Pandit Deendayal Energy University

School of Technology

Department of Computer Science and Engineering

Post Graduate Curriculum Handbook (Academic Year 2024-28)

M. Tech. (Data Science)
w. e. f. July, 2024.

Vision

"To contribute to the society by imparting transformative education and producing globally competent professionals having multidisciplinary skills and core values to do futuristic research & innovations."

Mission

- To accord high quality education in the continually evolving domain of Computer Engineering by offering state-of-the-art undergraduate, postgraduate, doctoral programmes.
- To address the problems of societal importance by contributing through the talent we nurture and research we do:
- To collaborate with industry and academia around the world to strengthen the education and multidisciplinary research ecosystem.
- To develop human talent to its fullest extent so that intellectually competent and imaginatively exceptional leaders can emerge in a range of computer professions.

Program Educational Objectives (PEOs)

PEO-1. Graduate will have successful professional career as innovators, entrepreneurs, and business professionals who will be able to adapt to an ever-changing world and its demands for computational and data analytic skills.

PEO-2. Graduate will undertake research work or pursue higher studies by acquiring in depth knowledge in data science and allied fields.

Program Outcomes (as per NBA-PG SAR Guidelines)

PO-1: An ability to independently carry out research /investigation and development work to solve practical problems.

PO-2: An ability to write and present a substantial technical report/document.

PO-3: Students should be able to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program.

PO-4: Students should be able to analyse and relate critically to different sources of information, datasets and data processes; and apply these to structure and formulate data-driven reasoning.

PO-5 : Students should be able to apply modern data science methods to the solution of real world business problems, communicate findings, and effectively present results using data visualization techniques for societal benefits.

PO-6: Recognize and analyse ethical issues in business related to intellectual property, data security, integrity, and privacy.

Course Outline (M.Tech-Data Science)

Semester 1

Subjects	Teaching Scheme	Total Credits
Probability and Statistics for Data Science	3-1-0	4
Foundations of Computational Data Science	3-0-0	3
Applied Machine Learning	3-0-0	3
Applied Machine Learning Laboratory	0-0-2	1
Data Engineering	3-0-0	3
Data Engineering Laboratory	0-0-2	1
Optimization and Bio Inspired Computing	3-0-0	3
Scientific Writing and Professional Ethics	2-0-0	2
Total Credits		**20**

Semester 2

Subjects	Teaching Scheme	Total Credits
Neural Network & Deep Learning	3-0-0	3
Neural Network & Deep Learning Laboratory	0-0-2	1
Department Elective I	3-0-0	3
Department Elective I Laboratory	0-0-2	1
Department Elective II	3-0-0	3
Department Elective III	3-0-0	3
Department Elective IV	3-0-0	3
Research Methodology	2-0-0	2
Seminar		1
Total Credits		**20**

Semester 3

Subjects	Teaching Scheme	Total Credits
Project Phase -I		13
Summer Internship/ IEP (6 Weeks)		1

Semester 4

Subjects	Teaching Scheme	Total Credits
Project Phase – II and Dissertation		16

List of Electives

		Program Elective	L	T	P	H	C
Industry Track	PE-1	Big Data & Cloud Computing	3	0	0	3	3
	PE-2	Big Data & Cloud Computing Laboratory	0	0	2	2	1
	PE-3	Graphs Algorithms and Mining	3	0	0	3	3
	PE-4	Financial Forecasting	3	0	0	3	3
	PE-5	MLOps	3	0	0	3	3
Research Track		**Program Elective**	**3**	**0**	**0**	**3**	**3**
	PE-1	Computer Vision and Image Processing	3	0	0	3	3
	PE-2	Computer Vision and Image Processing Laboratory	0	0	2	2	1
	PE-3	Speech Processing Applications in Data Science	3	0	0	3	3
	PE-4	Language Processing and Computational Linguistics	3	0	0	3	3
	PE-5	Data Science for Health Informatics	3	0	0	3	3

1st Semester

<table>
<tr><th colspan="5">Course Code</th><th colspan="6">Probability and Statistics for Data Science</th></tr>
<tr><th colspan="5">Teaching Scheme</th><th colspan="6">Examination Scheme</th></tr>
<tr><th rowspan="2">L</th><th rowspan="2">T</th><th rowspan="2">P</th><th rowspan="2">C</th><th rowspan="2">Hrs/Week</th><th colspan="3">Theory</th><th colspan="2">Practical</th><th rowspan="2">Total Marks</th></tr>
<tr><th>MS</th><th>ES</th><th>IA</th><th>LW</th><th>LE/Viva</th></tr>
<tr><td>3</td><td>1</td><td>0</td><td>4</td><td>4</td><td>25</td><td>50</td><td>25</td><td>-</td><td>-</td><td>100</td></tr>
</table>

PREREQUISITES: Basic knowledge of linear algebra.

COURSE OBJECTIVES

- To introduce the fundamental concepts and theorems of probability theory.
- Apply elements of stochastic processes for problems in real life.
- Understand elementary queuing concepts and apply elsewhere in Data science.
- Learn about Markov Chain Monte Carlo (MCMC) methods for parameter estimation.

UNIT 1: Fundamentals of Probability Axiomatic definition, properties, conditional probability, Bayes' rule and independence of events, Random variables, distribution function, probability mass and density functions, expectation, moments, moment generating function, Chebyshev's inequality, Special distributions: Bernoulli, binomial, geometric, negative binomial, hypergeometric, Poisson, exponential, gamma, Weibull, beta, Cauchy, double exponential, normal, Joint distributions, marginal and conditional distributions, moments, Mean and variance, Weak Law of large numbers and Central limit theorems.	**11 Hrs.**
UNIT 2: Statistical Inference Sampling Techniques and Estimation**:** Population, Sample, Parameters, Random sampling with and without replacement, Point estimation, maximum likelihood estimation (MLE) and method of moments, Confidence intervals for population parameters, Hypothesis Testing: Null and alternative hypotheses, Type I and Type II errors, Common hypothesis tests: t-tests, chi-squared tests, and ANOVA**.** Regression Analysis: Linear regression, Multiple regression, Residual analysis and diagnostics.	**11 Hrs.**
UNIT 3: Queuing Theory Elements of a queuing system, Standard notations and definitions, Little's Law, birth and death process models, Poisson process and its properties, M/M/1, M/M/m, M/G/1 queuing systems and their characteristics, Embedded Markov chains, other standard results from the literature, basic ideas of priority queuing systems, Modeling of computer systems, finite population models, Jackson networks.	**10 Hrs.**
UNIT 4: Bayesian Statistics **I**ndependence of random variables, covariance and correlation, Functions of random variables, Bayesian Inference: Bayes' theorem and posterior probability, Prior and posterior distributions, Maximum a posteriori estimation (MAP), Markov Chain Monte Carlo (MCMC) Methods: Introduction to MCMC, Metropolis-Hastings algorithm and Gibbs sampling, Practical implementation of MCMC for parameter estimation.	**10 Hrs.**
	42 Hrs.

COURSE OUTCOMES

On completion of the course, student will be able to

CO1 - Understand the basic principles of probability theory, including sample spaces, events, and probability axioms.
CO2 - Explain the concept of statistical inference and its role in data science.
CO3 - Apply statistical techniques to real-world data sets using software tools.
CO4 - Implement linear and logistic regression models for predictive analytics.
CO5 - Analyze the impact of outliers and influential observations on statistical results.
CO6 - Evaluate the goodness of fit for regression models and assess model assumptions.

TEXT/REFERENCE BOOKS

1. Robert V. Hogg, Joseph W. McKean and Allen T. Craig, "Introduction to Mathematical Statistics", Pearson Education.
2. Paul G. Hoel, Sidney Port, and Charles Stone, "Introduction to Probability Theory", Houghton Mifflin.
3. K. Rohatgi and A. K. Md. E. Saleh, "An Introduction to Probability and Statistics", John Wiley & Sons.
4. Dani Gamerman and Hedibert F. Lopes, "Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference", Second Edition, Chapman and Hall/CRC Press.
5. Boes, D. C., Graybill, F. A., Mood, A. M. "Introduction to the Theory of Statistics" McGraw-Hill International Book Company.

Course Code					Foundations of Computational Data Science					
Teaching Scheme					Examination Scheme					
L	T	P	C	Hrs/Week	Theory			Practical		Total Marks
					MS	ES	IA	LW	LE/Viva	
3	0	0	3	3	25	50	25	-	-	100

COURSE OBJECTIVES

- To understand and apply advanced principles in data structures, algorithms, and operating systems.
- Understand the utilization of parallel computing architectures and programming models for high-performance computing.
- Demonstrate proficiency in implementing CUDA for GPU programming and optimizing data science workflows.

UNIT 1: Advance Data Structures Advance data structures like B-trees, Fibonacci heaps, and interval trees, Disjoint-set data structures and their applications, Data structures for efficient storage and retrieval of large datasets, Different Types of Graphs, Euler Circuits, Hamiltonian Graphs and Chromatic Numbers.	**9 Hrs.**
UNIT 2: Advanced Algorithms Analysis of algorithm efficiency and complexity. Review of design techniques: greedy method, divide-and-conquer, dynamic programming with advanced applications and optimization. Algorithmic Optimization Techniques: Approximation Algorithms, Randomized Algorithms, Amortized Analysis. Other algorithms such as, convex hull algorithms, line segment intersection, closest pair points, algorithms for maximum flow and string matching.	**11 Hrs.**
UNIT 3: Advance Operating System Principles of modern operating systems: advanced process/thread management, Multithreading and concurrency, Inter-process communication mechanisms. File system organization and structure with security and access control. Distributed operating systems: principles, architectures, and communication protocols. Virtualization techniques: virtual machines, hypervisors, and containerization. Performance measurement and monitoring tools.	**11 Hrs.**
UNIT 4: Parallel Computing With Cuda Parallel computing architectures: shared-memory, distributed-memory, and hybrid architectures. Parallel programming models: threading models, OpenMP, MPI, and GPU programming with CUDA. Programming Models in high-performance computing architectures: (Examples: IBM CELL BE, Nvidia Tesla GPU, Intel Larrabee Micro architecture and Intel Nehalem micro architecture) Case studies and projects applying advanced techniques to optimize data science workflows and performance.	**11 Hrs.**
	42 Hrs.

COURSE OUTCOMES

Upon completion of this course, students will be able to:

CO1 - Understand the data structures and algorithms for solving complex computational problems.
CO2 - Demonstrate diverse computational methodologies for large-scale parallel applications.
CO3 - Apply the principles of modern operating systems, including process/thread management and concurrency.
CO4 - Analyze the efficiency of data science workflows through the application of CUDA for GPU programming.
CO5 - Compare appropriate virtualization techniques for efficient resource management.
CO6 - Evaluate the performance and scalability of parallel computing architectures and programming models in real-world applications.

TEXT/REFERENCE BOOKS

1. Ellis Horowitz, SartajSahni, “Data Structures, Algorithms and Applications in C++”, Universities Press/Orient Longman.
2. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest and Clifford Stein, “Introduction to Algorithms”, MIT Press.
3. Abraham Silberschatz, Peter B. Galvin, Greg Gagne, Operating System Concepts, Wiley.
4. Maekawa, Oldehoeft, OS: Advanced Concepts, Addison-Wesley.
5. Pacheco S. Peter, “Parallel Programming with MPI”, Morgan Kaufman Publishers.
6. Shane Cook, CUDA Programming: A Developer's Guide to Parallel Computing with GPUs, Morgan Kaufmann publishers.

<table>
<tr><th colspan="5">Course Code</th><th colspan="6">Applied Machine Learning</th></tr>
<tr><th colspan="5">Teaching Scheme</th><th colspan="6">Examination Scheme</th></tr>
<tr><th rowspan="2">L</th><th rowspan="2">T</th><th rowspan="2">P</th><th rowspan="2">C</th><th rowspan="2">Hrs/Week</th><th colspan="3">Theory</th><th colspan="2">Practical</th><th rowspan="2">Total Marks</th></tr>
<tr><th>MS</th><th>ES</th><th>IA</th><th>LW</th><th>LE/Viva</th></tr>
<tr><td>3</td><td>0</td><td>0</td><td>3</td><td>3</td><td>25</td><td>50</td><td>25</td><td>-</td><td>-</td><td>100</td></tr>
</table>

PREREQUISITES: Basic knowledge of linear algebra, probability, and statistics.

COURSE OBJECTIVES

- To provide students with a solid understanding of core concepts and techniques in machine learning.
- To equip students with the skills and knowledge necessary to apply machine learning techniques to real-world problems effectively.
- How to evaluate the performance of machine learning models using appropriate metrics and validation techniques

UNIT 1: Supervised Learning Introduction to Machine Learning: Machine Learning Introduction, Bias and Variance, Oversampling and Undersampling, Different Methods to Handle Overfitting and Underfitting, Cross-validation. Supervised Learning Techniques: Different types of Regression, Classification, Naive Bayes Classifier, KNN, Decision Boundary, SVM, Decision Tree, Logistic Regression, Discriminant Functions, Fisher's Linear Discriminant and Thresholding for Classification, Evaluation Metrics for Supervised Learning.	**12 Hrs.**
UNIT 2: Unsupervised Learning Clustering: Partitional Clustering, Hierarchical Clustering, Spectral Clustering, Clustering using Self-Organizing Maps, K-Means Clustering, Fuzzy K-Means Clustering, DBSCAN, PCA, EM Algorithm, Gaussian Mixture Models, Evaluation Metrics for Unsupervised Learning	**10 Hrs.**
UNIT 3: Ensemble and Semi-Supervised Learning Ensemble Methods: Bagging, Boosting and Staking, Random Forests, Empirical Comparison among Ensemble Methods, Holdout Method and Random Subsampling, Bootstrapping, Semi-Supervised Learning and its Advantages, Self-Training, Co-Training, Multi-View Learning Approaches, Label Propagation and Graph-Based Methods.	**08 Hrs.**
UNIT 4: Reinforcement Learning Overview of Reinforcement Learning (RL), Elements of Reinforcement Learning, Exploration vs. Exploitation trade-off, Function Approximation in Reinforcement Learning, Basics of Dynamic Programming, Q-learning, Temporal Difference Learning (TDA), SARSA-Learning, Generalization.	**12 Hrs.**
	42 Hrs.

COURSE OUTCOMES

On completion of this course, students will be able to:

CO1 - Describe the concepts of supervised, unsupervised and semi-unsupervised learning techniques.
CO2 - Understand advanced machine learning techniques such as ensemble methods and reinforcement learning.
CO3 - Identified appropriate ML algorithm according to the problem.
CO4 - Analyze the concepts of bias and variance in machine learning models and apply techniques such as oversampling, undersampling, and regularization to mitigate overfitting and underfitting.
CO5 - Analyze the strengths and weaknesses of ensemble learning methods, including bagging, boosting, and random forests, and compare their performance
CO6 - Evaluate the performance of the different Machine Learning approaches.

TEXT/REFERENCE BOOKS

1. Tom M. Mitchell "Machine Learning", McGraw-Hill
2. Christopher M. Bishop ," Pattern Recognition and Machine Learning ", O'Reilly Media, Inc.
3. Sebastian Raschka , Vahid Mirjalili, "Python Machine Learning", Packt publication
4. Richard S. Sutton and Andrew G. Barto, "Reinforcement Learning: An Introduction", MIT Press

Course Code					Applied Machine Learning Laboratory					
Teaching Scheme					Examination Scheme					
L	T	P	C	Hrs/Week	Theory			Practical		Total Marks
					MS	ES	IA	LW	LE/Viva	
0	0	2	1	2	--	--	--	50	50	100

PREREQUISITES:

- Basic knowledge of linear algebra, probability, and statistics.
- Familiarity with programming languages such as Python.

COURSE OBJECTIVES

- To familiarize students with popular tools and libraries used in applied machine learning.
- To develop problem-solving skills and critical thinking abilities to approach and tackle different types of machine learning problems effectively.
- How to evaluate the performance of machine learning models using appropriate metrics and validation techniques

LIST OF EXPERIMENTS

1.	**Regression** Implement linear and non-linear regression.
2.	**Classification** Implement, visualize and compare different classification models.
3.	**Imbalanced Data Handling** Work with imbalanced datasets and implement techniques such as SMOTE, ADASYN
4.	**Clustering Techniques** Implement, visualize and compare different types of clustering algorithms.
5.	**Ensemble Methods - Random Forest and Bagging** Implement and visualize Random Forest Tree, Bagging, Boosting.
6.	**Semi-Supervised Learning - Self-Training and Co-Training** Implement and compare self and co-training algorithms.
7.	**Advanced Clustering - Spectral Clustering and Self-Organizing Maps (SOM)** Implement spectral clustering algorithm and SOM for clustering high-dimensional data.
8.	**Reinforcement Learning - Q-Learning** Implement Q-learning algorithm and experiment with different exploration vs. exploitation strategies.
9.	**Project**

COURSE OUTCOMES

On completion of this course, students will be able to:

CO1 - Apply feature engineering on the given dataset.

CO2 - Implement and visualize different supervised machine learning algorithms for regression and classification.

CO3 - Implement and visualize different unsupervised machine learning algorithms for clustering, and anomaly detection.

CO4 - Implement and visualize different semi-supervised and reinforcement machine learning algorithms.

CO5 - Interpret the results of machine learning experiments to draw meaningful conclusions and insights from the data.

CO6 - Design solutions to address real-world problems using machine learning algorithms.

TEXT/REFERENCE BOOKS

1. Tom M. Mitchell “Machine Learning”, McGraw-Hill
2. Christopher M. Bishop," Pattern Recognition and Machine Learning ", O'Reilly Media, Inc.
3. Sebastian Raschka, Vahid Mirjalili, “Python Machine Learning”, Packt publication
4. Manohar Swamynathan, “Mastering Machine Learning with Python in Six Steps: A Practical Implementation Guide to Predictive Data Analytics Using Python”, Apress.
5. Richard S. Sutton and Andrew G. Barto, “Reinforcement Learning: An Introduction”, MIT Press

<table>
<tr><th colspan="5">Course Code</th><th colspan="6">Data Engineering</th></tr>
<tr><th colspan="5">Teaching Scheme</th><th colspan="6">Examination Scheme</th></tr>
<tr><th rowspan="2">L</th><th rowspan="2">T</th><th rowspan="2">P</th><th rowspan="2">C</th><th rowspan="2">Hrs/Week</th><th colspan="3">Theory</th><th colspan="2">Practical</th><th rowspan="2">Total Marks</th></tr>
<tr><th>MS</th><th>ES</th><th>IA</th><th>LW</th><th>LE/Viva</th></tr>
<tr><td>3</td><td>0</td><td>0</td><td>3</td><td>3</td><td>25</td><td>50</td><td>25</td><td>--</td><td>--</td><td>100</td></tr>
</table>

PREREQUISITES: Data Structure, statistics, familiarity with programming languages such as python.

COURSE OBJECTIVES

- To provide strong foundation for data science and application area related to it
- Understand the underlying core concepts and emerging technologies in data science.

UNIT 1: Data Representation Data definition, variety of data, data life cycle, data lake, data representation: Array, Matrix, List tuple, dictionary, stack, queue. Data preprocessing: data cleaning, missing values assimilation, data inconsistency handling, Data privacy and ethics.	**12 Hrs.**
UNIT 2: Data Visualization Introduction, Types of data visualization, Data for visualization: Data types, Data encodings, Retinal variables, mapping variables to encodings, Visual encodings, explore tools libraries: Matplotlib, seaboarn. Data visualization using tableau/Power BI	**10 Hrs.**
UNIT 3: Data Analysis Understanding the concept of data inconsistency, Common causes of inconsistent data in real-world datasets, Impact of inconsistent data on data analysis and machine learning models, Data normalisation and standardization process, Feature scaling, Dimensionality Reduction: Principal Component Analysis (PCA), Decision boundaries, t-distributed Stochastic Neighbor Embedding (t-SNE), Singular Value Decomposition (SVD)	**10 Hrs.**
UNIT 4: Time Series Data Handling Introduction: Examples, simple descriptive techniques, trend, seasonality, the correlogram. Probability models for time series: stationarity. Moving average (MA), Autoregressive (AR), ARMA and ARIMA models. Estimating the autocorrelation function and fitting ARIMA models. Forecasting: Exponential smoothing, Forecasting from ARIMA models. Stationary processes in the frequency domain: The spectral density function, the periodogram, spectral analysis.	**10 Hrs.**
	42 Hrs.

COURSE OUTCOMES

On completion of the course, student will be able to
CO1 - Understand the variety of data and their representations.
CO2 - Understand data analysis techniques for applications handling large data.
CO3 - Apply data visualization tools and libraries to present the data inference.
CO4 - Apply statistical models MA, AR, ARMA, ARIMA on the time series data.
CO5 - Analyze the ethics surrounding data privacy, data sharing and algorithmic decision-making.
CO6 - Design data driven solutions for the real-world problems.

TEXT/REFERENCE BOOKS

1. Joel Grus, Data Science from Scratch, O'Reilly Media, Inc.
2. Cathy O’Neil, Rachel Schutt, Doing Data Science, Straight Talk from The Frontline. O’Reilly Media Inc.
3. Introducing Data Science, Davy Cielen, Arno D. B. Meysman, Mohamed Ali, Manning Publications Co.
4. Storytelling with Data: A Data Visualization Guide for Business Professionals,v Nussbaumer Knaflic, Wiley
5. Time Series Analysis 1st Edition, James D. Hamilton, Princeton University Press.

<table>
<tr><th colspan="5">Course Code</th><th colspan="6">Data Engineering Laboratory</th></tr>
<tr><th colspan="5">Teaching Scheme</th><th colspan="6">Examination Scheme</th></tr>
<tr><th rowspan="2">L</th><th rowspan="2">T</th><th rowspan="2">P</th><th rowspan="2">C</th><th rowspan="2">Hrs/Week</th><th colspan="3">Theory</th><th colspan="2">Practical</th><th rowspan="2">Total Marks</th></tr>
<tr><th>MS</th><th>ES</th><th>IA</th><th>LW</th><th>LE/Viva</th></tr>
<tr><td>0</td><td>0</td><td>2</td><td>2</td><td>2</td><td>--</td><td>--</td><td>--</td><td>50</td><td>50</td><td>100</td></tr>
</table>

COURSE OBJECTIVES

- Provide a foundational understanding of database management systems and operations.
- Foster practical skills in data manipulation and analysis using Python.
- Develop proficiency in exploratory data analysis techniques for insightful visualization and interpretation.

Dataset Repository Source: Students may use publicly available datasets such as https://archive.ics.uci.edu/.

LAB EXPERIMENTS

1.	**SQL Commands** Create Relational Table for basic DDL, DML and DCL operation
2.	**Python Basics** Implementation of python data structures, collections and dataframes.
3.	**Data Preprocessing** Data Cleaning, Missing Values, Data Balancing SMOTE.
4.	**Data Visualization using graphical libraries** Data Visualization using matplotlib, pyplot, seaborn.
5.	**Data Visualization using Tableau** Collecting data, Plotting Graphs, and charts, Implement dashboard
6.	**Time Series Analysis** Implementation of ARMA, ARIMA models
7.	**Project** Design and Development of projects

COURSE OUTCOMES

On completion of the course, student will be able to

CO1 - Understand fundamental concepts of database management systems, including basic operations such as Data Definition Language (DDL), Data Manipulation Language (DML), and Data Control Language (DCL), demonstrated through the creation and manipulation of relational tables.
CO2 - Apply data handling techniques for structured data management using data frames and arrays.
CO3 - Apply data pre-processing techniques for cleaning, scaling to ensure data quality.
CO4 - Apply visualization techniques and tools using graphical libraries, allowing effective representation and interpretation of data.
CO5 - Analyze trends and seasonality in the data.
CO6 - Develop the solutions for real-time data driven problems.

TEXT/REFERENCE BOOKS

1. Joel Grus, Data Science from Scratch, O'Reilly Media, Inc.
2. Cathy O’Neil, Rachel Schutt, Doing Data Science, Straight Talk from The Frontline. O’Reilly Media Inc.
3. A Hands-On Introduction to Data Science, Chirag Shah, Cambridge University Press.
4. Storytelling with Data: A Data Visualization Guide for Business Professionals,v Nussbaumer Knaflic, Wiley.
5. Time Series Analysis 1st Edition, James D. Hamilton, Princeton University Press.

Course Code					Optimization and Bio-inspired Computing					
Teaching Scheme					Examination Scheme					
L	T	P	C	Hrs/Week	Theory			Practical		Total Marks
					MS	ES	IA	LW	LE/Viva	
3	0	0	3	3	25	50	25	-	-	100

PREREQUISITES: Basic knowledge of linear algebra, calculus, data structure, and algorithm.

COURSE OBJECTIVES

- To understand of optimization techniques, including linear and non-linear programming.
- To understand how to formulate real-world problems as optimization tasks.
- To equip students with the knowledge, skills, and ethical awareness necessary to utilize nature-inspired algorithms effectively in solving real-world optimization.

UNIT 1: Linear and Non Linear Optimization Introduction to optimization, Principle of Optimality, Classification of optimization techniques, Single variable optimization, Linear Programming: Simplex method, Revised simplex method, Duality in linear programming, Decomposition principle, Post-optimality analysis, Sensitivity analysis, Non Linear Programming: First and second order conditions, elimination methods, Direct root method.	**11 Hrs.**
UNIT 2: Unconstrained and Constrained Optimization Constrained Optimization Techniques: Standard form of the problem and basic terminology, Lagrange variables, Karush-Kuhn-Tucker conditions, Regular points, Sensitivity analysis, Sequential Linear Programming, Quadratic programming. Unconstrained Optimization Techniques: Introduction, Random search method, Univariate and pattern search method, Indirect search method-Steepest Descent (Cauchy) method, Conjugate gradient method, Newtons Method, Quasi-Newton Method. Convex/Non-convex Analysis: Convex Sets, Convex Functions, Convex programming problem, Duality-Lagrange and Conic duality, gradient based optimization, Generalised Reduced gradient method, stochastic gradient descent, Momentum, Lipschitzness, Adam and other variants, Bergman dynamics.	**15 Hrs.**
UNIT 3: Nature Inspired Optimization and Modelling Genetic programming, Evolution strategies, Genetic algorithms-Representation and Reproduction, Crossover and Mutation Operators, Crossover and Mutation rates, Selection mechanisms, Fitness proportionate, ranking and tournament selection, Building Block – Hypothesis and Schema Theorem.	**08 Hrs.**
UNIT 4: Swarm Intelligence Modelling Stigmergy, Competition and Cooperation, Particle Swarm Optimization, Anatomy of a particle, Velocity and Position updation, PSO topologies, Control parameters, Ant Colony Optimization: Pheromone updation and evaporation, Bee Algorithms, Hybridization of nature-inspired algorithms. Artificial immune systems, Cultural algorithms, Memetic algorithms, Physics-Based Algorithms, Bio-inspired computing (BIC) in artificial intelligence, Simulated annealing, Quantum-inspired algorithms, Firefly algorithm.	**08 Hrs.**
	42 Hrs.

COURSE OUTCOMES

On completion of the course, student will be able to

CO1 - Understand the fundamental concepts of optimization, including linear and non-linear programming.
CO2 - Evaluate the strengths, limitations, and potential applications of nature and biologically inspired algorithms in various domains.
CO3 - Apply convex and non-convex techniques to solve optimization problems.
CO4 - Evaluate the performance of optimization algorithms using appropriate metrics.
CO5 - Design and formulate optimization models for industrial engineering problems.
CO6 - Develop optimal solutions using nature-inspired algorithms for real-world problems.

TEXT/REFERENCE BOOKS

1. Kalyanmoy Deb, "Optimization for Engineering Design Algorithms and Examples", Prentice Hall of India.
2. Edwin K.P. Chong and Stanislaw H. Zak, "An Introduction to Optimization", Wiley-Interscience Series in Discrete Mathematics and Optimization.
3. S.S. Rao, "Optimization Theory and Applications", New Age International (P) Limited Publishers.
4. Jorge Nocedal and Stephen J. Wright, "Numerical Optimization: Springer Series in Operations Research and Financial Engineering", Springer Science & Business Media.
5. M. Asghar Bhatti, "Practical Optimization Methods: with Mathematics Applications", Springer Verlag Publishers.

Course Code: XXXXXX					Scientific Writing and Publication Ethics					
Teaching Scheme					**Examination Scheme**					
L	**T**	**P**	**C**	**Hrs./Week**	**Theory**			**Practical**		**Total Marks**
					MS	**ES**	**IA**	**LW**	**LE/Viva**	
2	**0**	**0**	**2**	**2**	**25**	**50**	**25**	**--**	**--**	**100**

COURSE OBJECTIVES

1. To comprehend the significance of scientific writing and to understand of the basic structure of a scientific paper.
2. To get familiarize with the process of selecting appropriate target journals and conferences.
3. To cultivate an awareness of publication ethics within the realm of scientific writing.
4. To get acquainted with the knowledge and tools necessary to identify, understand, and avoid plagiarism in scientific writing

UNIT-1: Introduction to Scientific Writing Importance of scientific writing in engineering, understanding the structure and components of a scientific paper, research paper writing style, referencing style	**07 Hrs.**
UNIT 2: Selecting Target Journals and Conferences Types of journals and conferences in engineering, open access journals, journal impact factors, conference rankings, manuscript submission process, responding to reviewer comments	**07 Hrs.**
UNIT 3: Publication Ethics Introduction and importance, publication misconduct, violation of publication ethics, falsification and/or fabrication of data, understanding of copyright form, collaboration issues (authorship), conflicts of interest issues, Committee on Publication Ethics (COPE)	**07 Hrs.**
UNIT 4: Avoiding Plagiarism Plagiarism – definition, reasons for plagiarism, types of plagiarism, avoiding plagiarism	**07 Hrs.**
TOTAL	**28 Hrs**

COURSE OUTCOMES

On completion of the course, student will be able to:

CO1- Describe the importance of scientific writing in engineering and identifying its role in knowledge dissemination and academic integrity
CO2 - Understand the structure and components of a scientific paper
CO3 - Evaluate and select suitable journals and conferences to submit their research work
CO4 - Understand publication ethics
CO5 - Define plagiarism, identify its different types and reasons, and apply techniques to avoid plagiarism
CO6 - Analyze and respond to reviewer comments for their research work

TEXT/REFERENCE BOOKS

1. Getting It Published: A Guide for Scholars and Anyone Else Serious about Serious Books by William Germano
2. Publish and Flourish: Become a Prolific Scholar by Tara Gray
3. Adil E. Shamoo, and David B. Resnik, Responsible Conduct of Research, Oxford University Press
4. Gary Comstock, Research Ethics: A Philosophical Guide to the Responsible Conduct of Research, Cambridge University Press
5. Tony Mayer, and Nicholas H. Steneck, Promoting Research Integrity in a Global Environment, World Scientific Publishing
6. Ethical Issues in Engineering Research, Publication, and Practice by Caroline Whitbeck

2nd Semester

Course Code					Neural Network and Deep Learning					
Teaching Scheme					**Examination Scheme**					
L	**T**	**P**	**C**	**Hrs/Week**	**Theory**			**Practical**		**Total Marks**
					MS	**ES**	**IA**	**LW**	**LE/Viva**	
3	**0**	**0**	**3**	**3**	**25**	**50**	**25**	**--**	**--**	**100**

PREREQUISITES: Probability and Random Processes, Linear Algebra

COURSE OBJECTIVES

- Understand the concepts and applications of neural network.
- Understand the concepts, usage and impact of deep learning algorithms in various domains.
- Discuss advanced deep-learning algorithms to solve real life problems.

UNIT 1: Neural Networks & Learning Process McCulloch Pitts Neuron, ADALINE network, Sigmoid Neurons, Feed Forward Neural Networks, Single Layer Perception: Introduction, Linear classifier, Simple perception, Perception learning algorithm, Modified Perception learning algorithm, Adaptive linear combiner, Continuous perception, Back propagation algorithm, learning in continuous perception, Limitation of Perception, Parameters v/s Hyper-parameters, Learning Rule: Hebb's Rule. Multi-Layer Perceptron Networks: Introduction, MLP with hidden layers, Simple layer of a MLP, Delta learning rule of the output layer, Multilayer feed forward neural network with continuous perceptions, Generalized delta learning rule.	**8 Hrs.**
UNIT 2: Deep Neural Network Neuro architectures as necessary building blocks for the DL techniques. Deep Convolutional Neural Networks, Feature extraction, Deep Belief Networks, Different deep CNN architectures LeNet, AlexNet, VGG, PlacesNet.	**11 Hrs.**
UNIT 3: Improving Deep Neural Networks Hyperparameter tuning, Regularization, Loss functions, Dropouts, Cross validations, Optimizations, Revisiting Gradient Descent, Momentum Optimizer, RMSProp, Adam, Batch Normalization	**11 Hrs.**
UNIT 4: Deep Learning Models Recurrent Neural Network, Long Short-Term Memory, Gated Recurrent Unit, Boltzmann Machine, Generative Adversarial Network, Transfer Learning.	**12 Hrs.**
	42 Hrs.

COURSE OUTCOMES

On completion of the course, student will be able to

CO1 - Understand the theoretical foundations, and practical applications of neural network and deep learning.
CO2 - Understand the mathematical foundation of neural network architecture.
CO3 - Develop single-layer and multiple-layer perceptron learning algorithms.
CO4 - Design layered networks using deep learning principles, exploring advanced architectures like CNNs and RNNs.
CO5 - Identify appropriate deep learning algorithms by considering algorithmic characteristics and task requirements.
CO6 - Design complex engineering solutions using advanced deep learning models.

TEXT/REFERENCE BOOKS

1. Simon Haykins, Neural Networks and Learning Machines, Pearson Prentice Hall.
2. Ian Goodfellow, Yoshua Bengio and Aaron Courville, Deep Learning, MIT Press.
3. Christopher M. Bishop. Neural Networks for Pattern Recognition. Oxford University Press.

Course Code					Neural Network and Deep Learning Laboratory					
Teaching Scheme					Examination Scheme					
L	T	P	C	Hrs/Week	Theory			Practical		Total Marks
					MS	ES	IA	LW	LE/Viva	
0	0	2	2	2	--	--	--	50	50	100

COURSE OBJECTIVES

- To build the foundation of deep learning.
- To understand how to build the neural network.
- To enable students to develop successful machine learning concepts

LAB EXPERIMENT

1.	**Understanding of Activation Functions** Implementation of different activation functions to train Neural Network.
2.	**Learning Rules** Implementation of different Learning Rules
3.	**Data Augmentation** Implement various data augmentation techniques
4.	**Implementation of Feed forward Network** Building the Simple Feed Forward Neural Network
5.	**Understanding of Back Propagation Techniques** Implement a complete neural network including back propagation techniques
6.	**Implementation of Convolution Neural Network** Implement convolution neural network, Convolution, Polling, Flattern, Fully Connected
7.	**Implementation of Recurrent Neural Network** Implementation of recurrent Neural Network
8.	**Boltzmann Machine** Implementation of Boltzmann Machine
9.	**Generative Adversarial Network** Implementation Generative Adversarial Network
10.	**Transfer Learning Techniques** Implementaion of ResNet, VGG pretrained models

COURSE OUTCOMES

On completion of the course, student will be able to

CO1 - Understand neural network basics, including activation functions, learning rules, and backpropagation techniques.
CO2 - Implement activation functions to optimize neural network training, assessing their impact on model performance.
CO3 - Build and train basic feedforward neural networks using backpropagation techniques.
CO4 - Apply convolutional neural networks (CNNs) for image processing, understanding their structure and applications.
CO5 - Implement recurrent neural networks (RNNs) for sequential data tasks like NLP and time series prediction.
CO6 - Integrate diverse deep learning techniques in a practical project, demonstrating proficiency in solving real-world problems effectively.

TEXT/REFERENCE BOOKS

1. Simon Haykins, Neural Networks and Learning Machines, Pearson Prentice Hall.
2. Ian Goodfellow, Yoshua Bengio and Aaron Courville, Deep Learning, MIT Press.
3. Christopher M. Bishop. Neural Networks for Pattern Recognition. Oxford University Press.

Course Code: XXXXXX					Research Methodology					
Teaching Scheme					Examination Scheme					
L	T	P	C	Hrs./Week	Theory			Practical		Total Marks
					MS	ES	IA	LW	LE/Viva	
2	0	0	2	2	25	50	25	--	--	100

COURSE OBJECTIVES

1. To understand the role of research in the field of engineering and get an overview of the research process.
2. To develop proficiency in literature review techniques.
3. To understand the process of formulating and solving research problems.
4. To understand different types of intellectual property rights.

UNIT I : Introduction to Research Role of research in engineering, research process overview, types of research, outcomes of research, characteristics of a researcher, research terminology	**06 Hrs.**
UNIT II : Literature Review Techniques Searching for the existing literature, reviewing the selected literature, developing a theoretical framework, developing a conceptual framework	**06 Hrs.**
UNIT III : Formulating and Solving a Research Problem Importance of formulating a research problem, sources of research problems, identifying a problem, formulation of research objectives and research questions, Need for research design, different research designs, experimental test-setups, data sampling, data collection, data analysis & interpretation	**08 Hrs.**
UNIT VI: Intellectual Property Rights Introduction and significance of intellectual property rights, types of intellectual property rights, introduction to patents, patent drafting and filing, copyright, trademarks, industrial design, geographical indicators	**08 Hrs.**
Total	**28 Hrs.**

COURSE OUTCOMES

On completion of the course, student will be able to:

CO1 - U**nderstand the role and significance of research in engineering**
CO2 - Develop understanding of the basic framework of research process and design
CO3 - Identify technical gaps in the literature and formulate a problem.
CO4 - Develop an understanding of various research designs and techniques.
CO5 - Develop an understanding of the ethical dimensions of conducting applied research
CO6 - E**valuate and apply intellectual property rights concepts to the research outcomes**

TEXT/REFERENCE BOOKS

1. Stuart Melville, Wayne Goddard, Research Methodology: An Introduction for Science and Engineering Students, Juta & Co. Ltd.
2. David V. Thiel, Research Methods for Engineers, Cambridge University Press, UK
3. Ranjit Kumar, Research Methodology: A Step by Step Guide for Beginners, Pearson
4. CR Kothari, Research Methodology (Methods and Techniques), New age Publications

<table>
<tr><th colspan="5">Course Code</th><th colspan="6">Big Data and Cloud Computing</th></tr>
<tr><th colspan="5">Teaching Scheme</th><th colspan="6">Examination Scheme</th></tr>
<tr><th rowspan="2">L</th><th rowspan="2">T</th><th rowspan="2">P</th><th rowspan="2">C</th><th rowspan="2">Hrs/Week</th><th colspan="3">Theory</th><th colspan="2">Practical</th><th rowspan="2">Total Marks</th></tr>
<tr><th>MS</th><th>ES</th><th>IA</th><th>LW</th><th>LE/Viva</th></tr>
<tr><td>3</td><td>0</td><td>0</td><td>3</td><td>3</td><td>25</td><td>50</td><td>25</td><td>-</td><td>-</td><td>100</td></tr>
</table>

PREREQUISITE: SQL and Programming Languages.

COURSE OBJECTIVES

- Able to understand concepts, techniques, and technologies of big data and cloud computing in terms of infrastructure and resource management.
- Learn about various big data technologies, tools, and frameworks, as well as the architectural components and deployment models of cloud computing.
- Learn to build and maintain reliable, scalable, distributed systems with Hadoop.
- Able to apply Hadoop ecosystem components and working with SPARK for data analytics.

UNIT 1: Introduction to Big Data and Hadoop Basics of Structured, Semi-structured and Unstructured Data, Hadoop Architecture, Data Serialization, Apache Hadoop & Hadoop Ecosystem, Analysing Data with Hadoop, MAP-Reduce paradigm, Hadoop Streaming, Eco System components: PIG, HIVE, KAFKA, SPARK, YARN. Introduction to NOSQL	**11 Hrs.**
UNIT 2: Data Analysis with Spark Introduction to Data Analysis with Spark, Programming with RDDs,Machine Learning with MLlib, SQL, GraphX, structured streaming, Data Visualization, Use cases for data analytics.	**10 Hrs.**
UNIT 3: Fundamentals of Cloud Computing Basic concepts and terminology, Goals and benefits of cloud technology, Risks and Challenges, Cloud Characters, Cloud Delivery Models, Cloud Deployment Models, Cloud Enabling Technologies: Broadband Networks and Internet Architecture, Data Center Technology, Virtualization Technology, Web Technology, Multitenant Technology, Containerization, Cloud Security: Basic Terms and Concepts, Threat Agents, Cloud Security Threats, Additional Considerations	**13 Hrs.**
UNIT 4: Industrial Platforms and New Developments Study of Cloud Computing Systems like Amazon EC2 and S3, Google App Engine, and Microsoft Azure, Build Private/Hybrid Cloud using open-source tools. MapReduce and its extensions to Cloud Computing. Cloud solutions for Big Data: AWS DocumentDB, Databricks, Hortonworks Data Cloud, Cloudera Data Platform. Case studies: Netflix, Airbnb, GE.	**8 Hrs.**
	Total 42 Hrs.

COURSE OUTCOMES

On completion of the course, student will be able to

CO1 - Explain the strengths and limitations of cloud computing and the possible applications for state-of-the-art cloud computing.
CO2 - Identify the architecture and infrastructure of cloud computing.
CO3 - Understand industrial cloud platforms, like Amazon web services, Google AppEngine, Microsoft Azure and Open-source cloud building tools.
CO4 - Apply the concepts of Hadoop for Map Reduce Programming
CO5 - Analyze and exploit different features of SPARK for big data applications
CO6 - Develop solutions for application development using big data and cloud computing technologies.

TEXT/REFERENCE BOOKS:

1. Thomas Erl, Zaigham Mahmood, Ricardo Puttini, Cloud Computing Concepts, Technology Architecture, PHI.
2. Rajkumar Buyya, James Broberg,Andrzej M Goscinski, Cloud Computing: Principles and Paradigms, Wiley.
3. Toby Velte, Anthony Velte, Cloud Computing: A Practical Approach, McGraw-Hill Osborne Media
4. Chris Eaton, Dirk deroos et al., “Understanding Big Data”, McGraw Hill.
5. Tom White, “HADOOP: The definitive Guide”, O Reilly.
6. Boris lublinsky, Kevin t. Smith, Alexey Yakubovich, “Professional Hadoop Solutions”, Wiley.

Course Code					Big Data and Cloud Computing Laboratory					
Teaching Scheme					Examination Scheme					
L	T	P	C	Hrs/Week	Theory			Practical		Total Marks
					MS	ES	IA	LW	LE/Viva	
0	0	2	1	2	-	-	-	50	50	100

COURSE OBJECTIVES

- Recognize the key concepts of Hadoop framework, MapReduce and SPARK.
- Apply the tools, techniques and algorithms for big data analysis.
- Develop a solution for performing big data analysis for structured and unstructured data.
- To provide an understanding of the key concepts of Cloud Computing technologies, applications, implementations and challenges in the area of cloud computing.

LAB EXPERIMENTS

1.	**Understanding MapReduce:** Perform Scala programming and run a basic Word Count MapReduce program to understand MapReduce Paradigm, RDD
2.	**Implementation of SQL in SPARK** Perform SQL operations in SPARK
3.	**Implementing Graph operations in SPARK** Perform Graph operations in SPARK
4.	**Implementation of Machine Learning operations in SPARK** Perform Machine learning operations in SPARK
5.	**Implement Big Data configurations in SPARK** Perform different configurations like HADOOP, PIG, HIVE, KAFKA, MONGODB and integrate with SPARK.
6.	**Implement Structured Streaming** Perform Structured Streaming in SPARK.
7.	**Demonstrate Virtualization** Implement containerisation using Docker, employing Dockerfile and Docker Compose for application development.
8.	**Implement Container networking** Perform container networking in Docker, deploying front-end and back-end applications in separate containers and establishing communication between them.
9.	**Exploring various AWS services like EC2, Auto-Scaling, Load Balancer and RDS.** Deploy multi-stack application using AWS services.
10.	**Introduction to MongoDB in AWS** Implement DocumentDB instances on AWS.

COURSE OUTCOMES

On completion of the course, student will be able to
CO1 - Understand SPARK and Hadoop related tools for big data analytics.
CO2 - Apply techniques to conduct different types of data analysis.
CO3 - Develop a solution for a given problem using suitable Big Data Techniques.
CO4 - Apply Map Reduce paradigm for Big Data Analysis.
CO5 - Demonstrate proficiency in containerization techniques using Docker, including creating Dockerfiles and Docker Compose scripts and docker networking.
CO6 - Create big data solutions using commercial cloud platform like AWS using DocumentDB, and RDS.

TEXT/REFERENCE BOOKS

1. Bill Chambers et al., Spark: The Definitive Guide, O'Reilly Media
2. Boris lublinsky et al., Professional Hadoop Solutions, Wiley
3. Rajkumar Buyya et al., Cloud Computing: Principles and Paradigms, Wiley publication
4. Toby Velte et al., Cloud Computing: A Practical Approach, McGraw-Hill
5. K. Chandrasekaran, Essentials of Cloud Computing, CRC Press
6. Aurobindo Sarkar et al., Learning AWS, Packt Publications.

Course Code					Graphs Algorithms and Mining					
Teaching Scheme					Examination Scheme					
L	T	P	C	Hrs/Week	Theory			Practical		Total Marks
					MS	ES	IA	LW	LE/Viva	
3	0	0	3	3	25	50	25	--	--	100

COURSE OBJECTIVES

- To impart the basic concepts of Graph Neural Networks.
- To provide insights into applications of Graph Neural networks.

UNIT 1: Fundamentals Of Graph for Prediction Introduction to Graphs in Machine Learning, Basic Graph Terminology, Applications of Graphs in Machine Learning, Introduction to Machine learning and Deep learning, graph representations, Machine learning on graphs, Node classification, Relation prediction, Clustering and community detection, Graph classification, regression, and clustering, Graph Statistics and Kernel Methods, Node-level statistics and features, Neighbourhood Overlap Detection, Graph Laplacians and Spectral Methods.	**12 Hrs.**
UNIT 2: Node Embedding Introduction to Node embedding, Neighbourhood Reconstruction Methods, An Encoder-Decoder Perspective, Optimizing an Encoder-Decoder Model, Factorization-based approaches, Random walk embeddings, Limitations of Shallow Embeddings, Multi-relational Data and Knowledge Graphs, Reconstructing multi-relational data, Loss functions	**12 Hrs.**
UNIT 3: Graph Neural Networks Graph Neural Networks, Graph Neural Network Model, Neural Message Passing, Message Passing Framework. Basic GNN, Message Passing with Self-loops, Generalized Neighbourhood Aggregation, Neighbourhood Normalization Aggregators, Neighbourhood Attention, Generalized Update Methods Concatenation and Skip-Connections, Gated Updates, Knowledge Connections Edge Features and Multi-relational GNNs, Relational Graph Neural Networks Attention and Feature Concatenation, Graph Pooling, Generalized Message Passing Graph	**10 Hrs.**
UNIT 4: Graph Neural Network In Practice Applications and Loss Functions, GNNs for Node Classification, GNNs for Graph Classification, GNNs for Relation Prediction, Pre-training GNNs, Efficiency Concerns and Node Sampling, Graph-level Implementations, Subsampling and Mini-Batching, Parameter Sharing and Regularization, Introduction to GCN, Applications of GCN: Social network analysis, Web mining.	**8 Hrs.**
	Total 42 Hrs.

COURSE OUTCOMES

On completion of the course, student will be able to

CO1 - Understand Fundamental Graph Theory Concepts in the Context of Machine Learning.
CO2 - Understand basic GNN models.
CO3 - Apply Node Embedding Techniques.
CO4 - Develop Graph Neural Networks in Practical Machine Learning Problems.
CO5 - Analyze graph algorithms and mining techniques to solve real-world prediction problems.
CO6 - Analyze Graph applications, such as social network analysis and web mining.

TEXT/REFERENCE BOOKS

1. William L. Hamilton, Graph Representation Learning, Springer.
2. David Easley and Jon Kleinberg, Networks, Crowds, and Markets: Reasoning About a Highly Connected World, Cambridge University Press.
3. Albert-László Barabás, Network Science Cambridge University Press.

Course Code					Financial Forecasting					
Teaching Scheme					**Examination Scheme**					
L	**T**	**P**	**C**	**Hrs/Week**	**Theory**			**Practical**		**Total Marks**
					MS	**ES**	**IA**	**LW**	**LE/Viva**	
3	**0**	**0**	**3**	**3**	**25**	**50**	**25**	**--**	**--**	**100**

COURSE OBJECTIVES

- Understand the fundamental concepts of financial forecasting.
- Explore statistical methods commonly used in financial forecasting.
- Gain proficiency in programming using Python for implementing forecasting models.
- Develop skills to evaluate and interpret the accuracy of financial forecasts.

UNIT 1: Introduction to Financial Forecasting Definition and importance of financial forecasting, Types of financial forecasts: short-term, medium-term, and long-term, Challenges in financial forecasting, Overview of Quantitative Forecasting Techniques. Model Building and Residual Analysis Time Series Regression: Modelling Trend by Using Polynomial Functions. Detecting Autocorrelation.	**10 Hrs.**
UNIT 2: ARMA & Exponential Smoothing Autoregressive-moving average models ARMA (p,q) Decomposition Methods: Multiplicative Decomposition. Additive Decomposition. The X-12-ARIMA Seasonal Adjustment Method. Exponential Smoothing: Simple Exponential Smoothing. Tracking Signals. Holt's Trend Corrected Exponential Smoothing. Holt-Winters Methods. Damped Trends and Other Exponential	**11 Hrs.**
UNIT 3: Financial Forecasts: Volatility Modeling Filters, risk-metrics, VaR, GARCH/ARCH models, Realized volatility, Risk control strategies, Long range value forecasts/mean reversion, Trend following and momentum, High/Low range analysis (Stochastics), Advance/decline ratios.	**11 Hrs.**
UNIT 4: Financial Data Analytics Financial modelling techniques, forecasting financial statements, Use cases: Stock market predictions, transactions, dynamic financial modeling, macro-economics.	**10 Hrs.**
	Total 42 Hrs.

COURSE OUTCOMES

On completion of the course, student will be able to

CO1 - Understand the principles and significance of financial forecasting, including different types and associated challenges.
CO2 - Apply quantitative forecasting techniques like Multiple Linear Regression and Time Series Regression for financial modeling.
CO3 - Apply ARMA models and decomposition methods for accurate financial forecasts.
CO4 - Analyze Exponential Smoothing techniques for tracking signals and making predictions in financial data.
CO5 - Interpret advanced volatility modeling techniques such as GARCH/ARCH models and risk-metrics like VaR for risk assessment and control in financial forecasting.
CO6 - Generate trend and seasonality identification, correlogram analysis, and ARIMA modeling for forecasting in financial analysis.

TEXT/REFERENCE BOOKS

1. Cathy O'Neil, Rachel Schutt, Doing Data Science, Straight Talk from The Frontline. O'Reilly.
2. Introducing Data Science, Davy Cielen, Arno D. B. Meysman, Mohamed Ali, Manning Publications Co.

Course Code					ML Ops					
Teaching Scheme					**Examination Scheme**					
L	**T**	**P**	**C**	**Hrs/Week**	**Theory**			**Practical**		**Total Marks**
					MS	**ES**	**IA**	**LW**	**LE/Viva**	
3	**0**	**0**	**3**	**3**	**25**	**50**	**25**	**-**	**-**	**100**

PREREQUISITES:

- Basic knowledge of linear algebra, probability, statistics and Machine Learning.
- Familiarity with programming languages such as Python.

COURSE OBJECTIVES

- To understand the principles of MLOps.
- Learning how an ML system works in production and insights about challenges.
- Identifying systems faults and applying strategies to identify root causes in ML systems.

UNIT 1: Introduction to Mlops Overview of MLOps, Importance of MLOps in Machine Learning Projects, MLOps Features, Machine Learning Lifecycle, Deployment Pipelines, Model Versioning and Reproducibility, Key Components of MLOps: Data Management, Model Development, Deployment, Monitoring, And Maintenance.	**10 Hrs.**
UNIT 2: Model Development and Management Familiarization with Popular MLOps Tools such as Docker, Kubernetes, Git, Jenkins, and MLflow, Data Preparation, Model Training, Version and Control Systems Tracking Using Tools Like Git and MLlflow, Deploying Models Using Containerization Techniques With Docker and Kubernetes, Blue-Green, Canary And A/B Testing, Challenges In Deploying Machine Learning Models.	**10 Hrs.**
UNIT 3: Continuous Integration and Delivery (Ci/Cd) For ML Understanding the Principles of CI/CD in the Context of Machine Learning, Designing and Automating End-To-End ML Pipelines Using Tools Like Apache Airflow or Kubeflow Pipelines, Automated Testing Strategies to Ensure the Reliability and Robustness of ML Models Throughout the Development Lifecycle, Integrating Model Deployment Into CI/CD Pipelines for Seamless Delivery to Production Environments.	**12 Hrs.**
UNIT 4: Monitoring, Governance, and Maintenance Monitoring Model Performance in Production, Setting up Alerts and Triggers for Model Issues, Feedback Loops for Model Retraining and Updates, Data Security and Privacy In ML Projects, Compliance with Regulations (GDPR, HIPAA, Etc.), Secure Model Deployment and Access Control, Cost Optimization and Resource Allocation, Strategies for Model Retraining, Updating and Version Management to Ensure Continued Performance and Relevance. Case Studies and Best Practices/MLOps: Analyzing Case Studies and Best Practices for Effective Monitoring, Governance, and Maintenance of ML Systems.	**10 Hrs.**
	Total 42 Hrs.

COURSE OUTCOMES

On completion of this course, students will be able to:
CO1 - Understand the fundamentals of MLOps and its importance in machine learning deployment.
CO2 - Describe the core concepts of Docker, including containers, images, Dockerfiles, and Docker Compose.
CO3 - Illustrate how to containerize machine learning applications using Docker.
CO4 - Explore Data Management in MLOps.
CO5 - Apply Docker-based MLOps techniques to real-world machine learning projects.
CO6 - Deploy ML models in production environments.

TEXT/REFERENCE BOOKS

1. Mark Treveil et al., “Introducing MLOps”, O'Reilly home.
2. Noah Gift, Alfredo Deza, “Practical MLOps”, O'Reilly home
3. Andrew P. McMahon and Adi Polak, “Machine Learning Engineering with Python: Manage the lifecycle of machine learning models using MLOps with practical examples”, Packt
4. Mikael Krief, “Learning DevOps: The complete guide to accelerate collaboration with Jenkins, Kubernetes, Terraform and Azure DevOps”, Packt

Course Code					Computer Vision and Image Processing					
Teaching Scheme					Examination Scheme					
L	T	P	C	Hrs/Week	Theory			Practical		Total Marks
					MS	ES	IA	LW	LE/Viva	
3	0	0	3	3	25	50	25	--	--	100

PREREQUISITE: Linear Algebra

COURSE OBJECTIVES

- To gain a comprehensive understanding of fundamental concepts in computer vision, including image processing, feature extraction, object detection, tracking, and image segmentation.
- To explore the integration of computer vision techniques with machine learning algorithms, including supervised and unsupervised learning methods, to enhance the capability and adaptability of vision systems.

UNIT 1: Digital Image Formation and Low-Level Processing Overview and State-of-the-art, Fundamentals of Image Formation, Pinhole Camera, Camera Parameters, Geometric Camera Calibration, Transformation: Orthogonal, Euclidean, Affine, Projective, etc; Fourier Transform, Convolution and Filtering, Image Enhancement, Restoration, Histogram Processing.	**6 Hrs.**
UNIT 2: Feature Extraction and Image Segmentation Edges - Canny, LOG, DOG; Line detectors (Hough Transform), Corners - Harris and Hessian Affine, Orientation Histogram, SIFT, SURF, HOG, GLOH, Scale-Space Analysis- Image Pyramids and Gaussian derivative filters, Gabor Filters and DWT, Background Subtraction and Modelling, Image Segmentation: Region Growing, Edge Based approaches to segmentation, Graph-Cut, Mean-Shift, K-Means, K-Medoids.	**10 Hrs.**
UNIT 3: Depth and Motion Estimation Depth Estimation and Multi-view Geometry: Perspective, Binocular Stereopsis, Camera and Epipolar Geometry; Homography, Rectification, DLT, RANSAC, 3-D Reconstruction Framework, Camera Calibration, Optical Flow, KLT, Spatio-Temporal Analysis, Dynamic Stereo, Motion parameter estimation, Shape from X, Light at Surfaces, Phong Model, Reflectance Map, Albedo estimation, Photometric Stereo, Use of Surface Smoothness Constraint, Shape from Texture.	**12 Hrs.**
UNIT 4: Computer Vision Using Deep Learning Convolutional Neural Network (CNN): Evolution of CNN Architectures, Convolution, Pooling, ReLU, Transfer Learning, CNNs for Recognition and Verification (Siamese Networks, Triplet Loss, Contrastive Loss, Ranking Loss); CNNs for Detection: Background of Object Detection, R-CNN, Fast R-CNN, Faster R-CNN, YOLO, SSD, RetinaNet; CNNs for Segmentation: FCN, SegNet, U-Net, Mask-RCNN, Deep Generative Models: Review of Deep Generative Models: GANs, VAEs, Introduction to Vision Transformers.	**14 Hrs.**
	Total 42 Hrs.

COURSE OUTCOMES

On completion of the course, student will be able to

CO1 - Recall basic concepts, terminology, theories, models, and methods within the domain of computer vision.

CO2 - Understand the object detection algorithm operations, including their underlying architectures, such as the single-shot nature of YOLO, the hierarchical feature extraction in CNNs, and the pixel-wise segmentation approach of FCN, SegNet, U-Net, and Mask-RCNN.

CO3 - Explain the fundamental methods of image processing including image filtering, restoration, edge detection, reconstruction, segmentation, classification, and representation.

CO4 - Compare different CNN architectures like R-CNN, Fast R-CNN, Faster R-CNN.

CO5 - Apply appropriate methods of computer vision to study stereo images, motion sequences, and recognize objects within images.

CO6 - Analyse the theoretical concepts and mathematical models behind depth and motion estimation algorithms, including disparity mapping, optical flow, and structure from motion.

TEXT/REFERENCE BOOKS

1. Richard Szeliski, Computer Vision: Algorithms and Applications, Springer-Verlag London Limited.
2. D. A. Forsyth, J. Ponce, Computer Vision: A Modern Approach, Pearson Education .
3. Richard Hartley, Andrew Zisserman, Multiple View Geometry in Computer Vision, e2, Cambridge University Press.
4. Rafael C. Gonza Lez, Richard E. Woods, Digital Image Processing, Pearson Education.

<table>
<tr><th colspan="5">Course Code</th><th colspan="6">Computer Vision and Image Processing Laboratory</th></tr>
<tr><th colspan="5">Teaching Scheme</th><th colspan="6">Examination Scheme</th></tr>
<tr><th rowspan="2">L</th><th rowspan="2">T</th><th rowspan="2">P</th><th rowspan="2">C</th><th rowspan="2">Hrs/Week</th><th colspan="3">Theory</th><th colspan="2">Practical</th><th rowspan="2">Total Marks</th></tr>
<tr><th>MS</th><th>ES</th><th>IA</th><th>LW</th><th>LE/Viva</th></tr>
<tr><td>0</td><td>0</td><td>2</td><td>1</td><td>2</td><td>-</td><td>-</td><td>-</td><td>50</td><td>50</td><td>100</td></tr>
</table>

COURSE OBJECTIVES

- Insight into image and video formation design, modelling and analysis.
- Ability to work with features above the pixel level.
- Develop ability to understand the difference in theory and practice of Computer Vision.

PREREQUISITE: Python Programming

LAB EXPERIMENTS

1.	**Basic Image and Video Processing:** Reading and writing Images, Image enhancement in spatial and frequency domain, Morphological operations.
2.	**Image and Video Segmentation:** Implementation of image segmentation algorithms such as, region growing, k-means, image thresholding, OTSU, Hough Transform.
3.	**Depth estimation:** Depth estimation using stereo images, depth and motion estimation using optical flow, structure from motion.
4.	**Object Detection and Tracking:** Object detection using HOG, Multiple objects tracking using multiple cameras.
5.	**Recognition:** Action recognition in videos, Face recognition for human identification, Human activity recognition.
6.	**Deep Learning based Computer Vision Applications:** Image segmentation using Mask-RCNN, Object detection using YOLO, Image classification and recognition, Image synthesis using generative adversarial networks.

COURSE OUTCOMES

On completion of the course, student will be able to

CO1- Implement fundamental operations on images and videos using MATLAB/OPENCV.

CO2 - Apply the knowledge of these object detection algorithms to solve specific tasks or problems in computer vision, such as object localization, classification, semantic segmentation, or instance segmentation.

CO3 - Implement generative adversarial networks as a pre-processing method for computer vision applications.

CO4 - Apply computer vision fundamentals for object tracking and human activity representation.

CO5 - Apply depth and motion estimation methods to analyze and interpret visual data in various applications such as 3D reconstruction, scene understanding, object tracking, and gesture recognition.

CO6 - Develop vision models utilizing deep neural networks by integrating various concepts and techniques learned throughout the course.

TEXT/REFERENCE BOOKS

1. Forsyth and Ponce, Computer Vision: A Modern Approach, Pearson Education.
2. Simon Prince, Computer Vision: Models, Learning, and Interface, Cambridge University Press.
3. Rajalingappaa Shanmugamani, Deep learning for Computer Vision, PACKT publishing.
4. Rosebrock Adrian, Deep Learning for Computer Vision with Python, PACKT Publishing.

<table>
<tr><th colspan="5">Course Code</th><th colspan="6">Speech Processing Applications in Data Science</th></tr>
<tr><th colspan="5">Teaching Scheme</th><th colspan="6">Examination Scheme</th></tr>
<tr><th rowspan="2">L</th><th rowspan="2">T</th><th rowspan="2">P</th><th rowspan="2">C</th><th rowspan="2">Hrs/Week</th><th colspan="3">Theory</th><th colspan="2">Practical</th><th rowspan="2">Total Marks</th></tr>
<tr><th>MS</th><th>ES</th><th>IA</th><th>LW</th><th>LE/Viva</th></tr>
<tr><td>3</td><td>0</td><td>0</td><td>3</td><td>3</td><td>25</td><td>50</td><td>25</td><td>--</td><td>--</td><td>100</td></tr>
</table>

PREREQUISITE: Linear Algebra, Machine Learning

COURSE OBJECTIVES

- Describe the physiological and acoustic aspects of speech production.
- Implement various feature extraction techniques and understand their role in speech processing.
- Implement the principles of speech coding and compression.
- Apply language modelling techniques for improving speech recognition accuracy.

UNIT 1: Introduction to Digital Speech Processing Overview of Speech Processing, Speech Production and Acoustics, Review of Digital Signal Processing Fundamentals, Digitization and Recording of speech signal, Speech Signal Sampling and Quantization, Short-Time Fourier Transform (STFT), Analysis: Fourier Transform view and Filtering view, Synthesis: Filter bank summation (FBS) Method.	**10 Hrs.**
UNIT 2: Human Speech Production and Source Filter Model Human Speech production, Acoustic Phonetics and Articulatory Phonetics, Different categories speech sounds and Location of sounds in the acoustic waveform and spectrograms, Uniform Tube Modeling of Speech Production, Speech Perception, Pitch and Formant Analysis, Mel-Frequency Cepstral Coefficients (MFCCs), Speech Feature Extraction Techniques, Segmental and Supra Segmental Features of Speech signal, Speech Prosody, Speech Prosody Modeling (Fujisaki Model).	**12 Hrs.**
UNIT 3: Speech Enhancement and Coding Noise Types and Models, Speech Enhancement Techniques (Wiener Filtering, Spectral Subtraction), Time Domain Methods in Speech Processing, Analysis and Synthesis of Pole-Zero Speech Models, Pulse Code Modulation (PCM) and Quantization, Linear Predictive Coding (LPC).	**10 Hrs.**
UNIT 4: Speech Recognition and Applications Introduction to Speech Recognition, Hidden Markov Models (HMMs), Acoustic Modeling in Speech Recognition, Language Modeling for Speech Recognition, Decoding Techniques (Viterbi Algorithm), Text-to-Speech (TTS) Systems, Speaker Identification, Principles of voice biometrics and speaker recognition, Feature extraction methods for speaker identification, Voice-based authentication and security applications, Speech analytics for extracting insights from spoken language data, Multimodal data fusion techniques integrating speech with other modalities, Natural language understanding (NLU) for processing spoken commands and queries.	**10 Hrs.**
	Total 42 Hrs.

COURSE OUTCOMES

On completion of the course, student will be able to

CO1 - Understand fundamental concepts of speech production, acoustics, and digital signal processing to represent speech signals.
CO2 - Distinguish feature extraction techniques and understand their role in speech processing applications.
CO3 - Illustrate fundamental comprehension of language models for implementing a speech recognition system.
CO4 - Implement the principles of speech coding techniques for compression.
CO5 - Demonstrate speech synthesis techniques for text-to-speech systems.
CO6 - Apply speech processing technologies including hidden Markov models (HMMs) and related systems such as text-to-speech (TTS), speaker identification, and speaker verification.

TEXT/REFERENCE BOOKS

1. Lawrence R. Rabiner, Ronald W. Schafer, An Introduction to Digital Speech Processing: Foundations and Trends® in Signal Processing, Now publishers Inc
2. A. Nejat Ince, Digital Speech Processing: Speech Coding, Synthesis and Recognition, The Springer International Series in Engineering and Computer Science.
3. Branko Kovacevic, Milan M. Milosavljevic, Mladen Veinović, Milan Marković, Robust Digital Processing of Speech Signals, Springer International Publishing AG.

<table>
<tr><th colspan="5">Course Code</th><th colspan="6">Language Processing and Computational Linguistics</th></tr>
<tr><th colspan="5">Teaching Scheme</th><th colspan="6">Examination Scheme</th></tr>
<tr><th rowspan="2">L</th><th rowspan="2">T</th><th rowspan="2">P</th><th rowspan="2">C</th><th rowspan="2">Hrs/Week</th><th colspan="3">Theory</th><th colspan="2">Practical</th><th rowspan="2">Total Marks</th></tr>
<tr><th>MS</th><th>ES</th><th>IA</th><th>LW</th><th>LE/Viva</th></tr>
<tr><td>3</td><td>0</td><td>0</td><td>3</td><td>3</td><td>25</td><td>50</td><td>25</td><td>--</td><td>--</td><td>100</td></tr>
</table>

PREREQUISITE: Probability and Statistics, Machine Learning

COURSE OBJECTIVES

- To understand the structure and basic operations of Natural Language Processing
- To understand the syntax of Natural languages for grouping local words for parsing
- To understand the concepts of linguistic rules and machine learning approaches for classification
- Apply language modelling techniques for improving application of NLP.

UNIT 1: Language Processing & Linguistic Fundamentals Basic concepts in linguistics relevant to NLP and CL, Multilingual NLP, Syntax, semantics, morphology, and phonology. Text Processing Techniques POS Tagging, Named Entity Recognition, Word Sense Disambiguation, Wordnet and Lexical Resources, Syntax – Constituency Parsing, Dependency Parsing, Distributional Semantics, Lexical Semantics,	**10 Hrs.**
UNIT 2: Statistical Methods in NLP Probability theory and its applications in NLP, N-grams and Language Models, Hidden Markov Models (HMMs) in Language Processing, Advanced smoothing for language modelling, Minimum Edit Distance, Models for Sequential Tagging – MaxEnt, CRF	**10 Hrs.**
UNIT 3: Deep Learning for NLP Representation Discovery: Word vectors, Word2Vec & GloVe, RNNs for Variable Length Sequences, Attention & Transformer, Neural & Pre-Trained Language Models, Advanced Pre-training for Language Models, GPT3 & Beyond: Few-Shot Learning, Prompt Learning.	**12 Hrs.**
UNIT 4: NLP Applications Machine Translation Systems, Application in Sarcasm, Text Summarization and Generation, Information Extraction and Text Mining, Question Answering Systems, Spelling Correction Ethical and Societal Implications of NLP - Bias and fairness in NLP systems	**10 Hrs.**
	Total 42 Hrs.

COURSE OUTCOMES

On completion of the course, student will be able to
CO1 - Understand the basic building blocks of NLP for text processing.
CO2 - Determine the Syntactical Structures of Natural Languages.
CO3 - Evaluate the Language Model for Natural Languages.
CO4 - Apply Information Retrieval Techniques to build search engines.
CO5 - Evaluate the performance of machine learning models for textual data.
CO6 - Demonstrate and present solution for a given problem using text analytics approach for Text Summarization, Sentiment Analysis.

TEXT/REFERENCE BOOKS

1. Yoav Goldberg Neural Network Methods for Natural Language Processing, Springer.
2. Dan Jurafsky and James Martin Speech and Language Processing, Pearson.
3. Christopher D. Manning and Hinrich Schütze, Foundations of Statistical Natural Language Processing, MIT Press
4. James Allen, Natural Language Understanding, Benjamin/Cummings.
5. Akshar Bharati, Vineet Chaitanya and Rajeev Sangal, Natural Language Processing: A Paninian Perspective, PHI.

<table>
<tr><th colspan="5">Course Code</th><th colspan="6">Data Science for Health Informatics</th></tr>
<tr><th colspan="5">Teaching Scheme</th><th colspan="6">Examination Scheme</th></tr>
<tr><th rowspan="2">L</th><th rowspan="2">T</th><th rowspan="2">P</th><th rowspan="2">C</th><th rowspan="2">Hrs/Week</th><th colspan="3">Theory</th><th colspan="2">Practical</th><th rowspan="2">Total Marks</th></tr>
<tr><th>MS</th><th>ES</th><th>IA</th><th>LW</th><th>LE/Viva</th></tr>
<tr><td>3</td><td>0</td><td>0</td><td>3</td><td>3</td><td>25</td><td>50</td><td>25</td><td>-</td><td>-</td><td>100</td></tr>
</table>

PREREQUISITES: Basic understanding of data analytics

COURSE OBJECTIVES

- To introduce the basic concepts and types of health components.
- To develop skills of using AI techniques for diagnosis and detection of health applications.
- To gain experience of developing and deploying health applications for possible prediction.

UNIT 1: Introduction To Health Informatics Introduction to health informatics and its significance, Definitions and key concepts in health informatics, Background disciplines, historical overview, and future challenges, knowledge hierarchy: Data, information, and knowledge, Types of healthcare data (e.g., structured, unstructured, clinical, administrative), Health information exchange (HIE) and interoperability standards.	**10 Hrs.**
UNIT 2: Health Data Management Maintain patient records, common issues related to healthcare data quality, biomedical research and publicly available health data resources, Electronic Health Records (EHR) and telemedicine, Privacy, security, and confidentiality. Health data standards (e.g., HL7, DICOM), Regulatory frameworks (e.g., HIPAA, GDPR) and their implications for health informatics, Ethical considerations in health data management and research.	**10 Hrs.**
UNIT 3: Healthcare Data Analysis Healthcare data Analysis: Data quality, integrity, and governance in healthcare, Health data acquisition, storage, and retrieval, Sources of the healthcare data, Pre-processing of the healthcare data, Handling of the healthcare data, Descriptive, predictive, and prescriptive analytics in healthcare, Creation of analysis-ready datasets, Healthcare datasets – Examples and Case studies.	**10 Hrs.**
UNIT 4: AI Transform Healthcare Opportunities and challenges in AI-driven healthcare, medical imaging modalities, Diagnosis and Detection using health data, Predictive modeling techniques such as KNN, SVM, Naïve Bayes, Decision Tree, and Ensemble techniques for health data prediction, and Deep learning applications in healthcare (e.g., pathology diagnosis, drug discovery).	**12 Hrs.**
	Total 42 Hrs.

COURSE OUTCOMES

On completion of the course, student will be able to

CO1 - Describe key concepts, terminology, and historical context of Health Informatics.
CO2 - Understand fundamental characteristics of data, information, and knowledge in the Health Informatics domain.
CO3 - Solve healthcare-related problems and improve processes using informatics principles.
CO4 - Apply the use of AI techniques for early prediction of diseases in health research data.
CO5 - Develop machine learning framework for diagnosis and detection of Health research data.
CO6 - Deployment of AI-enabled health applications for possible prediction and recommendations.

TEXT/REFERENCE BOOKS

1. Parag Suresh Mahajan, Artificial Intelligence in Healthcare, Academic Edition.
2. Kerrie L. Holley, Siupo Becker, AI-First Healthcare, O'Reilly Media, Inc.
3. Robert Shimonski, AI in Healthcare: How Artificial Intelligence Is Changing IT Operations and Infrastructure Services, Wiley.