



RAJALAKSHMI
ENGINEERING COLLEGE
An AUTONOMOUS Institution
Affiliated to ANNA UNIVERSITY, Chennai

REGULATION 2024

CURRICULUM AND SYLLABUS

M.Tech Data Science
(Based on Industry Partnership with L & T)

Department of Information Technology
Rajalakshmi Engineering College, Thandalam.

RAJALAKSHMI ENGINEERING COLLEGE
(An Autonomous Institution, Affiliated to Anna University, Chennai)
Choice Based Credit System (CBCS)

REGULATIONS 2024
CURRICULUM AND SYLLABUS

DEPARTMENT OF INFORMATION TECHNOLOGY

M.Tech Data Science

Vision

- To be a Department of Excellence in Information Technology Education, Research and Development.

Mission

- To train the students to become highly knowledgeable in the field of Information Technology.
- To promote continuous learning and research in core and emerging areas.
- To develop globally competent students with strong foundations, who will be able to adapt to changing technologies.

PROGRAMME OUTCOMES (POs)

PO1: Graduates should be able to learn how to interpret data, extracts meaningful information, and assesses findings.

PO2: Graduates should capable of demonstrating and developing a design of mastery over the key technologies in data science and Business Analytics such as structured/unstructured data mining, machine learning, visualization techniques, predictive modeling and statistics.

PO3: Graduates should be capable of applying ethical principles and responsibilities during Professional practice.

PO4: Graduates should be able to function effectively as a team member and to write/ present a substantial technical report / document.

PO5: Graduates should independently carry out research / investigation and development work to solve industry and organization-specific problems and challenges using advanced analytics and computational methods.

PO6: Graduates should be able to engage in independent and life-long learning in the broadest context of technological change.

Regulation 2024

SEMESTER I

S. Nos.	COURSE CODE	COURSE TITLE	Category	Contact Periods	L	T	P	C
THEORY COURSES								
1.	MH24114	Mathematical Foundations for Data Science	FC	4	2	2	0	4
2.	DS24111	Data Science with Python	PC	3	2	1	0	3
3.	DS24112	Machine Learning Practices	PC	3	2	1	0	3
4.	DS24113	Data Engineering	PC	3	2	1	0	3
5.		Professional Elective - I	PE	3	2	1	0	3
LABORATORY COURSES								
6.	GE24121	Professional Soft Skills - I	EEC	2	0	0	2	1
7.	DS24121	Applied Machine Learning Lab	PC	4	0	0	4	2
8.	DS24122	Data Science Lab	PC	4	0	0	4	2
NON-CREDIT COURSES								
9.	AC23111	English for Research Paper Writing	MC	3	3	0	0	0
Total				29	13	6	10	21

SEMESTER II

S. Nos.	COURSE CODE	COURSE TITLE	Category	Contact Periods	L	T	P	C
THEORY COURSES								
1.	DS24211	Deep Learning Applications	PC	3	2	1	0	3
2.	DS24212	Generative AI with Large Language Models	PC	3	2	1	0	3
3.		Professional Elective - II	PE	3	2	1	0	3
4.		Open Elective	OE	3	2	1	0	3
5.		Professional Elective - III	PE	3	2	1	0	3
LABORATORY COURSES								
6.	GE24221	Professional Soft Skills - II	EEC	2	0	0	2	1
7.	DS24221	Generative AI Applications	PC	4	0	0	4	2
8.	DS24222	Large Language Models Lab	PC	4	0	0	4	2
NON-CREDIT COURSES								
9.	AC23211	Constitution of India	MC	3	3	0	0	0
Total				28	13	5	10	20

SEMESTER III

S. Nos.	COURSE CODE	COURSE TITLE	Category	Contact Periods	L	T	P	C
THEORY COURSES								
1.		Professional Elective - IV	PE	3	2	1	0	3
LABORATORY COURSES								
2.	DS24321	Project / Thesis -I	EEC	24	0	0	24	12
Total				27	2	1	24	15

SEMESTER IV

S. Nos.	COURSE CODE	COURSE TITLE	Category	Contact Periods	L	T	P	C
1.	DS24421	Project /Thesis -II	EEC	32	0	0	32	16
Total				32	0	0	32	16

PROGRAM ELECTIVES (PE)								
S.Nos	Course Code	Course Title	Category	Contact Periods	L	T	P	C
Professional Elective I								
1.	DS24A11	Generative Adversarial Networks	PE	3	2	1	0	3
Professional Elective II								
2.	DS24B11	Ethics in Data Science	PE	3	2	1	0	3
3.	DS24B12	Computer Vision	PE	3	2	1	0	3
4.	DS24B13	Natural Language Processing	PE	3	2	1	0	3
Professional Elective III								
5.	DS24B14	Application Architecture and Deployment	PE	3	2	1	0	3
6.	DS24B15	Security for Data Engineering	PE	3	3	0	0	3
Professional Elective IV								
7.	DS24C11	Machine Learning Engineering for Production Specialization	PE	3	2	1	0	3
8.	DS24C12	Industry Specific Applications of Generative AI & Responsible AI	PE	3	2	1	0	3

TOTAL CREDITS : 72

Credit Distribution

Category	R2024
Mathematical courses FC	4
Professional core courses PC	23
Professional Elective Courses PE	12
Open Electives from other technical and /or emerging subjects OE	3
Project work, seminar and internship in industry or elsewhere EEC	30
Mandatory Courses [Environmental Sciences, Induction Program, Indian Constitution, The essence of Indian Knowledge Tradition] MC	0
Total	72

SUMMARY OF ALL COURSES

S.NO	Course Category	Credits per Semester				Total Credits
		1	2	3	4	
1.	FC	4				4
2.	PC	13	10			23
3.	PE	3	6	3		12
4.	OE		3			3
5.	EEC	1	1	12	16	30
6.	MC	0	0			0
	Total	21	20	15	16	72

Credit Distribution

Category	R2023	R2024
Mathematical courses FC / Basic Science BS	4	4
Professional core courses PC	30	23
Professional Elective Courses PE	12	12
Open Electives from other technical and /or emerging subjects OE	3	3
Project work, seminar and internship in industry or elsewhere EEC	18	30
Mandatory Courses [Environmental Sciences, Induction Program, Indian Constitution, The essence of Indian Knowledge Tradition] MC	3	0
Total	70	72

Subject Code	Subject Name (Theory Course)	Category	L	T	P	C
MH24114	Mathematical Foundations for Data Science	FC	2	2	0	4

Objectives:
<ul style="list-style-type: none"> Understand the basics of python and standard modules used for data science with hands-on.
<ul style="list-style-type: none"> Understand the data structures and visualization used for data science with hands-on.
<ul style="list-style-type: none"> Understand the machine learning libraries used for data science with hands-on.
<ul style="list-style-type: none"> Gain an understanding of essential machine learning algorithms and techniques using the Scikit-Learn library, and apply them to real-world data science problems
<ul style="list-style-type: none"> Apply Python programming and data science techniques to practical problems and projects, including data preprocessing, model building, evaluation, and interpretation of results.

UNIT-I	Linear Algebra	12
Systems of Linear Equations - Machine learning motivation - A geometric notion of singularity - Singular vs non-singular matrices - Linear dependence and independence - Matrix row-reduction - Row operations that preserve singularity - The rank of a matrix - Row echelon form - Reduced row echelon form- LU decomposition- Solving Systems of Linear Equations - Machine learning motivation - Solving non- singular systems of linear equations - Solving singular systems of linear equations - Solving systems of equations with more variables - Gaussian elimination.		
UNIT-II	Probability & Statistics	12
Introduction to probability - Concept of probability: repeated random trials - Conditional probability and independence - Random variables - Cumulative distribution function - Discrete random variables: Binomial distribution - Probability mass function - Continuous random variables: Uniform distribution - Continuous random variables: Gaussian distribution -Joint distributions - Marginal and conditional distributions - Independence - covariance - Multivariate normal distribution - Sampling and point estimates - Interval estimation -Confidence intervals – Confidence Interval for mean of population - Biased vs Unbiased estimates-Maximum likelihood estimation - Intuition behind maximum likelihood estimation - Hypothesis testing - Describing samples: sample proportion and sample mean - Two types of errors - Test for proportion and means - Two sample inference for difference between groups.		
UNIT-III	Bayesian Statistics & its applications in various fields	12
Bayesian statistics and its applications in various fields - Bayesian Learning: Bayes theorem - maximum likelihood and least squared error hypotheses – Naïve Bayes classifier- Bayesian belief networks- gradient ascent training of Bayesian networks- learning the structure of Bayesian networks- the EM algorithm- mixture of models- Markov models- hidden Markov models - Time series analysis and forecasting techniques - Basic Properties of time-series data: Distribution and moments- Stationarity- Autocorrelation- Heteroscedasticity- Normality- Survival Analysis.		
UNIT-IV	Non-Parametric Statistics	12
Non-parametric Statistics - Chi square test- Sign test -Wilcoxon signed rank test - Mann Whitney test - Run test - Kolmogorov Smirnov test - Spearman and Kendall’s test - Tolerance region.		
UNIT-V	Multivariate Statistical Methods for Analyzing Complex Datasets	12
Multivariate statistical methods for analysing complex datasets - Factor Analysis - Cluster Analysis- Regression Analysis - Discriminant Analysis.		
Total Contact Hours: 60		

Course Outcomes:
Students will be able to refresh the statistical knowledge learnt earlier with hands-on practical expertise.
<ul style="list-style-type: none"> Apply linear algebra techniques to model and solve complex problems in diverse domains.
<ul style="list-style-type: none"> Demonstrate proficiency with statistical analysis of data and to apply data science concepts and methods to solve problems in real-world contexts.
<ul style="list-style-type: none"> Model Bayesian forecasting techniques to time series data and other predictive modeling tasks.
<ul style="list-style-type: none"> Understands and apply various non-parametric statistics techniques
<ul style="list-style-type: none"> Formulate, test and interpret various nonparametric tests for solving various statistical problems.

Reference Books(s) / Web links:	
1.	James D. Miller, Statistics for Data Science, Packt Publishing, 2017
2.	IND James D. Hamilton, Time Series Analysis, Levant Books, 2012
3.	Bayesian statistics and its applications in various fields: INDIAN INSTITUTE OF TECHNOLOGY ROORKEE/Mehta Family School of Data Science and Artificial Intelligence/Course Title: Machine Learning.
4.	INDIAN INSTITUTE OF TECHNOLOGY ROORKEE/Mehta Family School of Data Science and Artificial Intelligence/Course Title: Time Series Data Analysis.
5.	M.R Anderberg, M.R. Anderberg, "Cluster Analysis for Applications", Academic Press

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
MH24114.1	3	-	-	-	-	-
MH24114.2	3	3	-	2	3	2
MH24114.3	3	3	2	2	3	2
MH24114.4	3	3	-	-	-	-
MH24114.5	3	3	2	2	3	2
Average	3	3	2	2	3	2

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium)

3: Substantial (High) No correlation : "-"

Subject Code	Subject Name (Theory Course)	Category	L	T	P	C
DS24111	Data Science with Python	PC	2	1	0	3

Objectives:
<ul style="list-style-type: none"> Understand the basics of python and standard modules used for data science with hands-on.
<ul style="list-style-type: none"> Understand the data structures and visualization used for data science with hands-on.
<ul style="list-style-type: none"> Understand the machine learning libraries used for data science with hands-on.
<ul style="list-style-type: none"> Gain an understanding of essential machine learning algorithms and techniques using the Scikit-Learn library, and apply them to real-world data science problems
<ul style="list-style-type: none"> Apply Python programming and data science techniques to practical problems and projects, including data preprocessing, model building, evaluation, and interpretation of results.

UNIT-I	Python - Data Structures, OOPS & Modules	9
<p>Data structures: Dictionaries - Maps - Hash Tables - Array Data Structures - Records - Structs - Data Transfer Objects - Sets and Multisets-Stacks (LIFOs) - Queues (FIFOs) ; Python : Python installation - Python OOPs - Polymorphism in OOPs programming - Python String Concatenation - Print Exception in Python - Python Libraries - Python Pandas - Python Matplotlib - Python Seaborn - Python SciPy - Chatbot in Python - Machine Learning using Python - Exploratory Data Analysis in Python - Open CV Python - Tkinter - Pythons Turtle Module - PyGame in Python - Pytorch - Scrapy - Web Scraping - Django - Python Programs - Types of Data structure in Python - Built in data structures - User defined data structures; Object Oriented Concepts and Design : APIs and Data Collection - Simple API - REST APIs & HTTP Requests - Web scraping - HTML for Web Scraping - file formats</p>		
UNIT-II	Python – Numpy, Pandas & DS Libraries	9
<p>Installation and setup : Anaconda Distribution - Anaconda Navigator to create a New Environment - Startup and Shutdown Process - Intro to the Jupyter Lab Interface - Code Cell - execution; Python : Basic datatypes - Operators - variables - Built in Functions - Custom Functions - String Methods - Lists - Index Positions and Slicing - Navigating Libraries using Jupyter Lab; Series : Create series object from a list and dictionary - The head and Tail methods - Passing Series to Python Built-In Functions – Methods for Data sorting ; Dataframe : Methods and Attributes between Series and DataFrames - Fill in Missing Values - Filtering data and methods in Dataframe - Data Extraction in dataframes - Working with Text Data - Merging Dataframes; Data Mining - Data Processing and Modelling - Data Visualization</p>		
UNIT-III	Visualization	9
<p>Introduction to Matplotlib - Matplotlib Basics - Matplotlib - Understanding the Figure Object - Matplotlib - Implementing Figures and Axes - Matplotlib - Figure Parameters - Matplotlib Styling - Legends - Matplotlib Styling - Colors and Styles - Advanced Matplotlib Commands - Introduction to Seaborn - Scatterplots with Seaborn – Distribution Plots - Part One - Understanding Plot Types - Distribution Plots - Part Two - Coding with Seaborn - Categorical Plots - Statistics within Categories - Understanding Plot Types - Categorical Plots - Statistics within Categories - Coding with Seaborn - Categorical Plots - Distributions within Categories - Understanding Plot Types - Categorical Plots - Distributions within Categories - Coding with Seaborn - Seaborn - Comparison Plots - Understanding the Plot Types - Seaborn - Comparison Plots - Coding with Seaborn - Seaborn Grid Plots - Seaborn - Matrix Plots.</p>		
UNIT-IV	Regression and Classification	9
<p>Introduction to Linear Regression : Cost Functions - Gradient Descent - Python coding Simple - Overview of Scikit-Learn and Python - Residual Plots - Model Deployment and Coefficient Interpretation - Polynomial Regression - Theory and Motivation - Creating Polynomial Features - Training and Evaluation - Bias Variance Trade-Off - Polynomial Regression Choosing Degree of Polynomial - Model Deployment - Feature Scaling; Introduction to Cross Validation : Regularization</p>		

Data Setup - Ridge Regression Theory - Lasso Regression - Background and Implementation - Elastic Net Feature Engineering and Data Preparation; Dealing with Outliers - Dealing with Missing Data - Evaluation of Missing Data - Filling or Dropping data based on Rows - Fixing data based on Columns - Dealing with Categorical Data - Encoding Options - Cross Validation - Test - Validation - Train Split - cross_val_score - cross validate - Grid Search; Linear Regression Project: The Logistic Function - Logistic Regression - Theory and Intuition; Linear to Logistic: Logistic Regression - Theory and Intuition - Linear to Logistic Math; Logistic Regression: Theory and Intuition Logistic Regression Model Training - Classification Metrics - Confusion Matrix and Accuracy - Classification Metrics - Precision, Recall, F1- Score - ROC Curves - Logistic Regression with Scikit-Learn - Performance Evaluation - Multi-Class Classification with Logistic Regression - Data and EDA – Model.

UNIT-V	Unsupervised and Advanced Classification Techniques	9
<p>Introduction to KNN Section: KNN Classification, KNN Coding with Python - Choosing K, KNN Classification Project Exercise; Introduction & history of Support Vector Machines- Hyperplanes and Margins, Kernel Intuition, Kernel Trick and Mathematics; SVM with Scikit-Learn and Python – Classification, Regression Tasks; Introduction to Tree Based Methods- Decision Tree, Understanding Gini Impurity; Constructing Decision Trees with Gini Impurity, Coding Decision Trees; Introduction to Random Forests-Key Hyperparameters, Number of Estimators and Features in Subsets, Bootstrapping and Out-of-Bag Error; Coding Classification with Random Forest Classifier, Coding Regression with Random Forest Regressor, Advanced Models. Introduction to K-Means Clustering Section; K-Means Color Quantization; K-Means Clustering Exercise Overview, Solution ; Introduction to Hierarchical Clustering, Coding - Data and Visualization, Scikit-Learn; Introduction to Principal Component Analysis(PCA)-Manual Implementation in Python-SciKit-Learn.</p>		
Total Contact Hours: 45		

Course Outcomes:
<ul style="list-style-type: none"> • Understand the basics of python and standard modules used for data science with hands-on.
<ul style="list-style-type: none"> • Understand the data structures and visualization used for data science with hands-on.
<ul style="list-style-type: none"> • Understand the machine learning libraries used for data science with hands-on
<ul style="list-style-type: none"> • Understand the Regression and Classification for data science with hands-on
<ul style="list-style-type: none"> • Design, develop and test various classification and clustering models

<p>SUGGESTED ACTIVITIES</p> <ul style="list-style-type: none"> • Problem solving sessions • Activity Based Learning • Implementation of small module
--

SUGGESTED EVALUATION METHODS

- Assignment problems
- Quizzes
- Class Presentation/Discussion

Text Book(s):

1. “Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython” by Wes McKinney, O'Reilly Media , 2nd Edition, 2017
2. “Data Science from Scratch First Principles with Python“ by Joel Grus, O'Reilly Media, Inc., Second Edition, Released May 2019.

Reference Books(s) / Web links:

1. Alvaro Fuentes, Become a Python Data Analyst – By Packt Publishing (2018)
2. Bharti Motwani, Data Analytics using Python – By Wiley (2020)
3. Jules S. Damji, Learning Spark: Lightning-Fast Data Analytics, Second Edition – By Shroff/O'Reilly (2020)

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
DS24111.1	3	3	-	2	-	2
DS24111.2	3	3	2	2	-	-
DS24111.3	3	3	-	2	2	2
DS24111.4	3	3	2	2	2	-
DS24111.5	3	3	2	2	-	2
Average	3	3	2	2	2	2

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium) 3: Substantial (High)

No correlation : “-”

Subject Code	Subject Name (Theory Course)	Category	L	T	P	C
DS24112	Machine Learning Practices	PC	2	1	0	3

Objectives:
<ul style="list-style-type: none"> To build and evaluate predictive models using supervised learning algorithms such as regression model for classification
<ul style="list-style-type: none"> To build a neural network for binary classification of handwritten digits . Focus on understanding the architecture, activation functions, and optimization algorithms, and interpret the results.
<ul style="list-style-type: none"> To apply techniques such as cross-validation, grid search, and regularization to optimize model performance and ensure robustness.
<ul style="list-style-type: none"> To apply unsupervised learning algorithms like K-Means clustering and anomaly detection to discover hidden patterns and structures in unlabeled data.
<ul style="list-style-type: none"> To design and develop recommender systems using collaborative filtering

UNIT-I	Supervised Learning	9
Implement and understand the cost function and gradient descent for multiple linear regression - Implement and understand methods for improving machine learning models by choosing the learning rate - plotting the learning curve performing feature engineering - applying polynomial regression - Implement and understand the logistic regression model for classification - Learn why logistic regression is better suited for classification tasks than the linear regression model is - Implement and understand the cost function and gradient descent for logistic regression - Understand the problem of overfitting - improve model performance using regularization - Implement regularization to improve both regression and classification models		
UNIT-II	Advanced Learning Algorithm	9
Build a neural network for binary classification of handwritten digits using TensorFlow - Gain a deeper understanding by implementing a neural network in Python from scratch - Optionally learn how neural network computations are vectorized to use parallel processing for faster training and prediction - Build a neural network to perform multi-class classification of handwritten digits in TensorFlow -using categorical cross-entropy loss functions and the SoftMax activation - Learn where to use different activation functions - ReLu - linear - sigmoid - SoftMax in a neural network - depending on the task you want your model to perform - Use the advanced Adam optimizer to train your model more efficiently - Discover the value of separating your data set into training - cross-validation -test sets - Choose from various versions of your model using a cross-validation dataset -evaluate its ability to generalize to real-world data using a test dataset - Use learning curves to determine if your model is experiencing high bias or high variance - learn which techniques to apply regularization - adding more data - adding or removing input features to improve your model's performance		
UNIT-III	Model Building	9
Learn how the bias-variance trade-off is different in the age of deep learning - and apply Andrew Ng's advice for handling bias and variance when training neural networks - Learn to apply the iterative loop of machine learning development to train - evaluate - tune your model - Apply data-centric AI to not only tune your model but tune your data using data synthesis or data augmentation to improve your model's performance - Build decision trees and tree ensembles - such as random forest and XGBoost - boosted trees - to make predictions - Learn when to use neural network or tree ensemble models for your task - as these are the two most commonly used supervised learning models in practice today.		
UNIT-IV	Unsupervised Learning	9
Use unsupervised learning techniques for unsupervised learning: including clustering and anomaly detection - Implement K-mean clustering - Implement anomaly detection - Learn how to choose between supervised learning or anomaly detection to solve certain tasks.		
UNIT-V	Recommender Systems	9
Build a recommender system using collaborative filtering - Build a recommender system using a content-based deep learning method - Build a deep reinforcement learning model (Deep Q Network)." - Histograms - Box Plots etc - use of frequency distributions - mean comparisons - cross tabulation - statistical inferences using chi square - t-test and ANOVA - Outlier Analysis and Detection - outlier analysis - density based and distance based.		
Total Contact Hours: 45		

<p>Course Outcomes: Students will be able to get the knowledge about all the tools and techniques you need to apply machine learning to solve business problems</p>
<ul style="list-style-type: none"> ● Explain the principles of supervised learning and apply linear regression, logistic regression to solve classification and regression problems.
<ul style="list-style-type: none"> ● Implement and tuning neural networks models, and deep learning architectures, to enhance model performance.
<ul style="list-style-type: none"> ● Learn about ML Model Development, Model Evaluation Techniques, Model Deployment and Inferences, Model Explainability
<ul style="list-style-type: none"> ● Apply unsupervised learning algorithms like K-Means clustering and anomaly detection to discover hidden patterns and structures in unlabeled data.
<ul style="list-style-type: none"> ● Design and implement collaborative filtering, content-based filtering, and hybrid approaches, to provide personalized recommendations in real-world scenarios.

<p>SUGGESTED EVALUATION METHODS</p> <ul style="list-style-type: none"> ● Assignment problems ● Class Presentation/Discussion

<p>SUGGESTED ACTIVITIES</p> <ul style="list-style-type: none"> ● Problem solving sessions ● Activity Based Learning ● Implementation of small module
--

<p>Text Book(s):</p>
<p>1. "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems" by Aurélien Géron, 2nd Edition, O'Reilly Media Publisher, 2019</p>
<p>2. "Introduction to Machine Learning with Python: A Guide for Data Scientists" by Andreas C. Müller and Sarah Guido, 2nd Edition, Shroff Publishers & Distributors Pvt. Ltd., 2016</p>
<p>3. "Machine Learning: A Probabilistic Perspective" by Kevin P. Murphy, MIT Press, 2012</p>

Reference Books(s) / Web links:
1. Hang Li, Machine Learning Methods - By Springer Nature Singapore (2023)
2. Dr. R. Nageswara Rao, Machine Learning in Data Science Using Python - By Dreamtech Press (2022)

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
DS24112.1	3	-	3	-	2	-
DS24112.2	3	3		-	2	2
DS24112.3	3	-	3	3	-	2
DS24112.4	3	3	-	3	2	-
DS24112.5	3	3	-	3	-	2
Average	3	3	3	3	2	2

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium) 3: Substantial (High)

No correlation : “-

Subject Code	Subject Name (Theory Course)	Category	L	T	P	C
DS24113	Data Engineering	PC	2	1	0	3

Objectives:
<ul style="list-style-type: none"> Understand the fundamental principles of data infrastructure and architecture, including data storage, processing, and retrieval mechanisms.
<ul style="list-style-type: none"> Learn to design, develop, and manage ETL (Extract, Transform, Load) processes to efficiently move and transform data across various systems.
<ul style="list-style-type: none"> Gain proficiency in using big data technologies and frameworks such as Hadoop, Spark, and Kafka to handle large-scale data processing and streaming.
<ul style="list-style-type: none"> Develop skills in managing and optimizing different types of databases (SQL and NoSQL) to ensure data integrity, performance, and scalability.
<ul style="list-style-type: none"> Learn to automate data pipelines using tools and scripts, ensuring data flows are reliable, repeatable, and scalable, with an emphasis on monitoring and error handling.

UNIT-I	Data Types & Formats	9
Introduction to Data Types and Formats - Types of Data - Structured vs. Unstructured Data - Formats of Data - Semi-Structured Data - Data Type Conversion and Transformation - Data Serialization - Choosing the Right Data Type and Format - Tools and Technologies for Data Types and Formats.		
UNIT-II	Data Ingestion Techniques	9
Introduction to Data Ingestion - Streaming Data Ingestion - Batch Data Ingestion - Hybrid Data Ingestion - Data Ingestion vs. Data Integration - Data Ingestion Challenges - Tools and Solutions for Data Ingestion - StreamSets DataOps Platform Benefits of Data Ingestion - Data Ingestion Framework.		
UNIT-III	Data Profiling & Visual Representation via Various Tools (Pandas)	9
Introduction to Data Profiling and Visualization - Exploratory Data Analysis (EDA) with Pandas - Steps Involved in Exploratory Data Analysis (EDA) Data Analysis (EDA) with Pandas - Market Analysis with Exploratory Data Analysis (EDA). Data Analytics and Its Future Scope - Data Analytics with Python - Top Business Intelligence Tools - Application of Data Analytics - Retrieving and Cleaning Data - Exploratory Data Analysis and Feature Engineering - Inferential Statistics and Hypothesis Testing - Descriptive Statistics - Types of Descriptive Statistics - Concepts of Populations, Samples, and Variables - Statistical Methods for Describing Data Characteristics - Real-World Applications of Descriptive Statistics using Excel - Types of Missing Data and Handling Techniques.		
UNIT-IV	Storage and Retrieval Methods	9
Introduction to Storage and Retrieval - Types of Data and Storage Methods - Local vs. Distributed Storage & Retrieval - Hardware Aspects of Storage & Retrieval - Choosing Storage Methods - Data Partitioning and Sharding - Data Replication and Redundancy - Data Compression and Encoding - Data Archiving and Retrieval - Backup and Disaster Recovery - Data Lifecycle Management.		
UNIT-V	Data Lineage Analysis	9
Introduction to Data Lineage Analysis - Building a Data Flow - ETL (Extract, Transform, Load) Process - Usage of Data Warehouse - Edge Intelligence in Data Flow - Understanding Data Lineage - How Data Lineage Works - Benefits of Data Lineage - Data Lineage Tool Features.		
Total Contact Hours: 45		

Course Outcomes:

Students to understand the fundamentals of data engineering and its importance in modern data-driven applications.

- Identify and explain different data formats and their use cases, including structured, semi-structured, and unstructured data.
- Describe various data ingestion techniques, such as ETL, and stream processing, and their advantages and limitations.
- Perform data profiling and analyze data quality metrics to ensure data accuracy, completeness, and consistency.
- Design and implement effective storage and retrieval methods for large-scale data sets, including relational databases, NoSQL databases, and distributed file systems.
- Apply data engineering principles to real-world scenarios, such as data warehousing, big data analytics, and machine learning.

SUGGESTED ACTIVITIES

- Problem solving sessions
- Activity Based Learning
- Implementation of small module

SUGGESTED EVALUATION METHODS

- Assignment problems
- Quizzes
- Class Presentation/Discussion

Text Book(s):

1. Charles M.Judd, " Data Analysis: A Model Comparison Approach To Regression, ANOVA, and Beyond", 3rd Edition , Routledge Publiser , 2017
2. Pierre-Yves Bonnefoy, Emeric Chaize, Raphaël Mansuy & Mehdi TAZI, "The Definitive Guide to Data Integration", 1st Edition, Packt Publishing, 2024.

Reference Books(s) / Web links:

1. Joe Reis and Matt Housley , "Fundamentals of Data Engineering Plan and Build Robust Data Systems ", First Ediction, Shroff Publishers, 2022.
2. Paul Crickard , "Data Engineering with Python Work with massive datasets to design data models and automate data pipelines using Python" , Kindle Edition, Packt Publishing, 2020.

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
DS24113.1	3	2	-	-	-	3
DS24113.2	3	2	-	-	-	3
DS24113.3	3	3	2	-	-	3
DS24113.4	3	3	2	3	3	-
DS24113.5	-	3	2	3	3	-
Average	3	2.6	2	3	3	3

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium)

3: Substantial (High) No correlation : “-”

Subject Code	Subject Name (Theory Course)	Category	L	T	P	C
DS24A11	Generative Adversarial Networks	PE	2	1	0	3

Objectives:
<ul style="list-style-type: none"> Grasp the basic principles of GANs, including their architecture, components and the adversarial training process
<ul style="list-style-type: none"> Gain hands-on experience in constructing simple GAN models using popular deep learning frameworks such as TensorFlow or PyTorch.
<ul style="list-style-type: none"> Learn advanced techniques to improve GAN performance, including network architecture modifications, loss function adjustments, and training stability methods
<ul style="list-style-type: none"> Apply GANs to generate realistic images and ensure high-quality outputs
<ul style="list-style-type: none"> Explore the application of GANs in processing and enhancing satellite images

UNIT-I	Basic Generative Adversarial Networks (GANs)	9
<p>Overview of GenAI - Intro to GANs - Learn about GANs and their applications, understand the intuition behind the basic components of GANs -build your very own GAN using PyTorch - Deep Convolutional GAN - Build a more sophisticated GAN using convolutional layers - Learn about useful activation functions - batch normalization - and transposed convolutions to tune your GAN architecture and apply them to build an advanced DCGAN specifically for processing images</p>		
UNIT-II	Build Basic Generative Adversarial Networks	9
<p>Wasserstein GANs with Normalization - Reduce instances of GANs failure due to imbalances between the generator and discriminator by learning advanced techniques such as WGANs to mitigate unstable training and mode collapse with a W-Loss and an understanding of Lipschitz Continuity - Conditional and Controllable GANs - Understand how to effectively control your GAN - modify the features in a generated image - and build conditional GANs capable of generating examples from determined categories.</p>		
UNIT-III	Build Better Generative Adversarial Networks	9
<p>GAN Evaluation - Understand the challenges of evaluating GANs - learn about the advantages and disadvantages of different GAN performance measures - and implement the Fréchet Inception Distance FID method using embeddings to assess the accuracy of GANs -GAN Disadvantages and Bias - Find out the disadvantages of GANs when compared to other generative models - discover the pros/cons of these models — plus - learn about the many places where bias in machine learning can come from - why it's important - and an approach to identify it in GANs - StyleGAN and Advancements - Understand how StyleGAN improves upon previous models and implements the components and the techniques associated with StyleGAN - currently the most state-of-the-art GAN with powerful capabilities.</p>		
UNIT-IV	Apply Generative Adversarial Networks to images	9
<p>GANs for Data Augmentation and Privacy Preservation - Explore the applications of GANs and examine them with respect to data augmentation, privacy, and anonymity Improve your downstream AI models with GAN-generated data - Image-to- Image Translation - Leverage the image-to-image translation framework and identify extensions – generalizations</p>		
UNIT-V	Apply Generative Adversarial Networks to satellite images	9
<p>GAN - to adapt satellite images to map routes with advanced U-Net generator -Patch GAN discriminator architectures - Image-to-Image Unpaired Translation - Compare paired image-to-image translation to unpaired image-to-image translation and identify how their key difference necessitates different GAN architectures - Implement CycleGAN- an unpaired image-to-image translation model, to adapt horses to zebras with two GANs in one.</p>		
Total Contact Hours: 45		

Course Outcomes: Students gain comprehensive understanding of deep learning techniques and generative AI models.
<ul style="list-style-type: none"> Understand generative models such as generative adversarial networks (GANs) and their advanced techniques.
<ul style="list-style-type: none"> To build sophisticated and robust GAN models using PyTorch & convolutional layers etc.,
<ul style="list-style-type: none"> Student to learn about the advantages and disadvantages of different GAN performance measures.
<ul style="list-style-type: none"> Students to explore and examine the applications of GANs
<ul style="list-style-type: none"> Students apply GAN model on images for data augmentation and image translation

SUGGESTED EVALUATION METHODS
<ul style="list-style-type: none"> Assignment problems Quizzes Class Presentation/Discussion

SUGGESTED ACTIVITIES
<ul style="list-style-type: none"> Problem solving sessions Activity Based Learning Implementation of small module

Text Book(s):
1. Jakub Langr & Vladimir Bok, "GANs in Action: Deep learning with Generative Adversarial Networks" Manning; 1st edition , 2019.
2. John Hany, "Hands-On Generative Adversarial Networks with PyTorch 1.x", Packt Publishing, 2019.

Reference Books(s) / Web links:
1. Jakub Langr and Vladimir Bok, "GANs in Action: Deep learning with Generative Adversarial Networks" , First Edition, Manning , 2019.
2. Rafael Valle , "Hands-On Generative Adversarial Networks with Keras: Your Guide to Implementing Next-Generation Generative Adversarial Networks" , Kindle Edition, Packt Publishing, 2019.

CO-PO MAPPING

PO \ CO	PO1	PO2	PO3	PO4	PO5	PO6
DS24A11.1	3	2	-	-	-	3
DS24A11.2	3	2	-	-	-	3
DS24A11.3	3	3	2	-	-	3
DS24A11.4	3	3	2	3	3	-
DS24A11.5	-	3	2	3	3	-
Average	3	2.6	2	3	3	3

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium) 3: Substantial (High)

No correlation : "-"

Subject Code	Subject Name (Laboratory Course)	Category	L	T	P	C
GE24121	Professional Soft Skills - I	EEC	0	0	2	1

Objectives:
<ul style="list-style-type: none"> Develop the ability to communicate clearly and effectively in both written and verbal forms, tailored to diverse professional settings and audiences.
<ul style="list-style-type: none"> Enhance skills in teamwork and collaboration, including understanding group dynamics, conflict resolution, and contributing positively to team goals
<ul style="list-style-type: none"> Cultivate critical thinking and problem-solving abilities, enabling students to analyze complex situations, make informed decisions, and implement effective solutions
<ul style="list-style-type: none"> Learn strategies for effective time management and organizational skills to prioritize tasks, manage workloads efficiently, and meet deadlines
<ul style="list-style-type: none"> Develop leadership and interpersonal skills, including emotional intelligence, networking, negotiation, and the ability to inspire and motivate others.

UNIT-I	Positive Attitude	9
<p>Attitude- Campus to Corporate attitude change, Recognizing Negative Attitude, Campus to Corporate attitude change; Attitude at work- Impact of Negative Attitude in the Workplace, Overcoming Negative Attitude, positive attitude, thought process, Building self-confidence and Assertiveness; Toxic positivity; 3Es, Motivation-Intrinsic and Extrinsic Motivation, Inspiration vs motivation; Emotional Intelligence-Intro to EI, Four clusters. Transactional Analysis (TA), SWOT analysis - Professional analysis</p>		
UNIT-II	Body Language	9
<p>Importance of Body Language, Five Cs of Body Language, Body language in different cultures, Positive Body Language; Voice Control- Pace. Pause and Pitch; Culture-Inclusivity and Proxemics across Global Cultures, Understanding POSH; Stress Management-What is Stress, Eustress, Reasons of stress (work/ personal); Stress Management Techniques</p>		
UNIT-III	Presentation Skills	9
<p>Self-introduction – Exercises, Why Give Presentations; Craft your message-Plan the visuals, Manage the Response; How to create an effective presentation - Virtual & Physical, Do's & Don'ts of Presentation Skills, Objection handling, Stage Fear – Causes and Cure, Practice the Delivery; Time Management-Common Time & Energy Wasters, Planning & Prioritizing Time Matrix & Analysis.</p>		
UNIT-IV	Listening & Questioning Skills	9
<p>Barriers to effective listening - how to overcome them; Exercises - Customer Call Flow – Role-play, Cust calls amongst the team; How to frame Questions, Different kinds of questions, asking appropriate questions; Spoken English- Introduction to Parts of Speech and its usage; Subject - Verb Agreement; Basic conversation skills-sentence construction -SVO.</p>		
UNIT-V	Team Work	9
<p>Teamwork and Ethics - Definition of TEAM - Team vs Groups. Difference b/w Healthy competition and cut throat competition, Importance of working in teams, Evolution of a TEAM, Benefits of team work; Virtual teams- Challenges and ways to overcome it, Diversity and Inclusion in a team; Development of Teams Stages of team development; Team dynamics-its importance & Interpersonal Skills Development Ethics- to enable students to identify and deal with ethical problems, develop their moral intuitions, which are implicit in everyday choices and actions; Conflict Management: Team building Activities- Predetermined/ Predesigned Indoor/ Outdoor activities to build a team, enhance language and inter personal skills</p>		
Total Contact Hours: 45		

Course Outcomes:

Students to understand professional, behavioral and presentation skills while working with team and practically experience important aspects of it.

- To help the students understand and implement positive outlook, interpret the body language of team members and stakeholders, better interpersonal relationships. Develop into self-motivated professionals with confidence. Practice Responding instead of Reacting.
- Create good Presentation and Present with confidence. Also, recognize and manage Stress, Prioritize and Plan.
- To listen to understand. To be able to ask good questions.
- To understand to be a good Team player, Team Dynamics and to understand the Business Ethics
- To be able to write and speak correctly, forming grammatically correct sentences.

SUGGESTED ACTIVITIES

- Problem solving sessions
- Flipped classroom
- Activity Based Learning

SUGGESTED EVALUATION METHODS

- Tutorial problems
- Assignment problems
- Quizzes
- Class Presentation/Discussion

Text Book(s):

1. Virender Kapoor , "The Soft Skills Handbook: Empowering Youth for Career Excellence ", Atlantic Publishers , 2024.
2. Frederick H. Wentz , "Soft Skills Training: A Workbook to Develop Skills for Employment" , 2012.
3. Stella Cottrell , "Skills for Success: Personal Development and Employability" , 4th Edition, Bloomsbury Publishing, 2021.

Reference Books(s) / Web links:

1. Allan Pease and Barbara Pease , "The Definitive Book of Body Language", RHUS, 2006
2. "Crucial Conversations: Tools for Talking When Stakes Are High" by Kerry Patterson, Joseph Grenny, Ron McMillan, and Al Switzler, Brilliance Audio , 2013

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
GE24121.1	3	3	-	-	-	-
GE24121.2	3	3	3	3	-	2
GE24121.3	3	3	3	3	-	2
GE24121.4	3	3	3	3	-	2
GE24121.5	3	3	-	-	-	2
Average	3	3	3	3	-	2

Correlation levels 1, 2 or 3 are as defined below:

- 1: Slight (Low) 2: Moderate (Medium)
3: Substantial (High) No correlation : “-”

Subject Code	Subject Name (Laboratory Course)	Category	L	T	P	C
DS24121	Applied Machine Learning Lab	PC	0	0	4	2

Objectives:
<ul style="list-style-type: none"> Develop Practical Skills in Data Preprocessing and Feature Engineering
<ul style="list-style-type: none"> Implement and Evaluate Machine Learning Algorithms
<ul style="list-style-type: none"> Apply Machine Learning Techniques to Real-World Problems
<ul style="list-style-type: none"> Optimize and Fine-Tune Machine Learning Models
<ul style="list-style-type: none"> Communicate Machine Learning Results Effectively

Description of the Experiments	Total Contact Hours: 6.0
1. Understanding "Mobile Price" dataset by doing feature analysis. Data is available at: https://www.kaggle.com/datasets/iabhishekoofficial/mobile-price-classification/data	
2. Execute data preprocessing step on the above dataset: perform outlier and missing data analysis towards building a refined dataset	
3. Build machine learning model/s to predict the actual price of the new mobile based on other given features like RAM, Internal Memory etc	
4. Calculate the prediction accuracy of the models used in Experiment 3 and do comparative analysis among them to identify the best technique.	
5. Understanding "Second Hand Car Prediction Price" dataset by doing feature analysis. Data is available at: https://www.kaggle.com/datasets/sujithmandala/second-hand-car-price-prediction	
6. Perform data preprocessing step on the above dataset: perform outlier and missing data analysis towards building a refined dataset.	
7. Perform Feature Engineering towards building new feature which is more impactful. Build machine learning model/s to predict the price of the car based on other given features like Brand, Model, Year, Fuel Type etc	
8. Calculate the prediction accuracy of the models used in Experiment 7 and do comparative analysis among them to identify the best technique.	
9. Plot the features (actual price and predicted price) in scatter plot to understand the variation.	
10. Understanding "Marketing Campaign Positive Response Prediction" dataset by analysing all the features. Data is available at: https://www.kaggle.com/datasets/sujithmandala/marketing-campaign-positive-response-prediction	
11. Perform exploratory data analysis on the above dataset: perform outlier and missing data analysis towards building a refined dataset. Show the outliers in box plot or through some statistical technique. Find the numerical and categorical features.	
12. Perform Feature Engineering towards building new feature which is more impactful than the existing ones. Build the correlation matrix and show visually the relationship among various features.	
13. Build machine learning model/s to predict the result of marketing campaign based on other given features like customer details, gender, annual income etc	
14. Calculate the prediction accuracy of the models used in Experiment 13 and do comparative analysis among them to identify the best technique.	
15. Please check whether you find imbalanced classes, overfitting, and data bias in the above two datasets. Please apply some technique to overcome it.	

Course Outcomes:

1. Expertise in Data Preparation and Feature Engineering
2. Hands-On Experience with Machine Learning Algorithms
3. Ability to Solve Real-World Problems
4. Skills in Model Optimization and Fine-Tuning
5. Effective Communication of Machine Learning Insights

SUGGESTED EVALUATION METHODS

- Experiment based viva
- Quizzes
- Mini Project

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
DS24121.1	3	3	-	-	-	2
DS24121.2	3	3	-	-	-	2
DS24121.3	3	3	2	3	3	2
DS24121.4	3	3	2	3	3	2
DS24121.5	3	3	2	3	3	2
Average	3	3	2	3	3	2

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium)

3: Substantial (High) No correlation : “-”

Subject Code	Subject Name (Laboratory Course)	Category	L	T	P	C
DS24122	Data Science Lab	PC	0	0	4	2

Objectives:
1. Develop Skills in Data Collection and Preparation
2. Gain Proficiency in Data Analysis and Visualization Tools
3. Apply Machine Learning Algorithms to Real-World Data
4. Develop and Evaluate GAN Models
5. Communicate Data Insights Effectively

Description of the Experiments	Total Contact Hours: 60
Case Study 1: Present your view on the different techniques you have employed to do outlier analysis, handling missing data, feature engineering, feature importance and improving the accuracy of the model both from a classifier as well as a regressor. Use any sample data and present your POV in a well-structured presentation.	
Case Study 2: Present your findings on different activation functions you have used and methods to improve the accuracy of the model using neural networks. You should be able to clearly articulate the advantage and disadvantage of each activation function. Use any sample data and present your POV in a well-structured presentation.	
Case Study 3: Present your findings on different techniques of anomaly detection and k means clustering. Use any sample data and present your POV in a well-structured presentation	
Case Study 4: Present your POV on how to generate synthetic data using GANs. You can assume a sample dataset from an IOT enabled machine where the failure rates are minimal.	
Case Study 5: Present your POV on Style related GANS. Explore the earliest models to the current models. Articulate the successive improvements in the models. Also articulate the future of GANs in generating realistic images.	
Case Study 6: Present your POV on GANs used for Deep Fakes. Articulate how we can identify the Deep Fake from the original.	

Course Outcomes:
1. Develop Competence in Data Cleaning and Preprocessing
2. Gain Proficiency in Machine Learning and Statistical Modeling
3. Develop Abilities in Data Synthesis and Augmentation
4. Communicate Data Science Insights Effectively
5. Apply GAN Techniques to Real-World Problems

SUGGESTED EVALUATION METHODS
<ul style="list-style-type: none"> ● Experiment based viva ● Quizzes ● Mini Project

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
DS24122.1	3	3	-	-	-	2
DS24122.2	3	3	-	-	-	2
DS24122.3	3	3	2	3	3	2
DS24122.4	3	3	2	3	3	2
DS24122.5	3	3	2	3	3	2
Average	3	3	2	3	3	2

Correlation levels 1, 2 or 3 are as defined below:

- 1: Slight (Low) 2: Moderate (Medium)
- 3: Substantial (High) No correlation : “-”

Subject Code	Subject Name	Category	L	T	P	C
AC23111	English for Research Paper Writing (Common to All branches of M.E. /M.Tech –Semester I)	MC	3	0	0	0

Objectives:
<ul style="list-style-type: none"> To facilitate the students to express technical ideas in writing
<ul style="list-style-type: none"> To train the students in using language structures appropriately
<ul style="list-style-type: none"> To enable students to plan and organize the research paper
<ul style="list-style-type: none"> To assist the students in understanding the structure and familiarize the mechanics of organized writing
<ul style="list-style-type: none"> To equip the students to improvise academic English and acquire research writing skills

UNIT-I	INTRODUCTION TO RESEARCH WRITING	9
Research – Types of Research - Selecting the Primary resources - Categorizing secondary sources - Discovering a researchable area and topic – Need Analysis - Research Question- Focus on the Research Problem- Developing Research Design – Framing the Hypothesis – Identifying the Scope of the Research - Writing – General and Academic Writing		
UNIT-II	LANGUAGE OF WRITING	9
Active reading – text mining – use of academic words – jargon's – ambiguities – use of expression – use of tense - proper voices – third person narration – phraseology – use of foreign words – use of quotes – interpreting quotes.		
UNIT-III	THE FORMAT OF WRITING	9
Types of Journals - different formats and styles - IEEE format - Structure – Margins - Text Formatting - Heading and Title - Running Head with Page Numbers - Tables and illustrations - Paper and Printing - Paragraphs - Highlighting – Quotation – Footnotes		
UNIT-IV	ORGANISING A RESEARCH PAPER	9
Title- Abstract – Introduction – Literature review - Methodology - Results –Discussion – Conclusion - Appendices - Summarizing - Citation and Bibliography		
UNIT-V	PUBLISHING PAPER	9
Finding the Prospective publication or Journal - analyzing the credits - Reviewing - Revising –Plagiarism Check - Proofreading - Preparing the Manuscript- Submitting - Resubmitting - Follow up - Publishing		
		Total Contact Hours : 45
Course Outcomes:		
On completion of the course, students will be able to		
1. Understand the basic structure of research work		
2. Apply proper use of language in writing paper		
3. Comprehend different formats of journal paper		
4. Follow the process of writing a research paper and write one		
5. Emulate the process of publishing journal paper and publish papers		

Text Book(s):
1. Adrian Wallwork: “English for Writing Research Papers”, Springer Science Business Media, Second Edition, LLC 2011
2. Stephen Howe and Kristina Henriksson: “Phrasebook for Writing Papers and Research in English”, The Whole World Company Press, Cambridge, Fourth edition 2007
3. The Modern Language Association of America: “MLA Handbook for Writers of Research Papers” 8th Edition, The Modern Language Association of America, 2016

Reference Books(s) / Web links:
1. Rowena Murray: The Handbook of Academic Writing: A Fresh Approach, Sarah Moore Open University Press, 2006
2. Stephen Bailey: Academic Writing: A Practical Guide for Students Routledge Falmer: 2003.
3. Joseph M. Moxley: Publish, Don't Perish: The Scholar's Guide to Academic Writing and Publishing, Praeger Publishers, 1992

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
AC23111.1	-	1	-	-	3	3
AC23111.2	-	-	-	-	3	1
AC23111.3	-	2	-	-	3	3
AC23111.4	-	1	-	-	2	3
AC23111.5	-	1	-	-	3	2
Average	0	1.25	0	0	2.8	2.4

Correlation levels 1, 2 or 3 are as defined below:

- 1: Slight (Low) 2: Moderate (Medium)
- 3: Substantial (High) No correlation : “-”

Subject Code	Subject Name (Theory Course)	Category	L	T	P	C
DS24211	Deep Learning Applications	PC	2	1	0	3

Objectives:

- Understand the Fundamentals of Deep Learning
- Develop and Implement a Facial Recognition System
- Evaluate and Optimize Facial Recognition Models
- Build a Weather Forecasting System Using Deep Learning
- Integrate a Chatbot for Weather Forecasting

UNIT-I	Convolutional Neural Network	9
Transfer learning - Data Augmentation - Image segmentation - Object detection - Video classification - Text and natural language processing - Structured data - Model optimization		
UNIT-II	Facial Recognition Model	9
Writing the code - Deploying the API as container - Consuming the API from Frontend and display- Preparing the image dataset - Creating and training the Model; Build and deploy Flask REST API on Docker: steps to dockerize your flask app; Docker: Docker Installation – Architecture – Working of Docker; Kubernetes: Overview – Architecture – Kubernetes Setup – Advanced Kubernetes; Flask : Overview – Environment – Application.		
UNIT-III	Facial Recognition System	9
Create Endpoints and UI to retrain the system with new data (faces) - Feedback system for face labels - Transfer Learning - Reusing the knowledge with additional learning; Technology: Flask, streamlit and Tensorflow - Create a multipage app - API reference - Advanced features - Components - Roadmap - Changelog - Cheat sheet - Streamlit community cloud.		
UNIT-IV	Recurrent Neural Network	9
Architecture - Technology and libraries - Application of RNN - Limitations of RNN - Improvement LSTM - RNN in time series - Build an RNN to predict time series in TensorFlow - Text generation with an RNN ; Chatbot : Working of chatbot - Types of Chatbot - Use cases of chatbots - Objective - End goal – constraints How to build a chatbot - A ten - minute introduction to sequence to sequence learning in keras - Chatbot using seq2seq LSTM models - Architecture of seq2seq model		
UNIT-V	Intelligent Chatbot	9
Using LSTM - Using NLP - LSTM Time series Analysis - LSTM weather - Create an Intelligent chatbot in Python using the spaCy NLP Library - Prerequisites - Setting up the environment - creating the city weather program - Creating the chatbot.		
Total Contact Hours: 45		

Course Outcomes:

Students to understand professional, behavioral and presentation skills while working with team and practically experience important aspects of it.

1. To understand the fundamentals of deep learning and its applications in computer vision, time series analysis and natural language processing.
2. To build a facial recognition system.
3. To build a weather forecasting system.
4. To build a chat bot for weather forecasting system.
5. Ability to evaluate and optimize deep learning models

SUGGESTED ACTIVITIES

- Problem solving sessions
- Flipped classroom
- Activity Based Learning

SUGGESTED EVALUATION METHODS

- Assignment problems
- Quizzes
- Class Presentation/Discussion

Objectives:

- Understand the Fundamentals of Generative AI
- Introduction to Large Language Models (LLMs)
- Fine-Tuning Large Language Models
- Reinforcement Learning with LLMs
- Developing LLM-Powered Applications

Text Book(s):

1. Simon J.D. Prince, Understanding Deep Learning, Illustrated Edition, MIT Press, 5 Dec 2023.
2. Stan Z. Li & Anil K. Jain, Handbook of Face Recognition Second Edition. Springer-Verlag, 2004.
3. Mohamed Elgendy, Deep Learning for Vision Systems, Manning, 2020.

Reference Books(s) / Web links:

1. James D. Hamilton, Time Series Analysis, Levant Books, 2012.
2. V Kishore Ayyadevara & Yeshwanth Reddy, Modern Computer Vision with PyTorch, Packt Publishing, 2020.
3. Stan Z. Li & Anil K. Jain, Handbook of Face Recognition Second Edition. Springer-Verlag, 2004 .

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
DS24211.1	3	2	-	-	-	2
DS24211.2	3	3	2	3	3	2
DS24211.3	3	3	2	3	3	2
DS24211.4	3	3	2	3	3	2
DS24211.5	-	3	2	3	3	2
Average	3	2.8	2	3	3	2

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium) 3:

Substantial (High) No correlation : “-”

Subject Code	Subject Name (Theory Course)	Category	L	T	P	C
DS24212	Generative AI with Large Language Models	PC	2	1	0	3

Objectives:
<ul style="list-style-type: none"> Understand the Fundamentals of Generative AI
<ul style="list-style-type: none"> Introduction to Large Language Models (LLMs)
<ul style="list-style-type: none"> Fine-Tuning Large Language Models
<ul style="list-style-type: none"> Reinforcement Learning with LLMs
<ul style="list-style-type: none"> Developing LLM-Powered Applications

UNIT-I	Introduction to Generative AI	9
Introduction Generative AI & LLMs - LLM use cases and tasks - Text generation before transformers - Transformers architecture - Generating text with transformers - Prompting and prompt engineering (CoT) – RAG Technique for retrieval - Generative configuration - Generative AI project lifecycle - Pre-training large language models - Computational challenges of training LLMs.		
UNIT-II	Fine Tuning	9
Instruction fine-tuning - Fine-tuning on a single task - multi-task instruction fine-tuning		
UNIT-III	Evaluation	9
Model evaluation –Benchmarks -Parameter efficient fine-tuning (PEFT) -PEFT techniques 1: LoRA - PEFT techniques 2: Soft prompts		
UNIT-IV	Reinforcement Learning	9
Aligning models with human values - Reinforcement learning from human feedback (RLHF) - RLHF: Obtaining feedback from humans - Reward model - Fine-tuning with reinforcement learning - Model optimizations for deployment		
UNIT-V	LLM-powered applications	9
Using the LLM in applications - Interacting with external applications - Helping LLMs reason and plan with chain-of-thought - Program-aided language models (PAL) - ReAct: Combining reasoning and action - LLM application architecture		
Total Contact Hours: 45		

Course Outcomes:
Students to get the knowledge to adapt pre-trained LLMs to more specialized tasks.
1. Understand Fundamentals of Generative AI and LLM
2. Learn the concept to Fine Tune Large Language Models
3. Learn the concept to evaluate Large Language Models
4. Integrated reinforcement learning methods with large language models
5. Design the power of large language models for real-world use cases.

SUGGESTED ACTIVITIES

- Problem solving sessions
- Flipped classroom
- Activity Based Learning

SUGGESTED EVALUATION METHODS

- Assignment problems
- Quizzes
- Class Presentation/Discussion

Text Book(s):

1. David Foster , "Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play", 2nd Edition, O'Reilly Media, 2023.

2. Maxim Lapan, "Deep Reinforcement Learning Hands-On" , Packt Publishing, 2018.

Reference Books(s) / Web links:

1. Edward R. DeForest, Prompt Engineering with Transformers and LLM – By Kindle (2024).

2. Altaf Rehmani, Generative AI for everyone – By Altaf Rehmani; 1st edition (2024).

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
DS24212.1	3	2	-	-	-	3
DS24212.2	3	2	-	-	-	3
DS24212.3	3	2	-	-	-	3
DS24212.4	3	3	3	3	3	-
DS24212.5	-	3	3	3	3	-
Average	3	2.4	2	3	3	3

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium)

3: Substantial (High) No correlation : “-”

Subject Code	Subject Name (Theory Course)	Category	L	T	P	C
DS24B11	Ethics in Data Science	PE	2	1	0	3

Objectives:
<ul style="list-style-type: none"> Develop a comprehensive understanding of the ethical principles and frameworks that guide data science practices, including privacy, fairness, and transparency..
<ul style="list-style-type: none"> Learn to identify and analyze ethical issues and dilemmas that arise in the collection, analysis, and use of data in data science projects..
<ul style="list-style-type: none"> Gain knowledge of relevant laws, regulations, and standards (such as GDPR and CCPA) that govern data privacy and protection, and understand their implications for data science.
<ul style="list-style-type: none"> Learn best practices for responsible data handling, including informed consent, data anonymization, and measures to prevent bias and discrimination in data-driven decisions.
<ul style="list-style-type: none"> Develop skills to make ethical decisions in data science by applying ethical frameworks and considering the broader social and cultural impacts of data science work.

UNIT-I	Introduction and Philosophical Frameworks for Assessing Fairness	9
Foundations of ethics - early theories of fairness (Utilitarianism etc.) - contemporary theories of fairness - significance of ethics in data science - ethics vs. law/compliance/public relations - cultural relativism - “professional” ethics in data science - individuals vs. collectives		
UNIT-II	Research Ethics	9
Understanding the difference between data ownership - data privacy and data anonymity - understanding the idea behind data surveillance - data privacy vs. data security		
UNIT-III	Data Ownership, Privacy and Anonymity	9
EU’s general data protection rules - GDPR - digital India policy - personal data protection bill - 2019 -PDP Bill- ethical issues on data privacy in context with India - case studies		
UNIT-IV	Algorithmic Fairness Policies on Data Protection	9
Data driven research, methods of collection of data - different types of data: qualitative and quantitative - overview of ethical issues in data-driven organizations - doing ethical data analysis - responsible use of research data - plagiarism - fake data and fabrication of data - creation of data base - Discrimination and algorithms- obscure and unintentional bias displayed by the algorithms - ethics of data scraping and storage- Mosaic data- found data- and designed data.		
UNIT-V	Responsible AI, Red Teaming on LLM & Case Study	9
Various dimensions of Responsible AI - Dimensions of Ethical AI - Bias Mitigation Techniques; Constitutional AI: Rules of Constitutional AI - How to create Constitutional AI complaint system - Model fine tuning for Constitutional AI - What are the vulnerabilities - How to attack those problems by Red Teaming		
Total Contact Hours: 45		

Course Outcomes:
1. Students will demonstrate a heightened awareness of ethical issues in data science and the ability to articulate these issues clearly.
2. Students will be able to identify and comply with relevant data privacy and protection regulations, ensuring their data practices are legally and ethically sound.
3. Students will be capable of identifying potential biases in data and algorithms and applying techniques to mitigate these biases to promote fairness and equity.
4. Students will apply best practices for ethical data handling, including data collection, storage, sharing, and analysis, to protect individuals' rights and privacy.
5. Students will use ethical decision-making frameworks to evaluate and address ethical dilemmas in data science, ensuring responsible and socially conscious data practices.

SUGGESTED ACTIVITIES

- Problem solving sessions
- Flipped classroom
- Survey on various storage technologies
- Activity Based Learning
- Implementation of small module

SUGGESTED EVALUATION METHODS

- Tutorial problems
- Assignment problems
- Quizzes
- Class Presentation/Discussion

Text Book(s):

1. David Martens , "Data Science Ethics Concepts, Techniques, and Cautionary Tales" , Oxford University Press, 2022
2. Mike Loukides, Hilary Mason, and DJ Patil, "Ethics and Data Science", O'Reilly Media, Inc, 2018.

Reference Books(s) / Web links:

1. Kirsten Martin, Ethics of Data and Analytics: Concepts and Cases,First Edition, Taylor & Francis, 2022.
2. Anne L. Washington, Ethical Data Science Prediction in the Public Interest, Oxford, 2023

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
DS24B211.1	3	2	3	-	-	2
DS24B211.2	3	3	3	-	-	2
DS24B211.3	3	3	3	2	-	2
DS24B211.4	3	3	3	3	3	2
DS24B211.5	3	3	3	3	3	2
Average	3	2.8	3	3.7	3	2

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium) 3:

Substantial (High) No correlation : “-”

Subject Code	Subject Name (Theory course)	Category	L	T	P	C
DS24B12	Computer Vision	PE	2	1	0	3

Objectives:
<ul style="list-style-type: none"> Understand and apply fundamental image processing technique Develop skills to extract and analyze significant features from images Implement and evaluate machine learning models for various computer vision applications Gain proficiency in 3D computer vision techniques to reconstruct and interpret three-dimensional scenes from images. Latest advancements and future trends in computer vision are explored for real time applications

UNIT-I	Image Processing Techniques	9
<p>Introduction to image processing: What is image processing? - Understanding about types of image processing- Visualization, Recognition, Sharpening & Restoration, Pattern Recognition, Retrieval; Image Transformation: Image Enhancement Techniques: Histogram Equalization, Contrast Stretching, Adaptive Enhancement - Image Restoration Methods: Deblurring, Denoising, Inpainting - Linear Filtering: Convolution, Gaussian Filtering, Edge Detection - Independent Component Analysis (ICA) - Pixelation and Its Applications; Image Generation Technique: Procedural Image Generation: Fractal Generation, Noise-based Generation - Generative Adversarial Networks (GANs) for Image Generation: Introduction to GANs- Understanding the architecture and training process of generative adversarial networks, Implementing GANs for generating realistic images, including applications in image-to-image translation and style transfer. - Applications of Image Generation Techniques: Data Augmentation, Creative Applications.</p>		
UNIT-II	Feature Extraction and Image Analysis	9
<p>Feature Detection: Introduction to feature detection - Object recognition techniques (key point detection, edge detection) - Image segmentation algorithms (region growing, thresholding, etc.) - Frequency domain processing (Fourier transform, frequency filtering) - Feature extraction methods (SIFT, SURF); Object Description: Introduction to fundamentals of moving object detection - Moving object description techniques (optical flow, background subtraction) - Camera geometry for object description (camera calibration, pose estimation).</p>		
UNIT-III	Machine Learning for Computer Vision	9
<p>Image Classification: Introduction to machine learning for computer vision - Image classification models (CNNs, transfer learning) - Object detection with machine learning (YOLO, SSD) - Labeling images for machine learning (annotation tools, data augmentation).</p>		
UNIT-IV	3D Computer Vision	9
<p>Depth Perception: Comparison of 2D and 3D computer vision - Real-world applications and trends in 3D computer vision - Classification of 3D data (point clouds, meshes).</p>		
UNIT-V	Advanced CV and Future Trends	9
<p>Advanced Computer Vision Applications: Brain Tumor Detection - Integrating Computer Vision in Autonomous Driving Systems - Computer Vision Applications in the Food Industry; Object Detection and Recognition: Visual Tracking - Semantic Segmentation - Human Recognition.</p>		
Total Contact Hours: 45		

Course Outcomes: Students to understand the computer vision techniques in depth with various applications of the same.
1. Understand what techniques are available to process the image.
2. Understand how to analyze the image and extract required features
3. Explain Machine Learning Concepts for Computer Vision
4. Explore 3D Computer Vision Concepts
5. Understand how computer vision solves real world problems

SUGGESTED ACTIVITIES
<ul style="list-style-type: none"> ● Problem solving sessions ● Flipped classroom ● Activity Based Learning

SUGGESTED EVALUATION METHODS
<ul style="list-style-type: none"> ● Assignment problems ● Quizzes ● Class Presentation/Discussion

Text Book(s):
1. V Kishore Ayyadevara & Yeshwanth Reddy, Modern Computer Vision with PyTorch, Packt Publishing, 2020
2. B Cyganek, An Introduction to 3D Computer Vision Techniques and Algorithms - By John Wiley & Sons Inc; 1st edition, 2009

Reference Books(s) / Web links:
1. Rafael C. Gonzalez and Richard E. Woods , "Digital Image Processing", Fourth Edition, Pearson, 2018.
2. Mark Nixon and Alberto S. Aguado , "Feature Extraction and Image Processing for Computer Vision" Third Edition, Academic Press, 2013.
3. Richard Szeliski , "Computer Vision: Algorithms and Applications", Second Edition, Springer, 2022.

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
DS24B12.1	3	2	-	-	-	2
DS24B12.2	3	2	-	-	-	2
DS24B12.3	3	2	-	3	-	2
DS24B12.4	3	2	-	-	-	2
DS24B12.5	2	3	3	3	3	-
Average	3	2.2	3	3	-	2

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium) 3:

Substantial (High) No correlation : “-”

Subject Code	Subject Name (Theory Course)	Category	L	T	P	C
DS24B13	Natural Language Processing	PE	2	1	0	3

Objectives:
<ul style="list-style-type: none"> Understand the fundamental concepts, theories, and algorithms that form the basis of Natural Language Processing, including syntax, semantics, and pragmatics.
<ul style="list-style-type: none"> Learn and apply various text processing techniques such as tokenization, stemming, lemmatization, and part-of-speech tagging to prepare textual data for analysis.
<ul style="list-style-type: none"> Develop skills in applying machine learning algorithms to NLP tasks, including text classification, sentiment analysis, and named entity recognition using libraries such as Scikit-learn and TensorFlow.
<ul style="list-style-type: none"> Gain proficiency in using deep learning techniques and models such as RNNs, LSTMs, and Transformers for advanced NLP applications like language modeling, machine translation, and text generation
<ul style="list-style-type: none"> Implement and evaluate real-world NLP applications, including chatbots, recommendation systems, and information retrieval systems, using Python and popular NLP libraries like NLTK, SpaCy, and Hugging Face

UNIT-I	NLP Need & Real-World Applications	9
What is NLP and its components? - Phases of NLP - Challenges of natural language - Applications of NLP - Industries using NLP - NLP programming languages - NLP libraries and Development environments - Use of AI in NLP - Basic Text Processing and Linguistic Concepts: Tokenization - Stemming - Lemmatization - Part-of-Speech Tagging.		
UNIT-II	Text Classification	9
Convolutional Neural Networks (CNNs) for NLP - Recurrent Neural Networks (RNNs) for NLP - Recursive Neural Networks - Hybrid Models for NLP		
UNIT-III	Deep Learning for NLP	9
Basics of Voice Recognition: Difference between speech and voice recognition - Use of NLP in voice recognition and transformation: Speech recognition using NLP models (HMM, DTW) - Acoustic modelling - Error correction in voice recognition		
UNIT-IV	Transfer Learning for NLP	9
Benefits of Text Classification - Types of Text classification - Challenges in text classification - Applications of text classification		
UNIT-V	Voice Recognition	9
Benefits of Transfer Learning for NLP - Fine Tuning techniques - Fine-Tune BERT for Spam Classification		
Total Contact Hours: 45		

Course Outcomes:
1. Understand the purpose of NLP and how to use it in real world applications with example.
2. Understand how to solve a classification problem.
3. Understand how deep learning is applied for NLP.
4. Understand the transfer learning concepts for reusability of knowledge.
5. Understand the applications of voice recognition system

SUGGESTED ACTIVITIES
<ul style="list-style-type: none"> Problem solving sessions Flipped classroom Activity Based Learning

SUGGESTED EVALUATION METHODS

- Assignment problems
- Quizzes
- Class Presentation/Discussion

Text Book(s):

- 1.Sowmya Vajjala, Bodhisattwa Majumder, Anuj Gupta & Harshit Surana, Practical Natural Language Processing, O'Reilly Media, 2020
- 2.Uday Kamath, John Liu & James Whitaker, Deep Learning for NLP and Speech Recognition, 1st edition, Springer 2019

Reference Books(s) / Web links:

1. Dan Jurafsky and James H. Martin, Speech and Language Processing, 2024.
2. Sowmya Vajjala , Bodhisattwa Majumder, Practical Natural Language Processing: A Comprehensive Guide to Building Real-World NLP Systems, Kindle Edition, 2020.

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
DS24B13.1	3	3	-	-	-	-
DS24B13.2	3	3	3	3	3	2
DS24B13.3	3	3	3	3	3	2
DS24B13.4	3	2	-	-	-	-
DS24B13.5	3	3	3	3	3	2
Average	3	2.8	3	3	3	2

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium) 3:
Substantial (High) No correlation : “-”

Subject Code	Subject Name (Theory Course)	Category	L	T	P	C
DS24B14	Application Architecture and Deployment	PE	2	1	0	3

Objectives:
<ul style="list-style-type: none"> Analyze the architectural differences between monolithic and micro services based applications
<ul style="list-style-type: none"> Design, implement, and secure robust APIs, to facilitate seamless communication and integration between different software components.
<ul style="list-style-type: none"> Gain foundational knowledge of containerization technology
<ul style="list-style-type: none"> Understand the core concepts and architecture of Kubernetes
<ul style="list-style-type: none"> Develop a comprehensive understanding of MLOPs practices from development to deployment

UNIT-I	Monolithic vs Microservices	9
Introduction to Software Architecture and its types - What is Monolithic Architecture and its Importance - Characteristics of Monolithic Architecture - Limitations of Monolithic Architecture - What are Microservices - Working of Microservices - Main Components of Microservices Architecture - Advantages of Microservices - Monolithic vs Microservices - Real World Example of Microservices - Challenges in Microservices.		
UNIT-II	Application Programming Interface	9
What is an API - How do an API Work - WEB APIs - LOCAL APIs - PROGRAM APIs - SOAP, REST API - What are REST APIs - HTTP methods (GET, POST, PUT, DELETE) - Status Codes and URI structure - SOAP vs REST - What is API testing - Types of Testing - Tools for API Testing - Authentication Mechanisms - Authorization Mechanisms - Role Based Access Control (RBAC)		
UNIT-III	Containers - An Introduction	9
What is Virtualization - Virtualization in Cloud Computing - Introduction to containerization - Container Lifecycle - Virtualization vs Containerization - Container Security - Serverless Containers - Introduction to Docker - Docker Architecture - Components of Docker - Concept of Docker Images - Docker Commands - Advantages of Docker - Introduction to Orchestration tools.		
UNIT-IV	Kubernetes - An Introduction	9
What is Kubernetes (K8s) - Why Kubernetes and not only docker - Kubernetes Components - Node - Control Plane - Networking in Kubernetes - Kubernetes Resources - Pod, Deployment, Service, Volume, Namespace, node, cluster - Storage - Security - Monitoring, Logging, Scaling - Writing YAML files.		
UNIT-V	ML Operations	9
Introduction to ML Operations - What is SDLC - Stages of SDLC - Waterfall Model - Agile Model - Iterative Model - Importance of Each Models - Model Training - Model Deployment.		
Total Contact Hours: 45		

Course Outcomes:
Students to understand how architect and AI Application deployment with important aspects to be taken care of.
<ul style="list-style-type: none"> Understand the differences between monolithic and micro services architecture and their respective advantages and disadvantages in AI applications.
<ul style="list-style-type: none"> Learnt to design, implement, and secure robust APIs for effective communication and integration between different software components.
<ul style="list-style-type: none"> Understand the basics of Kubernetes and how it can be used to manage and deploy AI models in a production environment.
<ul style="list-style-type: none"> Understand application programming interfaces (APIs) and their role in integrating AI models into larger systems.
<ul style="list-style-type: none"> Understand MLOps and how it can be used to streamline the machine learning lifecycle, from data preparation to model deployment and monitoring.

SUGGESTED ACTIVITIES

- Problem solving sessions
- Flipped classroom
- Activity Based Learning

SUGGESTED EVALUATION METHODS

- Assignment problems
- Quizzes
- Class Presentation/Discussion

Text Books:

1. Sam Newman , "Building Microservices: Designing Fine-Grained Systems", 1st Edition, O'Reilly, 2015
2. Leonard Richardson, Mike Amundsen, and Sam Ruby , "RESTful Web APIs: Services for a Changing World", GreyScale Indian Edition, O'Reilly Media, 2013
3. Jeff Nickoloff and Stephen Kuenzli, "Docker in Action" , Second Edition, 2019.

Reference Books(s) / Web links:

1. Scott Surovich & Marc Boorshtein, Kubernetes and Docker, Packt Publishing, 2021
2. Mark Treveil, Nicolas Omont & Clément Stenac, Introducing MLOps: How to Scale Machine Learning in the Enterprise, Grayscale Indian Edition , O'Reilly 2020.

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
DS24B14.1	3	2	2	2	-	2
DS24B14.2	2	3	3	3	3	2
DS24B14.3	3	3	3	2	-	2
DS24B14.4	3	3	3	3	3	2
DS24B14.5	3	3	3	3	3	2
Average	2.8	2.8	2.8	2.8	3	2

Correlation levels 1, 2 or 3 are as defined below:
1: Slight (Low) 2: Moderate (Medium) 3:
Substantial (High) No correlation : “-”

Subject Code	Subject Name (Theory Course)	Category	L	T	P	C
DS24B15	Security for Data Engineering	PE	3	0	0	3

Objectives:
• To Understand the Intersection of Cyber Security and Data Science
• To Identify and Mitigate Security Risks in Data Science Projects
• To Apply Data Privacy Regulations and Best Practices
• To Implement Secure Data Handling Practices
• To Utilize Machine Learning for Threat Detection and Security Analysis

UNIT-I	Introduction to Cyber Security and Data Science	9
Overview of Cyber Security and Data Science - Definitions and Concepts - Intersection of Cyber Security and Data Science - Cyber Threat Landscape - Types of Cyber Threats - Attack Vectors and Techniques - Impact of Cyber Attacks on Data Science Processes - Foundations of Data Science - Data Collection and Sources - Data Storage and Management - Data Processing and Analysis Techniques		
UNIT-II	Foundations of Cyber Security	9
Principles of Cyber Security - Confidentiality, Integrity, and Availability (CIA) - Authentication and Authorization - Encryption and Cryptography - Secure Data Handling - Data Classification and Sensitivity - Data Masking and Anonymization - Secure Data Transfer and Sharing - Data Privacy and Compliance - Privacy Regulations (GDPR, HIPAA) - Data Governance and Compliance Frameworks - Ethical Considerations in Data Science and Cyber Security		
UNIT-III	Data Privacy and Protection	9
Data Privacy and Protection -Secure Data Sharing and Transfer - Secure File Transfer Protocols - Secure Data Exchange Platforms - Securing Data Collection Systems - Best Practices for Secure Data Storage - Cloud Security and Data Privacy - Secure Data Transfer and Backup Strategies - Data Retention Policies and Compliance		
UNIT-IV	Threat Detection and Incident Response	9
Threat Detection and Incident Response - Security Information and Event Management (SIEM) - Log Management and Analysis - Real-time Threat Detection - Incident Response Frameworks - Preparation, Identification, Containment, Eradication, Recovery - Forensic Analysis Techniques - Machine Learning for Cyber Security - Threat Prediction and Classification - Behavioural Analysis and User Profiling		
UNIT-V	Advanced Topics in Cyber Security for Data Science	9
Advanced Topics in Cyber Security for Data Science - Adversarial Machine Learning - Evasion Attacks - Defence Mechanisms - Secure Machine Learning Models - Privacy-Preserving Machine Learning - Federated Learning - Ethical and Legal Considerations - Bias and Fairness in Cyber Security - Ethical Hacking and Responsible Disclosure		
Total Contact Hours:45		

Course Outcomes:
1. To Understand the Fundamentals of Cyber Security
2. To Implement Secure Data Handling Practices
3. To Analyze Security Risks in Data Science Projects
4. To Develop Threat Detection and Response Strategies
5. To Design Ethical and Privacy-Preserving Data Science Solutions

SUGGESTED ACTIVITIES

- Problem solving sessions
- Flipped classroom
- Activity Based Learning

SUGGESTED EVALUATION METHODS

- Assignment problems
- Quizzes
- Class Presentation/Discussion

Text Books

1. M Lakshmikanth "Indian Polity", McGraw Hill Education, 5th edition 2017.
2. Durga Das Basu, "Introduction to the Constitution of India ", Lexis Nexis, New Delhi., 21st edition, 2013.

Reference Books(s) / Web links:

1. "Data Science for Cyber-Security" by Nicholas A Heard , Niall M Adams , Patrick Rubin delanchy, Publisher: World Scientific Europe Ltd , 2018
2. "Cryptography and Network Security - Principles and Practice" by William Stallings 7th Edition , Pearson, 2017.
3. "Security Data Science: The Guide to Analyzing Threats and Attacks" by Julian Hillebrand and Heiko Tietze, Publisher: Springer
4. Security, Privacy, and Trust in Modern Data Management (Data-Centric Systems and Applications) by Milan Petkovic, Willem Jonker 2007th Edition, Springer

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
DS24B15.1	3	-	1	-	2	1
DS24B15.2	2	3	2	2	2	2
DS24B15.3	3	1	2	1	2	1
DS24B15.4	-	3	2	2	2	1
DS24B15.5	1	2	3	2	3	2
Average	2.25	2.25	2	1.75	2.75	1.75

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium)

3: Substantial (High) No correlation : "-"

Subject Code	Subject Name (Laboratory Course)	Category	L	T	P	C
GE24221	Professional Soft Skill - II	EEC	0	0	2	1

Objectives:

Students understand day in day out terms used in customer environment and demonstrate customer centric approach and practically experience the and important aspects of it.

- To understand what is spoken without distortion and respond appropriately.
- To participate productively in an official meeting keeping etiquette in mind.
- To communicate effectively through writing.
- To behave appropriately in an official environment.
- To be comfortable to dine with colleagues, clients, and leaders comfortably in a formal or informal setting.

UNIT-I	Accent Neutralization	9
Identifying and dealing with Mother Tongue Influence (MTI) – Pronunciation - Vowel Sounds and Consonant Sounds– Inflection – Pausing - Reducing rate of speech - Volume and tone – Pitch – Clarity - and enunciation.		
UNIT-II	Customer Service	9
Customer Service - Different types of customers - Difference between customer service and customer experience - Telephone Etiquette - Handling difficult customers.		
UNIT-III	Problem Solving and Decision Making	9
Define a Problem - Define Decision Making- Blocks in problem solving - Stereotyping and unconscious biases - The process of Problem Solving and decision-making - Problem Analysis- Decision Analysis - Potential Problem / Opportunity Analysis - Creative Thinking - Problem Solving process - Implementation of the solution.		
UNIT-IV	Business Email Etiquette and Chat	9
Emails Etiquette: Share format/ signature - Emails etiquette - dos and don'ts.		
UNIT-V	Basics of Finance	9
Accounting systems and how transactions are recorded - Financial statements: Profit & Loss account - balance sheet - cash flow statement - Fixed assets - depreciation and the capitalization of software development expense – Working capital and cash management - Using ratio analysis to assess corporate health and performance - Funding the business: equity - debt and other aspects - Budgeting & Forecasting – capex – apex - Designing a flexible budget - Capital expenditure appraisal and approval		
Total Contact Hours: 45		

Course Outcomes:

Students understand day in day out terms used in customer environment and demonstrate customer centric approach and practically experience the and important aspects of it.

1. To understand what is spoken without distortion and respond appropriately.
2. To participate productively in an official meeting keeping etiquette in mind.
3. To communicate effectively through writing.
4. To behave appropriately in an official environment.
5. To be comfortable to dine with colleagues, clients, and leaders comfortably in a formal or informal setting.

SUGGESTED ACTIVITIES

- Problem solving sessions
- Flipped classroom
- Activity Based Learning

SUGGESTED EVALUATION METHODS

- Assignment problems
- Quizzes
- Class Presentation/Discussion

Text Book(s):

1. Lisa Mojsin, "Mastering the American Accent wit online audio ", Second Edition, Barrons Educational Services, 2016
2. Lee Cockerell , "The Customer Rules: The 39 Essential Rules for Delivering Sensational Service", Crown Currency, 2013.
3. Ethan Rasiel and Paul N. Friga, "The McKinsey Mind: Understanding and Implementing the Problem-Solving Tools and Management Techniques of the World's Top Strategic Consulting Firm" , McGraw Hill, 2001

Reference Book(s)/Web link(s)

1. Gene Siciliano, "Finance for Non-Financial Managers" by, Second Edition, McGraw Hill, 2014.
2. Gary Blake and Robert W. Bly , "The Elements of Business Writing: A Guide to Writing Clear, Concise Letters, Mem", Pearson P T R, 2006.

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
GE24221.1	3	2	3	3	-	2
GE24221.2	2	3	3	3	-	2
GE24221.3	2	3	3	3	-	2
GE24221.4	2	3	-	3	2	2
GE24221.5	2	3	3	3	2	2
Average	2.8	2.8	3	3	2	2

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium)

3: Substantial (High) No correlation : "-"

Subject Code	Subject Name (Laboratory Course)	Category	L	T	P	C
DS24221	Generative AI Applications	PC	0	0	4	2

Objectives:
<ul style="list-style-type: none"> Gain a thorough understanding of generative AI concepts.
<ul style="list-style-type: none"> Develop practical skills in implementing and training various generative models
<ul style="list-style-type: none"> Acquire the ability to build, fine-tune, and optimize generative models to produce high-quality outputs
<ul style="list-style-type: none"> Understand and apply various metrics and techniques for evaluating the performance and quality of generative models
<ul style="list-style-type: none"> Engage in collaborative projects that involve building generative AI applications, fostering teamwork, problem-solving, and communication skills.

Description of the Experiments	Total Contact Hours: 6.0
1. Take any large language model (say GPT 3.5) and try to execute some query through it. Create a small program where you can change the parameter values of Temperature, Top P and Max Tokens. Please identify how you can make your answer more deterministic?	
2. Please identify what are the basic metrics to evaluate your large language model response? (As example, toxicity, biasness etc). Please write a short program where you can take model response as input and calculate the score for the above metrics to understand output quality.	
3. Please write a program where you can perform keyword-based search. Please take any text file as input and provide "keyword" dynamically and see whether your algorithm can search it effectively.	
4. Please write a program where you take perform embedding based search. Please take any vector database and use any embedding technique to search the answer of the query from the given input text file where query and text files are the inputs of your program.	
5. Please take 2/3 medical reports (may be blood reports) and store them in a place. Please write a program which can read all the files dynamically from the given locations. Please try to understand the metadata of the reports.	
6. Create a set of questions for which you want to retrieve information from the medical reports through large language models. Save it in some database and keep in the excel file.	
7. Apply large language model and Implement the RAG based approach to search the answer of the queries from the documents where two inputs will be taken: set of medical reports prepared in Experiment 5 and questions prepared in Experiment 6.	
8. Perform the evaluation based on RAG-triad (Context Relevance, Groundedness and Answer Relevance). Show the importance of "context" towards getting the optimized output.	
9. Use Palm 2 (or any other LLM) to perform automation of software development tasks which includes code generation, code debugging and test case generation.	
10. Use any diffusion model to generate images based on given prompt.	
11. Apply zero shot, one shot and few shot prompting and show how performance is improved in few shot prompting.	
12. Apply chain-of-thought (CoT) in prompting and see how output accuracy increases. Do a comparison between normal prompting and CoT based prompting from output performance perspective.	
13. Take a foundation model, create an instruction based fine tuning dataset, apply instruction fine tuning on the base model.	
14. Perform performance evaluation of the model response between foundation model and after fine tuning it.	
15. Explore various task specific benchmark datasets and try to create a new one.	

Course Outcomes:
1. A solid understanding of the principles, architectures, and techniques underlying generative AI models
2. Gained proficiency in training and fine-tuning generative AI models to specific applications and datasets.
3. Ability to design, develop, and deploy generative AI applications across various domains
4. Evaluated the performance of generative AI models using appropriate metrics
5. Integrated generative AI models into larger systems by applying AI solutions to real-world problems and business needs.

SUGGESTED EVALUATION METHODS
<ul style="list-style-type: none"> ● Experiment based viva ● Quizzes ● Mini Project

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
DS24221.1	3	2	-	-	-	2
DS24221.2	3	3	3	3	3	2
DS24221.3	3	3	3	3	3	2
DS24221.4	3	3	3	3	3	2
DS24221.5	3	3	3	3	3	2
Average	3	2.8	3	3	3	2

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium)

3: Substantial (High) No correlation : “-”

Subject Code	Subject Name (Laboratory Course)	Category	L	T	P	C
DS24222	Large Language Models Lab	PC	0	0	4	2

Objectives:
<ul style="list-style-type: none"> Understand the Evolution of Large Language Models
<ul style="list-style-type: none"> Compare different Fine-Tuning approaches of LLM
<ul style="list-style-type: none"> Investigate different quantization techniques for LLMs
<ul style="list-style-type: none"> Explore innovative transformer architectures to enhance training or inference efficiency
<ul style="list-style-type: none"> Explore Ethical and Trustworthy AI Principles

Description of the Experiments	Total Contact Hours: 6.0
<p>Case Study 1: Present your POV on the evolution of Large Language Models. Articulate their growth, architecture changes and application landscape</p>	
<p>Case Study 2: Present your POV on the different fine-tuning methodologies. Articulate the differences, the advantages, and disadvantages of each approach.</p>	
<p>Case Study 3: Present your POV on the constitutional AI, how it's different from RLHF.</p>	
<p>Case Study 4: Present your POV on the Quantization of LLMs, different techniques that are available, performance of the Quantized Models in comparison to the Original Models</p>	
<p>Case Study 5: Present your POV on innovative architectures in transformer model that can lead to savings in training or inference time. As an example, MoE from Mistral is one such unique architecture. Articulate the performance of new architectures compared to the original architectures and come up with some new architecture that can lead to savings</p>	
<p>Case Study 6: Present your POV on the Sustainable AI, Ethical AI, Trustworthy AI</p>	

Course Outcomes:
1. Understood the Historical Context of large language models (LLMs) from early to current state-of-the-art models.
2. Applied fine-tuning techniques to specific use cases and evaluated their effectiveness.
3. Applied quantization models and evaluated the performance trade-offs between quantized models and original models.
4. Investigated innovative architectures like Mixture of Experts (MoE) and other novel approaches that aim to reduce training and inference time.
5. Developed best practices and guidelines for ensuring AI systems are sustainable, ethical, and trustworthy.

SUGGESTED EVALUATION METHODS

- Experiment based viva
- Quizzes
- Mini Project

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
DS24222.1	3	2	-	-	-	2
DS24222.2	3	3	3	3	3	2
DS24222.3	3	3	3	3	3	2
DS24222.4	3	3	3	3	3	2
DS24222.5	3	3	3	3	3	2
Average	3	2.8	3	3	3	2

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium)

3: Substantial (High) No correlation : “-”

Subject Code	Subject Name (Theory Course)	Category	L	T	P	C
AC23211	CONSTITUTION OF INDIA (Common to M.E / M.Tech – Semester II)	MC	3	0	0	0

Objectives:
<ul style="list-style-type: none"> To inculcate the values enshrined in the Indian constitution.
<ul style="list-style-type: none"> To create a sense of responsible and active citizenship
<ul style="list-style-type: none"> To make the students aware of the Constitutional and the Non- Constitutional bodies
<ul style="list-style-type: none"> To help the students understand the relationships exist between union and states
<ul style="list-style-type: none"> To make the students understand the sacrifices made by the freedom fighters.

UNIT-I	INTRODUCTION	9
Historical Background - Constituent Assembly of India - Philosophical foundations of the Indian Constitution - Features - Basic Structure – Preamble.		
UNIT-II	UNION GOVERNMENT - EXECUTIVE, LEGISLATURE AND JUDICIARY	9
Union and its territory - Citizenship - Fundamental Rights - Directive Principles of State Policy (DPSP) - Fundamental Duties. President - Vice President - Prime Minister - Central Council of Ministers - Cabinet Committees - Parliament: Committees, Forums and Groups - Supreme Court.		
UNIT-III	STATE GOVERNMENT & UNION TERRITORIES: STATE GOVERNMENT : EXECUTIVE, LEGISLATURE AND JUDICIARY	9
Governor - Chief Minister - State Council of Ministers - State Legislature - High Court - Subordinate Courts - Panchayati Raj – Municipalities-Union Territories - Scheduled and Tribal Areas.		
UNIT-IV	RELATIONS BETWEEN UNION AND STATES	9
Relations between Union and States - Services under Union and States. Cooperative Societies - Scheduled and Tribal Areas - Finance, Property, Contracts and Suits - Trade and Commerce within Indian Territory – Tribunals.		
UNIT-V	CONSTITUTIONAL BODIES AND AMENDMENTS	9
Introduction to Constitutional & Non-Constitutional Bodies-Elections - Special Provisions relating to certain classes - Languages - Emergency Provisions - Miscellaneous - Amendment of the Constitution - Temporary, Transitional and Special Provisions - Short title, date of commencement, Authoritative text in Hindi and Repeals. Schedules of the Constitution of India - Appendices in the Constitution of India.		
		Total Contact Hours : 45
Course Outcomes :		
On completion of the course, students will be able to		
1. Appreciate the philosophical foundations of the Indian Constitution.		
2. Understand the functions of the Indian government.		
3. Apprehend and abide by the rules of the Indian constitution.		
4. Comprehend the functions of state Government and Local bodies.		
5. Gain Knowledge on constitution functions and role of constitutional bodies and amendments of constitution.		

SUGGESTED ACTIVITIES

- Online Quizzes
- Poster presentations
- Presentations
- Group Discussions
- Case study

Text Books

1. M Lakshmikanth "Indian Polity", McGraw Hill Education, 5th edition 2017.

2. Durga Das Basu, "Introduction to the Constitution of India ", Lexis Nexis, New Delhi., 21st edition, 2013.

Reference Books / Web links

1. Sharma, Brij Kishore, "Introduction to the Constitution of India", Prentice Hall of India, New Delhi, 7th edition, 2015.

2. Subhash Kashyap, "Our Constitution: An Introduction to India's Constitution and Constitutional Law", National Book Trust India, 1994.

3. Mahendra Prasad Singh and Himanshu Roy, "Indian Political System", Pearson India, 4th edition, 2017.

CO-PO MAPPING

PO CO	PO1	PO2	PO3	PO4	PO5	PO6
AC23211.1	3	3	2	2	1	2
AC23211.2	3	-	2	-	-	
AC23211.3	3	2	2	2	1	2
AC23211.4	3	2	-	2	-	-
AC23211.5	3	2	-	2	1	2
Average	3	2	2	2	1	2

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) 2: Moderate (Medium)

3: Substantial (High) No correlation : "-"