Revised Syllabus

M.Sc. CS (Big Data Analytics)

(2 Years Programme)

Syllabus Proposed to be implemented from July-2024 onwards

Department of Data Science and Analytics

School of Mathematics, Statistics and Computational Sciences

Central University of Rajasthan

NH-8 Jaipur- Ajmer Highway, Bandar Sindri

Kishangarh -305817

District-Ajmer, Rajasthan

Website: www.curaj.ac.in

M.Sc. CS (Big Data Analytics)

Program Outcomes (POs)

PO1	Data Science knowledge: Apply the knowledge of machine learning, statistics, and data science fundamentals to the solution of complex real-time problems.
PO2	Problem analysis: Identify, formulate, research literature, and analyze complex data analytics problems to provide conclusions using principles of machine learning, statistics, and data science fundamentals.
PO3	Design/development of solutions: Design solutions for complex real-time problems and design system components or processes that meet the specified needs for the industry and society.
PO4	Conduct research on complex problems: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
PO5	Modern data analytics tools: Create, select, and apply appropriate techniques, resources, and modern software tools including prediction and modeling to complex data with an understanding of the outcomes.
PO6	Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the data analytics practices.
PO7	Communication: Communicate effectively on complex data analytics activities with the data science community, and to improve professional communication skills.
PO8	Life-long learning: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological and information change.

Program Specific Outcomes (PSOs)

PSO1	To apply creativity in support of the design, simulation, implementation and inference of existing and advanced technologies in data analytics.
PSO2	To participate & succeed in data science oriented jobs/competitive examinations that offer inspiring & gratifying careers.
PSO3	To recognize the importance of professional developments by pursuing higher studies and positions.

LEVEL-4

Semester- I: M.Sc. CS (BDA)

<table>
<tr><td colspan="3">Course: FUNDAMENTALS OF PROBABILITY AND STATISTICS
Code: MBD401</td></tr>
<tr><td><u>TEACHING SCHEME:</u></td><td><u>EXAMINATION SCHEME:</u></td><td><u>CREDITS ALLOTED:</u></td></tr>
<tr><td>Theory: 3
Tutorial: 0
Practical: 1 (2 Hours)</td><td>Internal Assessment: 40 Marks
CIA-I: 20 Marks (Unit I & II)
CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce)

End Semester Examination: 60 Marks (Unit I- IV)</td><td>4</td></tr>
<tr><td colspan="3"></td></tr>
<tr><td colspan="3">Course Prerequisites: Basic Calculus</td></tr>
<tr><td colspan="3">Course Objectives: To develop the knowledge of statistics and probability distribution.</td></tr>
<tr><td colspan="3">Course Outcomes: After completing the course, the student shall be able to:

● Explain the concept of basic statistics tools and their application in data science.
● Express the features of discrete and continuous random variables.
● Solve the problems associated with the discrete/continuous distributions & sampling distributions.
● Understand applications of CLT in data science.</td></tr>
<tr><td colspan="3">Course Content:</td></tr>
<tr><td>Unit-I</td><td>Data and its Types, Frequency table, histogram, measures of location, measures of spread, skewness, kurtosis, percentiles, box plot. Probability function, conditional probability, Bayes theorem, random variables, discrete and continuous random variables, discrete distributions: binomial, Poisson, geometric, hypergeometric, negative binomial, etc.</td><td>12T + 4P
(20 hours)</td></tr>
<tr><td>Unit-II</td><td>Continuous distributions: uniform, normal, exponential, gamma, Weibull, Pareto, lognormal, Laplace, Cauchy, logistic distributions; properties and applications. Functions of random variables and their distributions using (i) transformation of rv (ii) MGF and (iii) method of Jacobian of transformation.</td><td>10T + 3P
(18 hours)</td></tr>
<tr><td>Unit-III</td><td>Concept of a sampling distribution. Sampling distributions of t, $\chi 2$ and F (central and noncentral), their properties and applications. Goodness of fit and Categorical data analysis, Contingency Table, Measure of Associations. Test for Correlation.</td><td>12T + 4P
(19 hours)</td></tr>
</table>

Unit-IV	Compound, truncated and mixture distributions. Convolutions of two distributions. Order statistics: their distributions and properties. Joint, marginal and conditional distribution of order statistics. The distribution of range and median. Convergence of RV, Central Limit Theorem, Weak law of large numbers.	**11T + 4P (20 hours)**

Text Book:

1. V. K. Rohatgi and A.K. MD. Ehsanes Saleh, "An Introduction to Probability Theory and Mathematical Statistics", John Wiley and Sons, 2nd Edition, 2001.
2. Sheldon M. Ross, "Introduction to Probability and Statistics for Engineers and Scientists", Academic Press, 5th Edition, 2014.

Reference Books:

1. Paul L. Mayer, "Introductory Probability and Statistical Applications", Addison-Wesley, 2nd Edition, 1970.
2. George Casella and Roger L. Beger Saleh, "Statistical Inference", Duxbury Press, 2nd Edition, 2001.

E-Resources:

1. https://archive.nptel.ac.in/courses/111/105/111105090/

Course: DATABASE MANAGEMENT SYSTEM
Code: MBD402

TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**
Theory: 3 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**4**

k

Course Prerequisites: No

Course Objectives: To develop the knowledge of Database and SQL queries

Course Outcomes: After completing the course, the student shall be able to:

- Draw Entity-Relationship diagrams to represent simple database application scenarios
- Write SQL queries for a given context in a relational database.
- Discuss normalization techniques with simple examples.
- Learn the concept of parallel and distributed database.

Course Content:

Unit-I	Basic Concepts: Need, purpose and goal of DBMS, Three-tier architecture, ER Diagram, data models- Relational, Network, Hierarchical. Database Design: Conceptual database design, concept of physical and logical databases, data abstraction and data independence, data aggregation.	**12T + 3P (18 hours)**
Unit-II	Relational Database: Relations, Relational Algebra, Theory of Normalization, Functional Dependency, Primitive and Composite data types.	**10T + 4P (18 hours)**
Unit-III	Application Development using SQL: DDL and DML, Host Language interface, embedded SQL programming, Stored procedures and triggers and views, Constraints assertions. NoSQL Databases-MongoDB	**11T + 4P (19 hours)**

Unit-IV	Internal of RDBMS: Physical data organization in sequential, indexed random and hashed files. Inverted and multilist structures, B trees, B+ trees, Query Optimization, Join Algorithm, Statistics and Cost Base optimization. Parallel and distributed database. Transaction Processing, concurrency control, and recovery management. Transaction model properties and state Serializability. Lock base protocols, two-phase locking.	**12T + 4P (20 hours)**

Text Books:

1. Elmasri Ramez and Navathe Shamkant, "Fundamentals of Database System", Pearson, 7th Edition, 2017.
2. Abraham Silberschatz, Henry F. Korth and S. Sudarshan, "Database System Concepts", McGraw Hill, 7th Edition, 2021.

Reference Book:

1. Ramon A Mata-Toledo and Pauline K. Cushman, "Database Management Systems", Schaum's Outlines, 1st Edition, 2014.

E-Resources:

1. https://archive.nptel.ac.in/courses/106/105/106105175/

Course: FOUNDATIONS OF DATA SCIENCE
Code: MBD403

TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**
Theory: 3 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**4**

Course Prerequisites: No

Course Objectives: To develop the knowledge of data collection, processing, analysis and visualization.

Course Outcomes: After completing the course, the student shall be able to:

- Cover the technical pipeline from data collection, to processing, analysis, and visualization.
- Understand the concept of graphs.
- Understand the concept to represent data into high dimensional space.
- Achieve the practical exposure of dimensionality reduction.

Course Content:

Unit-I	Data Collection: Concepts of measurement, scales of measurement, design of data collection formats with illustration, data quality and issues with data collection systems with examples from business, Data preparation, Data pre-processing, Data Integration, Data Transformation.	**12T + 3P (18 hours)**
Unit-II	Graph: Introduction to Graph, Application of Graphs, Finite and Infinite Graphs, Incidence & Degree, Isolated Vertex, Pendant Vertex, Null Graph. Paths & Circuits: Isomorphism. Subgraphs, Walks, Paths, and Circuits, Connected Graphs, Disconnected Graphs, and Components, Euler Graphs, Hamiltonian Paths and Circuits. Trees & Fundamental Circuits: Trees, Some properties of trees, Pendant Vertices in a Tree, Distance and Centers in a Tree, Rooted and Binary Trees, Spanning Trees. Cut-Sets & Cut-Vertices: Cut-Sets, Some Properties of a Cut-Set, All Cut-Sets in a	**10T + 4P (18 hours)**

	Graph, Fundamental Circuits and Cut-Sets, Connectivity and Separability, Network Flows. Random Graphs: Large graphs, G(n,p) model, Giant Component, Connectivity, Cycles, Non-Uniform models, Applications.	
Unit-III	High Dimensional Space: Properties, Law of large number, Sphere and cube in high dimension, Generation points on the surface of sphere, Gaussians in high dimension, Random projection, Practical applications of t-SNE, Multidimensional Scaling, Manifold Learning, UMAP.	**11T + 4P (19 hours)**
Unit-IV	The General Models for Massive Data Problems: Topic Models - Non-Negative Matrix Factorization, Latent Dirichlet Allocation (LDA). Singular Value Decomposition, Dimensionality Reduction: Principal Component Analysis, Independent Component Analysis, Linear Discriminant Analysis and Practical implementation.	**12T + 4P (20 hours)**

Text Books:

1. Avrim Blum, John Hopcroft and Ravindran Kannan, "Foundations of Data Science", Cambridge University Press, 1st Edition, 2020.
2. Narsingh Deo, "Graph Theory", PHI, 1st Edition, 2017.

Reference Books:

1. B. Uma Maheswari and R. Sujatha, "Introduction to Data Science: Practical Approach with R and Python", Wiley India Pvt Ltd., 1st Edition, 2021.

E-Resources:

1. https://onlinecourses.swayam2.ac.in/imb24_mg31/preview

Course: LINEAR ALGEBRA AND ITS APPLICATIONS
Code: MBD404

TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**
Theory: 2 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**

Course Prerequisites: No

Course Objectives: To develop the knowledge of matrices, Linear transformations, Eigenvalues and vectors.

Course Outcomes: After completing the course, the student shall be able to:

- Solve systems of linear equations using multiple methods, including Gaussian elimination and matrix inversion.

- How to apply all of the essential vector and matrix operations for machine learning and data science
- Carry out matrix operations, including inverses and determinants.
- Implement linear algebra concepts in scientific programming languages, and apply linear algebra concepts to real datasets

Course Content:

Unit-I	Matrices, System of Linear Equations, Elementary Row Operations, Row Reduced Matrices, Invertible Matrices. Fields, Vector spaces, Subspaces, Linear Combinations, Linear span, Linear dependence and Linear independence of vectors, Basis and Dimension, Ordered Basis, Finite dimensional vector spaces.	**7T + 4P (15 hours)**
Unit-II	Linear transformations and their representation as matrices, Kernel and Image of a linear transformation, Rank and Nullity Theorem, Change of Basis, Eigenvalues and Eigenvectors of a linear transformation (matrices), Eigendecomposition of symmetric matrices, Eigendecomposition of singular matrices,Characteristic polynomial and minimal polynomial, Diagonalization of linear operators, invariant subspaces, Jordan	**8T + 3P (14 hours)**
Unit-III	Normed space, Inner product spaces, Cauchy-Schwarz-inequality, Orthogonal vectors, Orthonormal sets and bases, Projections, decompose vector to orthogonal components, Gram-Schmidt algorithm. Positive Definite and Positive Semi-definite Matrices, Rayleigh quotient, Linear functional on an Inner Product Space and Adjoint of a Linear operator, Self-adjoint operators, Normal Operators, Spectral theory of matrix, Spectral Decomposition, The Trace Operator. Quadratic forms.	**7T + 4P (15 hours)**
Unit-IV	Numerical methods: Bisection Method, Newton Raphson, Steepest Ascent, method of conjugate gradients; Direct and iterative methods for solving a linear system of equations: Gaussian elimination, LU factorization, Cholesky method, QR factorization, Householder's matrices, Jacobi's method, Gauss-Seidel method, successive over-relaxation methods (SOR), conditioning of a problem.	**8T + 4P (16 hours)**

Text Book:

1. Sheldon Axler, "Linear Algebra Done Right (Undergraduate Texts in Mathematics)", Springer, 3rd Edition, 2015.
2. Lloyd N. Trefethen and David Bau III, "Numerical Linear Algebra", 1st Edition, SIAM,1997.

Reference Books:

1. Gilbert Strang, "Linear Algebra and learning from data", Wellesley-Cambridge Press, 2019.
2. Mike Cohen, "Practical Linear Algebra for Data Science", O'Reilly, 1st Edition, 2022

E-Resources:

1. https://archive.nptel.ac.in/courses/111/106/111106051/

Course: PROGRAMMING IN PYTHON
Code: MBD405

TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS**
Theory: 2 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**

Course Prerequisites: No

Course Objectives: To develop the knowledge of Python Programming and Data Structure

Course Outcomes: After completing the course, the student shall be able to:

- Understand the basic concepts of Python programming
- Create and use functions, including built-in functions, user-defined functions, and lambda functions.
- Employ various Python data structures such as matrices, data frames, arrays, tuples, lists, and dictionaries.
- Demonstrate proficiency in file handling, including opening, reading, and writing structured and text files.

Course Content:

Unit-I	Introduction to Python: Basics of Python, Identifiers, literals, keywords, indentation, Strings, input and output statements, basic operators (arithmetic, logical, relational, bitwise and assignment), global variables, iterations-for, while, do-while, Control structures-exit function, break, continue, if-else	**7T + 4P (15 hours)**
Unit-II	Functions and anonymous functions- built-in functions, user-defined functions, and lambda functions, recursion, scoping, and modules.	**8T + 3P (14 hours)**
Unit-III	Python Data structures: matrix, data frames, array, tuples, lists, dictionaries, lists, and mutability, string handling Object-oriented programming, abstract data types, classes, inheritance, encapsulation.	**8T + 3P (14 hours)**
Unit-IV	File Handling and Exceptions Handling: Introduction to Files, modes for file opening, reading and writing structured and text files, closing and deletion of files, errors, and exceptions. Essential Libraries: Pandas, NumPy, Scikit-learn, Matplotlib, SciPy, Statsmodels, Seaborn, etc.	**7T + 4P (15 hours)**

Text Books:

1. John V. Guttag, "Introduction to Computation and Programming using Python", PHI, 2022.
2. Jake Vander Plas, "Python Data Science Handbook – Essential Tools for Working with Data", O'Reilly Media Inc, 2016

Reference Books:

1. Michael T. Goodrich, "Data structures and Algorithms in Python", Wiley, 2020.

E-Resources:

1. https://nptel.ac.in/courses/106106145

Course: PROFESSIONAL COMMUNICATION
Code: MBD406

TEACHING SCHEME:		**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**
Theory: 2 **Tutorial: 0** **Practical: 0**		**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**2**

Course Prerequisites: No

Course Objectives: To develop the knowledge of communication theory and processes.

Course Outcomes: After completing the course, the student shall be able to:

- Understand and apply communication theory.
- Critically think about communication processes and messages.
- Write effectively for a variety of contexts and audiences
- Communicate confidently in professional environmen
- communicate with confidence in professional environment

Course Content:

Unit-I	**Fundamentals of Communication** ● Definition, Process, Importance of Communication ● Types of Communication ● Purpose of Professional Communication ● Barriers to Communication	**(7 Hours)**

Unit-II	**Oral Communication:** **Listening Skill –** ● Effective Listening ● Intensive Listening vs Extensive listening ● Techniques of Effective Listening ● Listening and Note Taking **Speaking Skills** ● **Paralinguistic features** ○ Rate ○ Pauses ○ Volume ○ Pitch/Intonation/Voice Modulation ○ Pronunciation and Articulation ● **Activities** ○ Group Discussions ○ Debates ○ Interviews ○ Public Speaking **Readings Skills** ● Effective Reading ● Types of Reading (Skimming, Scanning, Extensive Reading, Intensive Reading)	**(8 Hours)**
Unit-III	**Written communication:** ● Academic Writing ● Critical Thinking ● Technical Writing vs Creative Writing ○ Paragraph Writing (structure, construction, coherence and cohesion) ○ Business Letters (Acknowledgement letter, Appreciation letter, Order letter) ○ Business Reports ○ Research Papers ○ Advertising, Notices, Emails ○ Resume writing, Cover Letter	**(7 Hours)**
Unit-IV	**Soft Skills:** ● Body Language – Personal Appearance, Gesture, posture, facial expression, eye contact ● Proxemics/ Space Distance	**(8 Hours)**

	● Presentation Skills	

Text Books:

1. Banerjee Meera and Mohan Krishna, "Developing Communication Skills", Macmillan Publications, 1990.
2. Chaturvedi, P.D., "Business Communication", Pearson Publications,2013

Reference Books:

1. M.J. Mathew, "Business Communication", RBSA Publications,2005.
2. Taylor Shirley, "Communication for Business", Pearson Publications,2005

E-Resources:

1. https://nptel.ac.in/courses/102104061
2. https://nptel.ac.in/courses/102104061

Semester- II: M.Sc. CS (BDA)

<table>
<tr><td colspan="3">Course: MACHINE LEARNING
Code: MBD407</td></tr>
<tr><td><u>TEACHING SCHEME:</u></td><td><u>EXAMINATION SCHEME:</u></td><td><u>CREDITS ALLOTED:</u></td></tr>
<tr><td>Theory: 2
Tutorial: 0
Practical: 2 (4 Hours)</td><td>Internal Assessment: 40 Marks
CIA-I: 20 Marks (Unit I & II)
CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce)

End Semester Examination: 60 Marks (Unit I- IV)</td><td>4</td></tr>
<tr><td colspan="3"></td></tr>
<tr><td colspan="3">Course Prerequisites: No</td></tr>
<tr><td colspan="3">Course Objectives: To develop the knowledge of machine learning algorithms, data analytics solutions.</td></tr>
<tr><td colspan="3">Course Outcomes: After completing the course, the student shall be able to:
● Understand the types of machine learning algorithms.
● Understand the concept of different types of artificial neural network architectures.
● Understand the back-propagation neural network architecture and its training algorithms.
● Learn the deep learning architecture and algorithms.</td></tr>
<tr><td colspan="3">Course Content:</td></tr>
<tr><td>Unit-I</td><td>Basics: Introduction to Machine Learning, Different Forms of Learning-Supervised Learning, Semi-supervised Learning and Unsupervised Learning, Reinforcement Learning, Hypothesis space, inductive-bias, bias-variance trade-off, cross-validation methods.

Regression: Simple linear and multi-variable regression, Ridge Regression, Lasso, Bayesian Regression. Classification Methods: Instance-based classification, Logistic Regression, Naive Bayes Classifier, Large Margin Classification, Kernel Methods-Radial Basis function, Gaussian Kernel, Sigmoid, Polynomial, etc., Support Vector Machines, Multi-class Classification, Decision Trees.</td><td>7T + 8P (23 hours)</td></tr>
<tr><td>Unit-II</td><td>Artificial Neural Network & Classification: Fundamental of Artificial Neural Network, Evolution of Neural Networks, Basic Models of Artificial Neural Network, Important Terminologies of ANNs, McCulloch-Pitts Neuron, Linear Separability, Hebb Network.</td><td>8T + 7P (22 hours)</td></tr>
</table>

	Perceptron Networks: Theory, Perceptron Learning Rule, Architecture, Perceptron Training Algorithm for Single Output Classes, Perceptron Training Algorithm for Multiple Output Classes, Perceptron Network Testing Algorithm.	
Unit-III	Back-Propagation Network: Theory, Architecture, Training Algorithm, Learning Factors of Back-Propagation Network, Testing Algorithm of Back-Propagation Network.	**7T + 8P (23 hours)**
Unit-IV	Deep Learning: Introduction to Deep Learning, Vanishing Gradient, Overfitting, Computational Load. CNN: Padding and Strides, Convolution Operation, Pooling Operation, Convolution Layer. Introduction to AlexNet, VGGNet, Residual Network, Inception Network, and Transfer Learning.	**8T + 7P (22 hours)**

Text Books:

1. Tom Mitchell, "Machine Learning", McGraw-Hill, 1st Edition, 1997.
2. Christopher Bishop, "Pattern Recognition and Machine Learning", Springer, 1st Edition, 2016.

Reference Books:

1. Ryszard S. Michalski, Jaime G. Carbonell, Tom M. Mitchell and Morgan Kaufmann, "Machine Learning: An Artificial Intelligence Approach", Springer, 1st Edition, 1995.
2. Martin T. Hagan, Howard B. Demuth and Mark Beale, "Neural Network Design", Cenage Learning, 1st Edition, 1996.

E-Resources:

1. https://nptel.ac.in/courses/106106139

Course: DATA STRUCTURES AND ALGORITHMS **Code: MBD408**		
TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS**
Theory: 2 **Tutorial: 0** **Practical: 2 (4 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**4**
Course Prerequisites: No		
Course Objectives: To understand the concepts of algorithms and data structures.		
Course Outcomes: After completing the course, the student can ● Implement and empirically analyse data structures like Arrays, Stacks, linked list and queues.		

- Perform different search and sorting operations.
- Analyze the computational complexities.
- Design algorithms like divide and conquer,

Course Content:

Unit-I	Introduction to data structures: Arrays, Stacks and queues: insertion, deletion, and basic operations on stacks, arrays, queues, circular queues, and dequeue. Linked Lists: Single, double, and circular linked lists, insertion, deletion, and basic operations on Linked lists, Hash Tables.	**8T + 7P (22 hours)**
Unit-II	Sorting and searching techniques: linear search, binary search, insertion sort, selection sort, merge sort, quick sort, bubble sort. Computational Complexity: O-notation, Ω-notation, Θ-notation, Classes P, NP - Verifiability, NP-Hard - Reducibility, NP Complete.	**7T + 8P (23 hours)**
Unit-III	Advanced algorithm techniques: Divide & Conquer-Strassen's Matrix Multiplication, Counting; Inversion, Greedy Algorithms- Minimum Spanning Trees, Prims and Krushkal's; Dynamic Programming-rod cutting, Longest Common Subsequence	**7T + 8P (23 hours)**
Unit-IV	Trees and Graphs: definition and concepts, different types of trees. Graph Problems: Clique, Vertex Cover, Independent Set, Hamiltonian Cycle Problem, Travelling Salesman Problem, Graph Partitioning, Subgraph problem, Graph Isomorphism, Graph Coloring.	**8T + 7P (22 hours)**

Text Books:

1. T.H. Cormen, C.E. Leiserson, R.L. Rivest and C. Stein, "Introduction to Algorithms", MIT Press, 4th Edition, 2022.
2. Jon Kleinberg, Eva Tardos, "Algorithm Design", Pearson Education, 1st Edition, 2013.

Reference Book:

1. Seymour Lipschutz, "Data Structures", McGraw Hill Education, 2019.

E-Resources:

1. https://nptel.ac.in/courses/106102064

Course: ADVANCED STATISTICAL METHODS
Code: MBD409

TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**

Theory: 3 **Tutorial: 0** **Practical: 0**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**

Course Prerequisites: No

Course Objectives: To develop the knowledge of sample and population, hypothesis testing.

Course Outcomes: After completing the course, the student shall be able to:

- Draw the hypothesis testing decision matrix and explain the contents.
- Understand the categorical data and study
- Understand stochastic processes and type of Markov chains,
- Studying the different statistical methods to understand real life phenomena.

Course Content:

Unit-I	Estimation: Unbiasedness, Consistency, Efficiency and Sufficiency, UMVUE, Maximum likelihood estimates, Method of least squares, Confidence interval.	**(10 Hours)**
Unit-II	Test of Hypotheses: Types of errors, t-test statistic, parametric tests for equality of means & variances. one-way and two-way ANOVA	**(12 Hours)**
Unit-III	Stochastic Processes, Random Walk, Markov Process, Markov Chain, Discrete Time Markov chain, Long run behaviour of Markov Chain: Limiting Distributions, Stationary Distributions, Markov Chain Monte Carlo.	**(12 Hours)**
Unit-IV	Continuous time Markov chains, Poisson processes, birth and death processes, Kolmogorov differential equations.	**(11 Hours)**

Text Book:

1. Kishor S. Trivedi, “Probability and Statistics with Reliability, Queueing and Computer Science Applications”, John Wiley, 2nd Edition, 2001.

2. Sheldon M. Ross, “Introduction to Probability and Statistics for Engineers and Scientists”, Academic Press, 5th Edition, 2014.

Reference Book:

1. P. J. Bickel and K.A. Docksum, "Mathematical Statistics", Prentice Hall, 2nd Edition, 2000.

E-Resources:

1. https://nptel.ac.in/courses/111105090

Course: DIGITAL IMAGE PROCESSING AND COMPUTER VISION **Code: MBD410**		
TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**
Theory: 2 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**

Course Prerequisites: No

Course Objectives: To equip students with the theoretical knowledge and practical skills necessary to analyze, enhance, and manipulate digital images for a wide range of applications.

Course Outcomes: After completing the course, the student shall be able to:

- Understand Image Processing Principles.
- Proficient in Image Enhancement and Restoration.
- Competence in Image Segmentation and Feature Extraction.
- Apply Image Processing in Practical Scenarios and Proficient Use of Image Processing Tools and Software.

Course Content:

Unit-I	Intensity Transformation Functions: Logarithmic and Contrast-Stretching Transformations, Specifying Arbitrary Intensity Transformations, Some Utility functions for Intensity Transformations. Histogram Processing and Function Plotting: Generating and Plotting Image Histograms, Histogram Equalization, Histogram Matching (Specification). Spatial Filtering: Linear Spatial Filtering, Non-linear Spatial filtering.	**7T + 4P** **(15 hours)**
Unit-II	A Model of the Image Degradation/Restoration Process: Noise Models, Adding Noise to Images, Generating Spatial Random Noise with a Specified Distribution, Periodic Noise, Estimating Noise Parameters. Restoration in the Presence of Noise Only-Spatial Filtering: Spatial Noise Filters, Adaptive Spatial Filters, Periodic Noise Reduction Using Frequency Domain Filtering, Modeling the Degradation Function, Direct Inverse Filtering, Wiener Filtering, Constrained Least Squares (Regularized) Filtering.	**8T + 3P** **(14 hours)**
Unit-III	Preliminaries: Some Basic Concepts from Set Theory, Binary Images, Sets, and Logical Operators, Dilation and Erosion. Combining Dilation and Erosion: Opening and Closing, The Hit-or-Miss Transformation, Using Lookup Tables, Labeling Connected Components. Morphological Reconstruction: Opening by Reconstruction, Filling Holes, Clearing Border Objects. Gray-Scale Morphology: Dilation and Erosion, Opening and Closing Reconstruction.	**8T + 4P** **(16 hours)**

Unit-IV	Thresholding: Foundation, Basic Global Thresholding, Optimum Global Thresholding Using Otsu's Method, Using Image Smoothing to Improve Global Thresholding, Using Edges to Improve Global Thresholding, Variable Thresholding Based on Local Statistics, Image Thresholding Using Moving Averages. Region-Based Segmentation: Basic Formulation, Region Growing, Region Splitting and Merging.	**7T + 4P (15 hours)**

Text Books:

1. R. C. Gonzalez, R. E. Woods and S. L. Eddins, "Digital Image Processing using MATLAB", McGraw Hill, 6st Edition, 2008.
2. M. Sonka, V. Hlavac and R. Boyle, "Image Processing, Analysis and Machine Vision", Springer, 1st Edition, 2013.

Reference Book:

1. A. Marion, "Introduction to Image Processing", Springer, 1st Edition, 2013.

E-Resources:

1. https://nptel.ac.in/courses/117105135

Course: ECONOMETRICS AND FINANCE
Code: MBD411

TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**
Theory: 2 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**

Course Prerequisites: No

Course Objectives: To develop the knowledge of basic statistics, econometric computer packages, interpret linear-regression.

Course Outcomes: After completing the course, the student shall be able to:

- Comfortable with basic statistics and probability.
- To use a statistical/econometric computer package to estimate an econometric model and be able to report the results of their work in a non-technical and literate manner.
- Able to estimate and interpret linear regression models and be able to distinguish between economic and statistical importance.
- Able to critique reported regression results and interpret the results for someone who is not trained as an economist.

Course Content:

Unit-I	Introduction to Econometrics (using finance concepts): Assumptions of Classical Linear Regression Model, Ordinary Least Squares approaches, Autocorrelation, Heteroscedasticity, Multi collinearity, Dummy Variable approaches, and Distributed lag models.	**7T + 4 P (15 hours)**
Unit-II	A brief Introduction to Time Series and Panel data models, Components of time series, Stationary and non-stationary time series, ARMA and ARIMA models,	**8T + 4 P (16 hours)**

	Static panel data models: fixed effects and random effects.	
Unit-III	Basics of Finance: Time value of money, concept of present and future value analysis, stock and bond valuations, risk and return, Systematic and unsystematic risk, Diversification, cost of capital, capital structure,	**7T + 4 P (15 hours)**
Unit-IV	Dividend Discount Model, Portfolio Theory, Efficient Market Hypothesis (EMH), Capital asset pricing model (CAPM), Market Volatility, Options.	**8T + 3 P (14 hours)**

Text Books:

1. Richard A. Brealey, Stewart C. Myers, Franklin Allen and Pitabas Mohanty, "Principle of Corporate Finance", Tata McGraw Hill Education Pvt. Ltd., 10th Edition, 2012.
2. Chris Brooks, "Introductory Econometrics for Finance", Cambridge University Press, 1st Edition, 2014.

Reference Books:

1. Pamela Peterson Drake and Frank J. Fabozz, "The Basics of Finance: An Introduction to Financial Markets, Business Finance and Portfolio Management", John Wiley & Sons, Inc., 1st Edition, 2010.
2. Jeffery M. Wooldridge, "Introductory Econometrics: A Modern Approach", Cengage Learning, 1st Edition, 2015.
3. Damodar N. Gujarati, "Econometrics by Example", Tata McGraw Hill Education Pvt. Ltd., 1st Edition, 2011.

E-Resources:

1. https://archive.nptel.ac.in/courses/130/106/130106001/

LEVEL-4: ELECTIVE COURSES

Course: VISUALIZATION TOOLS
Code: MBD431

TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**
Theory: 1 **Tutorial: 0** **Practical: 2 (4 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**

Course Prerequisites: No

Course Objectives: To develop the knowledge of visualization.

Course Outcomes: After completing the course, the student shall be able to:

- Build professional-quality business intelligence reports from the ground up
- Design and implement the same tools used by professional analysts and data scientists
- Blend and transform raw data into beautiful interactive dashboards
- Explore powerful artificial intelligence tools and advanced visualization techniques

Course Content:

Unit-I	Principles of visualization: Data visualizations best practices, Purpose of visualization, visual perception, cognitive issues - evaluation, principles behind information visualization, Art of storytelling using visualization, different methods of presenting data in business analytics.	**3T + 7P (17 hours)**
Unit-II	MS Excel: Introduction to MS Excel, MS Excel fundamentals, working with basic Excel functions, Formatting data in an Excel worksheet, Excel templates, Excel list function, Importing & Exporting data, filters, sorting, data validation, Data validation, Freezing of rows/columns, Basic Excel charts & Graphs: Excel chart formatting & Customization., Excel Bar & Columns Charts, Excel Histograms & Pareto Charts, Line Charts & Trendlines, Excel Pie Charts, Scatter Plots, Excel Box & Whisker Charts, Excel Heat Maps with Conditional Formatting, Excel Surface & 3D Contour Charts, Geo-Spatial Mapping with Excel Power Map.	**4T + 7P (18 hours)**
Unit-III	Advanced Excel: Excel Formula, Powerpivot tools. Large sets of Excel data, Protection of sheets/columns, Conditional Functions, Lookup functions, Text Based function, Protecting worksheets, "What If" tools, Automating repetitive tasks , Excel macros ,VBA tools, VBA logic statements. Advanced Charts: Creating Custom Image Overlay Charts, Automating charts with named ranges, Creating Interactive Area Charts to Show Changes Over Time, Building a dynamic excel dashboard, Advanced Excel Data Visualization Technique Exploratory data analysis with excel : Data Collection, Data Cleaning, Data Exploration, Data Visualization, Data Interpretation,Reporting.	**4T + 8P (20 hours)**
Unit-IV	Power BI: Introduction to Microsoft Power BI Desktop, installing Power BI, connecting data with Power BI and creating and shaping data visualizations, Charts and graphs, creating data models, Working with calculations and expressions, Calculated field with Dax, Dashboard and stories. Application in AI, Optimization tools. Tableau: Introduction to Tableau, installing Tableau, connecting data with Tableau, Creating Data Extracts in Tableau, creating data visualizations, Charts and graphs, Working with Data Blending in Tableau, Parameters, Working with calculations and expressions, Advanced Data Preparation, Dashboard and stories, clusters, custom territories, design features. A project: using everything at once in a real-life project.	**4T + 8P (20 hours)**

Text Book:

1. Nussbaumer Knaflic and Cole, "Storytelling With Data: A Data Visualization Guide For Business Professionals", Wiley, 1st edition, 2015.
2. Alberto Ferrari and Marco Russo," Introducing Microsoft Power BI", Microsoft Press, 2017.
3. Marleen Meier and David Baldwin, "Mastering Tableau 2021: ImplementAdvanced Business Intelligence Techniques and Analytics with Tableau", 3rd Edition, Packt Publishing Limited, 2019.
4. Adam Ramirez, "Excel Formulas and Functions 2020: The Step by Step Excel Guide with Examples on How to Create Powerful Formulas (Excel Academy Book 1)", Caprioru, 2020.

Reference Books:

1. Ferrari Alberto and Russo Marco, "Analyzing Data with Power BI and Power Pivot for Excel (Business Skills)", Microsoft Press, 1st Edition, 2019.
2. George Mount , "Advancing into Analytics: From Excel to Python and R", O'Reilly, 1st Edition, 2021.
3. Jeremey Arnold, "Learning Microsoft Power BI: Transforming Data into Insights (Grayscale Indian Edition)", O'Reilly, 1st Edition, 2022.
4. Ryan Sleeper, "Practical Tableau: 100 Tips, Tutorials, and Strategies from a Tableau Zen Master (Greyscale Indian Edition)", O'Reilly, 2018.
5. Stefen flew, "Information Dashboard Design: Displaying Data for At-A-Glance Monitoring", Analytics Press, 2013.

E-Resources:

1. https://archive.nptel.ac.in/courses/110/107/110107157/

Course: PROGRAMMING IN JAVA **Code: MBD432**		
TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS**
Theory: 2 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**
Course Prerequisites: No		
Course Objectives: To understand the basic concepts of Java Programming.		
Course Outcomes: After completing the course, the student can ● Understand the basic concepts of Java Programming. ● Acquire object oriented skills in Java ● understand the concepts of inheritance and polymorphism ● learn how to handle program exceptions		

Course Content:		
Unit-I	Introduction to Java, Overview and Characteristics of Java, JVM, data types, primitive variables, arrays, operators, control statements, input, output, and main method.	**7T + 4P (15 hours)**
Unit-II	Object Oriented Programming concepts: Encapsulation and abstraction, designing classes, objects, instance variables and methods, class modifiers.	**8T + 4P (16 hours)**
Unit-III	Inheritance: types of inheritance, abstract classes, overloading, and overriding. working with packages and interfaces Polymorphism: overloading and overriding.	**8T + 4P (16 hours)**
Unit-IV	Constructors, use of this and super, garbage collection, static methods and variables, wrapper classes, Math and String, Exception Handling & applications: exception types, nested try-catch, throw, throws and finally statements. Multi Thread Programming: thread creation, synchronization and priorities. Input-output and file operations: Java.io, Object serialization and deserialization. Java Collections API: Arraylist, Set, list, Map, Hashtable, Comparator and comparable Database connectivity, Java Packages, creating a jar file	**7T + 3P (13 hours)**

Text Book:

1. E. Balagurusamy, "Programming with Java", McGraw Hill, 6th Edition, 2019.
2. James Gosling, Bill Joy, Guy L. Steele Jr, Gilad Bracha and Alex Buckley, "The Java Language Specification", Addison-Wesley, 7th Edition, 2013.

Reference Books:

1. Herbert Schildt, "Java-A Beginner's Guide", McGraw Hill, 8th Edition, 2020.

E-Resources:

1. https://nptel.ac.in/courses/106105191

Course: LINUX PROGRAMMING
Code: MBD433

TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED**
Theory: 1 **Tutorial: 0** **Practical: 2 (4 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**

Course Prerequisites: No

Course Objectives: To develop the knowledge of Linux Operating System and make effective use of Linux utilities, shell scripts, file utilities, and process utilities.

Course Outcomes: After completing the course, the student shall be able to:

- get introduced to basic Linux operating system commands
- write shell scripts to solve problems
- understand the file management system calls
- develop the skills required for process management, and signal management.

Course Content:

Unit-I	Introduction to Linux: What is Linux Operating systems, History of Linux, Linux vs. other OS, Architecture, Features, and Installation. Linux commands: printf, man, PATH, echo, passwd, who, uname, script, date, cp, rm, ls, mv, wc, more, cat, wc, lp, od, file handling utilities, security by file permissions, disk utilities, process utilities, df, du, unlink, find, mount, umount, unmask, ps, w. Text Processing utilities and backup utilities, head, tail, uniq, nl, sort, grep, fgrep, egrep, join, paste, cut, pg, cmp, comm, diff, awk, tr, cpio.	**4T + 8P (20 hours)**
Unit-II	Shell Scripting: Introduction to shells, Bourne Again SHell (BASH), standard streams, writing and executing the shell scripts, Shell variables, Shell commands, command substitution, arithmetic in shell, control structures.	**3T + 7P (17 hours)**
Unit-III	Files and Directories, Introduction to Files and directories, File Structures, File management System Calls– create, close, open, close, write, read, dup2, lseek, stat, lstat, fstat, chown, lchown, fchown, chmod, fchmod, link, unlink, symlink, cat, mkdir, opendir rmdir, chdir, readdir, closedir, unmask, getcwd, rewin functions	**4T + 7P (18 hours)**
Unit-IV	Process and Signals, Introduction to Process, environment variables - environ, getenv, setenv, process creation, process termination, thread and process, parent and child process, zombie process, orphan process, system calls - fork, vfrok, wait, exit, waitpid, exec, Introduction to signals, signal generation, signal functions - kill, alarm, pause, sleep, raise, abort, reliable and unreliable signals.	**4T + 8P (20 hours)**

Text Book:

1. Sumitabha Das, "UNIX – Concepts and Applications", Tata McGraw Hill, 4th Edition, 2006.
2. Robert Love. "Linux System Programming", O'Reilly, 2nd Edition, 2013.

Reference Books:

1. Behrouz A. Forouzan and Richard F. Gilberg, "Unix and Shell Programming", Cengage Learning, 1st Edition, 2014.
2. Richard Petersen. "The Complete Reference LINUX", McGraw Hill, 6th Edition, 2008.

E-Resources:

1. https://nptel.ac.in/courses/117106113

LEVEL-5

Semester-III: M.Sc. CS (BDA)

Course: ENABLING TECHNOLOGIES FOR DATA SCIENCE
Code: MBD502

TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS**

Theory: 3 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**4**

Course Prerequisites: No

Course Objectives: To develop fundamental understanding of the Technologies of Data Science including Hadoop, Pig and Hive.

Course Outcomes: After completing the course, the student can:

- Understand Big data and its analytics in the real world
- Work on Hadoop
- Develop applications in Hadoop.

Course Content:

Unit-I	Big Data and Hadoop: Hadoop architecture, versioning, single node and multinode Hadoop, Hadoop commands, Hadoop daemons.	**12T + 3 P (18 hours)**
Unit-II	Map Reduce: Framework, developing Map reduce program, Map reduce programs in local and pseudo-distributed mode, illustrations.	**10T + 4P (18 hours)**
Unit-III	Hive: Installation, data types, and illustrations, Spark and NoSQL.	**11T + 4P (19 hours)**
Unit-IV	Sqoop: Installation, importing, and exporting data, commands, and illustrations. Pig: Installation, commands, illustration.	**12 T + 4 P (20 hours)**

Text Book:

1. Chuck Lam, "Hadoop in Action", Dreamtech Press, 2020.
2. Tom White, "Hadoop: The definitive guide", O'Reilly, 3rd Edition, 2009.

Reference Books:

1. Jimmy Lin and Chris Dyer, "Data Intensive Text Processing with Map Reduce", Morgan and Claypool Publishers, 2010.

E-Resources:

1. https://archive.nptel.ac.in/courses/106/104/106104189/

Course: DATA MINING
Code: MBD503

TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED**
Theory: 2 **Tutorial: 0** **Practical: 2 (4 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**4**

Course Prerequisites: No

Course Objectives: To equip learners with the essential knowledge and skills needed to effectively extract meaningful insights from large datasets using various data mining techniques and algorithms in order to make informed decisions and solve real-world problems.

Course Outcomes: After completing the course, the student shall be able to:

- Understand the concept of data processing and visualization.
- Understand the concept of association rule mining.
- Learn the concept of data clustering and outlier analysis.
- Learn the concept of web mining and text mining.

Course Content:

Unit-I	Data Mining: Introduction, Techniques, Issues and challenges, applications, Data preprocessing, Knowledge representation, Measure of similarity & dissimilarity. Data Warehouses: Differences between Operational Database Systems and Data Warehouses, Data Warehouse Models, Extraction, Transformation, and Loading, Data Cube, Stars, Snowflakes, and Fact Constellations, OLAP Operations. Exploring Data: Visualization, OLAP & Multidimensional Data Analysis.	**8T + 7P (22 hours)**
Unit-II	Association Rule Mining: Introduction, Frequent Itemsets, Closed Itemsets, and Association Rules Methods to discover association rules, Association rules with item constraints. Association analysis-Advanced Concepts: Handling Categorical and Continuous attributes. Handling a concept Hierarchy, Sequential patterns, subgraph Patterns, Infrequent patterns.	**7T + 8P (23 hours)**
Unit-III	Cluster analysis: Introduction, clustering paradigms, Similarity and distance, Density, Characteristics of clustering algorithms, Center based clustering techniques, Hierarchical clustering, Density based clustering, other clustering techniques, Scalable clustering algorithms, Cluster evaluation. Advanced Clustering Techniques: Probabilistic Model-Based Clustering, Clustering High-Dimensional Data, Clustering Graph and Network Data, Clustering with Constraints.	**7T + 8P (23 hours)**
Unit-IV	Outlier Analysis: Introduction, Outlier Detection Methods, Statistical Approaches, Proximity-Based Approaches, Mining Contextual and Collective Outliers. Advanced techniques: Web mining - Introduction, Web content mining, Web structure mining, Web usage mining; Text mining- Unstructured text, Episode rule discovery from text, Text clustering; Temporal data mining – Temporal association rules, Sequence mining, Episode discovery, time series analysis; Spatial data mining – Spatial mining tasks, Spatial clustering, Spatial trends.	**8T + 7P (22 hours)**

Text Book:

1. P. N. Tan, M. Steinbach and V. Kumar, “Introduction to Data Mining”, Pearson Education India, 4th Edition, 2016.
2. J. Han, M. Kamber and J. Pei, “Data Mining Concepts and Techniques”, Morgan Kaufmann, 3rd Edition, 2012.

Reference Books:

1. M. Kantardzic, “Data mining: Concepts, Models, Methods, and Algorithms”, John Wiley & Sons, 1st Edition, 2011.

E-Resources:

1. https://archive.nptel.ac.in/courses/106/105/106105174/

Course: TIME SERIES AND FORECASTING **Code: MBD504**		
TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**

Theory: 2 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**

Course Prerequisites: No

Course Objectives: To develop the knowledge of time series with different structures, seasonality, and cyclical irregularity.

Course Outcomes: After completing the course, the student shall be able to:

- Explain time series with different structures.
- Explain trend, seasonality, cyclical irregularity.
- Construct and evaluate time series models.
- Understand the concept of multivariate time series.

Course Content:

Unit-I	Basics of Time series: A model Building strategy, Time series and Stochastic process, stationarity, Auto correlation, meaning and definition – causes of auto correlation - consequence of autocorrelation – test for auto – correlation. Study of Time Series model and their properties using correlogram, ACF and PACF. Yule walker equations	**7T + 3 P (13 hours)**
Unit-II	Time Series Models: White noise Process, Random walk, MA, AR, ARMA and ARIMA models, Box- Jenkins's Methodology fitting of AR(1), AR(2), MA(1), MA(2) and ARIMA(1,1) process. Unit root hypothesis, Co-integration, Dicky Fuller test unit root test, augmented Dickey – Fuller test.	**8T + 4 P (16 hours)**
Unit-III	Non-linear time series models, ARCH and GARCH Process, order identification, estimation and diagnostic tests and forecasting. Study of ARCH (1) properties. GARCH (Conception only) process for modelling volatility.	**7T + 4 P (15 hours)**
Unit-IV	Multivariate Linear Time series: Introduction, Cross covariance and correlation matrices, testing of zero cross correlation and model representation. Basic idea of Stationary vector Autoregressive Time Series with order one: Model Structure, Granger Causality, stationarity condition, Estimation, Model checking.	**8T + 4 P (16 hours)**

Text Books:

1. G. E. P. Box and G. M. Jenkins, "Time Series Analysis: Forecasting and Control", Holden day, San Francisco, 1st Edition, 1976.
2. Ruey S. Tsay, "Multivariate Time Series Analysis: with R and Financial Application", Wiley & Sons, 1st Edition, 2014.

Reference Books:

1. C. Chatfield, "Analysis of Time Series, An Introduction", CRC Press, 1st Edition, 2003.
2. Ruey S. Tsay, "Analysis of Financial Time Series", Wiley & Sons, 2nd Edition, 2005.
3. W.A. Fuller, "Introduction to Statistical Time Series", Wiley Series in Probability and Statistics, 1st Edition, 1996.

E-Resources:

1. https://nptel.ac.in/courses/103106123

Course: OPTIMIZATION TECHNIQUES **Code: MBD505**		
TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**
Theory: 2 **Tutorial: 0**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II)	**3**

Practical: 1 (2 Hours)	CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	

Course Prerequisites: No

Course Objectives: To develop the knowledge of linear programming and designing efficient algorithms.

Course Outcomes: After completing the course, the student shall be able to:

- Understand the linear programming techniques for designing efficient algorithms.
- Design and analyze algorithms for various combinatorial optimization problems.
- Develop a comprehensive understanding of optimization theory
- Understand queueing models and its applications in data analytics

Course Content:

Unit-I	Linear Programming, Convex Set, Simplex Method, Revised Simplex method, Duality, Dual Simplex, Interior Point Method, Transportation problem, Assignment Problem, Integer Linear Programming Problem.	**7T + 4 P (15 hours)**
Unit-II	Dynamic Programming, Network Flow Problem: Shortest Path problem, Max-Flow and Min-cut problem, Convex function, Concave function, Non-Linear Optimization Problem: Lagrange Multipliers, KKT Conditions, Quadratic programming Problems, Wolfe method, Multivariable functions and their maxima and minima	**8T + 4 P (16 hours)**
Unit-III	Unconstrained Optimization, Local & global Optimum, Direct Search, Steepest descent method, Conjugate Gradient method, Penalty Function Method, Barrier Function Method.	**7T + 4 P (15 hours)**
Unit-IV	Queueing Model: M/M/1, M/M/C, etc., Steady-state solutions of Markovian queues, Non-Markovian Queues. Queueing Decision models, Simulation of Queueing Systems.	**8T + 3 P (14 hours)**

Text Book:

1. Bemhard Korte and Jens Vygen. "Combinatorial Optimization-Theory and Algorithms", 2019.
2. Suresh Chandra, Jayadeva and Aparna Mehra, "Numerical Optimization with Applications", Narosa Publishing House, 2009.
3. Kishor S. Trivedi, "Probability and Statistics with Reliability, Queueing and Computer Science Applications", John Wiley, 2nd Edition, 2001.

Reference Books:

1. W.J. Cook, W.H. Cunningham, W.R. Pulleyblank and A. Schrijver, "Combinatorial Optimization", Wiley-Interscience, 1st Edition, 2011.
2. John Riordan. "Introduction to Combinatorial Analysis", Dover Publications Inc., 1st Edition, 2002.

E-Resources:

1. https://nptel.ac.in/courses/111105039

LEVEL-5: ELECTIVE COURSES

Course: SOFTWARE ENGINEERING

Code: MBD531		
TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED :**
Theory: 3 **Tutorial: 0** **Practical: 0**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**

Course Prerequisites: No

Course Objectives: To develop the knowledge of various phase of software lifecycle, choosing process model.

Course Outcomes: After completing the course, the student shall be able to:

- Decompose the given project in various phases of a lifecycle.
- Choose an appropriate process model depending on the user requirements.
- Perform various life cycle activities like Analysis, Design, Implementation, Testing and Maintenance.
- Know various processes used in all the phases of the product.
- Apply the knowledge, techniques, and skills in the development of a software product.

Course Content:

Unit-I	Software Development Life Cycle: Software Process, Software Development Life Cycle Models , Software Requirement Engineering: Requirement Engineering Process Function-oriented Design: Introduction to Structured Analysis, Data Flow Diagram, Process Specification, Entity Relationship (ER) Model, Structured Design Methodologies, Design Metrics	**(10 Hours)**
Unit-II	Object Oriented Concepts & Principles: Key Concepts, Relationships: Is-A Relationship, HasA Relationship, Uses-A Relationship; Modelling Techniques: Booch OO Design Model, Rumbaugh's Object Modelling Technique, Jacobson's model, The Unified Approach to Modelling, Unified Modelling Language (UML). Object Oriented Analysis & Design: UseCase Modelling, Use-Case Realization, Class Classification Approaches: Noun Phrase Approach, CRC Card Approach, Use-case Driven Approach, Identification of Classes, Relationship, Attributes and Method. System Context and Architectural Design, Principles of Class Design, Types of Design Classes	**(12 Hours)**
Unit-III	UML 2.0 diagrams: Structure diagrams, Behavior diagrams. Software coding and Testing: Coding standards and guidelines, Code review techniques, Testing Fundamentals, Verification & Validation, Black Box Testing, White Box Testing, Unit Testing, Integration Testing, System Testing, Object Oriented System Testing.	**(11 Hours)**
Unit-IV	Emerging Trends: Architecture styles, Service Oriented Architecture (SOA), CORBA, COM/DCOM; Web Engineering: General Web Characteristics, Emergence of Web Engineering, Web Engineering Process, Web Design Principles, Web Metrics	**(12 Hours)**

Text Book:

1. Roger S. Pressman, "Software Engineering: A Practitioner's Approach", McGraw Hill, 6th Edition, 2005.
2. Rajib Mall, "Fundamentals of Software Engineering", PHI, 3rd Edition, 2009.

Reference Books:

1. Grady Booch, James Rambaugh and Ivar Jacobson, "Unified Modeling Language User's Guide", Pearson, 2nd Edition, 2002.

E-Resources:

1. https://nptel.ac.in/courses/106105182

Course: NATURAL LANGUAGE PROCESSING
Code: MBD532

TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED**
Theory: 2 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**

Course Prerequisites: Knowledge of data structures and machine learning.

Course Objectives: To equip the students with theoretical and practical knowledge of NLP required to develop models for text classifications, information extraction, and chatbots.

Course Outcomes: After completing the course, the student shall be able to:

- Understand the roadmap to NLP
- Apply CNN and LSTM for Text Classification
- Extract information from the text and develop Chatbots
- Understand the applications and advancements in NLP

Course Content:

Unit-I	Basic Concepts: Introduction to NLP, Language, Building blocks of Language, Approaches to NLP (heuristic-based, machine learning, and deep learning), NLP Pipeline, Vector Space Models, Vectorization Approaches (One-Hot encoding, Bag of Words, Bag of N-Grams, TF-IDF)	**7T + 4P (15 hours)**
Unit-II	Text Classification: Word Embedding, Subword Embedding and fastText, document embeddings, CNN for Text Classification, LSTM for Text Classification, Pre-Trained Language Models, Interpreting Text Classification with LIME, Learning with Less data, Adapting to New Domains	**8T + 4 P (16 hours)**
Unit-III	Information Extraction from Text: IE Introduction, Applications, IE Task Pipeline, Key phrase Extraction, Named Entity Recognition, Named Entity Disambiguation Linking, Relationship Extraction	**7T + 4 P (15 hours)**
Unit-IV	Applications and Advancements Chatbots: Dialog Systems, Goal Oriented Dialog, Components of Dialog Systems, Rasa NLU Chatbots; Recent Advancements in NLP: Transformer-based Models like BERT, GPT-4, Multimodal NLP; Few-Shot and Zero-shot NLP; Multimodal NLP, Large Language Models; Applications of NLP: Language Translation, Healthcare, Education, Content Generation, etc.	**8T + 3 P (14 hours)**

Text Book:

1. Yoav Goldberg, "Neural Network Methods for Natural Language Processing", Morgan & Claypool Publishers, 1st Edition, 2017.
2. Steven Bird, Ewan Klein, and Edward Loper, "Natural Language Processing with Python", O'Reilly, 1st Edition, 2009.

Reference Books:

1. Daniel Jurafsky and James H. Martin, "Speech and Language Processing", Pearson Publications, 2nd Edition, 2013.
2. Ricardo Baeza – Yates, and Berthier Ribeiro – Neto, "Modern Information Retrieval: The concepts and Technology behind Search", ACM Press Books, 2nd Edition, 2011.

E-Resources:

1. https://nptel.ac.in/courses/106105158

Course: CLOUD COMPUTING **Code: MBD533**		
TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**
Theory: 2 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**
Course Prerequisites: No		
Course Objectives: To develop knowledge of various service models, changing infrastructure landscape worldwide.		
Course Outcomes: After completing the course, the student shall be able to: ● Gain knowledge of various service models. ● How cloud systems are changing the infrastructure landscape worldwide.		
Course Content:		
Unit-I	Overview of Computing Paradigm, Recent trends in Computing Grid Computing, Cluster Computing, Distributed Computing, Utility Computing, Cloud Computing Evolution of Cloud Computing Business driver for adopting Cloud Computing. Introduction to Cloud Computing. Cloud Computing (NIST Model) Introduction to Cloud Computing, History of Cloud Computing, Cloud service providers. Properties, Characteristics & Disadvantages Pros and Cons of Cloud Computing.	**7T + 4 P** **(15 hours)**
Unit-II	Cloud Computing Architecture, Comparison with traditional computing architecture (client/server), Services provided at various levels, How Cloud Computing Works, Role of Networks in Cloud computing, protocols used, Role of Web services	**8T + 4 P** **(16 hours)**

Unit-III	Service Models (XaaS), Infrastructure as a Service (IaaS), Platform as a Service(PaaS), Software as a Service(SaaS), Deployment Models , Public cloud, Private cloud, Hybrid cloud Community cloud, Infrastructure as a Service(IaaS). Introduction to virtualization, Different approaches to virtualization, Hypervisors, Machine Image, Virtual Machine (VM).	**7T + 4 P (15 hours)**
Unit-IV	Examples, Amazon EC2, Renting, EC2 Compute Unit, Platform and Storage, pricing, customers Eucalyptus, Platform as a Service (PaaS), Introduction to PaaS, What is PaaS, Service Oriented Architecture (SOA). Examples: Google App Engine, Microsoft Azure, SalesForce.com's Force.com platform Software as a Service (PaaS): Introduction to SaaS, Web services, Web 2.0, Web OS, Case Study on SaaS Cloud Security. Case Study on Open Source & Commercial Clouds.	**8T + 3 P (14 hours)**

Text Book:

1. Erl Thomas, Puttini Ricardo and Mahmood Zaigham, "Cloud Computing: Concepts, Technology and Architecture", Pearson Education India, 2014.
2. Barrie Sosinsky, "Cloud Computing Bible", Wiley, 1st Edition, 2010.

Reference Book:

1. Barrie Sosinsky, "Cloud Computing Bible", Wiley-India, 1st Edition, 2010.
2. Kamal Kant Hiran, "Cloud Computing: Master The Concepts, Architecture and Applications with Real-World Examples and Case Studies", BPB Publications, 1st Edition, 2019.

E-Resources:

1. https://archive.nptel.ac.in/courses/106/105/106105167/

Course: MULTIVARIATE STATISTICS **Code: MBD534**		
TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**
Theory: 2 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**

Course Pre-requisites: No

Course Objectives: To develop the knowledge of finding patterns and correlations between several variables simultaneously, analyze complex datasets.

Course Outcomes: After completing the course, the student shall be able to:

- Find patterns and correlations between several variables simultaneously.
- Analyze complex datasets, allowing to gain a deeper understanding of data and how it relates to real-world scenarios.

Course Content:

Unit-I	Review of Multivariate Normal Distribution (MVND) and related distributional results. Random sampling from MVND, Unbiased and maximum likelihood estimators of parameters of MVND, their sampling distributions, independence. Correlation matrix and its MLE. Partial and multiple correlation coefficients, their maximum likelihood estimators (MLE), Wishart distribution and its properties (only statement).	**7T + 4 P (15 hours)**
Unit-II	Hotelling's T2 and its applications. Hotelling's T2 statistic as a generalization of square of Student's statistic. Distance between two populations, Mahalnobis D2 statistic and its relation with Hotelling's T2 statistic.	**8T + 4 P (16 hours)**
Unit-III	Classification problem – two populations, two multivariate normal populations, several populations; Discriminant analysis - Fischer's method, Logistic Regression Principle component analysis – Introduction, population principal components, summarizing sample variation by principal components, graphing principal components.	**7T + 4 P (15 hours)**
Unit-IV	Canonical correlation – Introduction, canonical variates & correlations, interpreting canonical variables, Factor Analysis – Introduction, Orthogonal Factor model, Methods of Estimation, Factor Rotation & Scores, and Perspective & Strategy for Factor Analysis Cluster Analysis – Introduction, similarity measures, hierarchical & non-hierarchical clustering methods, multidimensional scaling, correspondence analysis	**8T + 3 P (14 hours)**

Text Book:

1. A. M. Kshirsagar, "Multivariate Analysis", Maral-Dekker, 1972.
2. D.F. Morrison, "Multivariate Statistical Methods", McGraw-Hill, 1976.

Reference Books:

1. Johnosn, R.A. and Wichern. D.W, "Applied Multivariate Analysis", Prentice –Hall, 5th Edition, 2002.
2. Anderson T. W., "An Introduction to Multivariate Statistical Analysis", John Wiely, 2nd Edition, 1984.

E-Resources:

1. https://nptel.ac.in/courses/111105091

Course: BIOINFORMATICS **Code: MBD535**		
TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**
Theory: 2 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**
Course Prerequisites: No		
Course Objectives: To develop the knowledge of basic principles and concept of biology, computer science and mathematics, extract information from large databases and use this information in computer modeling.		
Course Outcomes: After completing the course, the student shall be able to: ● Knowledge and awareness of the basic principles and concepts of biology, computer science and mathematics. ● Existing software effectively to extract information from large databases and to use this information in computer modeling.		
Course Content:		

Unit-I	Sequence Alignment problem & Algorithm, Pairwise & Multiple sequence Alignment, Advance Alignment Method.	**7T + 4 P (15 hours)**
Unit-II	Gibbs Sampling, Population Genomics, Genetic Mapping, Disease Mapping	**8T + 4 P (16 hours)**
Unit-III	Gene Recognition, Transcriptome & Evolution, Protein Structure, Protein Motifs.	**7T + 4 P (15 hours)**
Unit-IV	Hidden Markov Models, Lattice Model, Algorithms.	**8T + 3 P (14 hours)**

Text Book:

1. Arthur Lesk, "Introduction to Bioinformatics", Oxford University Press, 4th Edition, 2014.
2. Edward H. Shortliffe and James J. Cimino, "Biomedical Informatics: Computer applications in Health Care and Biomedicince" , Springer, 5th Edition, 2021.

Reference Book:

1. C Setubal and J Meidanis, "Introduction Computational Molecular Biology", PWS Publishing Boston, 1997.
2. Venkatarajan Subramanian Mathura and Pandjassarame Kangueane, "Bioinformatics: A Concept-Based Introduction", Springer, 1st Edition, 2010.

E-Resources

1. https://nptel.ac.in/courses/102106065

Course: INFORMATION RETRIEVAL

Code: MBD536

TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS**
Theory: 2 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**

Course Prerequisites: No

Course Objectives: To develop proficiency in techniques and principles of information retrieval for efficient access, evaluation, and utilization of diverse data sources

Course Outcomes: After completing the course, the student will be able to

- Apply information retrieval principles to retrieve the relevant records in large collections of data.
- Describe the techniques of indexing in retrieval i.e. Boolean retrieval and rank based retrieval.
- Evaluate the information retrieval system.
- Describe web characteristics, how indexing & crawling is managed in a web server.

Course Content:

Unit-I	Introduction to Information Retrieval: IR Concepts, Boolean Retrievals, Inverted Index, Processing Boolean Queries. The Term Vocabulary and Postings Lists: Document Delineation and Character Sequence Decoding, Determining the Vocabulary of Terms.	**7T + 4 P (15 hours)**
Unit-II	Dictionaries and Tolerant Retrieval: Search Structures for Dictionaries, Wildcard Queries, Spelling Correction, Phonetic Correction. Index Construction: Hardware Basics Blocked Sort-Based Indexing. Scoring, Term Weighting. Vector Space Model: Parametric and Zone Indexes, Term Frequency and Weighting, The Vector Space Model for Scoring.	**8T + 4 P (16 hours)**
Unit-III	Evaluation in Information Retrieval: Information Retrieval System Evaluation, Standard Test Collections, Evaluation of Unranked Retrieval Sets, Evaluation of Ranked Retrieval Results. Text Classification & Naïve Bayes: The Bernoulli model, Properties of Naïve Bayes, Feature Selection, Evaluation of text classification.	**7T + 4 P (15 hours)**
Unit-IV	Web Search Basics: Web Characteristics, Advertising as the Economic Model, The Search User Experience, Index Size and Estimation, Near-Duplicates and Shingling. Web Crawling and Indexes: Overview, Crawling, Distributing Indexes, Connectivity Servers. Link Analysis: The Web as a Graph, Page Rank, Hubs and Authorities.	**8T + 3 P (14 hours)**

Text Book:

1. C. Manning, P. Raghavan and H. Schütze, "Introduction to Information Retrieval", Cambridge University Press, 2011.
2. R. B. Yate and B. R.Neto, "Modern Information Retrieval", Addison Wesley, 2nd Edition, 2012

Reference Book:

1. S. Butcher and C.L.A. Clarke, "Information Retrieval – Implementing and Evaluating Search Engines" The MIT Press, 1st Edition, 2016.

Course: MODELLING IN OPERATIONS MANAGEMENT
Code: MBD537

TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**
Theory: 2 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment	**3**

	/Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	

Course Prerequisites: No

Course Objectives: To develop the knowledge of input-output framework, roles and responsibilities of operations managers.

Course Outcomes: After completing the course, the student shall be able to:

- Understand the input–process–output framework, the extensions of it, and apply them to a wide range of operations.
- Examine the types of transformation processes occurring within operations.
- Define the roles and responsibilities of operations managers and the challenges they face.
- Reflect on your own operations management responsibilities, if applicable.
- Understand the content of an operations strategy and the decisions involved.

Course Content:

Unit-I	Classify various operations management problems, Identify the nature of the information needed to be able to address the problem, translate these problems into the appropriate statistical and/or mathematical framework and interpret the results of the models in a verbal manner of the case study: **Venture Analysis, Banking analytics**	**7T + 4 P (15 hours)**
Unit-II	Classify various operations management problems, Identify the nature of the information needed to be able to address the problem, translate these problems into the appropriate statistical and/or mathematical framework and interpret the results of the models in a verbal manner of the case study**: Marketing analytics**	**8T + 4 P (16 hours)**
Unit-III	Classify various operations management problems, Identify the nature of the information needed to be able to address the problem, translate these problems into the appropriate statistical and/or mathematical framework and interpret the results of the models in a verbal manner of the case study: **Healthcare analytics, Retail analytics**	**7T + 4 P (15 hours)**
Unit-IV	Classify various operations management problems, Identify the nature of the information needed to be able to address the problem, translate these problems into the appropriate statistical and/or mathematical framework and interpret the results of the models in a verbal manner of the case study**: Supply chain analytics**	**8T + 3 P (14 hours)**

Text Book:

1. S. Anil Kumar and N. Suresh, "Operations Management", New Age International (P) Ltd. Publishers, 2009.
2. Lalit Kumar Awasthi, Sushendra Kumar Misra, Dilbagh Panchal and Mohit Tyagi, "Operations Management and Data Analytics Modelling", CRC Press, 1st Edition, 2021.

Reference Books:

1. William J. Stevenson, "Operations Management Models: A Problem-Solving Approach", McGraw-Hill Higher Education, 2002.
2. Handfield and Nichols, "Introduction to Supply Chain Management", Prentice Hall India Learning Private Limited, 1st Edition, 2015.

Course: DATA PRIVACY and SECURITY **Code: MBD538**		
TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**

Theory: 2 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**

Course Prerequisites: No

Course Objectives: To educate students on fundamental principles, laws, technologies and best practices in data privacy and security to effectively safeguard sensitive information and mitigate risks in diverse technological environments.

Course Outcomes: After completing the course, the student shall be able to:

- Understanding of Data Privacy Laws and Regulations
- Enabling students to analyze and compare real-world data sharing practices.
- To equip students with comprehensive understanding and practical skills in various facets of network and system security

Course Content:

Unit I	Introduction to Data Privacy: Types of privacy attacks, Data linking and profiling, access control models, role based access control, privacy policies, their specifications, privacy policy languages, privacy in different domains-medical, financial, etc.	**7T + 3 P (13 hours)**
Unit-II	Mathematical model for comparing real-world data sharing practices, computing privacy and risk measurements. Demographics and Uniqueness. Protection Models-Null-map, k-map, Wrong map. Survey of techniques-Protection models (null-map, k-map, wrong map), Disclosure control, Inferring entity identities, entry specific databases. Computation systems for protecting delimited data-Min Gen, Datafly, Mu-Argus, k-Similar.	**8 T + 4 P (16 hours)**
Unit-III	Introduction to Security: The OSI Security Architecture, Security Attacks, Services and Mechanisms, Model for Network Security, Number theory, Cryptographic Hash Functions, Digital Signatures, System Security, Symmetric Encryption and Message Confidentiality, Substitution ciphers, Stream ciphers, Public-key cryptography and Message Authentication, Key Distribution and Authentication.	**8 T + 4 P (16 hours)**
Unit-IV	Security metrics: Design, Data sources, Analysis of security metrics data, Measuring security cost and value, Different context for security process management. Acquisition and Duplication: Sterilizing Evidence Media, Acquiring Forensics Images, Acquiring Live Volatile Data, Data Analysis, Metadata Extraction, and File System Analysis.	**7 T + 4 P (15 hours)**

Text Book:

1. Ronald Leenes, Rosamunde van Brakel, Serge Gutwirth and De Hert Paul, “Data Protection and Privacy: The Age of Intelligent Machines (Computers, Privacy and Data Protection)”, Hart Publishing, 2017.
2. Milan Petkovic and Willem Jonker, "Security, Privacy, and Trust in Modern Data Management", Springer Science & Business Media, 1st Edition, 2007.

Reference Books:

1. William Stallings, “Cryptography and Network Security: Principles and Practice”, 7th Edition, Pearson Education, 2017.

E-Resources:

1. https://nptel.ac.in/courses/106106129

Course: ARTIFICIAL INTELLIGENCE
Code: MBD539

TEACHING SCHEME:	**EXAMINATION SCHEME:**	**CREDITS ALLOTED:**
Theory: 2 **Tutorial: 0** **Practical: 1 (2 Hours)**	**Internal Assessment: 40 Marks** CIA-I: 20 Marks (Unit I & II) CIA-II: 20 Marks (Written Exams/ Quizzes /Assignment /Presentations/ Viva-Voce) **End Semester Examination: 60 Marks (Unit I- IV)**	**3**

Course Prerequisites: No

Course Objectives: To develop the knowledge of intelligent agents, different search techniques, prolog language, formalization of knowledge, and reasoning.

Course Outcomes: After completing the course, the student shall be able to:

- Understand the concepts of rationality, nature of environment, and structure of agents.
- Apply the concepts and methods of different search and problem solving techniques.
- Equip with understanding of concepts of knowledge representation and reasoning in uncertainty.
- Understand the concepts of reinforcement learning and solve the problems using prolog programming.

Course Content:

Unit I	Introduction to AI and Intelligent Agents: What is AI, Foundations and History of AI, Turing Test, Applications of AI, Agents and Environments, Concept of Rationality, Structure of Agents (reflex, model-based, goal-based, utility-based, learning agents), Nature of Environments.	**7T + 3 P (13 hours)**
Unit-II	Problem Solving and Searching: Problem Space and Characteristics, Problem solving Agents, Production Systems, Uninformed Search Strategies- Depth First Search, Breadth First Search, Uniform Cost Search, Iterative Deepening, Heuristic Search Techniques- Best First Search, A* algorithm, AO* algorithm, Hill climbing and its Variations, Simulated Annealing, Genetic Algorithm, Problem in Game playing, Min-max algorithm, Alpha – Beta pruning, Constraint Satisfaction Problem.	**8 T + 4 P (16 hours)**
Unit-III	Knowledge Representation and Reasoning: Approaches of knowledge representation and issues, Propositional Logic, Semantic Nets, Representing knowledge using rules, Conceptual Dependencies, Production Rules, Conceptual Graphs, First Order Logic, Inference in first order logic, Unification and lifting, Backward Chaining, Resolution Principle, different types of uncertainty- degree of truth and degree of belief, Reasoning under uncertainty, Bayes Probabilistic Inference, Dempster-Shafer theory.	**8 T + 4 P (16 hours)**
Unit-IV	Reinforcement Learning: Introduction to reinforcement learning, active and passive reinforcement learning, generalization in reinforcement learning, applications of reinforcement learning. Prolog: Introduction, Goals, Prolog Terminologies, Variables, Control Structures, Matching in Prolog, Arithmetic Operators, Backtracking, Cuts, Recursion, Lists,	**7 T + 4 P (15 hours)**

Text Book:

1. S. Russell and P. Norvig, “Artificial Intelligence: A Modern Approach”, Pearson Education, 3rd Edition,

2015.
2. Elaine Rich and Kelvin Knight, “Artificial Intelligence”, Tata McGraw Hill, 3rd Edition, 2017.

Reference Books:

1. Nils J. Nilsson, “Artificial Intelligence: A new Synthesis”, Elsevier Publishers, 1st Edition, 2003.
2. W.F. Clocksin and C.S. Mellish, “Programming in PROLOG”, Springer, 5th edition, 2003.
3. Ivan Bratko, “Prolog Programming for Artificial Intelligence”, Pearson Education, 4th edition, 2011.

E-Resources:

1. https://nptel.ac.in/courses/106105077