INSTITUTE OF MATHEMATICS AND APPLICATIONS

COURSES OF STUDY

M.A./M.Sc. IN MATHEMATICS WITH DATA SCIENCE

(w.e.f. 2024-25)

UTKAL UNIVERSITY

BHUBANESWAR-751004

REVISED COURSE STRUCTURE FOR
M.A./M.Sc. IN MATHEMATICS WITH DATA SCIENCE

In C (T-P-Tu); C stands for number of credits for the paper, T, P and Tu, respectively stand for number of theory, practical and tutorial classes per week.

SEMESTER-I

Paper No.	Course Title	Category	Marks	C (T-P-Tu)
MDS 101	Analysis	Core	100	4 (3-0-1)
MDS 102	Computational Linear Algebra	Core	100	4 (3-2-0)
MDS 103	Probability, Statistics, and Stochastic Processes with Lab (in R)	Core	100	4 (3-2-0)
MDS 104	Programming (Python and C++)	Core	100	4 (3-2-0)
MDS 105	Discrete Mathematics and Graph Theory	Core	100	4 (3-0-1)
MDS 106	Numerical Optimization with Lab	Core	100	4 (3-2-0)

SEMESTER-II

Paper No.	Course Title	Category	Marks	Credits
MDS 201	Abstract Algebra	Core	100	4 (3-0-1)
MDS 202	Topology	Core	100	4 (3-0-1)
MDS 203	Data Structures and Algorithms with Lab (in C++)	Core	100	4 (3-2-0)
MDS 204	Multivariate Data Exploration and Analysis with Lab (in R)	Core	100	4 (3-2-0)
MDS 205	Machine Learning with Lab (in Python)	Core	100	4 (3-2-0)
MDS 206	Computer Systems for Data Science	Core	100	4 (3-2-0)

------------------------------- SUMMER *INTERNSHIP*---------------------------

SEMESTER-III

Paper No.	Course Title	Category	Marks	C (T-P-Tu)
MDS 301	Functional Analysis	Core	100	4 (3-0-1)
MDS 302	Advanced Machine Learning with Lab	Core	100	4 (3-2-0)
MDS 303	Complex Analysis	Core	100	4 (3-0-1)
MDS 304	Differential and Integral Equations	Core	100	4 (3-0-1)
MDS 305	Numerical Analysis with Lab	Core	100	4 (3-2-0)
MDS 306	Data Visualization with Lab + Summer Internship Evaluation	Core	100	3 (2-2-0) + 1

SEMESTER-IV

Paper No.	Course Title	Category	Marks	C (T-P-Tu)
MDS 401	Deep Learning & Reinforcement Learning with Lab	Core	100	4 (3-2-0)
MDS 402	Big Data Analytics with Lab	Core	100	4 (3-2-0)
MDS 403	Artificial Intelligence	Core	100	4 (3-2-0)
MDS 404	Elective I (To be chosen from a Group A)	Allied Elective	100	4 (3-2-0)/ (3-0-1)
MDS 405	Elective II (To be chosen from a Group B)	Allied Elective	100	4 (3-2-0)/ (3-0-1)
MDS 406	Dissertation with Viva-voce	Core	100	4

For Allied Elective papers a student can choose any two of the following courses from each group depending on the availability

Elective Papers Group-A **MDS-404**		Elective Papers Group-B **MDS-405**	
A1	Bioinformatics	B1	Causal Inference for Data Science
A2	Computational Fluid Dynamics	B2	Cyber Security
A3	Computer Vision	B3	Discrete Dynamical Systems
A4	Computational Finance	B4	High Performance Computing
A5	Computational Modelling of Financial Derivatives	B5	Internet of Things with Lab
A6	Image & Video Processing	B6	Natural Language Processing
A7	Number Theory and Cryptography with Lab	B7	Quantum Computing
A8	Numerical Solution of Partial Differential Equations	B8	Soft Computing Methods
A9	Object Oriented Software Engineering	B9	Time Series Analysis and Forecasting
A10	Probabilistic Graphical Models	B10	Theory of Computation & Compiler Design

DETAILED SYLLABUS

Semester–I

MDS 101

Analysis	**(Marks -100)**	**Credit 4 (3-0-1)**

Unit-I: Review of Calculus, Metric spaces, limit in metric spaces, open sets, closed sets, relative metric, Continuous functions, homeomorphisms, the space of continuous functions, connected sets, totally bounded sets, complete metric spaces, fixed points, completions, compact metric spaces, uniform continuity, the Baire category theorem. **12 Hours**

Unit-II: Sequence of functions, pointwise and uniform convergence of functions, uniform convergence & continuity, uniform convergence & differentiability, uniform convergence & integrability , series of functions, power series, Dini's theorem. Equicontinuity, Arzela-Ascoli theorem, The Stone- Weierstrass theorem. **12 Hours**

Unit-III: Linear transformations, differentiation in R^n, chain rule, partial derivatives,directional derivatives, Inverse function theorem, Implicit function theorem. Fourier series, Trigonometric series, Dirichlet kernel, convergence of Fourier series, Parseval's theorem. **12 Hours**

Unit-IV: Outer measure and its properties, Lebesgue measurable sets and Lebesgue measure, a non-measurable set. Measurable functions, Littlewood's three principles, Egoroff's theorem, Lusin's theorem. **12 Hours**

Unit-V: Simple functions, Lebesgue integral of a bounded function over a set of finite measure, bounded convergence theorem, Lebesgue integral of nonnegative functions, monotone convergence theorem, Fatou's Lemma, the general Lebesgue integral, Lebesgue convergence theorem. L^p -spaces, Holder's inequality and Minkowski's inequality, Completeness of L^p -spaces. **12 Hours**

Text Books:

1. N.L. Carothers: Real Analysis, Cambridge University Press, United Kingdom.
2. W. Rudin, Principles of Mathematical Analysis, Tata McGraw Hill Publishing Company, New Delhi.
3. Elias M.Stein and Rami Shakarchi, Real Analysis: Measure Theory, Integration, and Hilbert Spaces,, Princeton Lectures in Analysis, Princeton University Press.

Reference Books:

4. Richard R. Goldberg: Methods of Real Analysis
5. W. Rudin, Real and Complex Analysis, Tata McGraw Hill Publishing Company, New Delhi.
6. H. L. Royden, Real Analysis, McMillan, New York.
7. P. Halmos, Measure Theory, Van Nostrand, Princeton, New Jersey.

MDS 102

Computational Linear Algebra (Marks-100) Credit 4 (3-2-0)

Unit-I: Matrix Multiplication: Basic algorithm and notations, Structure and efficiency, Block matrices and algorithms, Fast matrix - vector products, Vectorization and locality, Parallel matrix multiplication. Kronecker Product, Triangular systems, Diagonal Dominance and symmetry, Banded systems, Symmetric indefinite systems, Block tridiagonal systems, Vandermonde systems, Visualization of different matrix computations and their Applications in Data Science. **13 Hours**

Unit-II: Vector Space & Subspaces, Solving AX = 0 & AX = b, Linear Independence, Basis and Dimension, The Four Fundamental Subspaces, Linear Transformation, Orthogonal Vectors, Projections to a line, Projections and Least Squares, Orthogonal Bases & Gram-Schmidt. **9 Hours**

Unit-III: Eigenvalues and eigenvectors, Diagonalization of a Matrix, Complex Matrices, Similarity Transformation, Test for positive definiteness, Vector norms, Matrix norms, Singular Value Decomposition.

Errors in Computations, Computing Norm, Inner product and solution of Triangular System, Efficiency and stability of an Algorithm, Conditioning, Perturbation Analysis, Perturbation Analysis of linear system. **13 Hours**

Unit-IV: LU Factorization Methods, Scaling, Effects of the condition number on accuracy, computing and estimating the condition number, Parallel LU, Orthogonalization and Least Square, Householder’s matrices and QR factorization, Classical and Modified Gram-Schmidt Algorithm for QR factorization, Solution of AX = b using QR Factorization, Projections Using QR Factorization, SVD and its computation. **12 Hours**

Unit-V: Existence and uniqueness of least square solutions, Pseudoinverse and the least square problem, sensitivity of the least square problem, Computational Methods for Over determined Problems, Computing selected eigenvalues and eigenvectors, Jacobi, Gauss-Seidel and SOR methods.

Functions of Matrices: Eigen value methods, Approximation methods, The matrix exponential, The sign, Square root and Log of a matrix, Tensors. **13 Hours**

Texts:

1. Gene. H. Golub, Charles F. Van Loan, Matrix Computations, Fourth Edition, The Johns Hopkins University Press, Baltimore 2013.
2. G. Strang, Linear Algebra for Everyone, Wellesley-Cambridge Press, 2020.

Reference Books:

1. G. Strang, Linear Algebra and Its Applications, 4th Edition, Cengage Learning, 2006.
2. J. W. Demmel, Applied Numerical Linear Algebra, SIAM, 1997.
3. P. G. Ciarlet, Introduction to numerical linear algebra and optimization
4. William Ford, Numerical linear algebra with applications, AP
5. Richard Bronson, G.B. Costa, Matrix Methods: Applied Linear Algebra, Third Edition, AP.

MDS 103

Probability, Statistics, and Stochastic Processes with Lab (Marks-100) Credit 4 (3-2-0)

Unit – I: Samples and Probability spaces, Properties of probability measures, Conditional probability and Bayes theorem, Continuity of probability measure, Independence, Borel sigma algebras and random variables, Distribution function, Proability density function and Probability mass function of a random variable.

Unit – II: Independence of events and independence of random variables, Expectation and Variance, Discrete random variables (Bernoulli, Binomial, Poisson, Geometric, Hypergeometric), Continuous random variables (Uniform, Exponential, Normal, Gamma, Beta), Functions of random variables, Joint distributions of discrete, continuous and independent random variables, Conditional distributions, The moment generating function.

Unit – III: Distributions derived from the normal distribution (X^2, t, and F), The sample mean and the sample variance, Independence weak laws of large numbers, Borel-Cantelli lemma, Strong law of large numbers, Types of convergence of random variables, Characteristic functions, Central limit theorems.

Unit – IV: Parameter estimation (the method of moments, the method of maximum likelyhood, Confidence intervals, Bayesian approach to parameter estimation, Efficiency and sufficiency of an estimator.

Unit – V: Testing of hypothesis**,** The Neyman-Person Paradigm, The duality of confidence intervals and hypothesis test, Generalized likelihood ratio test, Composite hypothesis and the t-test, Two-sample t-test and signed rank test, Rank sum test and permutation tests, Testing multiple hypothesis, The analysis of variance(one-way-layout), chisquare test of homogeneity and independence.

Conditional expectation**,** Matringles and the optional stopping time theorem, Stochastic processes, Markov chains, Examples of Markov chains with countable state-space (random walk, Birth-and-death chain, M|G|1 queue), Markov property, Continuous time markov processes.

Books Recommended

Texts:
1. Rick Durret, Probability: Theory & Examples, Fourth Edition, Cambridge University press, 2013.
2. John A Rice, Mathematical Statistics and Data Analysis, Third Edition, CENGAGE Learning, 2007.

References:
1. Vijay K. Rohatgi, A.K.Md. Ehsaries Sabh, An Introduction to Probability and Statistics, Wiley-Interscience, Second Edition, 2001.
2. LA. Wasserman-All of Statistics-SpringerScience, 2004.

MDS 104

Programming (Python and C++) **(Marks-100)** **Credit 4(3-2-0)**

Unit I: Introduction to flowchart and algorithm, Basics of Pseudo code. Problem solving using Pseudo code,examples: check for prime number, factorial of a number, GCD of two numbers, count the number of digits of an integers, convert a decimal number to binary, Fibonacci numbers, Armstrong numbers, evaluating of e^x and sin(x) series using their Taylor series expansion, find minimum and maximum among N numbers, linear search. sum and product of of N numbers, evaluation of polynomial. **6 Hours + 6 Labs (2 Hours each)**

Unit II: Getting Started with Python and C++: Introduction, Basic data types, Control Structures: if, if-elif-else, for, while, break, continue. Data Structures: Lists, Tuples, Sets, and Dictionaries. **6 Hours + 6 Labs (2 Hours each)**

Unit III: Functions: Defining functions, Calling functions, Passing arguments, Keyword arguments, Default arguments, Variable-length arguments, Anonymous functions, Function returning values, Scope of the variables in a function - global & local variables, User defined functions.Modules & Packages: Creating modules, Import statement, from import statement, name spacing; Creating user defined packages. **6 Hours + 6 Labs (2 Hours each)**

Unit IV: File Handling: Handling of csv file.Object Oriented Programming: Features, classes and objects, creating class and object, Using a class & its methods; Exception Handling: Errors, Types of exception, try, except and finally, assertion. **6 Hours + 6 Labs (2 Hours each)**

Unit V: Numpy: Introduction, Creating of arrays and matrices;Introduction to Panda: Creating a data frame, Dealing with row & columns, Indexing & selection data, Working with missing data, Iterating over rows and columns; Merging and joining DataFrame objects, Concatenation, Reshaping DataFrame objects, Pivoting, Data transformation, permutation & sampling, Data aggregation and GroupBy operations; Creating data frame from CSV file; Matplotlib: Creating effective visual representations of your data, Avoiding common pitfalls. **6 Hours + 6 Labs (2 Hours each)**

Text Books:

1. Donal.E.Knuth, Art of Computer Programming,Volume-1, Third Edition, 1997

2. R. N. Rao, Core Python Programming, 2nd Edition, Dreamtech Press, 2018.

3. J. V. Guttag, Introduction to Computation and Programming Using Python, with Application to Understanding Data, 2nd Edition, PHI Learning, 2016.

4. W. McKinney, Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Ipython, 2nd Edition, O'Reilly Media, 2017.

Reference Books:

1. P. Barry, Head First Python, 2nd Edition, O'Reilly Media, 2010.

2. A. Downey, Think Python : How to Think Like a Computer Scientist, 2nd Edition, Green Tea Press, 2015.

3. J. Zelle, Python Programming : An Introduction To Computer Science, 3rd Edition, Franklin, Beedle & Associates, 2016.

4. Brian W. Kernighan and Dennis M. Ritchie: The C Programming Language.

5. Bruce Eckel: Thinking in C++

MDS 105

Discrete Mathematics and Graph Theory (Marks-100) Credit 4 (3-0-1)

Unit – I: Principles of addition and multiplication, Permutations and Combinations, Counting with the help of mathematical induction, inclusion-exclusion principle and pigeon hole principle.

Unit – II: Recurrence relations, counting using recurrence relations, Solving linear homogeneous recurrence relations with constant coefficients. Stirlings Numbers Generating functions, solving recurrence relations with constant coefficients. solving recurrences using generating functions,

Unit – III: Partitions, partiality ordered sets, Lattices, Boolean algebra, Propositional and quantified logic.

Unit – IV: Graphs: Basic definitions, complement of graph, clique, independent set, bipartite graph. Graph isomorphism, subgraph, induced subgraph, path, cycle, walk, Petersen graph, connected component. Degree sum, adjacency matrix and number of walks, shortest path in a weighted graph.

Unit – V: Bipartite graphs and odd cycles. Strongly connected component, Eulerian trail in a graph. Edge subdivision, Planarity, K(3,3) and K(5). Kuratowski's theorem (statement only), Euler's theorem, chromatic number, five colour theorem.

Recommended Texts:

1. J. Matousek and J. Nesetril: *Invitation to Discrete Mathematics*, Clarendon press, Oxford.
2. Donald E. Knuth, Ronald L.Graham and O. Patashnik: *Concrete Mathematics*, Pearson Education.
3. C. L. Liu: *Elements of Discrete Mathematics*, Tata McGraw Hill.
4. Frank Harry: *Graph Theory*, Narosa Publishing House.
5. Reinhard Diestel: *Graph Theory*, Springer.
6. Douglas B.West: *Introduction to graph Theory*, Prentice Hall, India.
7. Martin J. Ericson: *Introduction to Combinatorics*, Wiley Interscience.
8. Rosen, *Discrete Mathematics and Its Applications*, Tata McGraw Hill.
9. G. A. Bondy & U. S. R. Murty: *Graph Theory*, Springer.
10. R. A. Brualdi, Introductory Combinatories, Pearson.
11. V. Krishnamurthy; Combinatorics: Theory and Applications.

MDS 106

Numerical Optimization with Lab (Marks-100) **Credit 4 (3-2-0)**

Unit-I: One Dimensional Optimization: Function comparison methods, Polynomial interpolation, Iterative methods.

Unit-II: Gradient Based Optimization Methods (I): Calculus on R^n , Method of steepest descend, Stochastic gradient descent, Conjugate gradient method, Conjugate gradient for non quadratic objective function, The generalized reduced gradient method, Gradient projection method.

Unit-III: Gradient Based Optimization Methods (II): Newton type methods(Newton's method, Marquardt's method), The Quasi-Newton method, BFGS, L-BFGS.

UNIT-IV: Convex sets, convex functions, First and second order conditions for convexity, Jensen`s inequality, Convex optimization problems, Characterization of optimal solutions, Necessary optimal conditions for constrained optimization problems(Fritz John and KKT), Lagrange multiplier method.

Unit-V: Sufficient conditions for optimality, KKT sufficient conditions, Constraint qualification, Second order optimality conditions, KKT second order optimality necessary conditions, Penalty function techniques, Primal dual method.

Books Recommended

Texts

1. M. C. Joshi, K. M. Moudgalya, Optimization: Theory and Practice, Narosa Publishing House, 2004.

2. Stephen Boyd, Lieven Vandenberghe, Convex Optimization, Cambridge University Press

References

1. J. A. Snyman, Practical Mathematical Optimization, Springer Sciences, 2005.

2. Nocedal J, Wright S.J., Numerical Optimization, Springer.

3. Convex Optimization Overview, Lecture notes by Zico Kotler, Stanford University.

4. Optimality Conditions for Constrained Optimization Problems, Lecture notes by Robert M. Freund, MIT.

5. Vivek S Borkar: Elementary Convexity with Optimization, Hindustan Book Agency.

Semester-II

MDS 201

Abstract Algebra	(Marks-100)	Credit 4(3-0-1)

Unit-I: Groups: Review of normal subgroups, quotient groups and homomorhism theorems. Group actions with examples, class equations and their applications, Sylow's Theorems and their applications; groups and symmetry. Direct sum and free Abelian groups. **12 Hours**

Unit-II: Commutative rings with unity: examples, ring homomorphisms, ideals, quotients, isomorphism theorems. Prime and maximal ideals, Zorn's Lemma and existence of maximal ideals.Product of rings, ideals in a finite product, Chinese Remainder Theorem. Prime and maximal ideals in a quotient ring and a finite product. **12 Hours**

Unit-III: Field of fractions of an integral domain. Irreducible and prime elements; PID and UFD. Polynomial rings over UFD's, Criteria for irreducibility of polynomials over UFD's. **12 Hours**

Unit-IV: Modules: Basic definitions and examples, Submodules and Direct sums, Quotient modules, Homomorphism and Isomorphism theorems, Cyclic modules, Free modules. **12 Hours**

Unit-V: Fields: Fields and their extensions; algebraic and transcendental extensions; algebraic closure; splitting fields and normal extensions; separable, inseparable and purely inseparable extensions; finite fields. Galois extensions and Galois groups, Fundamental theorem of Galois theory, cyclic extensions, solvability by radicals, constructibility of regular n-gons, cyclotomic extensions. **12 Hours**

Text Book:

1. D. S. Summit and R. M. Foote: Abstract Algebra, John Wiley.

Reference Books:

2. P. B. Bhattacharya, S. K. Jain, and S. R. Naipaul: Basic Abstract Algebra, Cambridge University Press.

3. N. Jacoboson: Basic Algebra Vol. I, W.H. Freeman and Co, 1985.

4. I. N. Herstein: Topics in Algebra", John Wiley & Sons (2nd Edition) 1999.

5. J. J. Rotman: An Introduction to the theory of groups, GTM (148), Springer-Verlag, 2002.

6. N. S. Gopalakrishnan: University Algebra, Wiley Eastern, 1986.

7. J. Rotman: Galois theory, Springer-Verlag, 1998.

8. S. Roman: Field Theory, Springer-Verlag, 1995.

9. M. Artin: Algebra, 2nd Edition, Pearson(Indian Edition).

MDS 202

Topology **(Marks-100)** **Credit 4(3-0-1)**

Unit-I: Infinite sets and Axiom of Choice, Well-ordered sets, Well-ordering theorem, Topological Spaces Basis and sub-basis for a topology, Order and product topologies, Closed sets, limit and interior points, Continuous functions. **14 hours**

Unit-II: The metric topology, Connected spaces, Connected subspaces of the real line, Components and local connectedness, Arc wise connectedness, Arc wise connectedness in Euclidean spaces.

10 hours

Unit-III: Compact spaces, Compact subspaces of the real line, Compactness and finite intersection properties, Compactness in metric spaces, Limit point compactness, Sequential compactness and its equivalence in metric spaces,Local compactness and One point compactification. **12 hours**

Unit-IV: Separation and Countability axioms, Topological spaces based on these axioms such as T_0, T_1, T_2 etc., Lindeloff space, Separable space and Normal spaces and their basic properties. **12 hours**

Unit-V: The Uryshon Lemma, The Uryshonmetrization theorem, The Tychonoff theorem, Completely regular spaces. **12 hours**

Prescribed Text Book:

1. James R. Munkres: Topology, Second Edition, Prentice-Hall.

   Relevent topics prescribed above from chapters 1,2,3,4, and 5.

Reference Books:

2. W. J. Pervin: Foundations of General Topology, Academic Press 1964.
3. J. Dugundji: Topology, UBS, 1999.
4. G. F. Simmons: Introduction to Topology and Modern Analysis, Tata McGraw-Hill, 1963.
5. M. A. Armstrong: Basic Topology, Springer, 1983.

6. G. E. Bredon: Topology and Geometry, Springer GTM 139, 1995.

MDS 203

Data Structures and Algorithms with Lab (in C++) **(Marks-100)** **Credit 4 (3-2-0)**

Unit I: The role of algorithms in computing, Introduction to design and analysis of algorithms, complexity of algorithms, growth of functions, asymptotic notations, Recurrences, Solving recurrences using Master Method, Introduction to data structures, classification of data structures, abstract data types, Arrays: introduction, representation of arrays (row and column major), basic operations on array (traverse, insert, delete, search), sparse matrix, representation of sparse matrix using triplet form, operation on sparse matrix (addition, transpose). **9 Hours**

Unit-II: Stack: stack model, representation using array, basic operations, and applications; Queue: queue model, representation using array, basic operations, circular queue; Linked List: introduction, types of linked list, representation in memory, operations on linked list (traverse, search, insert, delete), Representation of polynomial and its operations (addition, multiplication), implementation of stack and queue using linked list. **9 Hours**

Unit-III: Sorting Algorithms: Bubble sort, Selection sort, Insertion sort; Tree: terminology, representation, Binary tree: traversal algorithms, Binary search tree, Operations on Binary search tree, Height balanced tree; Divide and conquer strategy for designing algorithms, Merge Sort, Quick Sort; FFT, Strassen’s matrix multiplications, Heaps, Types of Heap, Maintaining the heap property, Building a Heap, The Heapsort algorithm, Lower bounds for sorting, counting sort, Priority Queue, Data structures for disjoint sets. **10 Hours**

Unit-IV: Dynamic Programming, Elements of dynamic programming, Matrix Chain Multiplication, Longest Common Subsequence; Greedy algorithms, Elements of Greedy strategy, Fractional Knapsack problem, Huffman codes; String matching algorithms (Naive, Rabin-Karp). **10 Hours**

Unit-V: Graphs: terminology, representation, graph traversal (BFS, DFS), Minimum spanning trees, Kruskal’s algorithm, Dijkstra’s algorithm, Warshall’s algorithm; Backtracking (8-queens problem, sum of subsets), Max flow, min cuts, Ford Fulkerson method, Introduction to NP completeness (Polynomial time, Polynomial time verification, NP completeness & reducibility), Examples of NP complete problems (without proof); Introduction to Approximation algorithms. **12 Hours**

Lab Work using C++ / Python **10 Lab classes**

1. Operation on array – insert, delete, merge.

2. Linear Search and Binary search.

3. Representation of sparse matrix, addition and transpose of sparse matrix.

4. Implementation of stack using array.

5. Conversion of infix to postfix expression.

6. Evaluation of postfix expression.

7. Operations of queue using array.

8. Operations of circular queue.

9. Single linked list operations.

10. Double linked list operations.

11. Stack using linked list.

12. Queue using linked list.

13. Selection Sort, Bubble sort.

14. Binary Search Tree operations.

15. Quick Sort, Merge Sort, and Heap Sort.

16. Priority Queue using min-Heap.

17. Longest Common Subsequence.

18. Fractional Knapsack Problem.

19. Rabin-Karp String matching algorithm.

20. Graph Traversal using BFS and DFS.

21. Kruskal's Algorithm for Minimum Spanning Tree.

22. Dijkstra's Single source shortest path algorithm.

23. Warshall's all pair shortest path algorithm.

24. Mini Project

Prescribed Text Books:

1. M. Weiss, Data Structures and Algorithm Analysis in C, 2nd Edition, Pearson Education, 2002.
2. T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, Introduction to Algorithms, 3rd Edition, PHI Learning, 2014.

Reference Books:

3. Dexter Kozen, *Algorithms*, Springer Verlag, 1992.
4. Jon Kleinberg and Eva Tardos, *Algorithm Design*, Pearson/ Addison-wesley, 2006.
5. E. Horowitz, S. Sahni, and S. Rajasekharan, *Fundamental of Computer Algorithms*, Galgotia Publications, 2002.
6. A. Aho, J. Hopcroft and J. Ullman: *Data Structures and Algorithms*.
7. T. A. Standish: *Data Structure Techniques*, Addison Wesley.
8. Y. Kanetkar, Data Structures Through C, 2nd Edition, BPB Publication, 2010.
9. E. Horowitz, S. Sahni, and S. Anderson-Freed, Fundamentals of Data Structures in C, 2nd Edition, Universities Press, 2008.

10. A. Tenenbaum, Data Structures Using C, 3rd Edition, Pearson Education, 2007.
11. S. Lipchitz, Data Structures, 1st Edition, Tata McGraw-Hill, 2005.
12. J. Kleinberg and E. Tardos, Algorithm Design, 1st Edition, Pearson Education, 2013.

MDS 204

Multivariate Data Exploration and Analysis with Lab (in R) **(Marks-100)**
Credit 4(3-2-0)

Unit I: Data organization, summary statistics, Data Preprocessing (Data cleaning, data amputation, Data integration, Data transformation and reduction), Graphical methods for data exploration(Histogram, scatter plots, Box plots, Quantile plots, Bag plots, Glyph plots, coplots, Dot charts, Biplots, Plotting Points as curves, Graphs of growth plots). **10 hours**

Unit II: Random vectors and matrices, Mean vectors and co-variant matrices, sample geometry, Random sampling, Expectation of sample mean and co-variance matrix, Generalized variance, Matrix operations for descriptive statistics.

Multivariate normal density and its properties and sampling from multivariate normal distribution, Maximum likelihood estimation. **10 hours**

Unit III: Sampling distribution sample mean and co-variance, Large sample behaviour of sample mean and co-variance, Assessing normality assumption, detecting outliers and transformation to near normality, and data cleaning.

Inferences about a mean vector, Hotelling's T-square, confidence regions for a mean vector, Likelihood ratio test for a mean vector, Inference about mean vector when observation is missing, large sample inference. **10 hours**

Unit IV: Comparing mean vectors from two populations, Comparison of several multivariate means(One-way MANOVA), Simultaneous confidence intervals for treatment effects, Testing equality of co-variance matrices.

Principal component analysis,Population Principal components, Summarizing the sample variable by principal components, Graphing principal components. **10 hours**

Unit V: Factor analysis (orthogonal factor model, Methods of estimation, Factor rotation, prospective and strategy for factor analysis, factor scores, Data example).

Canonical correlation analysis (canonical variables and canonical correlations, Interpreting the population canonical variables, sample canonical variables and sample canonical correlation), Large sample inference. **10 hours**

Lab Work: Implementation of the following using R **10 lab classes**

1. Computing the summary statistics of a Data set.

2. Data Computation and Data Transformation.

3. Plotting the graphs of the various plots and curves mentioned in Unit I.

4. Computation of co-variance and correlation matrices.

5. Computation of multivariate normal density function and bi variate normal.

6. Obtaining the maximum likelihood estimate of the mean μ and co-variance matrix of a multivariate normal distribution.

7. Detecting outliers in the data and cleaning the data.

8. Assessing the normality of a given data set.

9. Constructing large sample simultaneous confidence intervals.

10. Comparing several multivariate population mean vector using MANOVA.

11. Computing the population Principal Components and summarizing sample variation by Principal Components.

12. Finding Sample Principal Components from standardized data and plotting the Principal Components.

13. Performing factor analysis of a given data set.

14. Computing Canonical variates and Canonical correlations for standardized variables.

15. Performing canonical correlation analysis of a given data set.

Books recommended:

Texts:

1. R. A. Johnson and D. W. Wichern, Applied Multivariate Statistical Analysis, Pearson Education, 6th edition, 2015.

Chapters :1(1.1-1.4), 2(2.1-2.5), 3(3.5,3.6), 4(4.1-4.8), 5(5.1-5.7), 6(6.3-6.6), 8(8.1-8.4), 9(9.1-9.6), 10(10.1-10.4,10.6).

2. W. L. Martinez, A R Martinez, J. L. Solka, Exploratory data analysis, 2nd edition, CRC Press, 2010.

Chapters 1,9,10

Reference Book:

1. T. W. Anderson, An Introduction to Multivariate Statistical Analysis, Wiley, 2003.

2. B.S.Everitt and T.Hothorn, An Introduction to Applied Multivariate Analysis with R, Springer Verlag, 2011.

MDS 205

Machine Learning with Lab (in Python)	(Marks-100)	Credit 4 (3-2-0)

Unit I: Linear Methods for Regression and Classification: Overview of supervised learning, Linear regression models and least squares, Multiple linear regression, The Gauss-Markov theorem, Subset selection, Ridge regression, least angle regression and Lasso, Elastic net, Linear Discriminant Analysis(LDA), Reduced rank LDA, Logistic regression for two or more classes. Comparison of KNN, LDA, and logistic regression. **13 hours**

Unit II: Model Assessment and Selection: Bias,Variance,and model complexity,Bias-variance trade off, Optimism of the training error rate ,Estimate of In-sample prediction error,Effective number of parameters, Bayesian approach and BIC, Cross-validation ,Boot strap methods,Performance of classification algorithms (confusion matrix, precision and recall, ROC curve). **8 hours**

Unit III: Additive Models,Trees, and Boosting: Generalized additive models**,** Regression and classification trees, Boosting methods-exponential loss and AdaBoost, Numerical Optimization via gradient boosting, Examples (Spam data, California housing, Demographic data), XGboost. **9 hours**

Unit IV: Support Vector Machines(SVM), and K-nearest Neighbor: Basis expansion and regularization, Kernel smoothing methods, SVM for classification, Reproducing Kernels, SVM for regression, K-nearest –Neighbour classifiers (Image Scene Classification) **10 hours**

Unit V: Unsupervised Learning and Random forests: Cluster analysis(similarity and diasimilarity measures, K-means and K-medoids clustering, Hierarchial clustering, DBSCAN, EM algorithm, Gaussian mixture model and its use in clustering and anomaly detection. Random forests and analysis. **10 hours**

Lab work **10 lab classes**

Implementation of following using Python

1. Multiple linear regression,
1. Logistic regression,
2. Linear discriminant analysis,
3. Ridge regression, LASSO,
4. Cross-validation,
5. Boot strap,
6. Fitting classification and regression trees,
7. K-nearest neighbours,
8. SVM for classification,
9. K-means clustering.
10. Gaussian mixture, and
11. Random forest.

Prescribed Text Books

1.Trevor Hastie, Robert Tibshirani,Jerome Friedman , *The Elements of Statistical Learning-Data Mining, Inference,and Prediction*,Second Edition, Springer Verlag, 2009. [relevant topics prescribed above from chapters : 1 to 10, 13, 14, 15.]

2.Simon Haykin, Neural Networks- a comprehensive foundation, Ch.6, Second edition ,Pearson Education, 2001.

3. Stuart Russell and Peter Norving, “Artificial Intelligence: A Modern Approach”, Third Edition, 2010, Pearson Education, New Delhi.

Reference Books

1. G. James, D.Witten,T.Hastie, R.Tibshirani- *An introduction to statistical learning with applications in R*, Springer,2013.

2. C. M. Bishop –Pattern Recognition and Machine Learning, Springer, 2006.
3. M. J. ZAKI, Wagner M,Jr, Data mining and Machine Learning-Fundamental Concepts and Algorithms, Cambridge Univ.Press, second edition,2020
4. Joseph Giarratano and Gary Riley,"*Expert Systems: Principles and Programming*", Fourth Edition, CENGAGE Learning India Pvt. Ltd., New Delhi.
5. Elanie A. Rich and Kevin Knight, "*Artificial Intelligence*", 3rd Edition, 2009, McGraw-Hill Education (India), New Delhi.
6. Nills J. Nilsson, "*Artificial Intelligence: A New Synthesis*" 2nd Edition, 2000,Elsiver India Publications, New Delhi.
7. Michael Negnevitsky, "*Artificial Intelligence: A Guide to Intelligent Systems*", Second Edition, 2005, Pearson Education, Inc. New Delhi.
8. Dan W. Patterson, "*Introduction to Artificial Intelligence and Expert Systems*",1st Edition, 1996, PHI Learning Pvt. Ltd., New Delhi.
9. Ben Coppin, "*Artificial Intelligence Illuminated*", 2005, Narosa Publication, New Delhi. ISBN: 978-81-7319-671-3.

Texts 1 and 2 and references 1 and 3 are available on line.

MDS 206

Computer Systems for Data Science (Marks-100) **Credit 4 (3-2-0)**

Unit I: Computer Organisation: Introduction to computer architecture, Instruction set, Instruction set principles, Computer arithmetic, ALU design, Assessing computer system performance, Amdahl's Law. Pipelining, Pipeline control, pipeline hazards, Basics of caching, Cache mapping.

Unit II: Operating system: Introduction to Processes, CPU Scheduling, Dead lock- prevention Memory management- paging & segmentation, virtual memory, secondary storage structure & Disk Scheduling.

Unit III: Computer Networks: Introduction, Network models, Physical layer (Signals, digital and analogue transmission, Multiplexing, circuit switching), Data Link Layer (Error detection and correction, Data link control and Protocols, PPP, Local area networks. Network Layer (Host-to-host delivery - inter networking, addressing, and routing), Network layer protocols), Transport Layer (Process to process Delivery : VDP and TCP, Congestions control and quality of service), Application Layer (Client server models, Domain name system, SMTP, FTP, HTTP and WWW), Security.

Unit IV: Database Management System: Overview of database systems, Introduction to database design, The relational model, Relational algebra Functional dependencies, Normal forms, Indexing, Tree - structured indexing, Overview of query evaluation, Evaluating relational operators, query optimization.

Unit V: SQL lab **5 Labs (2 Hours each)**

Recommended Texts:

1. J. L. Hennessy, and D. A. Patterson, Computer Architecture: A Quantitative Approach, Morgan Kaufman Publishers 1990.
2. A. Silberschatz & P. B. Galvin, Operating System concepts, John Wiley.
3. Elmarsi, Navathe: Fundamentals of Database System, Addison Wesley.
4. B. A. Forouzan- *Data Communications and Networking*, TMH, 3rd Edition, 2004.
5. W. Stallings, *Computer Networking with Internet Protocols and Technology* Prentice Hall.
6. A. Silberschata, H. Korth, S. Sudarshan, Database System Concepts, Mc Grawhill.
7. Ivan Bayross, SQL, PL/SQL: The Programming Language of Oracle, 3rd Edition, BPB Publications.

Semester-III

MDS 301

Functional Analysis **(Marks-100)** **Credit 4(3-0-1)**

Unit-I: Normed linear spaces, Riesz lemma, characterization of finite-dimensional spaces, Banach spaces. Operator norm, continuity and boundedness of linear maps on a normed linear space. Baire's theorem and its consequence.

Unit-II: Hahn-Banach theorems, uniform boundedness principle, divergence of Fourier series, closed graph theorem, open mapping theorem and some applications.

Unit-III: Duals spaces, weak and weak* convergence, adjoint of an operator.

Unit-IV: Inner product spaces, orthonormal set, Gram-Schmidt ortho-normalization, Bessel's inequality, orthonormal basis, separable Hilbert spaces. Orthonormal complements, orthogonal projections, projection theorem, Riesz representation theorem.

Unit-V: Adjoint, normal, unitary, self-adjoint operators, compact operators. Spectral theorem for compact self-adjoint operators, statement of the spectral theorem for bounded self-adjoint operators.

Text Book:

1. Rajendra Bhatia: Notes on Functional Analysis, Hindustan Book Agency.

Reference Books:

2. J. B. Conway: A course in functional analysis, Graduate Texts in Mathematics, 96., Springer-Verlag, 1990.
3. Erwin Kreyszig: Introductory functional analysis with applications, Wiley India.
4. Christopher Heil: Metrics, Norms, Inner Products, and Operator Theory, Birkhäuser.
5. W. Rudin: Functional analysis, McGraw-Hill, Inc., 1991.
6. George Bachman and Lawrence Narici: Functional Analysis, Dover Publications.
7. K. Yosida: Functional analysis, Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, 1980.

MDS 302

Advanced Machine Learning with Lab **(Marks-100)** **Credit 4 (3-2-0)**

Unit-I: Generative models for discrete data (Bayesian concept learning, Naïve Bayes classifier), Gaussian discriminant analysis, Inference in jointly Gaussian distributions, Bayesian statistics, Bayesian linear regression, General linear models and exponential family. Softmax regression.

Unit-II: Mixture models and EM algorithm, Sparse linear models, Review of SVM for classification and regression, Multiclass SVM, Basic kernels, Kernels for texts and strings. kernels for generative models, Gaussian Processes for regression.

Unit-III: Probabilistic Graphical Models: DIRECTED Graphical models(Bayesian networks), Markov and Hidden Markov Models, Markov Random fields, Conditional Random fields, Exact inference for graphical models-variable elimination.

Unit-IV: Neural Network: Model of a neuron, Least Mean Square Algorithm, Perceptron(Learning algorithm, Convergence theorem), Multilayer Perceptron and back propagation(BPA), Methods of acceleration of convergence of BPA.

Unit-V: Dimensionality reduction (Kernel PCA, Independent Component analysis, LLE, t-SNE, Feature selection), Spectral clustering. Collaborative filtering methods.

Lab work **10 lab classes (2 Hours each)**

Implementation of the following using python:

1. Naïve Bayes classifier
2. Softmax regression
3. EM algorithm and Gaussian mixture model for anomaly detection ,and clustering
4. SVM for multiclass classification (ex. Document classification, sequence classification, etc.)
5. SVM for non-linear regression
6. Bayesian networks
7. Hidden Markov models
8. Sum-product variable elimination algorithm
9. Perceptron learning algorithm
10. Back propagation algorithm
11. Kernel PCA and feature selection algorithm
12. Spectral clustering
13. Collaborative filtering method.

Books Recommended:

1. Kevin P. Murphy, Probabilistic Machine Learning – an Introduction , MIT Press, 2022.
2. Kevin P. Murphy, Probabilistic Machine Learning – advanced topics, MIT Press, 2022.
3. Christopher Bishop, Pattern Recognition and Machine Learning, Springer, 2007.
4. Simon Haykin, Neural Network – a comprehensive foundation, Pearson Education-1994.
5. Tom Mitchel, Machine Learning – McGraw Hill Science,1997.

MDS 303

Complex Analysis (Marks-100) **Credit 4(3-0-1)**

Unit-I: Complex plane, topology of complex plane, open connected sets in the complex plane, the Riemann sphere. Holomorphic functions, Cauchy-Riemann Equations, examples of holomorphic functions, power series, radius of convergence, exponential and trigonometric functions.

Unit-II: Complex integration along piecewise smooth curves, Goursat's theorem, Local existence of primitives and Cauchy's theorem in a disc, evaluation of some integrals, Cauchy's integral formulas, Cauchy inequalities, power series expansion of holomorphic functions, Liouville's theorem, Morera's theorem, holomorphic functions defined in terms of integrals.

Unit-III: Zeros of holomorphic functions, Uniqueness theorem, Laurent series, Singularities, poles, residue theorem, Riemann's theorem on removable singularities, Casorati-Weierstrass theorem, meromorphic functions, argument principle, Rouche's theorem, Open Mapping theorem, Maximum modulus principle.

Unit-IV: Homotopies and simply connected domains, the complex logarithm. Jensen's formula, functions of finite order, Infinite products, product formula for the sine function, Weierstrass infinite products, Hadamard's factorization theorem.

Unit-V: Conformal map, conformal equivalence and examples, Schwarz lemma, automorphisms of the disc, upper plane, Riemann sphere and the complex plane. The Riemann mapping theorem and Picard's theorem.

Text Book:

1. Elias M.Stein and Rami Shakarchi: Complex Analysis, Princeton Lectures in Analysis, Princeton University Press.

Reference Books:

2. W. Rudin: Real and Complex Analysis, Tata McGraw-Hill Publishing Company Ltd., New Delhi, Chapters 10, 11, 12, 13, 14, 15 and 16.
3. J. B. Conway: Functions of one complex Variable, Springer International Student Edition, Second Edition.
4. L.V. Ahlfors, Complex Analysis, Third Edition, International Student Edition, McGraw-Hill Kogakusha, Ltd.

MDS 304

Differential and Integral Equations (Marks-100) **Credit 4(3-0-1)**

Unit-I: Existence and uniqueness of solutions: Lipschitz condition, Gronwall inequality, successive approximation, Picard's theorem, continuation and dependence on initial conditions, Existence of solutions in the large, Existence and uniqueness of solutions of systems, Fixed point method, Systems of linear differential equations: n^{th} order equation as a first order system, System of first order equations, Existence and uniqueness theorem, fundamental matrix, Non-homogeneous linear systems, Linear equations with constant coefficients

Unit-II: Non-linear Differential Equations: Existence theorem, Extremal solutions, Upper and lower solutions, Monotone iterative method and method of quasi linearization, Stability of Linear and Non-linear systems: Critical points, System of equations with constant coefficients, Linear equations with constant coefficients, Lyapunov stability.

Unit-III: Boundary value problem for ordinary differential equations: Sturm-Liouville problem, Eigenvalue and eigen functions, Expansion in eigen functions, Green's function, Picard's theorem for boundary value problems. Series solution of Legendre and Bessel equations.

Unit-IV: The Laplace equation: Boundary value for Laplace's equation, fundamental solution, Integral representation and mean value formula for harmonic functions, Green's function for Laplace's equation, solution of the Dirichlet problem for a ball, solution by separation of variables,

Unit-V: The wave equation and its solution by method of separation of variables, D'Alembert solution of the wave equation, Solution of wave equation by Fourier transform method. **Linear Integral Equations:** Linear integral equation of the first and second kind of Fredholm and Volterra type, Solutions with separable kernels. Characteristic numbers and eigenfunctions, resolvent kernel.

Books for reference

1. S. D. Deo, V. Lakshmikantham and V. Raghavendra: Text Book of Ordinary Differential Equations, Second Edition, Tata McGraw-Hill Publishing Company Ltd., New Delhi, Chapters 4(4.1 – 4.7), 5, 6(6.1 – 6.5), 7(7.5), 9(9.1 – 9.5).
2. Earl A. Coddington, Norman Levinson, The Theory of Ordinary Differential Equations
3. J. Sinha Roy, S. Padhy, A Course on Ordinary and Partial Differential Equations, Kalyani Publishers.
4. Tyn Mint-U, Lokenath Debnath, Linear Partial Differential Equations for Scientists and Engineers, Fourth Edition, Birkhauser.
5. Lawrence Evans, Partial Differential Equations, Second Edition, Graduate Studies in Mathematics, AMS.
6. Abdul J. Jerry, Introduction to Integral Equations with Applications, 2nd Ed., Clarkson University Wiley Publishers, 1999.
7. Chambers, Ll. G., Integral Equations: A short Course,International Text Book Company Ltd., 1976.
8. R. P. Kanwal, Linear Integral Equations,2nd Ed., Birkhauser Bosten,1997.
9. Hochstadt Harry, Integral Equations, John Wiley & Sons,1989.

MDS 305

Numerical Analysis with Lab	(Marks-100)	Credit 4 (3-2-0)

Unit I: Solution of Equations in One and Several Variables: Fixed point iteration method, Accelaration of convergence, Zeros of polynomials and Muller's method, Fixed points for functions of several variables, Newton's method for for non-linear system of equations. **7 hours**

Unit II: Interpolation: Hermite interpolation, Cubic spline interpolation, B-spline: interpolation and approximation, Parametric curves- Bazier curves. **7 hours**

Unit III: Least Square Approximation: Discrete least square approximation, Orthogonal polynomials, The Chebyshev polynomials and economization, Rational function approximation (Pade rational approximation, Chebyshev rational approximation) **7 hours**

Unit IV: Numerical Integration: Elements, Composite integration, Romberg integration, The Gauss quadrature. APPROXIMATION OF Multiple Integrals: Product rules, Rules exact for monomials, The Radon formula for integration of integrals in two dimensions. **7 hours**

Unit V: Numerical solution of ordinary differential Equations: Higher order Taylor series method, Runge-Kutta methods, Error control and Runge-Kutta-Fehlberg method, Multistep methods, Variable step size multistep methods-predictor-corrector method, Stability of numerical methods for solving IVP, Shooting and finite difference methods for boundary value problems (BVP) for ODE. Finite difference method for parabolic PDE. **9 hours**

Lab Work **8 lab classes**

Implementation of the following using Python/ Matlab

1. Finding zeros of a polynomial using Muller's method.

2. Solution of a non-linear system of equations using Newton's method.

3. Construction of cubic spline interpolant of a function.

4. Construction of a cubic Bazier corve given in parametric form.

5. Obtaining Pade rational approximation for a given function.

6. Approximate evaluation of a definite integral using (i) Composite Simpson rule (ii) Gaussian quadrature.

7. Approximate evaluation of a double integral using Guassian quadrature.

8. Finding approximate Solution of an IVP using (i) Runge-Kutta 4th order method (ii) Runge-Kutta-Fehlberg method.

9. Obtaining approximate solution of an IVP using Adms 4th order Predictor-corrector method

10. Obtaining approximate solution of a second order BVP using (i) shooting method (ii) finite difference method.

Prescribed Text Book

1. R. L. Burden and J. D. Faires, Numerical Analysis, Cengage Learning, 9th edition, 2011.

Chapters- 2 (2.2, 2.4-2.6), 3(3.4, 3.5, 3.6), 4(4.3, 4.4, 4.5, 4.7, 4.8), 5(5.3-5.7), 8(8.1-8.4), 10(10.1, 10.2),11(11.1-11.4),12(12.2)

Reference Books

1. P. J. Davis and Rabinowiwitz, Methods of Numerical Integration, A. P., Fourth Edition.

2. W.Cheney and D.Kincoid ,Numerical Mathematics and Computing, Cengage Learning,7th edition, 2013.

MDS 306

Data Visualization with Lab + Summer Internship Evaluation (Marks: 100)
Credit 3 (2-2-0) + 1

Unit-I: What is Visualization? The visualization process, Seven stages of data visualization, Type of data, Perception, Eight visual variables: position, shape, size, brightness, color, orientation, texture, motion.

Unit-II: Visualization techniques for Multivariate data, Tree, Graph, Networks, Text, Documents.

Unit-III: Visualization techniques for Spatial data, Geospatial data, Time-oriented data, Evaluating Visualizations.

Unit-IV: Introduction to D3, Working with data, Data-binding, Data-driven design and interaction, General charting principles, creating an axis, line charts and interpolations, Layouts (Histograms, Pie charts, Stack layout), Visualization with Scalable Vector Graphics (SVG), Drawing, Transformations, Building Charts with SVG.

Unit-V: Visualization libraries in R/Python: Matplotlib (Histograms, Bar charts, Line plots, Pie charts, Box plots, Scatter plots), Seaborn (Box, Violin plots, Regression plots, Heatmaps), Bokeh, ggplot2, Creating Dashboards with Plotly and Dash.

Experiments:

1. Excel: Use Excel charts for presentation of quantitative data.
2. Tableau: Import and prepare data, Perform data cleaning and harmonization.
3. Tableau: Create data visualizations.
4. Tableau: Use advanced data visualization to discover trends in data sets.
5. D3: Creating lines and circles with select and append.
6. D3: Loading data, displaying it as a bar chart and creating a scatterplot.
7. D3: Draw histograms, violin plots, pie charts and ring charts.
8. Matplotlib: Basic plotting with matplotlib library.
9. Seaborn: Create regression plots and heatmap using seaborn library.
10. Bokeh: Create few standalone interactive plots.
11. ggplot: Data visualization with ggplot2.
12. Plotly: Build a dashboard with Plotly and Dash.
13. Power BI: Data visualization with Power BI.

Text Books:

1. M. Ward, G. Grinstein, and D. Keim, *Interactive Data Visualization: Foundations, Techniques, and Applications*, 2nd Edition, CRC Press, 2015.
2. E. Meeks, *D3.js in Action: Data Visualization with JavaScript*, 2nd Edition, Manning Publications, 2018.
3. A. C. Telea, *Data Visualization Principles and Practice*, 2nd Edition, CRC Press, 2015.
4. H. Guerrero, *Excel Data Analysis: Modeling and Simulation*, Springer, 2010.
5. B. Jones, *Communicating Data with Tableau*, O'Reilly Media, 2014.

Books for Reference:

1. B. Fry, *Visualizing Data*, O'Reilly Media, 2007.
2. S. Murray, *Interactive Data Visualization for the Web*, 2nd Edition, O'Reilly Media, 2017.
3. K. Sosulski, *Data Visualization Made Simple: Insights into becoming Visual*, Routledge, 2018.
4. K. Healy, *Data Visualization: A Practical Introduction*, Princeton University Press, 2019.
5. A. Pajankar, *Practical Python Data Visualization: A Fast Track Approach to Learning Data Visualization with Python*, 1st Edition, Apress, 2020.
6. L. Ryan, *Visual Data Storytelling with Tableau*, 1st Edition, Addison-Wesley, 2018.

Semester-IV

MDS 401

Deep Learning & Reinforcement Learning with Lab (Marks: 100) Credit 4 (3-2-0)

Unit I: Exact inference for graphical models, Variational inference,Monte Carlo inference,MCMC inference,Learning undirected Gaussian graphical models

Unit II: Reinforcement learning and control- MDP, Bellman equations,value iterations and policy iteration, Linear quadratic regulation, LQG, Q-learningValue function approximation, Policy search, Reinforce POMDPs.

Unit III: Review of backpropagation. **Regularization for Deep Learning**: Parameter Norm Penalties, Norm Penalties as Constrained Optimization, Regularization and Under-Constrained Problems, Dataset Augmentation, Noise Robustness, Semi-Supervised Learning, Multitask Learning, Early Stopping, Parameter Tying and Parameter Sharing, Sparse Representations, Bagging and Other Ensemble Methods, Dropout, Adversarial Training, Tangent Distance, Tangent Prop and Manifold Tangent Classifier. **Optimization for Training Deep Models**: How Learning Differs from Pure Optimization, Challenges in Neural Network Optimization, Basic Algorithms, Parameter Initialization Strategies, Algorithms with Adaptive Learning Rates, Approximate Second-order Methods, Optimization Strategies and Meta-Algorithms.

Unit IV: Convolutional Networks: The Convolution Operation, Motivation, Pooling, convolution and Pooling as an infinitely strong prior, Variants of the Basic Convolution Function, Structured Outputs, Data Types, Efficient convolution Algorithms, Random or Unsupervised Features, The Neuroscientific Basis for Convolutional Networks, Convolutional Networks and the History of Deep Learning. **Sequence Modeling : Recurrent and Recursive Nets** : Unfolding Computational Graphs, Recurrent Neural Networks, Bidirectional RNNs, Encoder-Decoder Sequence-to-Sequence Architecture, Deep recurrent Networks, Recursive Neural Networks, The Challenge of Long-Term Dependencies, Echo State Networks, Leaky Units and Other Strategies for Multiple Time Scales, The Long Short-Term Memory and Other Gated RNNs, Optimization for Long-Term Dependencies, Explicit Memory.

Unit V: Practical Methodology: Performance Metrics, Default Baseline Models, Determining Whether to Gather More Data, Selecting Hyperparameters, Debugging Strategies, Example-Multi-Digit Number Recognition.

Linear Factor Models: Slow Feature Analysis, Sparse Coding, **Autoencoders**: Undercomplete Autoencoders, Regularized Autoendcoders, Representational Power, Layer Size and Depth, Stochastic Encoders and Decoders, DenoisingAutoencoders, Learning Manifolds with Autoencoders, Contractive Autoencoders, Predictive Sparse Decomposition, Applications of Autoencoders, **Deep Generative Models** : Boltzmann Machines, Restricted Boltzmann Machines, Deep Belief Networks.

Implementaion of the following algorithms:

i. Convolution Neural network (CNN)
ii. Recurrent Neural Network (RNN)
iii. Autoencoder
iv. Deep Belief Network

Books for reference:

1.Ian Goodfellow, YoshuaBengio, and Aaron Courville, Deep Learning, The MIT Press, 2016

2. Kevin P. Murphy, Machine Learning-a probabilistic prospective, MIT Press, 2012

3. Tom Mitchel, Machine Learning, McGraw Hill.

MDS 402

Big Data Analytics with Lab (Marks: 100) Credit 4 (3-2-0)

Unit-I Evolution of data management, Introduction to Big data, Introduction to Databases, Relational Model, and SQL, Data exploration and reproducibility, Data quality.

Unit-II Introduction to map reduce , Map reduce algorithm patterns and relations, Parallel data bases vs Map Reduce, Storage solutions, Introductions to Spark, Hadoop, Hive, Pig-Latin.

Unit-III Big Data Algorithms/Mining techniques- 1(Finding similar items and spark, exploring spatio-temporal data, mining association rules, EM algorithm for text processing).

Unit-IV Big Data Algorithms/Mining techniques-2(Page Rank and K-means in Pig and Spark, Efficient regularized SGD, Hash kernels for logistic regression, Matrix factorization, Matrix factorization with SGD, DGMs for naive Bayes, Fast sampling for LDA.

Unit-V Large scale graph processing –Pregel, Large-Scale visualisation.

Books for Reference:

1. Mining Massive Datasets, Anand Rajaraman, J. Leskovec, and Jeff Ullman- http://www.mmds.org
2. Data-Intensive Text Processing with nMapReduce, J. Lin and C. Dyer, http://lintool.github.com/MapReduceAlgorithms/index.html

MDS 403

Artificial Intelligence (Marks-100) Credit 4(3-2-0)

Unit I: Artificial Intelligence: Introduction, Intelligent Agents: Agents & Environments, Concept of Rationality, Nature & Structure of Agents.

Unit II: Problem Solving: Solving Problems by Searching, Classical Search, Adversarial Search, Constraint Satisfaction Problems. Knowledge, Reasoning and planning: Logical agents.

Unit III: First order logic, Inference in First order logic. Classical planning, Knowledge Representation: Uncertain Knowledge and Reasoning: Probabilistic Reasoning, Learning from Examples, Knowledge in learning;

Unit IV: Natural Language Processing: Language models, Text Classification, information retrieval, information extraction.

Unit V: Natural Language for Communication: Phrase structure Grammars, Syntactic Analysis, Augmented grammars and semantic interpretation, Machine translation, Speech recognition; Perception; Expert Systems: Introduction, Design of Expert systems.

Prescribed Text Books

1. Stuart Russell and Peter Norving, "Artificial Intelligence: A Modern Approach", Third Edition, 2010, Pearson Education, New Delhi.

Reference Books

1. Elanie A. Rich and Kevin Knight, "*Artificial Intelligence*", 3rd Edition, 2009, McGraw-Hill Education (India), New Delhi.
2. Nills J. Nilsson, "*Artificial Intelligence: A New Synthesis*" 2nd Edition, 2000,Elsiver India Publications, New Delhi.
3. Michael Negnevitsky, "*Artificial Intelligence: A Guide to Intelligent Systems*", Second Edition, 2005, Pearson Education, Inc. New Delhi.
4. Dan W. Patterson, "*Introduction to Artificial Intelligence and Expert Systems*",1st Edition, 1996, PHI Learning Pvt. Ltd., New Delhi.
5. Ben Coppin, "*Artificial Intelligence Illuminated*", 2005, Narosa Publication, New Delhi. ISBN: 978-81-7319-671-3.

MDS-404: Elective Papers Group-A

A1. Bioinformatics (Marks: 100) Credit 4(3-2-0)

Unit I: Molecular Biology Primer. Restriction mapping algorithm,

Unit II: Motif finding problem, Algorithms for motif finding (brute force, branch-and-bound and greedy)

Unit III: Edit distance and alignment, Global and local sequence alignment, multiple sequence alignment, Gene Prediction.

Unit IV: DNA sequencing, shortest superstring problem, sequencing by hybridization (SBH), SBH as Hamiltonian and Eulerian path problem, Fragment Assembly problem.

Unit V: Distance and character based methods for phylogenetics.

Recommended Texts:

1. N. C. Jones and P. A. Bavzner, An introduction to Bioinformatic algorithms, MIJ Press,2004. Relevant parts of chapters 2,4,5,6,8 and 10.
2. Dan E Krane, M. L. Baymer, Fundamental concepts of Bioinformatics, Pearson Education, 2003.
3. D. Gusfied, Algorithms on strings, Tress and sequences, Computer Science and computational Biology, Cambridge University Press, 1997.
4. M. S. Waterman, Introduction to computational Biology, Chapman Hall, 19995.

A2. Computational Fluid Dynamics (Marks:100) Credit 4(3-2-0)

Unit I: Introduction: Conservation equation; mass; momentum and energy equations; convective forms of the equations and general description.

Classification and Overview of Numerical Methods: Classification into various types of equation; parabolic elliptic and hyperbolic; boundary and initial conditions; over view of numerical methods.

Unit II: Finite Difference Technique: Finite difference methods; different means for formulating finite difference equation; Taylor series expansion, integration over element, local function method; treatment of boundary conditions; boundary layer treatment; variable property; interface and free surface treatment; accuracy of f.d. method.

Finite Volume Technique: Finite volume methods; different types of finite volume grids; approximation of surface and volume integrals; interpolation methods; central, upwind and hybrid formulations and comparison for convection-diffusion problem.

Unit III: Finite Element Methods: Finite element methods; Rayleigh-Ritz, Galerkin and Least square methods; interpolation functions; one and two dimensional elements; applications.

Methods of Solution:Solution of finite difference equations; iterative methods; matrix inversion methods;ADI method; operator splitting; fast Fourier transform.

Unit IV: Time integration Methods:Single and multilevel methods; predictor corrector methods; stability analysis;Applications to transient conduction and advectiondiffusion problems.

Numerical Grid Generation: Numerical grid generation; basic ideas; transformation and mapping.

Unit V: Navier-Stokes Equations:Explicit and implicit methods;SIMPLE type methods; fractional step methods.

Turbulence modeling: Reynolds averaged Navier-Stokes equations, RANS modeling, DNS and LES.

Text Books:

1. Ferziger, J. H. and Peric, M.(2003). Computational Methods for Fluid Dynamics. Third Edition,SpringerVerlag,Berlin.
2. Versteeg, H.K. and Malalasekara, W.(2008). Introduction to Computational Fluid Dynamics: The Finite Volume Method.Second Edition (Indian Reprint) Pearson Education.

3. Anderson, D.A., Tannehill, J.C. and Pletcher, R.H.(1997). Computational Fluid Mechanics and Heat Transfer. Taylor & Francis.

A3. Computer Vision (Marks 100) Credit 4(3-2-0)

Unit-I Introduction to computer vision, Human vision, color spaces, and transformations, Image formation: Geometric primitives, 2D transformations, 3D transformations, 3D rotations, 3D to 2D projections, Lens distortions, Parametric transformations, Mesh-based warping, Image processing: Point operators, Linear filtering, More neighborhood operators, Fourier transformations.

Unit-II Interpolation and Optimization: Pyramids and wavelets, Image blending, Model fitting and optimization: Scattered data interpolation, Varational methods and regularization, Markov random field and application. Convolutional neural networks: Application to digit classification, Network architectures, model zoos, Visualizing weights and activations, Adversarial examples, self-supervised learning. Recognition: Object detection, Semantic segmentation.

Unit-III Feature detection and matching: Points and patches, Edges and contours with applications, Lines and vanishing points: Hough transformation, Segmentation: Graph-based segmentation, Mean shift, Image alignment and stitching: Pairwise alignment and application, Image stitching and application.

Unit-IV Motion estimation: Translational alignment, parametric motion and application, Spline-based motion and application, Optical flow: deep learning approaches and application, Layered motion and application. Depth estimation: Epipolar geometry, Sparse correspondence, Dense correspondence, Local methods, Global optimization, Deep neural networks, Multi-view stereo.

Unit-V Structure from motion and SLAM : Geometric intrinsic calibration, Pose estimation, Two-frame structure from motion, Multi-frame structure from motion, simultaneous localization and mapping(SLAM) , 3D reconstruction: Shape from X , 3D scanning, Computational photography: Photometric calibration, High dynamic range imaging, Super-resolution, denoising, and blur removal. Image matting and composing, Hole filling and inpainting.

Text Books:

1. R Szeliski, **Computer vision: algorithms and applications**, 2nd edition,2021,springer, szeliski.org/book

Reference Books:

1. D.A. Forsyth, J Ponce, **Computer vision: a modern approach,** 2nd Edition, Pearson, 2012.
2. R Hartley, A Zisserman, **Multiple view geometry in Computer vision**, 2nd Edition, Cambridge University Press, 2005.
3. R. C. Gonzalez, R.E. Woods, R.E., Digital **image processing**, 3rd Edition, Pearson Education, 2002.

Online Reference Material(s):

1. https://www.cs.cornell.edu/courses/cs6670/2011sp/lectures/lectures.html
2. http://www.cs.cmu.edu/~16385/
3. http://web.stanford.edu/class/cs223b/syllabus.html
4. https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/classes/cs280_s99/
5. https://onlinecourses.nptel.ac.in/noc21_ee23/preview

A4. Computational Finance (Marks 100) Credit 4(3-2-0)

Unit-I: Introduction: Investment and markets, Typical investment problems; MeanVariance Portfolio Theory: Asset return, random variables, random returns, portfolio mean and variance, the feasible set, the Markowitz model, the two-fund theorem, inclusion of risk-free asset, the one fund theorem.

Unit-II: Capital Asset Pricing Model: Market equilibrium, the capital market line, Pricing Model, Security Market Line, Divestment implications, Performance evaluation, CAPM pricing formula; Models and Data: Factor model, CAPM as a Factor Model, Arbitrage Pricing theory, data and statistics.

Unit-III: Forwards and Futures: Forward contracts, Forward prices, Value of forward contract, Basics of futures contracts, Futures prices, Relation to expected spot price, Perfect hedge, Minimum variance hedge. Options, The No Arbitrage Principle, option and futures market. Option Pricing by Monte Carlo Simulation.

Unit-IV: Stochastic Differential Equations: Stochastic Ito Processes, Wiener Process, Stochastic differential equation, Ito integral, Ito Lemma, and Applications in Stock Market, The Black-Scholes Equation: Derivation of the Black-Scholes Equation, Solution of the Black-Scholes Equation, Closed Form Solutions of European Call and Put Options, Hedging Portfolios: The Greeks, hedging, Implied Volatility.

Unit-V: Option Pricing by Partial Differential Equations: Classification of PDEs, Finite Difference Methods for Parabolic Equations (An Explicit Method, An Implicit method, Crank-Nicolson Method), Option Pricing by the Heat Equation, Pricing American Options (Projected SOR Method for American Options).

Lab Work **10 lab Classes**

Implementations of the following:

I. (a) Markowitz model

(b) Computing the minimum variance portfolio using two fund theorem.

(c) Computing the minimum variance portfolio using one fund theorem.

II.(a) Computing the beta of a portfolio CAPM and hence draw the security market line.

(b) Computing the market return using Multi factor model

(c) Computing the price of an asset using arbitrage pricing theory.

III. (a) Computing the value of forwards and futures using appropriate data.

(b) Computing the option price using Monte Carlo Option pricing model.

IV. (a) Computing the price of European put and call option using Black-Scholes(BS) formula

(b) Computing implied volatility using BS formula

(c) Computing the Greeks and using those hedging a portfolio.

V. (a) Computing the pricing of a European option using explicit finite difference method.

(b) Computing the pricing of a European option using implicit finite difference method.

(c) Computing the pricing of a European option using Crank-Nicholson finite difference method.

(d) Computing the pricing of an American option using projected SOR method.

Text Books:

(For Units 1, 2)

T1: D. G. Luenberger, Investment Science, 2nd Edition, Oxford University Press, 2014.

For unit 3 : T1 and T2 :

T2: J. C. Hull and S. Basu, Options, Future and Other Derivatives, 10th Edition, Pearson, 2018.

(For units 4 and 5)

T3: O. Ugur, An Introduction to Computational Finance, 1st Edition, Imperial College Press, 2009.

Reference Books:

R1. E. J. Elton, M. J. Gruber, S. J. Brown, and W. N. Goetzmann, Modern Portfolio Theory and Investment Analysis, 7th Edition, John Wiley & Sons, 2007.

R2. Z. Bodie, A. Kane, A. J. Marcus, and P. Mohanty, Investments, 11th Edition, McGraw-Hill Education, 2019.

R3. A. Hirsa, Computational Methods in Finance, Chapman & Hall/CRC Press, 2013.

R4. P. Wilmott, Option Pricing : Mathematical Models and Computation, 1st Edition, Oxford Financial Press, 1993.

R5. L. Clewlow and C. Strickland, Implementing Derivative Models, John Wiley & Sons, 1998.

A5. Computational Modelling of Financial Derivatives (Marks100) Credit 4(3-2-0)

Unit-I: Construction of finite difference schemes for Laplace and heat equation. Solution of Tri diagonal system of linear equations, iterative methods (Jacobi Successive over relaxation) Introduction to stability and convergence of finite difference method, crank Nicholson and ADI methods.

Unit-II: The Black Scholes (B.S.) Model: Derivation of B.S. equation using no arbitrage argument, options dividend paying equities, Derivation of B.S. formula for calls, Puts and simple digitals, obtaining formulae for Greeks and their numerical computation, Finite Difference method for solution of B.S. equation.

Unit-III: Simple generalization of B. S. model (Dividend payments, time dependent parameters), Early exercise and American options (Perpetual American call and put, general payoff, local solutions), American option problem as free boundary problems, Finite difference method for American options, Monte-Carlo simulation.

Unit-IV: Exotic and path dependent options: General introduction, Barrier options, Strongly path dependent options, Asian options, Look back options, Multi Asset options, Numerical implementation.

Unit-V: Extensions of B-S. Model: Defects in B-S Models, Discrete hedging, Transaction costs, stochastic volatility. Jump diffusion, Crash modelling.

Books Recommended:

Text: Paul Wilmott: Paul Wilmott on quantitative finance John Wiley – 2006

Reference Books:

1. *Quantitative Methods in Derivative pricing - an introduction to computational Finance, D. A. Tavella , John Wiley – 2002.*
2. *The concepts and practice of mathematical finance, M. S. Joshi, Cambridge Univ. Press - 2003.*
3. *Derivative Securities and Difference Methods, Y. Zhu, X. Wu, F. Chern, Springer Verlag- 2004.*
4. *Numerical Solution of PDE, G.D. Smith, CUP.*

A6. Image & Video Processing (Marks: 100) Credit 4(3-0-1)

Unit I: Computer Graphics: Development of computer Graphics: Raster Scan and Random Scan graphics storages, displays processors and character generators, colour display techniques, interactive input/output devices. Points, lines and curves: Scan conversion, line-drawing algorithms.

Unit II: Circle and ellipse generation, polygon filling anti aliasing. Two-dimensional viewing: Coordinate systems, linear transformations, line and polygon clipping algorithms.

Unit III: Image Processing: Fundamentals of Image processing and Image Transforms: Basic steps of Image processing system sampling and quantization of an Image – Basic relationship between pixels Image Transforms: 2 – D Discrete Fourier Transform, Discrete Cosine Transform (DCT), Discrete Wavelet transforms. Image Processing Techniques: Image Enhancement: Spatial Domain methods: Histogram Processing, Fundamentals of Spatial Filtering, Smoothing Spatial filters, Sharpening Spatial filters Frequency Domain methods: Basics of filtering in frequency domain, image smoothing, image sharpening, selective filtering

Unit IV: Image Segmentation: Segmentation concepts, point, line and Edge detection, Thresholding, region based segmentation Image Compression Image compression fundamentals – coding Redundancy, spatial and temporal redundancy. Compression models : Lossy and Lossless, Huffmann coding, Arithmetic coding, LZW coding, run length coding, Bit Plane coding, transform coding, predictive coding , wavelet coding, JPEG standards

Unit V: Video Processing: Basic Steps of Video Processing: Analog video, Digital Video, Time varying Image Formation models : 3D motion models, Geometric Image formation , Photometric Image formation, sampling of video signals, filtering operations 2-D Motion Estimation: Optical flow, general methodologies, pixel based motion estimation, Block matching algorithm, Mesh based motion Estimation, global Motion Estimation, Region based motion estimation, multi resolution motion estimation. Waveform based coding, Block based transform coding, predictive coding, Application of motion estimation in video coding.

Books Recommended:

1. D. Hearn and M. P. Baker, Computer Graphics, 2nd Ed., Prentice–Hall of India, 2004.
2. J. D. Foley, A van Dam, S. K. Feiner and J. F. Hughes, Computer Graphics: Principles and ractices, 2nd Ed., Addison-Wesley, MA, 1990.
3. D. F. Rogers, Procedural Elements in Computer Graphics, 2nd Ed., McGraw Hill Book Company, 2001.

4. D. F. Rogers and A. J. Admas, Mathematical Elements in Computer Graphics, 2nd Ed., McGraw Hill Book Company, 1990.
5. Gonzaleze and Woods ,"Digital Image Processing ", 3rd edition , Pearson.
6. Yao Wang, Jorn Ostermann, Ya-Qin Zhang, 'Video Processing and Communications', Prentice Hall, 2002.
7. Alan C.Bovik, 'The Essential Guide to Video Processing', Elsevier Science, edition 2, 2009.
8. A. Murat Tekalp, 'Digital Video Processing', Prentice Hall, edition 1, 1996.
9. M. Tekalp ,"Digital video Processing", Prentice Hall International
10. Relf, Christopher G.,"Image acquisition and processing with LabVIEW", CRC press
11. Aner ozdemi R, "Inverse Synthetic Aperture Radar Imaging with MATLAB Algorithms", John Wiley & Sons
12. Chris Solomon, Toby Breckon, "Fundamentals of Digital Image Processing A Practical Approach with Examples in Matlab", John Wiley & Sons.

A7. Number Theory and Cryptography with Lab (Marks100) Credit 4(3-2-0)

Unit-I: Basic number theory: Divisibility, and primality, Greatest Common Divisor, Unique factorization and its consequences, Congruences(Basic properties, solving linear congruences, residue classes, Euler's phi function, Fermat's little theorem), Computing with large integers (basic integer arithmetic, integer exponentiation, computing in Z_n, Faster integer arithmetic), Euclid's algorithm (basic and extended), Computing modular inverses and Chinese remainder theorem, Hensel lifting, orders and primitive roots.

Unit-II: Overview of Cryptography, Affine cipher and its cryptanalysis, Shannon ciphers and perfect security, Computational ciphers and semantic security.
Stream ciphers: Pseudorandom generators(PRG), encryption with a PRG, stream cipher limitations, Composing PRGs-the Blum Micali method, unpredictability for a PRG, Case studies – linear generators, and the cryptanalysis of the RC4 stream cipher.

Block Ciphers: basic definitions and properties, constructing block ciphers in practice-Data Encryption Standard(DES)(Algorithm, Challenges, drawbacks), triple DES, Advanced Encryption Standard(AES)(algorithm and its security), Using Block ciphers(pseudo random permutations(CRP) and functions (CRF)—definition and properties, security against chosen plain text attack (CPA), Modes of operation-CBC and counter modes.

Unit-III: Message integrity(Message Authentication Code(MAC), Secure MACs, MAC security with verification queries, Constructing MACs from pseudorandom functions(PRFs), From prefix-free secure PRF to fully secure PRF: CMAC, Converting a blockwise PRF to bitwise PRF, Case studies: CBC-MAC, CMAC.

Collision resistant hashing, The Merkle-Damgard paradigm, Building compression functions-(Davies-Meyer), Collision resistance of Davies-Meyer. Case studies: SHA,HMAC, SHA-3.

Authentication encryption(definition, implication: every AE-secure system is CCA-secure), Constructions from ciphers and MACs, Case study TLS 1.2, CBC padding attacks, Attacking non-atomic decryption. Diffie -Hellman key exchange protocol and its security.

Unit-IV: Principles of Public-Key Cryptosystems, The RSA cryptosystem, Primality testing(Quadratic residues, Legendre and Jacobi symbols, The Solovay-Strassen algorithm,The-Miller

Rabin algorithm, AKS algorithm), square roots modulo n, Factoring algorithms and their attack on RSA, other attacks on RSA and semantic security of RSA, The Rabin cryptosystem, Pailler encryption schemes.

Discrete Logarithm Problem(DLP), ElGamal Crypto System, Algorithms for DLP, Security of ElGamal Systems, Schnorr signature scheme, The ElGamal signature scheme, The digital signature algorithm, Provable secure signature schemes.

Unit-V: Elliptic curve over the reals, Finite fields, Elliptic curves modulo a prime, Properties of Elliptic curves, Point compression and ECIES, Computing point multiples on Elliptic curves, Elliptic curve digital signature algorithm, Elliptic curve factorization, Elliptic curve primality test.

Lab Work **10 lab classes (2 hours each)**

Implementation of the following using C /PARI (free software)

I. Solving linear congruences, Computing with large integers(basic integer arithmetic, modular exponentiation, computing in Z_n, Faster integer arithmetic), Euclid's algorithm (basic and extended),Computing modular inverses and Chinese remainder theorem.

II. Algorithms for stream ciphers and Block Ciphers

III. Constructing MAC's from Pseudorandom functions, Building compression functions, AE-secure system construction, Diffie-Hellman key exchange protocol implementation.

IV. RSA crypto system, The Solovey-Strassen algorithm, Miller-Robin algorithm, The Rabin crypto system, Elgamal crypto system, Elgamal signature scheme, The digital signature algorithm.

V. Elliptic curve digital signature algorithm, Elliptic curve primality test.

Books Recommended:

Texts:
For Unit 1
1.Victor Shoup, A Computational introduction to number theory and algebra,Cambridge University Press, 2005. Relevent sections from chapters 1, 2, 3, and 4.

For Units 2 and 3:,
2. Dan Boneh and Victor Shoup, A Graduate Course in Applied Cryptography, verson 0.5, 2020, available online. http://toc.cryptobook.us/
Relevant sections from Chapters 2,3,4,5,6,8,9
For Units 4 and 5
3. Douglas Stinson , Cryptography : Theory and Practice, CRC Press, 2006, Chapters: 5, 6, 7.

References:

1. J. Katz, and Y. Lindell, Introduction to Modern Cryptography, Chapman and Hall/CRC, second edition.
2. A. Menezes, P. C. Van Oorschot, and S. A. Vanstans – Handbook of Applied Cryptography, CRC Press, 1997.
3. N. Koblitz – A Course in Number Theory and Cryptography, Springer Verlag
4. A. DAS, Computational Number Theory, CRC press.
5. D. M. Burton, Elementary Number Theory, McGrawHill Education, 2012.
6. A. Das, C. E.VENI Madhavan, Public Key Cryptography-Theory and Practice, Dorling Kindersley, 2009.

A8. Numerical Solution of Partial Differential Equations (Marks100) Credit 4(3-0-1)

Unit I: Solution of first order ordinary differential equations Approximate Solution: Picard Iteration Method, Taylor Series method.

Unit II: Numerical Solution: Euler method; Algorithm; Example; analysis.

Unit III: Modified Euler Method: Algorithm; Example; analysis

Unit IV: Runge Kutta methods, Second Order methods

Unit V: Fourth Order Runge Kutta methods, Higher Order Runge Kutta methods

References:

1. Bradie B A Friendly Introduction to Numerical Anaysis Pearson Education,2007
2. Burden RL, Faires J D Numerical Analysis Cengage Learning, 2007
3. Chapra SC, Canale, R P Numerical Methods for Engineers Tata McGraw Hill, 2003
4. Gerald C.F., Wheatley P O Applied Numerical analysis, Addison Wesley, 1998

A9. Object Oriented Software Engineering (Marks100) Credit 4(3-0-1)

Unit-I: Introduction, Software life cycle modules, Requirements Analysis and specification, software design, Function oriented software design.

Unit-II: Coding and Testing, Software reliability and quality management, Computer Aided Software Engineering, Software maintenance.

Unit-III: Introduction to object oriented Analysis and Design, Iterative Development and the unified process, case study-The next-Gen POS, Inception, Understanding Requirements, Use case Model, Identifying other requirements.

Unit-IV: Elaboration, Use case Model, Drawing system sequence diagrams, Visualizing concepts, Adding Associations, adding attributes, adding details with operation contracts, interaction diagram notation.

Unit-V: PATTERNS, GRASP, Creating design class diagrams, GOF Design pattern Planning and project queues comments on iterative development and the UP, Rational Unified Process.

Recommended Text Books:

1. *Crag Larman: Applying UML and Patterns-An introduction OOAP & D and the Unified process, Pearson Education Asia.*
2. *Rajib Mail: Fundamentals of Software Engineering, PHI.*

A10. Probabilistic Graphical Models (Marks100) Credit 4(3-0-1)

Unit I: Introduction, Probability Theory, Bayesian Networks

Unit II: Undirected models, Learning Bayes Nets, Exact Inference; Message Passing

Unit-III: Sampling, MAP Inference; Structured prediction

Unit-IV: Parameter Learning, Bayesian Learning; Structure Learning,

Unit-V: Exponential families; variational inference

Recommended Text Books:
1. Probabilistic Graphical Models: Principles and Techniques by Daphne Koller and Nir Friedman. MIT Press.

References:

1. *Modeling and Reasoning with Bayesian networks* by Adnan Darwiche.
2. *Pattern Recognition and Machine Learning* by Chris Bishop.
3. *Machine Learning: a Probabilistic Perspective* by Kevin P. Murphy.
4. *Information Theory, Inference, and Learning Algorithms* by David J. C. Mackay.
5. *Bayesian Reasoning and Machine Learning* by David Barber.
6. *Graphical models, exponential families, and variational inference* by Martin J. Wainwright and Michael I. Jordan.

MDS-405: Elective Papers Group-B

B1. Causal Inference for Data Science (Marks: 100) Credit 4(3-2-0)

Unit I: Introduction, Statistics and probability review, Introduction to Causation, Ignorability and unconfoundedness, Randomized Experiments, Randomization Inference.

Unit II: Randomized Experiments and contextual covariates: Treatment Effect Heterogeneity, Blocking and Stratification.

Unit III: Observational inference: Introduction to Observational Inference, Directed acyclic graphs, Weighting Methods The Propensity Score, Introduction to Matching, Matching Methods.

Unit IV: Regression: Regression Adjustment, Causal Inference with grouped data, Difference in differences, Instrumental Variables: assumptions and motivation Instrumental Variables Estimators.

Unit V: Regression Discontinuity Designs: Introduction to RDD and Examples, RDD: Estimators and Inference, Sensitivity Analysis, Causal Inference and Machine Learning, Causal Inference and Ethics.

References Text:
1. Mastering' Metrics: The Path From Cause to Effect. Princeton University Press, 2014. Joshua D. Angrist and Jörn-Steffen Pischke.

2. Causal Inference: What If Miguel A. Hern´ an, James M. Robins January 2, 2024

Recommended Text Books:
1. Steiner et al. 2017 Graphical models for quasi-experimental designs Sociological Methods and Research.

2. G. W. Imbens (2020) Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics Arxiv 1907.07271v2 (A graph skeptical overview of this week's topics from an economist)

3. A. S. Gerber and D. P. Green (2012) 'Field experiments: Design, analysis, and interpretation' Norton. (highly recommended for those new to field experiments)

4. EGAP's methods guides on many of the topics from lecture. Well worth bookmarking. Mostly development examples.

5. Meng, X.-L. (2018) Statistical paradises and paradoxes in big data (I): Law of large populations, big data paradox, and the 2016 US presidential election Annals of Applied Statistics.

B2. Cyber Security (Marks: 100) Credit 4(3-2-0)

Unit I: Introduction - Introduction to cyber security, Confidentiality, integrity, and availability. Foundations - Fundamental concepts, CIA, CIA triangle, data breach at target.Security management, Governance, risk, and compliance (GRC)- GRC framework, security standards.

Unit II: Contingency planning - Incidence response, Disaster Recovery, BCP. Cyber security policy - ESSP, ISSP, SYSSP. Risk Management - Cyber Risk Identification, Assessment, and Control.

Unit III: Cyber security: Industry perspective - Defense Technologies, Attack, Exploits. Cyber security technologies - Access control, Encryption, Standards. Foundations of privacy - Information privacy, Measurement, Theories.

Unit IV: Privacy regulation - Privacy, Anonymity, Regulation, Data Breach. Privacy regulation in Europe, Privacy: The Indian Way - Data Protection, GDPR, DPDP, Aadhar.

Unit V: Information privacy: Economics and strategy, Economic value of privacy, privacy valuation, WTA and WTC, Business strategy and privacy, espionage, Privacy vs safety.

References Text:

1. Michael E. Whitman, Herbert J. Mattord, (2018). Principles of Information Security, 6th edition, Cenage Learning, N. Delhi.
2. Darktrace, "Technology" https://www.darktrace.com/en/technology/#machine-learning, accessed November 2018.
3. Van Kessel, P. Is cyber security about more than protection? EY Global Information Security Survey 2018-2019.
4. Johnston, A.C. and Warkentin, M. Fear appeals and information security behaviors: An empirical study. MIS Quarterly, 2010.
5. Arce I. et al. Avoiding the top 10 software security design flaws. IEEE Computer Society Center for Secure Design (CSD), 2014.
6. Smith, H. J., Dinev, T., & Xu, H. Information privacy research: an interdisciplinary review. MIS Quarterly, 2011.
7. Subramanian R. Security, privacy and politics in India: a historical review. Journal of Information Systems Security (JISSec), 2010.
8. Acquisti, A., John, L. K., & Loewenstein, G. What is privacy worth? The Journal of Legal Studies, 2013.

9. Xu H., Luo X.R., Carroll J.M., Rosson M.B. The personalization privacy paradox: An exploratory study of decision making process for location-aware marketing. Decision Support Systems, 2011.

B3. Discrete Dynamical Systems (Marks: 100) Credit 4(3-2-0)

Unit-I: Representations of Dynamical Systems, Dynamical systems with discrete time. Types of trajectories, fixed and periodic points, Attracting, repelling and saddle fixed and periodic points and Chaos, Lyapunov Exponent.

Logistic Map, Hennon Map, Lorenz Model, Rossler Map, Horseshoe map, Complex Quadratic Map, Tent Map, Duffing Map.

Unit-II: Fractals: Cantor Set, Sierpinski Triangle, Koch Curve, Fractal dimension, Similarity Dimension, Box-counting dimension.

Unit-III: Cellular Automata: Components of Cellular Automata, Basic types, Space time diagram and state transition diagram of CA rules, Rule Space, Boolean functions, Matrix and Polynomial representation of Cellular Automata rules, Classification of Cellular Automata rules, Density Classification Task, Relation of Cellular Automata with other areas.

Unit-IV: L-System, Tilings, Iterated Function System, Mathematical Morphological Operations: Erosion, Delition, Opening and Closing and Integral Value Transformations.

Unit-V: Pattern Formation, Pattern Recognition, Pattern Classification and Pattern Analysis, & Applications.

Lab work

Implementation of following methods using R or Python or Matlab:

1. Program for Fractal formation and calculating its fractal dimension.

2. Program for obtaining the space time diagram and STD for CA rules.

3. Program for finding patterns for different IVTs.

4. Any other program recommended by the instructor.

Recommended Text

1. Barnsley, Michael F.; and Rising, Hawley; Fractals Everywhere. Boston: Academic Press Professional, 1993. ISBN 0-12-079061-0
2. Mandelbrot, Benoit B.; The Fractal Geometry of Nature. New York: W. H. Freeman and Co., 1982. ISBN 0-7167-1186-9
3. S. Wolfram, A New kind of Science, Wolfram Publisher, 2002.
4. Grzegorz Rozenberg and Arto Salomaa. The mathematical theory of L systems (Academic Press, New York, 1980). ISBN 0-12-597140-0.
5. Grzegorz Rozenberg, Arto Salomaa – Lindenmayer Systems: Impacts on Theoretical Computer Science, Computer Graphics, and Developmental Biology ISBN 978-3-540-55320-5
6. Mathematical morphology: from theory to applications, Laurent Najman and Hugues Talbot (Eds). ISTE-Wiley. ISBN 978-1-84821-215-2. (520 pp.) June 2010.
7. Image Analysis and Mathematical Morphology, Volume 2: Theoretical Advances by Jean Serra, ISBN 0-12-637241-1 (1988)

8. Discrete Dynamical Systems, Oded Galor, Springer.
9. *Nonlinear Dynamics and Chaos with Applications to Physics, Biology, Chemistry and Engineering* by Steven H. Strogatz, published by Addison Wesley (1994).

B4. High Performance Computing (Marks:100) Credit 4(3-2-0)

Unit I: Program execution: Program, Compilation, Object files, Function call and return, Address space, Data and its representation.

Unit II: Computer organization: Memory, Registers, Instruction set architecture, Instruction processing

Unit III: Pipelined processors: Pipelining, Structural, data and control hazards, Impact on programming. Virtual memory: Use of memory by programs, Address translation, Paging, Cache memory: Organization, impact on programming, virtual caches.

Unit IV: Operating systems: Processes and system calls, Process management, Program profiling. File systems: Disk management, Name management, Protection.

Unit V: Parallel architecture: Inter-process communication, Synchronization, Mutual exclusion, Basics of parallel architecture, Parallel programming with message passing using MPI.

References Text:
1. J. L. Hennessy and D. A. Patterson, Computer Architecture: A Quantitative Approach, Morgan Kaufmann.
2. A. Silberschatz, P. B. Galvin, G. Gagne, Operating System Concepts, John Wiley.
3. R. E. Bryant and D. R. O'Hallaron, Computer Systems: A Programmer's Perspective, Prentice Hall.

B5. Internet of Things with Lab (Marks-100) Credit 4 (3-2-0)

Unit I: Introduction to IoT: Sensing, Actuation, Basics of Networking, Communication Protocols, Sensor Networks, Machine-to-Machine Communications.

Unit II: Interoperability in IoT. Introduction to Arduino Programming, Integration of Sensors and Actuators with Arduino,

Unit III: Introduction to Python programming, Introduction to Raspberry Pi, Implementation of IoT with Raspberry Pi, Introduction to SDN, SDN for IoT, Data Handling and Analytics.

Unit IV: Cloud Computing, Sensor-Cloud, Fog Computing, Smart Cities and Smart Homes, Connected Vehicles, Smart Grid, Industrial IoT, Case Study: Agriculture, Healthcare, Activity Monitoring.

Unit V: Different projects on IoT.

Books Recommended:

1. B. A. Forouzan- *Data Communications and Networking*, TMH, 3rd Edition, 2004.

2. W. Stallings, *Computer Networking with Internet Protocols and Technology* Prentice Hall.

3. "The Internet of Things: Enabling Technologies, Platforms, and Use Cases", by Pethuru Raj and Anupama C. Raman (CRC Press)

4. "Internet of Things: A Hands-on Approach", by Arshdeep Bahga and Vijay Madisetti (Universities Press)

B6. Natural Language Processing (Marks:100) Credit 4(3-2-0)

Unit-I: Introduction to Natural Language Processing, various applications, issues and processing complexities, Regular Expressions, Text Normalization, Edit Distance, N-gram Language Models, Smoothing techniques.

Unit-II: Naive Bayes and Sentiment Classification, Other text classification task and logistic regression, Vector Semantics and Embeddings, Sequence Labeling for Parts of Speech and Named Entities.

Unit-III: Deep Learning Architectures for Sequence Processing, Recurrent neural networks applied to language problem, Contextual Embeddings, Machine Translation and Encoder-Decoder Models.

Unit-IV: Constituency Grammars, Constituency Parsing, Dependency Parsing, Logical Representations of Sentence Meaning, Computational Semantics and, Semantic Parsing, Information Extraction.

Unit-V: Word Senses and WordNet, Coreference Resolution, Discourse Coherence, Question Answering, Chatbots & Dialogue Systems, Automatic Speech Recognition, Text-to-Speech.

Text Books:

1. D. Jurafsky and J. H. Martin, ***Speech and Language Processing – An introduction to language processing, computational linguistics, and speech recognition***, Pearson Education, 3rd Edition, Dec, 2020 (Draft Copy: https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf)
2. Christopher D. Manning and Hinrich Schütze, ***Foundations of Statistical Natural Language Processing***, MIT Press, 2nd Edition, 2000

Reference Books:

1. T. Siddiqui and U. S. Tiwary, ***Natural language Processing and Information Retrieval***, Oxford University Press, 1st Edition, 2008
2. Charu C. Aggarwal, ***Machine Learning for Text***, Springer, 1st Edition, 2018
3. J. Allen, ***Natural Language Understanding***, Pearson Education, 2nd Edition, 2008

Online Reference Material(s):

1. https://nptel.ac.in/courses/106/101/106101007/ : by Prof. P. Bhattacharyya, IIT Bombay
2. https://nptel.ac.in/courses/106/105/106105158/: by Prof. P. Goyal, IIT Kharagpur

https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-863j_naturallanguage-and-the-computer-representation-of-knowledge-spring-2003/lecture-notes/

B7. Quantum Computing (Marks: 100) Credit 4(3-2-0)

Unit I: Introduction to Quantum Computing, Postulates of Quantum Mechanics, Qubits and Bloch Sphere, Basic Quantum Gates, Quantum Circuits.

Unit II: Quantum No Cloning Theorem and Teleportation. Dense coding, Density Matrix, Projective measurement, POVM, EPR and Bell's Inequalities.

Unit III: Deutsch Algorithm, Deutsch-Jozsa Algorithm, Simon Problem, Grover's Search Algorithm, Quantum Fourier Transform, Period Finding, Method of Continued Fraction.

Unit IV: Shor's Factorization Algorithm, Quantum Error Correction Codes. Classical Information Theory, Shannon entropy, Von Neumann Entropy,

Unit V: Classical Cryptography, RSA Algorithm, Quantum Cryptography, BB 84 protocol, B-92 and Eckart protocol, Practical realization of a quantum computer.

References Text:

1. MichaelA. Nielsen and Issac L. Chuang,"Quantum Computation and Information, Cambridge (2002).
2. Mikio Nakahara and Tetsuo Ohmi,"Quantum Computing", CRC Press (2008).
3. N. David Mermin,"Quantum Computer Science",Cambridge (2007).

B8. Soft Computing Methods (Marks:100) Credit 4(3-2-0)

Unit-I Neural Network topologies, activation functions and learning methods, perception training algorithm, The multilayer perception (MLP), Back Propagation learning algorithm, financial applications.

Unit-II Self organization maps, Support vector machine for classification and regression, application to finance.

Unit-III Genetic algorithm (GA), MLP-GA, SVM-GA hybrid methods and financial applications.

Unit-IV Elements of fuzzy set theory, Fuzzy logic and approximate reasoning, Neurofuzzy and Fuzzy –GA hybrid methods, Rough set theory financial applications.

Unit-V The particle swarm optimization algorithm, Discrete PSO, MLP-Swarm Hybrids, Ant colony optimization methods, financial applications.

Books Recommended:

1. S. Haykin-Neural Networks: a comprehensive foundation, Pearson Education, 2001

2. Brabazan, M. O' Neill-Biologically Inspired Algorithm for Financial ModellingSpringer-2006

3. R. A. Aliev, B. Fazlollahi, R.R. Aliev-Soft Computing and its Applications in Bussines and Economic, Springer Verlag, 2004.

4. Relevant research papers an use of soft computing methods for financial problems.

B9. Time Series Analysis and Forecasting (Marks: 100) Credit 4(3-2-0)

Unit-I: Linear Time Series Analysis and its Applications: Stationary, Correlation and Autocorrelation Function. White Noise and Linear Time Series, Simple Autoregressive Models, Simple Autoregressive Models.

Unit-II: Conditional Hoteroscedastic Models: Characteristic of Volatility, Structure of a Model, Model Building, The Arch MODEL, The GARCH Model, The Integrated GARCH Model, The GARCH-M Model, The Exponential GARCH Model, The threshold GARCH Model, The CHARAM Model, Random Coefficient Autoregressive Models, The stochastic Volatility Model, Application.

Unit-III: Nonlinear Models and Their Applications: Nonlinear Models, Nonlinearity Tests, Modelling, Forecasting, Applications.

Unit-IV: High- Frequency Data Analysis and Market Microstructure: Nonsynchronous Trading, Bid-Ask Spread, Empirical Characteristics of Transactions Data, Models for price changes, Duration Models, Nonlinear Duration Models, Bivariate Models for Price Change and Duration.

Unit-V: Principle Component Analysis and Factor Models: A Factor Model, Macro econometric Factor Models, Fundamental Factor Models, Principal Component Analysis, Statistical Factor Analysis, Asymptotic Principal Component Analysis.
Recommended Book:

Analysis of Financial Time Series, by Ruey S. Tsay, Wiley Series in Probability and Statistics.

B10. Theory of Computation & Compiler Design (Marks: 100) Credit 4(3-0-1)

Unit I: Regular Languages: Finite automata, Nondeterminism, regular expressions, Non regular Languages. Applications to Compiler Design.

Unit II: Context-Free Languages: Context free grammars, Push down automata, Non-context-free languages. Applications to Compiler Design.

Unit III: The Church Turing Thesis: Turing machines, Variants of Turing machines, the definition of algorithm.

Unit IV: Decidability: Decidable languages, The Halting problem. Reducibility: Undecidable problems from language theory, Mapping reducibility.

Unit V: Complexity Theory: Measuring complexity, The P and NP completeness.

Recommended Texts:

1. M. Sipser - *Introduction to the theory of computation* - Thomson Learning - 2001, Chapter: 1, 2, 3, 4, 5, 7.
2. J. E. Hopcroft, Motwani, J. D. Ullman - *Introduction to Automata Theory, Languages and Computation*, 2nd Edition, Pearson Education, 2001.

MDS-406

Dissertation with Viva-voce (Marks-100) Credit 4(3-2-0)