LTIMindtree

LTIM Recommendations

Syllabus for MTech (AI & DS)

Document Version / Details: Ver. 1.5 / 12-06-2024

Course Overview

Today's modern world is majorly driven Artificial Intelligence (AI) which has made inroads almost every aspect of life. It was way back in 1959 Arthur Samuel, a pioneer in the field of machine learning (ML) defined it as the "field of study that gives computers the ability to learn without being explicitly programmed". Machine Learning is one of the fundamental blocks of AI. The term "AI" could be attributed to John McCarthy of MIT (Massachusetts Institute of Technology), which Marvin Minsky (Carnegie-Mellon University) defines as "the construction of computer programs that engage in tasks that are currently more satisfactorily performed by human beings because they require high-level mental processes such as: perceptual learning, memory organization and critical reasoning.

AI has come a long way since then. It was during the late 2000 era that we saw the rebirth of neural networks in the form of Deep Learning, thanks to the availability of large compute power through cloud and massive availability of data. Deep Learning is a term to describe deep and complex neural networks that magically solved at ease the problems around structured and unstructured data (vision and language). A few years apart came the Generative Adversarial Networks followed by Large Language Models, now heading towards Large Multimodal Models, Large Action Models. The rise of AI has been exponential to say the least and is likely to keep going. It's no wonder, we see robots cleaning the floor, dev co-pilots generating code across languages for software developers, devices that human language and follow instructions etc., the list is almost endless. Hence its pertinent that we are prepared for the future that embraces AI.

We offer M. Tech DS which is a specially designed two-year post-graduate programme for engineers to leverage the advantages of emerging trends in market in the domain of artificial intelligence and machine learning and get them skilled to do their job as soon as they complete their Postgraduate.

We focus on building concepts on Mathematics and Statistics required for DS domain, programming, applied machine learning, generative ai and its applications, advanced deep learning, fine tuning large language models & building applications and responsible AI. This M. Tech offering is an amalgam of the fundamentals, essentials of DS and the intricate deep practical knowledge of engineering AI & DS applications.

The curriculum and syllabus are designed and developed by LTIMindtree Industry practitioners, chief architects, and experts who have more than 2 decades of experience in this field with rich technical and domain knowledge, and it will be integrated with a real time lab simulation to help students get industry exposure while learning.

The outcome of this course is to shape a student's career in emerging technologies for designing and implementing the DS applications end-to-end.

Total programme consists of theory, theory + integrated lab, exclusive laboratory, soft skills, and projects.

Course Structure

PG program M.Tech (AI & DS) Total: 70 credits

Semester I							
Subject	**Course Code**	**Hours**	**Type**	**L**	**T**	**P**	**Credits**
Mathematical Foundations		60	C	4	0	0	4
Python for Data Science		45	C	3	0	0	3
Applied Machine Learning		45	C	3	0	0	3
Data Engineering		45	C	3	0	0	3
Professional Elective - I		45	C	3	0	0	3
Professional Skills - I		30	P	0	0	2	1
Applied ML Lab		60	P	0	0	4	2
Data Science Lab		60	P	0	0	4	2
Total		390					21
Professional Elective - I: Generative Adversarial Networks (Mandate) / CS62101/ CS60308							

Semester II							
Subject	**Course Code**	**Hours**	**Type**	**L**	**T**	**P**	**Credits**
Deep Learning Applications		45	C	3	0	0	3
Generative AI with Large Language Models		45	C	3	0	0	3
Professional Elective - II		45	C	3	0	0	3
Open Elective		45	C	3	0	0	3
Professional Elective - III		45	E	3	0	0	3
Professional Skills - II		30	P	0	0	2	1
GenAI based Applications Lab		60	P	0	0	4	2
Large Language Models Lab		60	P	0	0	4	2
Total		375					20
Professional Elective - II: Ethics in Data Science (Mandate) / Computer Vision / NLP							
Professional Elective - III: Application Architecture and Deployment / Security for Data Science							

Semester III						
Subject	**Course Code**	**Type**	**L**	**T**	**P (Hours)**	**Credits**
Professional Elective - IV		C	3	0	0	3
Project - I		P	-	-	-	16
Total						19
Professional Elective - IV: Machine Learning Engineering for Production (MLOps) Specialization / Industry Specific Applications of GenAI & Responsible AI						

Semester IV						
Subject	**Course Code**	**Type**	**L**	**T**	**P**	**Credits**
Project -II		P	-	-	-	20
Total						20

C-Core, E-Elective, A- Audit, O- Open elective, P- Practical, L- Lecture, T- Tutorial

Semester - I

Title	Mathematical Foundations	Code	
Prerequisite		Credits | Total Hours	4-0-0 [4] | 60

Course Outcome:
Students will be able to refresh the statistical knowledge learnt earlier with hands-on practical expertise.

CO1: Refresh the mathematics knowledge with respect to Linear algebra, Vectors, Projections, Principal Component Analysis and Generative Models
CO2: Refresh the mathematics knowledge with respect to Matrices, Gradient Calculus, Optimization models.
CO3: Refresh the mathematics knowledge with respect to probability, statistics.

Unit 1: 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 2: 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 3: 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 4: Non-Parametric Statistics 12

Non-parametric Statistics - Chi square test- Sign test -Wilcoxon signed rank test - Mann Whitney test - Run test - Kolmogorov Smirnov test - Spearmann and Kendall's test - Tolerance region.

Unit 5: Multivariate Statistical Methods for Analyzing Complex Datasets 12

Multivariate statistical methods for analysing complex datasets - Factor Analysis - Cluster Analysis- Regression Analysis - Discriminant Analysis.

Reference Books:

1. James D. Miller, Statistics for Data Science - By Packt Publishing (2017)
2. IND James D. Hamilton, Time Series Analysis – By Levant Books (2012)

Title	Python for Data Science	Code	
Prerequisite		Credits |Total Hours	3-0-0 [3] | 45

Course Outcome:

Students will be able to learn about two widely used programming languages in the field of data science and how to go about choosing any language.

CO1: Understand the basics of python and standard modules used for data science with hands-on.
CO2: Understand the data structures and visualization used for data science with hands-on.
CO3: Understand the machine learning libraries used for data science with hands-on.

Unit 1: Python - Data Structures, OOPS & Modules 9

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

Unit 2: Python – Numpy, Pandas & DS Libraries 9

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

Unit 3: Visualization 9

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.

Unit 4: Regression and Classification 9

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 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 - Precison, Recall, F1-Score - ROC Curves - Logistic Regression with Scikit-Learn - Performance Evaluation - Multi-Class Classification with Logistic Regression - Data and EDA – Model.

Unit 5: Unsupervised and Advanced Machine Learning 9

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.

Reference Books:

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)

Title	Applied Machine Learning	Code	
Prerequisite		Credits |Total Hours	3-0-0 [3] | 45

Course Outcome:
Students will be able to get the knowledge about all the tools and techniques you need to apply machine learning to solve business problems.

CO1: To know about Supervised Learning, Support Vector Machines, Unsupervised Learning
CO2: Get the knowledge about Feature Engineering, Statistical Data Analysis, Outlier Analysis and Detection
CO3: Learn about ML Model Development, Model Evaluation Techniques, Model Deployment and Inferences, Model Explainability

Unit 1: Supervised Learning **15**

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 2: Advanced Learning Algos **15**

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 - 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 3: Unsupervised Learning and Recommender Systems **15**

Use unsupervised learning techniques for unsupervised learning: including clustering and anomaly detection - Build recommender systems with a collaborative filtering approach and a content-based deep learning method - Build a deep reinforcement learning model - Implement K-mean clustering - Implement anomaly detection - Learn how to choose between supervised learning or anomaly detection to solve certain tasks - 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.

Reference Books:

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)

Title	Data Engineering	Code	
Prerequisite		Credits |Total Hours	3-0-0 [3] | 45

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

CO1: Identify and explain different data formats and their use cases, including structured, semi-structured, and unstructured data.
CO2: Describe various data ingestion techniques, such as ETL, and stream processing, and their advantages and limitations.
CO3: Perform data profiling and analyze data quality metrics to ensure data accuracy, completeness, and consistency.
CO4: Design and implement effective storage and retrieval methods for large-scale data sets, including relational databases, NoSQL databases, and distributed file systems.
CO5: Apply data engineering principles to real-world scenarios, such as data warehousing, big data analytics, and machine learning.

Unit 1: 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 2: 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 3: 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 4: 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 5: 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.

Reference Books:

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

Title	Generative Adversarial Networks (Professional Elective – I)	**Code**	
Prerequisite		**Credits |Total Hours**	3-0-0 (3) | 45

Course Outcome:
Students gain comprehensive understanding of deep learning techniques and generative AI models.

CO1: Understand generative models such as generative adversarial networks (GANs) and their advanced techniques.
CO2: To build sophisticated and robust GAN models using PyTorch & convolutional layers etc.,
CO3: Student to learn about the advantages and disadvantages of different GAN performance measures.
CO4: Students to explore and examine the applications of GANs

Unit 1: Build Basic Generative Adversarial Networks (GANs) **15**

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 - 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.

Unit 2: Build Better Generative Adversarial Networks (GANs) **15**

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.

Unit 3: Apply Generative Adversarial Networks (GANs) **15**

GANs for Data Augmentation and Privacy Preservation - Explore the applications of GANs and examine them wrt 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 - applications of this framework to modalities beyond images -Implement Pix2Pix - a paired image-to-image translation 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.

Reference Books:

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

Title	Professional skill - I	**Code**	
Prerequisite		**Credits |Total Hours**	0-0-2 [1] | 30

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

CO1: 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.
CO2: Create good Presentation and Present with confidence. Also, recognize and manage Stress, Prioritize and Plan.
CO3: To listen to understand. To be able to ask good questions.
CO4: To understand to be a good Team player, Team Dynamics and to understand the Business Ethics
CO5: To be able to write and speak correctly, forming grammatically correct sentences.

Unit 1: Positive Attitude 6

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

Unit 2: Body Language 6

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

Unit 3: Presentation Skills 6

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

Unit 4: Listening & Questioning skills 6

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

Unit 5: Teamwork 6

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

Applied Machine Learning Lab

Experiment No	Experiment Details	Type
Experiment 1	Understanding "Mobile Price" dataset by doing feature analysis. Data is available at: https://www.kaggle.com/datasets/iabhishekofficial/mobile-price-classification/data	Individual
Experiment 2	Execute data preprocessing step on the above dataset: perform outlier and missing data analysis towards building a refined dataset	Individual
Experiment 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	Individual
Experiment 4	Calculate the prediction accuracy of the models used in Experiment 3 and do comparative analysis among them to identify the best technique.	Individual
Experiment 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	Individual
Experiment 6	Perform data preprocessing step on the above dataset: perform outlier and missing data analysis towards building a refined dataset.	Individual
Experiment 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	Individual
Experiment 8	Calculate the prediction accuracy of the models used in Experiment 7 and do comparative analysis among them to identify the best technique.	Individual
Experiment 9	Plot the features (actual price and predicted price) in scatter plot to understand the variation.	Individual
Experiment 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	Individual
Experiment 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 categorial features.	Individual
Experiment 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.	Individual
Experiment 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	Individual
Experiment 14	Calculate the prediction accuracy of the models used in Experiment 13 and do comparative analysis among them to identify the best technique.	Individual
Experiment 15	Please check whether you find imbalanced classes, overfitting, and data bias in the above two datasets. Please apply some technique to overcome it.	Individual

Data Science Lab

Experiment No	Experiment Details	Type
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.	Group Project
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.	Group Project
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	Group Project
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.	Group Project
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.	Group Project
Case Study 6	Present your POV on GANs used for Deep Fakes. Articulate how we can identify the Deep Fake from the original.	Group Project

Semester - II

Title	Deep Learning Applications	Code	
Prerequisite		Credits |Total Hours	3-0-0 [3] | 45

Course Outcome:
Students learn both theoretical and practical aspects of deep learning applying to build real world applications.

CO1: To understand the fundamentals of deep learning and its applications in computer vision, time series analysis and natural language processing.
CO2: To build a facial recognition system.
CO3: To build a weather forecasting system.
CO4: To build a chatbot.

Unit 1: Building a Facial Recognition System - Part 1 9

Convolutional Neural Network (CNN): Transfer learning - Data Augmentation - Image segmentation - Object detection - Video classification - Text and natural language processing - Structured data - Model optimization

Unit 2: Building a Facial Recognition System - Part 2 9

Facial recognition model: 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 3: Building a Facial Recognition System - Part 3 9

Facial Recognition system: 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 4: Building a Weather Forecasting System with Chatbot - Part 1 9

Recurrent Neural Network: 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 5: Building a Weather Forecasting System with Chatbot - Part 2 9

Intelligent Chatbox: 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.

Reading Materials:

1. James D. Hamilton, Time Series Analysis – By Levant Books (2012)

2. V Kishore Ayyadevara & Yeshwanth Reddy, Modern Computer Vision with PyTorch - By Packt Publishing (2020)

Title	Generative AI with Large Language Models	Code	
Prerequisite		Credits |Total Hours	3-0-0 [3] | 45

Course Outcome:
Students to get the knowledge to adapt pre-trained LLMs to more specialized tasks.

CO1: Understand Fundamentals of Fine Tuning, Types of fine-tuning Techniques.
CO2: Reinforcement learning and LLM-powered applications.

Unit 1: Introduction to Generative AI **15**

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 retrival - Generative configuration - Generative AI project lifecycle - Pre-training large language models - Computational challenges of training LLMs.

Unit 2: Fine Tuning and Evaluation **15**

Instruction fine-tuning - Fine-tuning on a single task - multi-task instruction fine-tuning - Model evaluation – Benchmarks -Parameter efficient fine-tuning (PEFT) -PEFT techniques 1: LoRA - PEFT techniques 2: Soft prompts.

Unit 3: Reinforcement learning and LLM-powered applications **15**

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 - Generative AI Project Lifecycle - 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 architectures.

Reading Material:

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).

Title	Ethics in Data Science（Professional Elective – II）	Code	
Prerequisite		Credits | Total Hours	3-0-0 [3] | 60

Course Outcomes:
Students to get the knowledge of Ethics in Data Science.

CO1: Understand Philosophical frameworks for assessing fairness.

CO2: Get knowledge on Data ownership, privacy and anonymity.

Unit 1: Introduction and Philosophical frameworks for assessing fairness 7

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 2: Research Ethics 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.

Unit 3: Data ownership, privacy, and anonymity 9

Understanding the difference between data ownership - data privacy and data anonymity - under- standing the idea behind data surveillance - data privacy vs. data security.

Unit 4: Algorithmic fairness 9

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 5: Policies on data protection 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 6: Responsible AI 8

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.

Unit 7: Red Teaming on LLM & Case study 9

What are the vulnerabilities - How to attack those problems by Red Teaming.

Title	Natural Language Processing (Professional Elective – II)	**Code**	
Prerequisite		**Credits | Total Hours**	3-0-0 [3] |45

Course Outcome:
Students to understand natural language processing in depth with various factors involved in it.

CO1: Understand the purpose of NLP and how to use it in real world applications with example.
CO2: Understand how to solve a classification problem.
CO3: Understand how deep learning is applied for NLP.
CO4: Understand the transfer learning concepts for reusability of knowledge.
CO5: Understand the applications of voice recognition system.

Unit 1: 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 2: Text Classification 9

Benefits of Text Classification - Types of Text classification - Challenges in text classification - Applications of text classification

Unit 3: Deep Learning for NLP 9

Convolutional Neural Networks (CNNs) for NLP - Recurrent Neural Networks (RNNs) for NLP - Recursive Neural Networks - Hybrid Models for NLP

Unit 4: Transfer Learning for NLP 9

Benefits of Transfer Learning for NLP - Fine Tuning techniques - Fine-Tune BERT for Spam Classification

Unit 5: Voice Recognition 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.

Reading Materials:

1. Sowmya Vajjala, Bodhisattwa Majumder, Anuj Gupta & Harshit Surana, Practical Natural Language Processing – By Shroff/O'Reilly (2020)
2. Uday Kamath, John Liu & James Whitaker, Deep Learning for NLP and Speech Recognition – By Springer 1st ed. (2019)

Title	Computer Vision (Professional Elective – II)	Code	

Prerequisite		Credits |Total Hours	3-0-0 [3] | 45

Course Outcome:
Students to understand the computer vision techniques in depth with various applications of the same.

CO1: Understand what techniques are available to process the image.
CO2: Understand how to analyze the image and extract required features.
CO3: Understand how computer vision solves real world problems.

Unit 1: Image Processing Techniques 9

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.

Unit 2: Feature Extraction and Image Analysis 9

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).

Unit 3: Machine Learning for Computer Vision 9

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).

Unit 4: 3D Computer Vision 9

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).

Unit 5: Advanced CV and Future Trends 9

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.

Reading Materials:

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

Title	Professional skill - II	Code	
Prerequisite		Credits |Total Hours	0-0-2[1] |30

Course Outcome:
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.

CO1: To understand what is spoken without distortion and respond appropriately.
CO2: To behave professionally.
CO3: To participate productively in an official meeting keeping etiquette in mind.
CO4: To communicate effectively through writing.
CO5: To behave appropriately in an official environment.
CO6: To be comfortable to dine with colleagues, clients, and leaders comfortably in a formal or informal setting.

Unit 1: Accent Neutralization 6

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 2: Customer Service 6

Customer Service - Different types of customers - Difference between customer service and customer experience - Telephone Etiquette - Handling difficult customers.

Unit 3: Problem Solving and Decision Making 6

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 4: Business Email Etiquette and Chat 6

Emails Etiquette: Share format/ signature - Emails etiquette - dos and don'ts.

Unit 5: Basics of Finance 6

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

Title	Security for Data Science (Professional Elective – III)	Code	
Prerequisite		Credits | Total Hours	3-0-0 [3] | 45

Course Outcome:
Students will be able to refresh the Fundamentals of Cyber Security

CO1: Understand the Fundamentals of Cyber Security.
CO2: Implement Secure Data Handling Practices
CO3: Analyse Security Risks in Data Science Projects
CO4: Develop Threat Detection and Response Strategies.
CO5: Design Ethical and Privacy-Preserving Data Science Solutions.

Unit 1: 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 2: 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 3: 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 4: 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 5: 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.

Title	Application Architecture & Deployment (Professional Elective – III)	Code	
Prerequisite		Credits | Total Hours	3-0-0 [3] | 45

Course Outcome:
Students to understand how architect and AI Application deployment with important aspects to be taken care of.

CO1: Understand the differences between monolithic and microservices architecture and their respective advantages and disadvantages in AI applications.
CO2: Understand the basics of Kubernetes and how it can be used to manage and deploy AI models in a production environment.
CO3: Understand application programming interfaces (APIs) and their role in integrating AI models into larger systems.
CO4: Understand MLOps and how it can be used to streamline the machine learning lifecycle, from data preparation to model deployment and monitoring.

Unit 1: 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 2: 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 3: 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 4: 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 5: 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.

Reading Materials:

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

Gen AI Based Applications lab

Experiment No	Experiment Details	Type
Experiment 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?	Individual
Experiment 2	Please identify what are the basic metrices 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 metrices to understand output quality.	Individual
Experiment 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.	Individual
Experiment 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.	Individual
Experiment 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.	Individual
Experiment 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.	Individual
Experiment 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.	Individual
Experiment 8	Perform the evaluation based on RAG-triad (Context Relevance, Groundedness and Answer Relevance). Show the importance of "context" towards getting the optimized output.	Individual
Experiment 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.	Individual
Experiment 10	Use any diffusion model to generate images based on given prompt.	Individual
Experiment 11	Apply zero shot, one shot and few shot prompting and show how performance is improved in few shot prompting.	Individual
Experiment 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.	Individual
Experiment 13	Take a foundation model, create an instruction based fine tuning dataset, apply instruction fine tuning on the base model.	Individual
Experiment 14	Perform performance evaluation of the model response between foundation model and after fine tuning it.	Individual
Experiment 15	Explore various task specific benchmark datasets and try to create a new one.	Individual

Large Language Models Lab

Experiment No	Experiment Details	Type
Case Study 1	Present you POV on the evolution of Large Language Models. Articulate their growth, architecture changes and application landscape	Group Project
Case Study 2	Present your POV on the different fine-tuning methodologies. Articulate the differences, the advantages, and disadvantages of each approach.	Group Project
Case Study 3	Present your POV on the constitutional AI, how it's different from RLHF.	Group Project
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	Group Project
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 tne performance of new architectures compared to the original architectures and come up with some new architecture that can lead to savings	Group Project
Case Study 6	Present your POV on the Sustainable AI, Ethical AI, Trustworthy AI	Group Project

LTIMindtree

Let’s get to the future, faster. Together.