MAHATMA GANDHI INSTITUTE OF TECHNOLOGY (Autonomous)

B. Tech. Minor in ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Scheme of Instruction and Examination

Applicable from the Academic Year 2025-26 Batch

(MR-25 Regulation)

S. No.	Year/ Semester	Course Code	Course Title	Instruction Hours per week			Credits
				L	T	P	
1	IV Semester	M154AA	Fundamentals of AI	3	0	0	3
2	IV Semester	M15401	Artificial Intelligence Laboratory	0	0	2	1
3	V Semester	M155AA	Machine Learning	3	0	0	3
4	V Semester	M15501	Machine Learning Lab	0	0	2	1
3	VI Semester	M156AA M156AB M156MA M156MB	Prompt Engineering / Deep Learning / MOOCS (Prompt Engineering) / MOOCS (Deep Learning)	3	0	0	3
6	VII Semester	 M157AA M157AB M157AC M157AD M157AE	Elective: Any **ONE** of the following subjects Natural Language Processing Generative AI Data Stream Mining Reinforcement learning Predictive Analytics	3	0	0	3
7	VII Semester	M15701	Project/ Experiential Learning	0	0	8	4
Total Credits							**18**

M154AA: FUNDAMENTALS OF AI

B.Tech. IV Semester.

L	T	P	C
3	0	0	3

Course Objectives:

1. To learn the difference between optimal reasoning Vs human like reasoning
2. To understand the notions of state space representation, exhaustive search, heuristic search along with the time and space complexities
3. To learn different knowledge representation techniques
4. To understand the applications of AI namely, Game Playing, Theorem Proving, Expert Systems, Machine Learning and Natural Language Processing
5. To learn the AI application in computer vision in Robotics

Course Outcomes:

CO1. Gain the knowledge of what is AI, risks and benefits of AI, limits of AI and the ethics involved in building an AI application.
CO2. Understand the nature of environments and the structure of agents.
CO3. Possess the ability to select a search algorithm for a problem and characterize its time and space complexities.
CO4. Possess the skill for representing knowledge using the appropriate technique
CO5. Gain an understanding of the applications of AI

Unit – I

Foundations of AI: Introduction to AI, History of AI, Strong and Weak AI, The State of the Art, Risks and Benefits of AI

Philosophy, Ethics and Safety of AI: The Limits of AI, Machine thinking capability, The Ethics of AI

Intelligent Agents: Agents and Environments, Good Behavior: The Concept of Rationality, The Nature of Environments, The Structure of Agents.

Unit – II

Solving Problems by Searching: Problem – Solving Agents

Uninformed Search Strategies: Best-First Search, Breadth-First Search, Uniform-Cost Search, Depth-First Search, Iterative Deepening Search and Bidirectional Search

Informed Search Strategies: Greedy Best-First Search, A* Search

Unit – III

Logical Agents: Knowledge-based agents, Propositional Logic, Propositional Theorem Proving

First-Order Logic: Syntax and Semantics of First-Order Logic

Inference in First-Order Logic: Propositional Vs. First-Order Inference, Unification and First-Order Inference, Forward Chaining, Backward Chaining

Knowledge Representation: Ontological Engineering, Categories and Objects, Events

Unit – IV

Quantifying Uncertainty: Basic Probability Notation, Inference Using Full-Joint Distributions, Independence, Bayes' Rule and its Use, Naive Bayes Models

Probabilistic Reasoning: Representing Knowledge in an Uncertain Domain, The semantics of Bayesian Networks, Exact Inference in Bayesian Networks

Unit – V

Learning from Examples: Forms of Learning, Supervised Learning, Learning Decision Trees, Model Selection, Linear Regression and Classification, Ensemble Learning

Natural Language Processing: Language Models, Grammar, Parsing, Complications of Real Natural Language, Natural Language Tasks

Robotics: Robots, Robot Hardware, Kind of Problems solved, Application Domains

Computer Vision: Simple Image Features, Using Computer Vision

TEXT BOOK:

1. "Artificial Intelligence a Modern Approach", Fourth Edition, Stuart J. Russell & Peter Norvig – Pearson.

REFERENCE BOOKS:

1. "Artificial Intelligence", Elaine Rich, Kevin Knight & Shivashankar B Nair McGraw Hill Education.
2. Artificial Intelligence, 3rd Edn, E. Rich and K.Knight (TMH)
3. Artificial Intelligence, 3rd Edn., Patrick Henny Winston, Pearson Education.
4. Artificial Intelligence, Shivani Goel, Pearson Education.
5. Artificial Intelligence and Expert systems – Patterson, Pearson Education

M15401: ARTIFICIAL INTELLIGENCE LAB

B.Tech. IV Semester.

L	T	P	C
0	0	2	1

Course Objectives:

1. To understand the fundamental concepts of Artificial Intelligence and problem-solving techniques using Python programming.
2. To understand and implement search algorithms such as Breadth First Search (BFS) and Depth First Search (DFS).
3. To develop programs for classical AI problems including Water Jug, 8-Puzzle, Tower of Hanoi, and Monkey Banana problems.
4. To apply game-playing and decision-making techniques using Tic-Tac-Toe and Alpha-Beta Pruning algorithms.
5. To analyze and solve optimization and constraint satisfaction problems such as Travelling Salesman Problem (TSP) and 8-Queens Problem.

Course Outcomes:

CO1. Design and implement AI-based programs using Python for various search and problem-solving techniques.

CO2. Apply BFS and DFS algorithms to explore and solve state-space search problems effectively.

CO3. Develop solutions for classical AI problems such as 8-Puzzle, Water Jug, Tower of Hanoi, and Monkey Banana problems.

CO4. Implement game-playing strategies and decision-making algorithms using Tic-Tac-Toe and Alpha-Beta Pruning.

CO5. Analyze optimization and constraint satisfaction problems and develop suitable AI solutions for TSP and 8-Queens Problem.

LIST OF EXPERIMENTS:

Write a Program to Implement the following using Python.

1. Breadth First Search
2. Depth First Search
3. Tic-Tac-Toe game
4. 8-Puzzle problem
5. Water-Jug problem
6. Travelling Salesman Problem
7. Tower of Hanoi
8. Monkey Banana Problem
9. Alpha-Beta Pruning
10. 8-Queens Problem

TEXT BOOK:

1. Artificial Intelligence A Modern Approach, Third Edition, Stuart Russell and Peter Norvig, Pearson Education.

REFERENCE BOOKS:

1. Artificial Intelligence, 3rd Edn, E. Rich and K. Knight (TMH)
2. Artificial Intelligence, 3rd Edn., Patrick Henny Winston, Pearson Education.
3. Artificial Intelligence, Shivani Goel, Pearson Education

M155AA: MACHINE LEARNING

B.Tech. V Semester.

L	T	P	C
3	0	0	3

Course objectives:

1. To introduce students to the basic concepts and techniques of Machine Learning.
2. To have a thorough understanding of the Supervised and Unsupervised learning techniques.
3. To study the various probability-based learning techniques.
4. To analyze classification, regression, and clustering techniques.
5. To understand advanced Machine Learning concepts Markov Models, Bayesian

Course Outcomes:

CO1. Distinguish between, supervised, unsupervised and semi-supervised learning

CO2. Understand algorithms for building classifiers applied on datasets of non-linearly separable classes

CO3. Understand the principles of evolutionary computing algorithms

CO4. Design an ensemble to increase the classification accuracy

CO5. Analyze and develop intelligent systems using Reinforcement Learning, Bayesian Networks, Markov Models, and probabilistic learning approaches

UNIT - I

Learning – Types of Machine Learning – Supervised Learning – The Brain and the Neuron – Design a Learning System – Perspectives and Issues in Machine Learning – Concept Learning Task – Concept Learning as Search – Finding a Maximally Specific Hypothesis – Version Spaces and the Candidate Elimination Algorithm – Linear Discriminants: – Perceptron – Linear Separability – Linear Regression.

UNIT - II

Multi-layer Perceptron– Going Forwards – Going Backwards: Back Propagation Error – Multi-layer Perceptron in Practice – Examples of using the MLP – Overview – Deriving Back-Propagation – Radial Basis Functions and Splines – Concepts – RBF Network – Curse of Dimensionality – Interpolations and Basis Functions – Support Vector Machines.

UNIT - III

Learning with Trees – Decision Trees – Constructing Decision Trees – Classification and Regression Trees – Ensemble Learning – Boosting – Bagging – Different ways to Combine Classifiers – Basic Statistics – Gaussian Mixture Models – Nearest Neighbor Methods – Unsupervised Learning – K means Algorithms.

UNIT - IV

Dimensionality Reduction – Linear Discriminant Analysis – Principal Component Analysis – Factor Analysis – Independent Component Analysis – Locally Linear Embedding – Isomap – Least Squares Optimization

Evolutionary Learning – Genetic algorithms – Genetic Offspring: - Genetic Operators – Using Genetic Algorithms

UNIT - V

Reinforcement Learning – Overview – Getting Lost Example

Markov Chain Monte Carlo Methods – Sampling – Proposal Distribution – Markov Chain Monte Carlo – Graphical Models – Bayesian Networks – Markov Random Fields – Hidden Markov Models – Tracking Methods

TEXT BOOK:

1. Stephen Marsland, ―Machine Learning – An Algorithmic Perspective, Second Edition, Chapman and Hall/CRC Machine Learning and Pattern Recognition Series, 2014.

REFERENCE BOOKS:

1. Tom M Mitchell, Machine Learning, First Edition, McGrawHill Education, 2013.
2. Peter Flach, ―Machine Learning: The Art and Science of Algorithms that Make Sense of Data, First Edition, Cambridge University Press, 2012.
3. Jason Bell, ―Machine learning – Hands on for Developers and Technical Professionals, First Edition, Wiley, 2014
4. Ethem Alpaydin, ―Introduction to Machine Learning 3e (Adaptive Computation and Machine Learning Series), Third Edition, MIT Press, 2014

M15501: MACHINE LEARNING LAB

B.Tech. V Semester.

L	T	P	C
0	0	2	1

Course Objectives:

1. The objective of this lab is to get an overview of the various machine learning Techniques and can demonstrate them using python.
2. To gain practical knowledge of Python libraries such as Statistics, Math, NumPy, SciPy, Pandas, and Matplotlib for data analysis and visualization.
3. To implement supervised learning techniques such as Linear Regression, Decision Trees, K-Nearest Neighbours, and Logistic Regression using Sklenar.
4. To develop skills in unsupervised learning techniques such as K-Means Clustering for data grouping and pattern discovery.
5. To analyze and evaluate the performance of different Machine Learning classification algorithms using real-world datasets.

Course Outcomes:

CO1. Understand modern notions in predictive data analysis
CO2. Select data, model selection, model complexity and identify the trends
CO3. Understand a range of machine learning algorithms along with their strengths and weaknesses
CO4. Build predictive models from data and analyze their performance
CO5. Evaluate and compare the performance of Machine Learning algorithms on datasets and develop mini projects for real-world problem solving

List of Experiments:

1. Write a python program to compute Central Tendency Measures: Mean, Median Mode , Measure of Dispersion: Variance, Standard Deviation
2. Study of Python Basic Libraries such as Statistics, Math, Numpy and Scipy
3. Study of Python Libraries for ML application such as Pandas and Matplotlib
4. Write a Python program to implement Simple Linear Regression
5. Implementation of Multiple Linear Regression for House Price Prediction using sklearn
6. Implementation of Decision tree using sklearn and its parameter tuning
7. Implementation of KNN using sklearn
8. Implementation of Logistic Regression using sklearn
9. Implementation of K-Means Clustering
10. Performance analysis of Classification Algorithms on a specific dataset (Mini Project)

TEXT BOOK:

1. Machine Learning – Tom M. Mitchell, - MGH

REFERENCE BOOK:

1. Machine Learning: An Algorithmic Perspective, Stephen Marshland, Taylor & Francis

M156AA: PROMPT ENGINEERING

B.Tech. VI Semester.

L	T	P	C
3	0	0	3

Course Objectives:

1. To introduce the principles and techniques of effective prompt engineering for generative AI models.
2. To understand the architecture, capabilities, and evolution of large language models such as GPT-3.5, GPT-4, Gemini, and LLaMA.
3. To explore standard practices in structured and unstructured text generation using tools like ChatGPT.
4. To apply chunking, tokenization, and formatting techniques for improving text generation and manipulation.
5. To understand the role of embeddings, vector databases (FAISS, Pinecone), and Retrieval-Augmented Generation (RAG) in modern NLP systems.

Course Outcomes: After completion of the course, the student should be able to

CO1. Explain and apply the core principles of prompt engineering for guiding generative AI outputs effectively.
CO2. Describe the underlying architecture and functionality of state-of-the-art large language models (LLMs).
CO3. Generate and manipulate structured outputs (JSON, YAML, CSV) using ChatGPT with advanced prompting techniques.
CO4. Implement text chunking, tokenization, and format control using tools like SpaCy, Tiktoken,and Python.
CO5. Utilize vector databases such as FAISS and Pinecone in Retrieval-Augmented Generation (RAG) pipelines for efficient information retrieval.

UNIT – I:

Fundamentals and Principles of Prompting

Overview of the Five Principles of Prompting: Give Direction, Specify Format, Provide Examples, Evaluate Quality, Divide Labor.

UNIT – II:

Introduction to Large Language Models for Text Generation

What are Text Generation Models, Vector Representations: The Numerical Essence of Language, Transformer Architecture: Orchestrating Contextual Relationships, Probabilistic Text Generation: The Decision Mechanism, Historical Underpinnings: The Rise of Transformer Architectures, OpenAI's Generative Pretrained Transformers, GPT-3.5-turboand ChatGPT, GPT-4, Google's Gemini, Meta's Llama and Open Source.

UNIT – III:

Standard Practices for Text Generation with ChatGPT- Part-A

Generating Lists, Hierarchical List Generation, When to Avoid Using Regular Expressions, Generating JSON, YAMLFiltering YAML Payloads, Handling Invalid Payloads in YAML, Diverse Format Generation with ChatGPT, Mock CSV Data, Universal Translation Through LLMs, Ask for Context, Text Style Unbundling, Identifying the Desired Textual Features, Generating New Content with the Extracted Features, Extracting Specific Textual Features with LLMs.

UNIT – IV:

Standard Practices for Text Generation with ChatGPT- Part-B

Chunking Text, Benefits of Chunking Text, Scenarios for Chunking Text, Poor Chunking Example, Chunking Strategies, Sentence Detection Using SpaCy, building a Simple Chunking Algorithm in Python, Sliding Window Chunking, Text Chunking Packages, Text Chunking with Tiktoken, Encodings, Understanding the Tokenization of Strings.

UNIT – V:

Vector Databases with FAISS and Pinecone

Retrieval Augmented Generation (RAG), Introducing Embeddings, Document Loading Memory Retrieval with FAISS, RAG with LangChain, Hosted Vector Databases with Pinecone, Self- Querying, Alternative Retrieval Mechanisms.

TEXTBOOK:

1. Phoenix J, Taylor M. Prompt engineering for generative AI. " O'Reilly Media, Inc."; 2024 May 16.

REFERENCE:

1. Tunstall L, Von Werra L, Wolf T. Natural language processing with transformers. " O'Reilly Media, Inc."; 2022

M156AB: DEEP LEARNING

B.Tech. VI Semester.

L	T	P	C
3	0	0	3

Course Objectives:

1. To understand deep Learning algorithms and their applications in real-world data.
2. To understand the importance of regularization techniques in improving the generalization performance of deep learning models.
3. To understand the fundamentals of Convolutional Neural Networks (CNNs) and the convolution operation used in deep learning.
4. To understand the Neural Network Models in Deep Learning.
5. To understand Deep learning Applications.

Course Outcomes:

CO1. Understand machine learning basics and neural networks
CO2. Understand optimal usage of data for training deep models
CO3. Apply CNN and RNN models for real-world data
CO4. Evaluate deep models
CO5. Develop deep models for real-world problems

UNIT -I

Machine Learning Basics

Learning Algorithms, Capacity, Overfitting and Underfitting, Hyperparameters and Validation Sets, Estimators, Bias and Variance, Maximum Likelihood Estimation, Bayesian Statistics, Supervised Learning Algorithms, Unsupervised Learning Algorithms, Stochastic Gradient Descent, Building a Machine Learning Algorithm, Challenges Motivating Deep Learning **Deep Feedforward Networks** Learning XOR, Gradient-Based Learning, Hidden Units, Architecture Design, Back-Propagation and Other Differentiation Algorithms

UNIT -II

Regularization for Deep Learning

Parameter Norm Penalties, Norm Penalties as Constrained Optimization, Regularization and Under- Constrained Problems, Dataset Augmentation, Noise Robustness, Semi-Supervised Learning, Multi- Task Learning, Early Stopping, Parameter Tying and Parameter Sharing, Sparse Representations, Bagging and Other Ensemble Methods, Dropout, Adversarial Training, Tangent Distance, Tangent Prop, and Manifold Tangent Classifier, Optimization for Training Deep Models, Learning vs Pure Optimization, Challenges in Neural Network Optimization, Basic Algorithms, Parameter Initialization Strategies, Algorithms with Adaptive Learning Rates

UNIT-III

Convolutional Networks

The Convolution Operation, Motivation, Pooling, Convolution and Pooling as an Infinitely Strong Prior, Variants of the Basic Convolution Function, Structured Outputs, Data Types, Efficient Convolution Algorithms, Random or Unsupervised Features

UNIT -IV

Recurrent and Recursive Nets

Unfolding Computational Graphs, Recurrent Neural Networks, Bidirectional RNNs, Encoder-Decoder Sequence-to-Sequence Architectures, Deep Recurrent Networks, Recursive Neural Networks, The Challenge of Long-Term Dependencies, Echo State Networks, Leaky Units and Other Strategies for Multiple Time Scales, The Long Short-Term Memory and Other Gated RNNs, Optimization for Long- Term Dependencies, Explicit Memory

UNIT -V

Practical Methodology: Performance Metrics, Default Baseline Models, Determining Whether to Gather More Data, Selecting Hyperparameters, Debugging Strategies, Example: Multi-Digit Number Recognition

Applications: Large-Scale Deep Learning, Computer Vision, Speech Recognition, Natural Language Processing, Other Applications.

TEXT BOOK:

1. Deep Learning by Ian Goodfellow, Yoshua Bengio and Aaron Courville, MIT Press.

REFERENCES:

1. The Elements of Statistical Learning.Hastie, R.Tibshirani, & J.Friedman, Springer.
2. Probabilistic Graphical Models. Koller, and N. Friedman, MIT Press.
3. Bishop, C.M, Pattern Recognition and Machine Learning, Springer, 2006.
4. Yegnanarayana, B., Artificial Neural Networks PHI Learning Pvt. Ltd, 2009.
5. Golub, G.H., and Van Loan, C., F., Matrix Computations, JHU Press, 2013.

M157AA: NATURAL LANGUAGE PROCESSING

(Minors Elective)

B.Tech. VII Semester.

L	T	P	C
3	0	0	3

Prerequisites: Data structures and compiler design

Course Objectives:

1. To understand the structure of words and documents using morphological analysis and document processing techniques in Natural Language Processing.
2. To study syntactic structure representation and parsing algorithms for analyzing natural language syntax.
3. To analyze ambiguity resolution methods in parsing and understand semantic interpretation and word sense analysis.
4. To explore semantic parsing concepts including predicate-argument structures and meaning representation systems.
5. To understand language modeling techniques including N-Gram models, Bayesian estimation, and multilingual language models for NLP applications

Course Outcomes:

CO1. Show sensitivity to linguistic phenomena and an ability to model them with formal grammars.

CO2. Understand and carry out proper experimental methodology for training and evaluating empirical NLP systems

CO3. Manipulate probabilities, construct statistical models over strings and trees, and estimate parameters using supervised and unsupervised training methods.

CO4. Design, implement, and analyze NLP algorithms; and design different language modelling Techniques.

CO5. Ability to analyze, evaluate, and apply language modelling techniques such as N-Gram models

UNIT - I

Finding the Structure of Words: Words and Their Components, Issues and Challenges, Morphological Models

Finding the Structure of Documents: Introduction, Methods, Complexity of the Approaches, Performances of the Approaches, Features

UNIT - II

Syntax I: Parsing Natural Language, Treebanks: A Data-Driven Approach to Syntax, Representation of Syntactic Structure, Parsing Algorithms

UNIT – III

Syntax II: Models for Ambiguity Resolution in Parsing, Multilingual Issues

Semantic Parsing I: Introduction, Semantic Interpretation, System Paradigms, Word Sense

UNIT - IV

Semantic Parsing II: Predicate-Argument Structure, Meaning Representation Systems

UNIT - V

Language Modelling: Introduction, N-Gram Models, Language Model Evaluation, Bayesian parameter estimation, Language Model Adaptation, Language Models- class based, variable length, Bayesian topic based, Multilingual and Cross Lingual Language Modelling

TEXT BOOKS:

1. Multilingual natural Language Processing Applications: From Theory to Practice Daniel M. Bikel and Imed Zitouni, Pearson Publication

REFERENCES:

1. Speech and Natural Language Processing - Daniel Jurafsky& James H Martin, Pearson Publications
2. Natural Language Processing and Information Retrieval: Tanvi Siddiqui, U.S. Tiwary

M157AB: GENERATIVE AI

(Minors Elective)

B.Tech. VII Semester.

L	T	P	C
3	0	0	3

Course Objectives:

1. To introduce the foundations, evolution, and core concepts of AI, ML, DL, NLP, and Generative AI.
2. To develop understanding of advanced neural architectures and generative models such as GANs, VAEs, and Transformers.
3. To explore Large Language Models, prompt engineering, and their real-world applications.
4. To familiarize learners with frameworks, multimodal applications, and ethical considerations in Generative AI.
5. To understand Generative AI, retrieval techniques.

Course Outcomes:

CO1. Demonstrate knowledge of AI foundations, generative models, and advanced neural architectures.

CO2. Apply generative AI techniques to create solutions for text, image, video, and multimodal tasks.

CO3. Design, fine-tune, and optimize Large Language Models for specific applications.

CO4. Evaluate ethical, social, and legal implications of Generative AI deployments and propose mitigation strategies.

CO5. Ability to develop multimodal Generative AI applications

UNIT I

Foundations of AI and Generative Models

Introduction and historical evolution to Artificial Intelligence (AI), Machine Learning (ML), Natural Language Processing (NLP) and Deep Learning (DL), Structure of Artificial Neural Networks (ANNs), Mathematical and computational foundations of generative modelling, Overview of generative models and their applications across various domains; Importance of Generative AI in modern applications, Transfer learning and in advancing Generative AI

UNIT II

Advanced Neural Architectures for Generative AI

Variational Autoencoders (VAEs): principles and applications, Generative Adversarial Networks (GANs): architecture and working principles; Transformer architecture and attention mechanisms (in detail); Long Short-Term Memory Networks (LSTMs) and the limitations of traditional RNNs/LSTMs, Advanced Transformer architectures and techniques, Pre - training and transfer learning strategies for generative models

UNIT III

Large Language Models and Prompt Engineering

Overview of Large Language Models (LLMs), GPT architecture, variants, and working principles, Pré- training and fine-tuning GPT models for applications (e.g., chatbots, text generation), Case study: GPT- based customer support chatbot, BERT architecture, pre-training objectives, and fine-tuning, Prompt Engineering: Designing effective prompts, controlling model behaviour, and improving output quality, Fine-tuning language models for creative writing and chatbot development

UNIT IV

Multi-Agent Systems and Generative AI Applications

Introduction to Multi-Agent Systems (MAS), Types of agents: reactive, deliberative, hybrid, and learning agents, Multi - agent collaboration and orchestration for generative tasks, Use cases: autonomous research assistants, cooperative creative generation, distributed problem-solving, Frameworks and tools: AutoGen, CrewAI, Hugging GPT for LLM-powered multi-agent systems, Generative AI

Applications: Art, Creativity, Image/Video generation, Music composition, Healthcare, Finance, Real- world case studies and deployment challenges

UNIT V

Frameworks, Multimodal Applications, and Ethics

LangChain framework: components and LLM application development, Retrieval-Augmented Generation (RAG), Embeddings, Indexing networks, and Vector databases, Generative AI across modalities: Text, Code, Image, and Video generation, Image and Video generation using GANs and VAEs, Multimodal Generative AI: integration and training strategies, Ethical considerations: bias, fairness, trust, and responsible AI deployment, Social and legal implications of Generative AI, Risk mitigation strategies and real-world ethical case studies

TEXT BOOKS:

1. Altaf Rehmani, Generative AI for Everyone: Understanding the Essentials and Applications of This Breakthrough Technology.
2. Charu C. Aggarwal, Neural Networks and Deep Learning: A Textbook. Joseph Babcock, Raghav Bali, Generative AI with Python and TensorFlow 2, 2024.

REFERENCE BOOKS:

1. Josh Kalin, Generative Adversarial Networks Cookbook.
2. Jesse Sprinter, Generative AI in Software Development: Beyond the Limitations of Traditional Coding, 2024.

ONLINE REFERENCES:

1. Fabian Gloeckle et al., Better & Faster Large Language Models via Multi-token Prediction, arXiv:2404.19737v1, 2024.Vaswani et al., Attention Is All You Need, NeurIPS 2017.

M157AC: DATA STREAM MINING

(Minors Elective)

B.Tech. VII Semester.

L	T	P	C
3	0	0	3

Prerequisites: A basic knowledge of "Data Mining"

Course Objectives:

1. To understand the fundamentals of stream mining
2. To understand Hoeffding Trees and Hoeffding Bound concepts
3. To understand prediction strategies, Naïve Bayes leaves, adaptive hybrids, and ensemble methods in data stream mining.
4. To understand evolving data streams, change detection mechanisms, adaptive stream mining, and sliding window techniques.
5. To understand ADWIN-based stream mining techniques

Course Outcomes:

CO1. Understand how to formulate a knowledge extraction problem from data streams.
CO2. Ability to apply methods/algorithms to new data stream analysis problems.
CO3. Evaluate the results and understand the functioning of the methods studied.
CO4. Demonstrate decision tree and adaptive Hoeffding Tree concepts
CO5. Ability to develop real-time stream mining applications.

UNIT- I

MOA Stream Mining, Assumptions, Requirements, Mining Strategies, Change Detection Strategies, MOA Experimental Settings, Previous Evaluation Practices, Evaluation Procedures for Data Streams, Testing Framework, Environments, Data Sources, Generation Speed and Data Size, Evolving Stream Experimental Setting.

UNIT-II

Hoeffding Trees, The Hoeffding Bound for Tree Induction, The Basic Algorithm, Memory Management, Numeric Attributes, Batch Setting Approaches, Data Stream Approaches.

UNIT-III

Prediction Strategies, Majority Class, Naïve Bayes Leaves, Adaptive Hybrid, Hoeffding Tree Ensembles, Data Stream Setting, Realistic Ensemble Sizes.

UNIT-IV

Evolving Data Streams, Algorithms for Mining with Change, A Methodology for Adaptive Stream Mining, Optimal Change Detector and Predictor, Adaptive Sliding Windows, Introduction, Maintaining Updated Windows of Varying Length.

UNIT-V

Adaptive Hoeffding Trees, Introduction, Decision Trees on Sliding Windows, Hoeffding Adaptive Trees, Adaptive Ensemble Methods, New methods of Bagging using trees of different size, New method of bagging using ADWIN, Adaptive Hoeffding Option Trees, Method performance.

TEXT BOOK:

1. DATA STREAM MINING: A Practical Approach by Albert Bifet and Richard Kirkby.

REFERENCE BOOKS:

1. Knowledge discovery from data streams by Gama João. ISBN: 978-1-4398-2611-9
2. Machine Learning for Data Streams by Albert Bifet, Ricard Gavalda; MIT Press, 2017.

M157AD: REINFORCEMENT LEARNING
(Minors Elective)

B.Tech. VII Semester.

L	T	P	C
3	0	0	3

Course Objectives:

1. To understand the fundamentals of probability.
2. To study Markov Decision Processes and reward model.
3. To learn Reinforcement Learning concepts including prediction.
4. To understand temporal difference learning and model-free control algorithms.
5. To Understand advanced reinforcement learning techniques.

Course Outcomes:

CO1. Apply stochastic bandit algorithms and regret minimization techniques for decision-making problems.
CO2. Understand basics of RL.
CO3. Understand RL Framework and Markov Decision Process.
CO4. Analyzing Ning through the use of Dynamic Programming and Monte Carlo.
CO5. Understand TD (0) algorithm, TD(λ) algorithm.

Unit I

Basics of probability and linear algebra, Definition of a stochastic multi-armed bandit, Definition of regret, Achieving sublinear regret, UCB algorithm, KL-UCB, Thompson Sampling.

Unit II

Markov Decision Problem, policy, and value function, Reward models (infinite discounted, total, finite horizon, and average), Episodic & continuing tasks, Bellman's optimality operator, and Value iteration
& policy iteration

Unit III

The Reinforcement Learning problem, prediction and control problems, Model-based algorithm, Monte Carlo methods for prediction, and Online implementation of Monte Carlo policy evaluation

Unit IV

Bootstrapping; TD (0) algorithm; Convergence of Monte Carlo and batch TD(0) algorithms; Model-free control: Q-learning, Sarsa, Expected Sarsa.

Unit V

n-step returns; TD(λ) algorithm; Need for generalization in practice; Linear function approximation and geometric view; Linear TD(λ). Tile coding; Control with function approximation; Policy search; Policy gradient methods; Experience replay; Fitted Q Iteration; Case studies.

TEXTBOOKS:

1. "Reinforcement learning: An introduction," First Edition, Sutton, Richard S., and Andrew G. Barto, MIT press 2020.
2. "Statistical reinforcement learning: modern machine learning approaches," First Edition, Sugiyama, Masashi. CRC Press 2015.

REFERENCES:

1. "Bandit algorithms," First Edition, Lattimore, T. and C. Szepesvári. Cambridge University Press, 2020.
2. "Reinforcement Learning Algorithms: Analysis and Applications," Boris Belousov, Hany Abdulsamad, Pascal Klink, Simone Parisi, and Jan Peters First Edition, Springer 2021.
3. Alexander Zai and Brandon Brown "Deep Reinforcement Learning in Action," First Edition, Manning Publications 2020.

M157AE: PREDICTIVE ANALYTICS
(Minors Elective)

B.Tech. VII Semester.

L	T	P	C
3	0	0	3

Prerequisite: Data Science/ Data analytics

Course Objectives:

1. To understand the fundamentals, applications, and problem formulation techniques in predictive analytics.
2. To study prediction deployment strategies.
3. To learn machine learning techniques and predictive models such as decision trees, logistic regression.
4. To understand ensemble learning methods including bagging, boosting, random forests
5. To understand Text Mining.

Course Outcomes:

CO1. Understand the processing steps for predictive analytics.
CO2. Construct and deploy prediction models with integrity.
CO3. Explore various techniques (machine learning/data mining, ensemble) for predictive analytics.
CO4. Apply predictive analytics to real world examples.
CO5. To Implement Ensemble effect.

UNIT I

Introduction – types of analytics, applications of predictive analytics, overview of predictive analytics. Setting up the problem - processing steps, business understanding, objectives, data for predictive modelling, columns as measures, target variables, measures of success for predictive models.

UNIT II

Prediction effect, deployment of prediction model, ethics and responsibilities The Data effect

UNIT III

Machine Learning for prediction

Predictive modelling – decision trees, logistic regression, neural network, KNN, Bayesian method.

Regression model

Assessing Predictive Models - Batch Approach to Model Assessment, Percent Correct Classification, Rank-Ordered Approach to Model Assessment, Assessing Regression Models

UNIT IV

Ensemble effect

Model ensembles – motivation, wisdom of crowds, Bagging, Boosting, Random forests, stochastic gradient boosting, heterogeneous ensembles.

UNIT V

Case studies: Survey analysis, question answering– challenges in text mining, persuasion by the numbers

TEXT BOOKS:

1. Eric Siegel, Predictive Analytics: The Power to predict who will Click, buy, lie, or die, John Wiley & Sons, 2013.
2. Dean Abbott, Applied Predictive Analytics - Principles and Techniques for the Professional Data Analyst, 2014.

REFERENCES:

1. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning-Data Mining, Inference, and Prediction, Second Edition, Springer Verlag,2009.
2. G. James, D. Witten, T. Hastie, R. Tibshirani- An introduction to statistical learning with Applications in R, Springer, 2013
3. E. Alpaydin, Introduction to Machine Learning, Prentice Hall of India, 2010