COURSE STRUCTURE AND EVALUATION SCHEME
BS-MS (MATHEMATICS AND DATA SCIENCE)

Year III, Semester- V

S. No.	Course Type	Course Title	Subject Code	Credits	Periods			Sessional Marks				ESE	Total Marks
					L	T	P	MSE	TA	Lab	Total		
1	PCC	Principles of Data Science	NMA311	4	3	0	2	15	20	15	50	50	100
2	PCC	Artificial Intelligence and Machine Learning	NMA313	4	3	0	2	15	20	15	50	50	100
3	PCC	Modern Algebra	NMA315	3	3	0	0	30	20	-	50	50	100
4	PCC	Topology & Geometry	NMA317	3	3	0	0	30	20	-	50	50	100
5	PCC	Computational Statistics	NMA319	4	3	0	2	30	20	-	50	50	100
6	PCC	Data Science Lab-3	NMA321	2	0	0	4	15	20	15	50	50	100
7	HSMC	Entrepreneurship Development		2	2	0	0	30	20	-	50	50	100
	Total Credits: 22												**700**

Year III, Semester- VI

S. No.	Course Type	Course Title	Subject Code	Credits	Periods			Sessional Marks				ESE	Total Marks
					L	T	P	MSE	TA	Lab	Total		
1	PCC	Deep Learning	NMA312	4	3	0	2	15	20	15	50	50	100
2	PCC	Big Data Analytics	NMA314	4	3	0	2	15	20	15	50	50	100
3	PCC	Multivariate Data Analysis	NMA316	3	3	0	0	30	20	-	50	50	100
4	PCC	Functional Analysis	NMA318	3	3	0	0	30	30	-	50	50	100
5	PCC	Fundamentals of Computing Algorithms	NMA320	3	3	0	0	30	20	-	50	50	100
6	PEC-I			3	3	0	0	30	20	-	50	50	100
7	OEC-I			2	2	0	0	30	20	-	50	50	100
	Total Credits: 22												**700**

PEC-I: *Mathematical Modeling and Numerical Simulation (NMA-322), Statistical Computing (NMA-324, Cloud Computing for Data Science (NMA-326)*

PRINCIPLES OF DATA SCIENCE

Course Code: NMA311 **L-T-P-C: 3-0-2-4**

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have knowledge/skills/competency to-

CO 1	apply measures of association and graphical techniques to perform exploratory data analysis on bivariate and multivariate datasets
CO 2	develop, compare, and interpret regularized regression models including Ridge, Lasso, Elastic Net, and polynomial regression
CO 3	implement and evaluate Logistic Regression, Naïve Bayes, KNN, and LDA using accuracy, precision, recall, F1-score, and AUC-ROC metrics
CO 4	apply and validate K-means, hierarchical, DBSCAN, and DECLUE clustering algorithms using appropriate cluster validation measures
CO 5	design neural network architectures and explain forward-backward propagation and weight update mechanisms

Unit-I Exploratory Data Analysis (EDA)
Analytical Methods: Measures of association for quantitative and qualitative data, Correlation coefficient, Rank correlation coefficient, Bivariate frequency distribution, Measures of association for discrete data.
Graphical Methods: Bivariate and Multivariate graphics.

Unit-II Regression Models:
Simple and multiple linear regression analysis, Polynomial regression, Ridge and Lasso regression, Variants of Lasso regression, Elastic net regression.

Unit-III Classification Models:
Classification problems/models in data science, Logistic regression, Naïve Bayes, K-nearest neighbors, LDA classification evaluation metrics: accuracy, precision recall, F1-score, AUC-ROC.

Unit-IV Cluster Analysis:
Introduction of clustering, k-means/k-medoid, hierarchical clustering, top-down, bottom-up: single linkage, multi-linkage, DBSCAN and DECLUE algorithms, clustering validation.

Unit-V Neural Networks (NN):
Basic concepts, components, NN types with structures, SLP, MLP, ANN, CNN, RNN, LSTM etc. Working of NN system, forward propagation, back propagation and iteration, multi-layer perceptron.

Lab Work: Development and execution of any 2-3 programs from each unit.

Textbook

1. Robert I. Kabacoff, Modern Data Visualization with R, CRC Press.
2. Benjamin S. Baumer, Daniel T. Kaplan, Nicholas J. Horton, Modern data science with R, CRC Press.
3. Introduction to Linear Regression Analysis by Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining (Wiley).
4. Applied Regression Analysis by Norman R. Draper, Harry Smith (Wiley)

Reference Books

5. Joel Grus, Data Science from Scratch, O'Reilly Media, Inc., 2019.
6. Kevin P. Murphy, Machine learning – a Probabilistic Perspective, MIT Press, 2012

ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

Course Code: NMA313 **L-T-P-C: 3-0-2-4**

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have knowledge/skills/competency to-

CO 1	understand AI fundamentals and apply basic informed and uninformed search methods.
CO 2	apply advanced search techniques and use logical knowledge representation for reasoning.
CO 3	build and evaluate supervised ML models for regression and classification.
CO 4	apply unsupervised learning and dimensionality reduction techniques with proper evaluation.
CO 5	formulate reinforcement learning problems and analyse solution methods & algorithms.

Unit-I Artificial Intelligence: Introduction and Fundamentals: Characteristics, Types and Applications, AI Problems and their Types, AI Modelling and Solution Techniques, AI Agents and Environments, Structure of Agents, Problem Solving Agents, Tools and Techniques for AI Problem Solving. Basic Search Algorithms: Informed and Uninformed search.

Unit-II Advanced Search: Constructing Search Trees, Stochastic search, Adversarial search: Minimax search, Alpha-Beta Pruning.
Knowledge Representation and Reasoning: Introduction, Semantic Networks, Semantic Web, Logical Agents, Propositional Logic, First Order logic, Forward Chaining and Backward Chaining with common Approaches and Applications, Expert System.

Unit-III Machine Learning: Overview, Deterministic and Probabilistic Models, Generative and Discriminative Models, Examples.
Supervised Machine Learning - Regression: Decision tree and Random Forest Regression, Information gain, Gini index, Gain ratio, Bayesian Regression, Evaluation Metrics, Tuning and Cross Validation.
Classification: Decision Trees, Random Forest, Support Vector Machines, Confusion Matrix, Log Loss, Gradient Boosting, Classification Metrics: Bias Variance Trade-OFF, Cross-validation Methods such as Leave-One-Out (LOO) Cross-validation.

Unit-IV Unsupervised Machine Learning: Apriori Algorithm for Association Rule Learning, Basic concepts of Clustering Algorithms, Dimensionality Reduction Algorithms: PCA, ICA, Factor Analysis, Evaluation Metrics for Dimensionality Reduction.

Unit-V Reinforcement Learning: Introduction, Types, Elements, Approaches, RL Problem: Components, Formulation, Bellman Equation, Optimality Principle, MDP and POMDP: Components, Mathematical Framework, Solution Methods, Value iteration, Policy Evaluation and Policy Iteration, Common RL Algorithms with their features and relevance.

Lab Work: Development and execution of any 2-3 programs from each unit.

Textbook

1. Elaine Rich, Kevin Knight and Shivashankar B Nair, Artificial Intelligence, Tata McGraw-Hill, 2009.
2. Tom Mitchell, Machine Learning – McGraw Hill Science, 1997.
3. Jeeva Jose, Introduction to Machine Learning using Python, Khanna Book Publishing Co. (P) LTD, 2020.

Reference Books

1. Kevin P. Murphy, Machine learning – a Probabilistic Perspective, MIT Press, 2012
2. Stuart J. Russell and Peter Norvig, Artificial Intelligence: A Modern Approach, Pearson Education Limited, 2022.

MODERN ALGEBRA

Course Code: NMA315 **L-T-P-C: 3-0-0-3**

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have knowledge/skills/competency to-

CO 1	demonstrate an understanding of the concepts of various types of groups and subgroups with their structure and properties.
CO 2	construct homomorphism and isomorphism maps between two groups and understand Sylow theorems to find the inherent properties of a group.
CO 3	examine a given set with two binary operations for different types of rings, ideals, and domains by using relevant axioms.
CO 4	understand polynomial division in an integral domain and apply irreducibility tests to check the irreducibility of a polynomial.
CO 5	construct field extensions and apply Galois theory to determine Galois group.

Unit I: Binary Operation, Group, Examples of Groups, Elementary Properties of Groups, Order of a group and Element, Subgroups, One and Two step subgroup tests, Center of a group, Centralizer of an element, Cyclic Groups, Fundamental theorem of cyclic groups, Permutation Groups, Cosets and Lagrange's theorem. Normal Subgroups, Quotient Groups.

Unit II: Homomorphism, Isomorphism, Cayley's theorem, First Isomorphism Theorem for Group, Finite abelian groups, Fundamental Theorem of Finite Abelian Groups, Internal Direct Products, External Direct Products, Class Equation, Sylow Subgroups, Sylow theorems, Finite Simple Groups.
Special Groups: Pauli group, Clifford group, Lie group.

Unit III: Rings, Examples of Rings, Properties of Rings, Subrings, Integral Domains, Characteristic of a Ring, Ideals, Quotient Rings, Prime and Maximal ideals, Ring Homomorphisms, Properties of Ring Homomorphisms, First Isomorphism Theorem for Rings.

Unit IV: Polynomial Rings, Division Algorithms, Principal ideal domains, Irreducibility Tests for polynomials of order 2 or 3, Mod-p irreducibility tests, Eisenstein Criterion, Irreducible and prime elements in an integral domain, Euclidean domains, Unique Factorization Domains.

Unit V: Fields, Fundamental Theorem of Field Theory, Field extensions, Splitting fields, Perfect Fields, Algebraic extensions, Finite Extension, algebraically closed fields, Finite fields.
Galois Theory: Galois group, Determination of Galois groups, fundamental theorem of Galois theory.

Textbooks

1. Gallian, J. A. (1999) Contemporary Abstract Algebra (4th edition), Narosa Publishing House, New Delhi.
2. Dummit, D.S. and Foote, R.M (2003) Abstract Algebra, John Wiley & Sons.

Reference books

1. Herstein, I. N. (2003) Topics in Algebra (4th edition), Wiley Eastern Limited, New Delhi.
2. Fraleigh, J. B. (2002) A First Course in Abstract Algebra (4th edition), Narosa Publishing House, New Delhi.

TOPOLOGY AND GEOMETRY

Course Code: NMA317 **L-T-P-C: 3-0-0-3**

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have knowledge/skills/competency to-

CO 1	define and analyze metric spaces, different inequalities and various topological spaces.
CO 2	classify spaces by verifying connectedness, compactness, countability, and understand separation axioms.
CO 3	construct manifolds, compute tangent vectors/differentials, different bundles and understand the basics of Lie groups.
CO 4	understand homology groups with computation of degrees for complexes and homotopy theory.
CO 5	understand the concept of topological data analysis vis persistent homology and Mapper algorithms in Python.

UNIT-I Metric Space and Topological Space: Metric Space, Pseudo metric, Limiting Point, Interior Point, Open Set, Closed Set, Open Sphere, Closed Sphere, Holder's Inequality, Minkowski's Inequality, Cauchy-Swartz Inequality, Convergent Sequence, Cauchy Sequence in Metric space, Complete Matric Space.
Topological Space, Closure, Interior and Exterior of a Set, Discrete and Indiscrete Topology, Union and intersection of two Topologies, Weak and Strong Topology, Basis and Sub-basis for a topology, Subspace topology, Order topology, Product topology, Quotient topology, Metric topology, Homeomorphism, Continuity

UNIT-II Connectedness, Compactness and Separation axioms: Connectedness, Examples, Local connectedness, Path connectedness, connected subsets of real line, Compact Spaces, Examples, locally compacts spaces, sequential compactness, limit point compactness, compact subsets of real line,
Countability axioms, First and second countable spaces, separable and Lindelof spaces, Separation axioms, Regular & completely regular space, normal spaces, Urysohn's lemma

UNIT-III Differentiable Manifolds: Differentiable Manifolds, Local Coordinates, Induced Structures and Examples, Tangent Vectors and Differentials, Sard's Theorem and Regular Values, Vector Fields and Flows, Tangent Bundles, Embedding in Euclidean Space, Tubular Neighborhoods and Approximations, Classical Lie Groups, Fiber Bundles, Induced Bundles and Whitney Sums, Transversality.

UNIT-IV Homology Theory: Homology Groups, Betti number, The Zeroth Homology Group, The First Homology Group, Functorial Properties, Homological Algebra, Axioms for Homology, Computation of Degrees, CW-Complexes, Conventions for CW-Complexes, Cellular Homology, Cellular Maps, Products of CW-Complexes, Euler's Formula, Homology of Real Projective Space, Singular Homology, The Cross Product, Subdivision, Simplicial Complexes, Simplicial Maps, Introduction and applications of homotopy.

UNIT-V Computational Topology: Persistent Homology: Introduction, working algorithms, Illustration with examples using Python.
Topological Data Analysis: Mapper Algorithm, Illustration of Mapper algorithm using Python, Reeb Graphs and Morse Theory, Statistical Inference in TDA.

Textbooks

1. Topology by James R. Munkres, 2nd edition, Pearson.
2. Topology and Geometry by Glen E. Bredon, Springer.

Reference Books

1. J Milnor. Topology from the Differentiable Viewpoint. rev. ed. Princeton University Press, 1997.

COMPUTATIONAL STATISTICS

Course Code: NMA319 **L-T-P-C: 3-0-2-4**

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have knowledge/skills/competency to-

CO 1	use Monte Carlo inference techniques such for probabilistic estimation.
CO 2	apply probabilistic to perform exact inference using algorithms like Belief Propagation and Junction Tree.
CO 3	develop and apply MCMC algorithms for Bayesian inference and model estimation.
CO 4	apply Bootstrapping for model evaluation and statistical inference.
CO 5	apply Density Estimation and Smoothing techniques for statistical inference.

UNIT I Monte Carlo Inference:
Inverse CDF method. Rejection-Sampling, Adaptive Rejection sampling, Importance Sampling, Adaptive Importance Sampling, Resampling, Particle Filtering, Sequential Importance Sampling.

UNIT II Gaussian Processes:
Introduction, Markov and Hidden Markov models, Markov Random Fields, Conditional Random Fields, Expectation-Maximization (EM) Algorithm: EM Variants, Exact Inference for Graphical models: Enumeration, Belief Propagation (BP), Variable Elimination Junction-Tree Algorithms, Bayesian Networks.

UNIT III Markov Chains Monte Carlo (MCMC) Inference:
Metropolis-Hastings, Gibbs Sampling implementation, MCMC in Bayesian Inference, Adaptive MCMC, Reversible Jump MCMC, other metropolis-Hastings Algorithms.

UNIT IV Bootstrapping:
Introduction, Applications, working parametric and non-parametric bootstrapping, Bootstrap confidence Intervals and Significance Tests; Bootstrapping for Regression. Bootstrapping in classification tasks with decision tree, Bootstrapping Bais correction, Bootstrap Inference, Implementing Bootstrapping in Python.

UNIT V Density Estimation and Smoothing:
Parametric and non-parametric density estimation, Techniques Kernel Density Estimation, Density Estimation in Higher Dimensions, Kernel smoothing, spline smoothing.
NPDE: measures of performances, KDE: Non-kernel methods, Multivariate Methods, Bivariate Smoothing, Multivariate Smoothing.

Lab work: Implementation of any 2-3 techniques from each **UNIT** (selected by Instructor) using R/Python.

Text Books

1. Geof H. Givens and Jennifer A. Hoeting, Computational Statistics, John Wiley & Sons, Inc., 2012.
2. "Simulation" by Sheldon M. Ross (Academic Press, Fourth Edition), 2006.
3. Bootstrap from "An Introduction to the Bootstrap" by B. Efron and R.J. Tibshirani (Chapman and Hall), 1994.

Reference Books

1. **(For lab only)** –G. James, D. Witten, T. Hastie, R. Tibshirani - An introduction to statistical learning with applications in R, Springer, 2013.
2. Cluster Analysis from "Cluster Analysis" by B.S. Everitt, S. Landau, M. Leese, D. Stahl, (Wiley), 2011.
3. E.M. Algorithm from "The EM Algorithm and. Extensions" by G. M. McLachlan and T. Krishnan, (Wiley), 1997.

DATA SCIENCE LAB – 3

Course Code: NMA321 **L-T-P-C: 0-0-4-2**

1 Develop a database management system using Python/SQL/Power BI.
2 Perform exploratory data analysis (EDA) using Python/SQL/Power BI.
3 Implement informed search: A* and Hill Climbing algorithm in Python/R.
4 Implement the Minimax algorithm with Alpha-Beta pruning in Python/R.
5 Implement the Expectation-Maximization (EM) algorithm applied to Gaussian mixture models in Python/R.
6 Compare different variational inference techniques using their Python/R codes.
7 Compare different regression models using hyper parameter tuning and performance metrics with Python/R codes.
8 Compare different classification models using hyper parameter tuning and performance metrics with Python/R codes.
9 Compare different Monte Carlo (MC) inference methods using Python/R.
10 Compare MC inference techniques: Metropolis-Hastings and Gibbs sampling algorithms using Python/R codes.
11 Compare different clustering algorithms with their Python/R codes.
12 Compare different dimensionality reduction algorithms with their Python/R implementation.
13 Illustrate 4 non-parametric density estimation methods through their Python/R implementation.
14 Implement at least 3 reinforcement learning algorithms in Python/R.
15 Implement the concept of persistent homology in Python/R.
16 Implement Mapper algorithm of Topological data analysis using Python.
17 Build a neural network with real-life example and Python/R code.
OR
Implement multilayer perceptron using TensorFlow.

LAB WORK: 10-12 codes are to be developed and executed out of the following list.

Reference Books

1. **Christopher M. Bishop**, *Pattern Recognition and Machine Learning*, Springer, 2006.
2. **Sebastian Raschka and Vahid Mirjalili**, *Python Machine Learning (3rd Edition)*, Packt Publishing, 2020.
3. **Andrew Rosen**, *Topological Data Analysis with Python: Harness the Power of TDA*, Packt Publishing, 2024.

DEEP LEARNING

Course Code: NMA312 **L-T-P-C: 3-0-2-4**

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have knowledge/skills/competency to-

CO 1	apply optimization algorithms and strategies for effectively training deep neural networks.
CO 2	use regularization techniques to reduce overfitting and improve deep model generalization.
CO 3	explain and implement key concepts and operations of Convolutional Neural Networks.
CO 4	apply sequence modeling techniques using RNNs, LSTMs, and related architectures for sequential data.
CO 5	understand and apply Transformer architectures and Autoencoders for modern deep learning tasks.

UNIT-I Optimization for Training Deep Models: Optimization Problems in Deep Learning, 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-II Regularization for Deep Learning: Overfitting, Regularization and its need, Regularization Techniques: L2 and L1 Regularization, Dropout, Data Augmentation, Early Stopping, Batch Normalization.

UNIT-III Convolutional Neural Networks (CNNs): Convolution Operation, Sparse Interactions, Parameter Sharing, Equivalence, Pooling: Variants of convolution: Strided, Tiled, Transposed and Dilated (Altrous) Convolutions; Key Concepts of a CNN and CNN Learning.

UNIT-IV Sequence Modeling: Recurrent and Recursive Nets
Unfolding Computational Graphs, Recurrent Neural Networks (RNNs), Bidirectional RNNs (BRRNs), Encoder-Decoder: Sequence-to-Sequence Architectures (RNNs), Deep Recurrent Networks, Challenge of Long-Term Dependencies, Long Short-Term Memory (LSTMs), Optimization for Long-Term Dependencies, Explicit Memory.

UNIT-V: Transformers and Autoencoders: Introduction to Transformers, Self-Attention mechanism, Types, Architecture and Workflow steps of Transformer models, Transformer-based models Applications, Introduction and types of AEs with their preferrable uses, architectures and working/learning Processes.

Textbooks

1. Ian Goodfellow, Yashua Bengio, and Aaron Courville, Deep Leaning, The MIT Press, 2016.
2. Kevin P. Murphy, Machine Leaning-a probabilistic prospective MIT Press,2012.

Reference Book

1. Charu C. Aggarwal, Neural Networks and Deep Learning, Springer, 2018.

BIG DATA ANALYTICS

Course Code: NMA314 **L-T-P-C: 3-0-2-4**

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have knowledge/skills/competency to-

CO 1	understand the fundamentals of Big Data Analytics, its architecture, technologies, and real-world applications in diverse domains.
CO 2	explain the Hadoop ecosystem, its components, and implement distributed data processing using HDFS and MapReduce.
CO 3	apply NoSQL concepts and manage large-scale data using databases such as MongoDB and Cassandra.
CO 4	develop data analysis workflows using MapReduce, Hive, and Pig for efficient data querying and transformation.
CO 5	utilize Apache Spark and MLIB for scalable data analysis, machine learning, and data visualization tasks.

UNIT-I: INTRODUCTION TO BIG DATA ANALYTICS

Introduction, Big Data Scalability and Parallel Processing, Designing Data Architecture and Sources, Quality, Pre-Processing and Storing, Big Data Technologies, Data Storage and Analysis, Big Data Analytics: Applications and Case studies, Convergence in Big Data Analytics, Mobile Business Intelligence, Web Analytics, Edge Analytics, Crowd Sourcing Analytics, Inter- and Trans-Firewall Analytics.

UNIT-II: INTRODUCTION TO HADOOP

Introduction, Hadoop and its Ecosystem, Scaling Out, Data Formats, Steps of Data Analysis with Hadoop, Design of Hadoop Distributed File System (HDFS), HDFS Concepts, Java Interface, Data Flow, Hadoop I/O, Data Integrity Compression Serialization Avro, File-based Data Structures, MapReduce Framework and Programming Model, Hadoop YARN, Hadoop Ecosystem Tools, Cassandra-Hadoop Integration.

UNIT-III: NoSQL, BIG DATA MANAGEMENT, MONGODB, AND CASSANDRA

Introduction, NoSQL data store, NoSQL Data Architecture Patterns, Types of NoSQL Databases with Features and Popular Databases and use cases, NoSQL to manage Big Data, Shared-Nothing Architecture for Big Data Tasks, Mongo DB Database, Cassandra Database.

UNIT-IV: MAPREDUCE, HIVE AND PIG

MapReduce: Introduction, MapReduce Map Tasks, Reduce Tasks and MapReduce Execution Composing MapReduce for calculations and Algorithms, MapReduce Types, MapReduce Formats, MapReduce with Examples, MapReduce Programming models.

Hive: Architecture, Installation and Comparison with RDBMS, Data Types and File Formats Data Model, Hive Integration and Workflow Steps, Built-in functions, User defined Functions, HiveQL Data Definition Language, HiveQL Data Manipulation Language, HiveQL for Querying, The data aggregation,

Pig: Apache Pig-Grunt Shell, Installing Pig, Pig Latin Data Model, Pig Latin, Developing and Testing Pig Latin Scripts.

UNIT-V: SPARK AND BIG DATA ANALYTICS

Introduction, Spark, Introduction to Data Analysis with Spark, Downloading Spark and Programming using RDDs and MLIB. Data ETL Process, Introduction to Analytics, Reporting and Visualizing.

Textbook

1. Seema Acharya, Subhashini Chellappan, Big Data Analytics, 1st Edition, Wiley, 2015.

Reference Books

1. Boris Lublinsky, Kevin t. Smith, Alexey Yakubovich, Professional Hadoop Solutions, 1st Edition, Wrox, 2013.
2. Chris Eaton, Dirk Deroos et. al., Understanding Big data, Indian Edition, McGraw Hill, 2015.
3. Tom White, HADOOP: The definitive Guide, 3rd Edition, O Reilly, 2012**.**
4. Vignesh Prajapati, "Big Data Analytics with R and Hadoop", 1st Edition, Packet Publishing Limited, 2013.

MULTIVARIATE DATA ANALYSIS

Course Code: NMA316 **L-T-P-C: 3-0-0-3**

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have knowledge/skills/competency to-

CO 1	understand and apply concepts of random vectors, covariance structures, and similarity measures for multivariate data representation and analysis.
CO 2	analyse multivariate normal distributions and Wishart Distribution.
CO 3	understand the distribution of Simple, Partial, Multiple correlation coefficients
CO 4	apply dimensionality reduction and interrelationship analysis techniques such as PCA, Canonical Correlation, and Discriminant Analysis.
CO 5	implement General Linear Models (GLMs) such as multivariate regression, MANOVA, and MANCOVA.

UNIT-1 Random Vectors and Matrices

Definitions, Joint, marginal and conditional probability distributions, Operations, Expectation and Properties, Covariance, Correlations, Orthogonality, Independence

Multivariate Data Organization and Metrics: multivariate data matrix

Summary measures: Geometry of Shape, Euclidean distance, Hamming or Manhattan distance, General Minkowski distance, Mahalanobis distance, Common measures of similarity, Common Scaling Techniques.

UNIT-2 Multivariate Distributions

Multivariate Normal Distribution (MND), Properties of MND, Estimation of Parameter of MND, Marginal and Conditional Distributions, Wishart Distribution.

UNIT-3 Correlation Coefficients

Distribution of Simple, Partial, Multiple correlation coefficients and related tests.

UNIT-4 Multivariate Hypothesis Testing

Discriminant Analysis, Principal Component Analysis, Canonical Correlation Analysis, testing of hypothesis related to mean vector, Hotelling T^2 statistics.

UNIT-5 General Linear Models

One-way MANOVA and MANCOVA, Two-Way MANOVA and MANCOVA, Exploratory Factor Analysis, Confirmatory Factor Analysis, Wilks' Lambda.

Textbooks

1. Johnson, R. A., & Wichern, D. W. (2019). Applied multivariate statistical analysis (7th ed.), Pearson.
2. Anderson, T. W. (2003). An introduction to multivariate statistical analysis (3rd ed.). Wiley.

Reference Books

1. Joseph F. Hair Jr., William C. Black, Barry J. Babin, & Rolph E. Anderson (2021), Multivariate Data Analysis (8th ed.), Cengage Learning.

FUNCTIONAL ANALYSIS

Course Code: NMA318 **L-T-P-C: 3-0-0-3**

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have knowledge/skills/competency to-

CO 1	understand the basics of normed spaces and continuity of linear maps.
CO 2	understand the concept of Hahn Banach extension and Banach spaces.
CO 3	the openness and closedness of linear maps in Banach spaces and describe weak and weak* topologies in dual spaces.
CO 4	construct orthonormal bases in Hilbert spaces using the Gram-Schmidt process and apply the Riesz Representation Theorem to characterize continuous linear functionals.
CO 5	classify bounded operators on Hilbert spaces and apply the Spectral Theorem to decompose self-adjoint operators into spectral measures.

Unit I: Normed spaces, Examples of Normed Spaces, Subspaces of Normed Spaces, Quotient Normed Spaces, Riesz Lemma, Finite-Dimensional Normed Spaces, Convex Subsets of Normed Spaces, Linear Maps Between Normed Spaces, Continuity of linear maps, Various Criterion for Continuity of Linear Maps, Operator Norm of Bounded Linear Maps.

Unit II: Hahn Banach theorems: Geometric and extension forms and their applications, Uniqueness of the Hahn-Banach Extension, Banach Spaces, Subspaces of Banach Spaces, Quotient Banach Spaces, Product of Banach Spaces, Schauder Basis, Uniform Bounded Principle and its Applications, Banach-Steinhaus Theorem.

Unit III: Closed Maps, Closed graph theorem, Linear Projections, Open Maps, Quotient Maps, Open Mapping Theorem and its Applications, Bounded Inverse Theorem.
Dual spaces and adjoint of an operator: Duals of classical spaces, weak and weak* convergence.

Unit IV: Hilbert spaces: Inner product spaces, orthonormal set, Gram Schmidt orthonormalization, Bessel inequality, Orthonormal basis, Separable Hilbert spaces. Projection and Riesz representation theorem: Orthonormal complements, orthogonal projections, projection theorem, Riesz representation theorem.

Unit V: Bounded operators on Hilbert spaces: Adjoint, normal, unitary, self-adjoint operators, compact operators, eigen values, eigen vectors, Banach algebras. Spectral theorem: Spectral theorem for compact self-adjoint operators, Spectral theorem for bounded self-adjoint operators.

Textbooks

1. B.V. Limaye, Functional Analysis, New Age International Ltd. (Second Edition).
2. J. B. Conway, A course in functional Analysis, GTM 96, Springer-Verlag` 1990.

Reference Books

1. E. Kreyzig, Introductory Functional Analysis with Applications, John Wiley & Sons, New York, 1989.
2. W. Rudin, Functional Analysis, Tata McGraw-Hill Publishing Company Ltd., New Delhi.
3. K. Yoshida, Functional Analysis, GTM,123, Springer-Verlag, 1980.

FUNDAMENTALS OF COMPUTING ALGORITHMS

Course Code: NMA320 **L-T-P-C: 3-0-0-3**

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have knowledge/skills/competency to-

CO 1	understand and apply foundational quantum algorithms and principles of quantum computation to solve problems using quantum oracles, QFT, Grover's algorithm, and Shor's algorithm.
CO 2	analyze NP-hard and NP-complete problems and evaluate approximation techniques, including PTAS, FPTAS, and probabilistically good algorithms.
CO 3	design and analyze parallel algorithms using the PRAM computational model for tasks such as selection, sorting, graph problems, and convex hull computation.
CO 4	apply mesh-based parallel computational models to implement efficient algorithms for routing, sorting, graph problems, and geometric computations.
CO 5	develop and evaluate algorithms on hypercube architectures, including routing, merging, sorting, graph algorithms, and convex hull computation.

UNIT-I: Quantum Algorithms
Axioms of quantum mechanics Quantum computer science Key concepts: Quantum oracle, Foundational Quantum Algorithms: Deutsch-Jozsa, Bernstein-Vazirani Algorithms, Quantum Fourier Transform (QFT), Grover's Algorithm, Quantum Algorithms & Shor's Algorithm

UNIT-II: NP-HARD AND NP-COMPLETE PROBLEMS AND APPROXIMATION ALGORITHMS
Basic Concepts, Cook's Theorem, NP-Hard Graph Problems, NP-Hard Scheduling Problems, NP-Hard Code Generation Problems, NP-Hard Approximation Algorithms: Introduction, Absolute Approximations, ε-Approximation Algorithms, Polynomial time Approximation schemes, fully polynomial time Approximation schemes, probabilistically good algorithms.

UNIT-III: PRAM ALGORITHMS
Introduction, Computational model, Fundamental Techniques and Algorithms, Selection, merging, Sorting, Graph Problems, Computing the Convex Hull. Lower Bounds

UNIT-IV: MESH ALGORITHMS
Introduction, Computational model, Packet Routing, Fundamental Algorithms, Selection, merging, Sorting, Graph Problems, Computing the Convex Hull

UNIT-V: HYPERCUBE ALGORITHMS
Introduction, Computational model, PPR Routing, Fundamental algorithms, Selection, merging, Sorting, Graph Problems, Computing the Convex Hull

Textbook

1. Ellis Horowitz, Sartaj Sahni, Sanguthevar Rajasekaran, "Computer Algorithms"
2. Ellis Horowitz, Sartaj Sahani, "Fundamentals of Computer Algorithms".

Reference Books

1. Nielsen, M. A., & Chuang, I. L. (2010). Quantum computation and quantum information (10th anniversary ed.). Cambridge University Press.
2. Garey, M. R., & Johnson, D. S. (1979). Computers and intractability: A guide to the theory of NP-completeness. W. H. Freeman.
3. JáJá, J. (1992). An introduction to parallel algorithms. Addison-Wesley.

CLOUD COMPUTING FOR DATA SCIENCE

Course Code: NMA326 **L-T-P-C: 3-0-0-3**

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have knowledge/skills/competency to-

CO 1	explain the fundamental concepts of cloud computing, including its evolution, architecture and delivery mechanisms.
CO 2	demonstrate an understanding of virtualization technologies and differentiate between various virtualization types, cloud deployment models.
CO 3	analyse cloud storage systems, key-value databases, cloud data warehouses along with their applications.
CO 4	evaluate cloud-based machine learning services and platforms along with their benefits and limitations.
CO 5	design, train, deploy, and monitor machine learning models on cloud platforms by selecting appropriate ML services, handling ETL/ELT pipelines etc.

UNIT-I: INTRODUCTION TO CLOUD COMPUTING
Definition, Evolution of Cloud Computing (from mainframes to clouds), service-Oriented Architecture, Web Services, Grid Computing, Utility computing, characteristics of a cloud computing; cloud computing Architecture: Basic Components: front-end platform, back-end platform, Networking, Cloud-based delivery; Cloud Service Models: Software as a Service (SaaS), Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Continuous delivery using PaaS.

UNIT-II: VIRTUALIZATION AND CLOUD COMPUTING
Introduction to Virtualization, types of Virtualizations, Application of Virtualization, Network Virtualization, Server Desktop Virtualization, Storage Virtualization, Server Virtualization, Data Virtualization, Cloud Deployment models: Public, Private, Community, Hybrid Role of Cloud Computing in Data Science, Advantages of Cloud Computing in Machine Learning.

UNIT-III: CLOUD STORAGE
Introduction, Benefits of using Cloud storage, Use cases (Backup, Archiving, Disaster recovery, Data Processing, Content delivery); Cloud storage system: Block-Based, File-Based, Object-Based storages; Key-Value databases: Introduction, features, Limitations, Batch data and Streaming data in machine learning, Cloud data warehouse- AWS Redshift, Various Cloud-based tools used for Data Science on ML - GCP Big Query.

UNIT-IV: CLOUD COMPUTING FOR DATA SCIENCE
Machine Learning in the Cloud; Benefits and Limitations; Types of cloud- based machine learning services: Artificial Intelligence as a service (AIaaS), GPU as a Service (GPUaaS); Introduction to various ML systems and benefits of using managed ML platforms; Cloud Machine Learning Platform: AWS SageMaker, Azure Machine Learning Studio, Google Cloud

AutoML.

UNIT-V: TRAINING AND DEPLOYMENT OF ML ON CLOUD

Factors for Selection of Cloud Machine Learning Platform, support for ETL or ELT Pipelines, support for Scale-Up and Scale-Out Training, Support for Machine Learning Frameworks, Pre-Tuned AI Services, monitor Prediction Performance: Training Machine Learning Projects in the Cloud; Steps to Train Machine Learning Project in the Cloud, Identify and Understand Data Sources, Engineer the Features, Train and Validate model, Deploy and Monitor Model.

Textbooks

1. Anand Nayyar, "Handbook of Cloud Computing", BPB Publication
2. Toby Velte, Anthony Velte, Robert C., "Cloud Computing: A Practical Approach", McGraw Hill Professional

Reference Books

1. Noah Gift, Alfredo Deza, "Cloud Computing for Data Analysis", Pragmatic AI Labs
2. Valliappa Lakshmanan, "Data Science on the Google Cloud Platform: Implementing End-to-End Real Time Data Pipelines", O'Reilly
3. Abhishek Mishra, "Machine Learning in the AWS Cloud: Add Intelligence to Applications with Amazon SageMaker", Wiley Publication