programming and develop structured programs.
2. Gain proficiency in operators, control structures, and modular programming with functions.
3. Learn the use of arrays, pointers, and structures to manage data effectively.
4. Acquire object-oriented programming skills using C++ (classes, objects, constructors, destructors).

Unit I –Basic structure of a C program: main() function, header files, Compilation, and Execution of a C program. Data types, Variables, Constants, Input/Output operations.

Unit II –: Operators and Expressions: Arithmetic, Relational, Logical, Assignment, Increment/Decrement, Introduction to Bitwise Operators. Control Flow: Conditional Statements – if, else if, else, switch. Loops – for, while, do-while. Break and Continue statements. Functions: Prototype, Definition, and Calling.

Unit III – Arrays: One-dimensional arrays – declaration, initialization, access. Two-dimensional arrays. Pointers (Basics): Concept of memory addresses, Pointer declaration, initialization, and dereferencing. User-defined types: Structures and Unions

Unit IV –Basic C++ Program Structure: Classes, Objects, Scope resolution operator (::). Constructors and Destructors: Default, Parameterized, Copy constructor, Constructor overloading, Destructor.

Text Books / Reference Books

1. Balagurusamy, E. (2019). Programming in ANSI C. McGraw-Hill.
2. Lafore, R. (2018). Object-Oriented Programming in C++. Pearson Education.
3. Kernighan, B. W., & Ritchie, D. M. (1988). The C Programming Language. Prentice Hall
4. Schildt, H. (2017). C++: The Complete Reference. McGraw-Hill.
5. Kanetkar, Y. (2018). Let Us C. BPB Publications.

Web References

1. NPTEL – Programming in C and C++ by Prof. Deepak Gupta.
2. GeeksforGeeks – C and C++ Programming Tutorials.
3. Harvard University CS50 – Introduction to Computer Science (Modules on C).

FYUGP Data Science

Course Title: Applied Linear Algebra	L	T	P	S	Semester: 3 rd
Course Code: MTHS203MN	3	x	2	x	Max Marks: 100
Credits: 4					

Course Objectives: The course aims to introduce the fundamental concepts of linear algebra and develop problem solving skills through computational techniques using Python programming, enabling students to apply linear algebraic methods effectively in the fields of data science, artificial intelligence and machine learning.

Course Outcomes: After completing the course students will be able to

1. Perform matrix operations and applications using Python libraries for computational and visualization purposes.
2. Understand and apply the concepts of vector spaces, basis, dimension and rank of matrices in linear algebra.
3. Analyze and solve systems of linear equations using analytical and computational methods in Python and interpret and visualize their solutions.
4. Understand and compute linear transformations, eigenvalues and eigenvectors and understand the role of linear algebra in machine learning and AI models.

Prerequisites: An integrated computational lab for hands-on practice and students should also have basic Python programming skills.

Unit I: Creating matrices using Python libraries: NumPy, SymPy and Pandas, performing matrix arithmetic operations: addition, subtraction, scalar multiplication and matrix multiplication using python libraries, Creation of special matrices using Numpy Library: Symmetric, Skew-Symmetric, Hermitian and Skew-Hermitian matrices, creation of upper triangular and lower triangular matrices, finding determinant of a matrix using `linalg.det()` function and inverse of a matrix using `numpy.linalg.inv()` function, demonstrating applications of matrices through Python-based computations and visualizations.

Unit II: Introduction to vector spaces, examples of vector spaces, linear combination, spanning sets, subspaces with examples, linearly dependent and independent sets of vectors, basis and dimension of a vector space, dimension and subspaces (basic relationship), application to matrices, rank of a matrix.

Unit III: Review of systems of linear equations: homogeneous and non-homogeneous systems, Matrix representation of systems: $AX=B$, augmented matrix, row echelon form. solution, Gaussian elimination method, Cramer's Rule (for small systems), existence and uniqueness of solutions: consistent vs. inconsistent systems, role of rank of a matrix in solvability (using `numpy.linalg.matrix_rank`), computational implementation using Python, Solving systems using `numpy.linalg.solve`, `scipy.linalg` and symbolic solving with `sympy`, Interpreting solutions: unique, infinite or no solution cases, visualizing 2D and 3D systems of equations using `matplotlib.pyplot`.

FYUGP Data Science

Unit IV: Definition of linear mappings (linear transformations) and examples, kernel and image of a linear mapping, rank and nullity of a linear mapping with examples. Eigen values and eigen vectors of a matrix, trace and determinant of a matrix expressed in terms of eigenvalues, computation of eigen values and eigenvectors of a matrix in Python, Matrix diagonalization and its computation using Sympy Library.

Textbooks/ References

1. Gilbert Strang, Linear Algebra and its Applications, Fourth Edition, Cengage India Private Limited by, 2005.
2. S. Lipschutz & M. Lipson, Linear Algebra, Schaum's outline series, Tata McGraw-Hill, 4th Edition 2009.
3. M. Tsukada, Y. Kobayashi, H. Kaneko, Sin-E. Takahasi, K. Shirayanagi and M.Noguchi, Linear Algebra with Python-Theory and Applications, Springer 2023