SHALAKA FOUNDATION'S

Keystone School of Engineering

Pune, Maharashtra, India

An Autonomous Institute, Affiliated to SPPU, Pune

Faculty of Science and Technology

National Education Policy (NEP)- 2020 Compliant Curriculum



Master of Technology

Department of Artificial Intelligence and Data Science

2026 Pattern

(With effect from Academic Year 2026-27)

Program Outcomes (POs)

PO1	Ability to independently carry out research/investigation and development work to solve practical problems.
PO2	Ability to write and present a substantial technical report/document.
PO3	Ability to demonstrate a degree of mastery over the area of specialization (Artificial Intelligence and Data Science) at a level higher than the bachelor's program.



Abbreviations

Programme Core Course	PCC
Programme Elective Course	PEC
Laboratory Courses	LC
Technical Seminar	TS
Research Methodology	RM
On Job Training/Internship	OJT/ INT
Research Project	RPROJ
Theory	TH
Practical	PR
Comprehensive Continuous Evaluation	CCE
End Semester Examination	ESE
Term work	TW
Practical	PR
Oral	OR
Bloom's Taxonomy	BL
Course Outcome	CO
Program Outcome	PO

Artificial Intelligence & Data Science
KSE
Pune

Shalaka Foundation's

Keystone School of Engineering, Pune

Credit Distribution Summary for Master of Technology (M.Tech.) Programme (2026 Pattern)

Department of Artificial Intelligence and Data Science

Semester I to Semester IV

Semester	I	II	III	IV	Total
Programme Core Course	16	12	-	-	28
Programme Elective Course	3	6	-	-	9
Laboratory Courses	3	2	-	-	5
Technical Seminar	-	2	4	4	10
Research Methodology	-	-	4	-	4
On Job Training/Internship	-	-	5	-	5
Research Project	-	-	9	18	27
Total	**22**	**22**	**22**	**22**	**88**



Shalaka Foundation's

Keystone School of Engineering, Pune

Curriculum Structure for Master of Technology (M.Tech.) Programme (2026 Pattern)

First Year Master of Technology (F.Y. M.Tech.)

Department of Artificial Intelligence and Data Science

Semester I

Courses			Teaching Scheme Hrs/Week		Examination Scheme						Credit		
Course Code	Course Type	Title of Course	TH	PR	CCE	ESE	TW	PR	OR	Total	TH	PR	Total
PCC-26501 -MTAD	PCC	Mathematics for Data Science	4	-	50	50	-	-	-	100	4	-	4
PCC-26502 -MTAD	PCC	Natural Language Understanding	4	-	50	50	-	-	-	100	4	-	4
PCC-26503 -MTAD	PCC	Machine Learning Techniques	4	-	50	50	-	-	-	100	4	-	4
PCC-26504 -MTAD	PCC	Generative Artificial Intelligence	4	-	50	50	-	-	-	100	4	-	4
LC-26505-MTAD	LC	Computational Laboratory-I	-	4	-	-	25	-	25	50	-	2	2
PEC-26506-MTAD	PEC	Program Elective-I	3	-	50	50	-	-	-	100	3	-	3
LC-26507-MTAD	LC	Skill-Based Laboratory-I	-	2	-	-	25	-	25	50	-	1	1
Total			**19**	**6**	**250**	**250**	**50**	**-**	**50**	**600**	**19**	**3**	**22**
Total Hours/Week											**22**		

Program Elective-I	
PEC-26506A-MTAD	Cloud Computing
PEC-26506B-MTAD	Deep Learning
PEC-26506C-MTAD	Blockchain Technology
PEC-26506D-MTAD	Recommendation Systems



Shalaka Foundation's

Keystone School of Engineering, Pune

Curriculum Structure for Master of Technology (M.Tech.) Programme (2026 Pattern)

First Year Master of Technology (F.Y. M.Tech.)

Department of Artificial Intelligence and Data Science

Semester II

Courses			Teaching Scheme Hrs/Week		Examination Scheme						Credit		
Course Code	Course Type	Title of Course	TH	PR	CCE	ESE	TW	PR	OR	Total	TH	PR	Total
PCC-26508 -MTAD	PCC	Explainable Artificial Intelligence	4	-	50	50	-	-	-	100	4	-	4
PCC-26509- MTAD	PCC	AI in Cyber Security	4	-	50	50	-	-	-	100	4	-	4
PCC-26510 -MTAD	PCC	Data Visualization Techniques	4	-	50	50	-	-	-	100	4	-	4
LC-26511- MTAD	LC	Computational Laboratory-II	-	4	-	-	25	-	25	50	-	2	2
PEC-26512 -MTAD	PEC	Elective–II	3	-	50	50	-	-	-	100	3	-	3
PEC-26513- MTAD	PEC	Elective–III	3	-	50	50	-	-	-	100	3	-	3
TS-26514- MTAD	TS	Technical Seminar I	-	4	-	-	25	-	25	50	-	2	2
Total			**18**	**8**	**250**	**250**	**50**	**-**	**50**	**600**	**18**	**4**	**22**
Total Hours/Week											**22**		

Program Elective-II		Program Elective-III	
PEC-26512A-MTAD	Social Media Analytics	PEC-26513A-MTAD	Federated Learning
PEC-26512B-MTAD	Quantum Computing	PEC-26513B-MTAD	Big Data Analytics
PEC-26512C-MTAD	Business Intelligence	PEC-26513C-MTAD	Reinforcement Learning
PEC-26512D-MTAD	Spatial Data Analytics	PEC-26513D-MTAD	Game Theory and Applications



Shalaka Foundation's

Keystone School of Engineering, Pune

Curriculum Structure for Master of Technology (M.Tech.) Programme (2026 Pattern)

Second Year Master of Technology (S.Y. M.Tech.)

Department of Artificial Intelligence and Data Science Semester III

Courses			Teaching Scheme Hrs/Week		Examination Scheme						Credit		
Course Code	Course Type	Title of Course	TH	PR	CCE	ESE	TW	PR	OR	Total	TH	PR	Total
RM-26601-MTAD	RM	Research Methodology	4	-	50	50	-	-	-	100	4	-	4
OJT/INT-26602-MTAD	OJT/INT	On-Job Training/Internship	-	10	-	-	100	-	-	100	-	5	5
TS-265603-MTAD	TS	Technical Seminar II	-	8	-	-	25	-	25	50	-	4	4
RPROJ-26604-MTAD	RPROJ	Research Project I	-	18	-	-	25	-	25	50	-	9	9
Total			**4**	**36**	**50**	**50**	**150**	**-**	**50**	**300**	**4**	**18**	**22**
Total Hours/Week											**22**		



Shalaka Foundation's

Keystone School of Engineering, Pune

Curriculum Structure for Master of Technology (M.Tech.) Programme (2026 Pattern)

Second Year Master of Technology (S.Y. M.Tech.)

Department of Artificial Intelligence and Data Science

Semester IV

Courses			Teaching Scheme Hrs/Week		Examination Scheme						Credit		
Course Code	Course Type	Title of Course	TH	PR	CCE	ESE	TW	PR	OR	Total	TH	PR	Total
TS-26605-MTAD	TS	Technical Seminar-III	-	8	-	-	50	-	50	100	-	4	4
RPROJ-26606-MTAD	RPROJ	Research Project-II	-	36	-	-	150	-	50	200	-	18	18
Total			**4**	**36**	**-**	**-**	**200**	**-**	**100**	**300**	**-**	**22**	22
Total Hours/ Week											**22**		



Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

First Year M.Tech. Artificial Intelligence and Data Science (AI&DS)
Curriculum with effect from A.Y. 2026-27
Semester – I

Mathematics for Data Science (PCC-26501-MTAD)		
Teaching Scheme: **Theory: 04 Hrs/Week**	**Credit Scheme:** **TH: 04**	**Examination Scheme:** **Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks** **Total Marks: 100 Marks**

Course Introduction:

Mathematics for Data Science is the rigorous mathematical foundation of the postgraduate curriculum. It systematically integrates the basic principles of linear and multivariate calculus, probability theory, mathematical statistics and optimization to solve complex analytical problems. The whole syllabus is analyzed to help you to build the foundations of theory to manipulate high-dimensional data with vectors and matrices, analyze multi-dimensional functions with differential and integral calculus and perform structured decision-making under uncertainty with statistical inference and probability distributions. Moving from theory to practice, it bridges the gap between mathematical theory (eigenvalues, gradient descent, hypothesis testing, convex optimization) and the design, regularization, and implementation of robust machine learning algorithms and neural networks.

Prerequisites:

- Discrete Mathematics

Course Objectives: The objective of this course is:

1. To develop a foundation in linear algebra, probability, statistics, and optimization used in data science
2. To study mathematical tools for data representation, pattern recognition, and model building
3. To prepare students to apply mathematical methods to machine learning and AI problems

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Apply linear algebra concepts to manipulate datasets and model transformations
CO2: Use calculus for understanding optimization in machine learning algorithms
CO3: Understand and apply concepts of probability and distributions in data analysis
CO4: Perform statistical analysis and hypothesis testing on real-world data
CO5: Solve optimization problems relevant to data fitting and predictive modeling

Course Contents

Unit I	Linear Algebra for Data Science	12 Hours

Overview of vectors, matrices, tensors, matrix operations, properties, types of matrices, identity, diagonal, orthogonal, inverse, dot product, cross Product, norms, eigenvalues and eigenvectors. Principal component analysis (PCA) – Mathematical Intuition, dimensionality reduction, image data transformation.



Unit II	Calculus and Functions	12 Hours
Functions, limits, continuity, differential calculus, derivatives, chain rule, gradient, jacobian, partial derivatives and directional derivatives, integral Calculus - area under curves, expected values, optimization techniques - maxima, minima, gradient descent, cost function minimization in ML models.		
Unit III	**Probability Theory**	**12 Hours**
Overview of probability, sample space, events, axioms, conditional probability, Bayes' Theorem, random variables – Discrete and Continuous, probability distributions - Bernoulli, Binomial, Poisson, Normal, exponential joint, marginal distributions, Central Limit Theorem, applications for Classification, decision-making under uncertainty.		
Unit IV	**Statistics & Data Analysis**	**12 Hours**
Overview of descriptive Statistics - mean, median, mode, variance, standard deviation, correlation, covariance, sampling, estimation, confidence intervals, hypothesis testing - null and alternative hypotheses, p-values, Chi-square test, t-test, ANOVA, applications for statistical inference, feature importance, A/B testing, tools - NumPy, Matplotlib, Scikit-learn, Jupyter Notebooks.		
Unit V	**Applications in Data Science**	**12 Hours**
Convex sets, convex functions, unconstrained and constrained optimization, Lagrange multipliers, linear programming, stochastic gradient descent, regularization - L1, L2 norms, applications for model fitting, regression, neural network training.		

Learning Resources

Textbooks

T1. Marc Peter Deisenroth, A. Aldo Faisal, Cheng Soon Ong, "Mathematics for Machine Learning", Cambridge University Press, 2020

T2. Sheldon M. Ross, "Introduction to Probability and Statistics for Engineers and Scientists", Elsevier Academic Press, Third Edition, ISBN: 0-12-598057-4.

Reference Books

R1. Biswadip Basu Mallik, Kirti Verma, Rahul Kar, Ashok Kumar Shaw, Sardar M. N. Islam, "Advanced Mathematical Applications in Data Science", Bentham Books Publisher, ISBN: 978-981-5124-85-9.

R2. Gilbert Strang, "Linear Algebra and Its Applications", Fourth Edition.

Web Resources

1. MIT OpenCourseWare: Mathematics for Computer Science



Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Natural Language Understanding (PCC-26502-MTAD)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory: 04 Hrs/Week	TH: 04	Comprehensive Continuous Evaluation (CCE): 50 Marks End Semester: 50 Marks Total Marks: 100 Marks

Course Introduction:

The aim of Natural Language Understanding is to give postgraduate students a deep, systematic understanding of the linguistic principles and computational frameworks needed to process, interpret and generate human language. This course bridges classical computational linguistics and state-of-the-art NLP by systematically walking students through the hierarchy of language processing, from foundational tokenization and syntax-based structures (CFGs and dependency parsing) to distributional vector semantics and modern transformer-based deep learning architectures. Through detailed units covering language modeling, syntactic structure parsing, word representations, and deep attention mechanisms (such as BERT and GPT models), this course equips students to develop, evaluate, and optimize scalable NLP pipelines for real-world downstream applications, including conversational artificial intelligence (AI), intent detection, and information extraction.

Prerequisites:

- Discrete Mathematics
- Machine Learning

Course Objectives: The objective of this course is:

1. Introduce foundational concepts of language structure and computational linguistics.
2. To apply data science and AI techniques to language understanding tasks.
3. Provide knowledge of statistical, neural, and transformer-based NLP models.
4. Develop the ability to solve real-world language problems.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Understand the structure and meaning of natural language using linguistic and statistical techniques.
CO2: Apply machine learning and deep learning models to core NLP tasks.
CO3: Analyze and compare traditional NLP pipelines with modern transformer-based approaches.
CO4: Evaluate and optimize NLP systems for accuracy and performance.
CO5: Design and evaluate scalable NLP systems for AI applications.

Course Contents

Unit I	Introduction to Natural Language Processing	12 Hours

Foundation of Natural Language Understanding, Natural Language Processing, linguistic levels, morphology, syntax, semantics, pragmatics, tokenization, lemmatization, POS tagging, corpus creation, annotation formats, CoNLL, JSONL, overview of NLP data pipelines, annotation tools like Prodigy, Doccano.



Unit II	Syntax and Structure Based Understanding	12 Hours
Introduction to context-free grammars, parsing algorithms, top-down & bottom-up parsing, dependency parsing, constituency parsing, feature structure, unification, POS tagging with Hidden Markov Models, CRFs Shallow Parsing, deep parsing, statistical parsing, Chunking, role of downstream ML tasks.		
Unit III	**Semantic Understanding and Representation**	**12 Hours**
Introduction: word sense disambiguation, semantic role labeling, logic based meaning representation, lexical semantics, distributional semantics, vector semantics, TF-IDF, Word2Vec, GloVe, knowledge graphs, WordNet, ConceptNet, word embedding, embedding language data into machine-readable vectors for modeling, semantic search engine.		
Unit IV	**Deep Learning and Transformer Architectures**	**12 Hours**
Traditional neural networks for text classification, neural networks for RNN, LSTM, GRU, attention mechanisms, overview of transformer architecture - BERT, GPT, transfer learning, fine tuning, pre-trained models for NLP, building NLP pipelines, overview of tools like HuggingFace.		
Unit V	**Applications of NLP**	**12 Hours**
Dialogue systems, Intent detection, slot filling, text classification & summarization, question answering, information extraction, explainability, transformers for building domain-specific chatbot, analysis of gender bias in NLP datasets, case studies for real AI systems like ChatGPT, Alexa, Google BERT-based search.		

Learning Resources

Textbooks

T1. Daniel Jurafsky, James H. Martin, "Speech and Language Processing", Third Edition, 2025.
T2. Yoav Goldberg, "Neural Network Methods for Natural Language Processing", Springer, 2017.
T3. Denis Rothman, "Transformers for Natural Language Processing", Packt, 2021.
T4. Steven Bird, Ewan Klein, Edward Loper, "Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit".

Reference Books

R1. Steven Bird, Ewan Klein, Edward Loper, " Security Essentials, Natural Language Processing with Python", Orelly, 2009.
R2. Soyeon Caren Han , Henry Weld , Yan Li , Jean Lee , Josiah Poon, " Natural Language Understanding in Conversational AI with Deep Learning", Springer, 2025.

Web Resources

1. James Allen, "Natural Language Understanding". https://www.scribd.com/document/537550277/James-Allen-Natural-Language-Understanding.



Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Machine Learning Techniques (PCC-26503-MTAD)		
Teaching Scheme: **Theory: 04 Hrs/Week**	**Credit Scheme:** **TH: 04**	**Examination Scheme:** **Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks** **Total Marks: 100 Marks**

Course Introduction:

The Machine Learning Techniques course provides students with an advanced understanding of key machine learning concepts, algorithms, and methodologies. It covers supervised, semi-supervised, and unsupervised learning, along with ensemble methods, regularization, dimensionality reduction, anomaly detection, representation learning, active learning, and transfer learning. The course emphasizes practical implementation and evaluation of robust and scalable machine learning models for applications across domains such as NLP, computer vision, and speech recognition.

Prerequisites:

- Design of Analysis & Algorithms

Course Objectives: The objective of this course is:

1. To design robust predictive models using ensemble methods, regularization, and hyper parameter tuning.
2. To investigate semi-supervised learning for improved performance with limited labeled data.
3. To implement and evaluate unsupervised learning for clustering, dimensionality reduction, and anomaly detection.
4. To understand and apply representation learning (embeddings, deep metric learning) for meaningful data representations.
5. To employ active learning strategies for efficient data labeling.
6. To leverage transfer learning to adapt pre-trained models to new tasks with limited data.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Apply supervised learning algorithms to build and evaluate predictive models.
CO2: Use semi-supervised learning methods to enhance model performance with limited labeled data.
CO3: Implement and evaluate unsupervised learning techniques for clustering, dimensionality reduction, and anomaly detection.
CO4: Develop and analyze data representations using embeddings and deep learning models.
CO5: Design and apply active learning strategies to minimize labeling efforts in machine learning tasks.



Course Contents		
Unit I	**Supervised and Semi-Supervised Learning**	**12 Hours**
Supervised Learning: Ensemble Methods: Bagging (e.g., Random Forests), Boosting (AdaBoost, XGBoost, LightGBM). Model Calibration, Regularization Techniques: L1, L2, Elastic Net for preventing overfitting, Cross-validation, Hyper parameter Tuning. Semi-Supervised Learning: Self-training, Co-training, Graph-based Methods, Generative Models for Semi-supervised Learning (e.g., Semi-supervised GANs), Applications in NLP and Computer Vision.		
Unit II	**Unsupervised Learning Algorithms**	**12 Hours**
Clustering Techniques: Partitional Clustering: K-Means and its variations, Density-Based Clustering: DBSCAN, HDBSCAN for identifying clusters of varying shapes, Hierarchical Clustering: Agglomerative and divisive approaches, Model-Based Clustering: Gaussian Mixture Models (GMM) and Expectation-Maximization (EM) algorithm. Dimensionality Reduction: Linear & Non-linear Methods, Auto-encoders (as an unsupervised method for learning compressed representations). Anomaly Detection Techniques, Evaluation Metrics: Silhouette Score, Davies-Bouldin Index for assessing clustering quality.		
Unit III	**Representation Learning**	**12 Hours**
Embedding Methods: Word Embeddings: Word2Vec (Skip-gram, CBOW), Doc2Vec for representing words and documents in a vector space, GloVe (Global Vectors for Word Representation). Contrastive Learning: SimCLR (Simple Framework for Contrastive Learning of Visual Representations) & MoCo (Momentum Contrast for Unsupervised Visual Representation Learning). Deep Metric Learning: Triplet Loss, Siamese Networks. Representation in Graphs: Node embeddings, Graph Neural Networks (GNNs).		
Unit IV	**Active Learning**	**12 Hours**
Query Strategies: Pool-Based Sampling, Stream-Based Sampling, Membership Query Synthesis (brief overview), and Uncertainty Sampling: Entropy, Margin Sampling, and Least Confidence. Query by Committee (QBC), Expected Model Change, and Active Learning in Deep Learning: Bayesian neural networks, Monte Carlo (MC) Dropout for uncertainty estimation, Human-in-the-loop Systems: The role of human experts in the active learning process. Tools/Frameworks: mod AL (Python library for modular active learning), PyTorch/TensorFlow (for model training), Label Studio (for annotation).		
Unit V	**Transfer Learning**	**12 Hours**
Feature-based Transfer Learning: Extracting and reusing learned features from a source model, Fine-tuning Pre trained Networks: Adapting large-scale models (CNNs like ResNet, EfficientNet; Transformers like BERT, large language models) to new tasks, Domain Adaptation: Techniques for bridging the gap between different data distributions (source and target domains), Domain Generalization: Learning models that can generalize well to unseen target domains, Few-shot and Zero-shot Learning: Learning with very limited or no labeled data in the target domain, Multi-task Learning: Training a single model to perform multiple related		

Artificial Intelligence & Data Science
KSE
Pune

tasks simultaneously. Applications: Natural Language Processing (BERT and other large language models), Computer Vision, Speech recognition.

Learning Resources

Textbooks

T1. Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow, by Aurélien Géron.
T2. Pattern Recognition and Machine Learning, by Christopher Bishop.
T3. The Hundred-Page Machine Learning Book, by Andriy Burkov.

Reference Books

R1. Artificial Intelligence: A Modern Approach, by Stuart Russell and Peter Norvig.
R2. Deep Learning with Python, by François Chollet.
R3. Introduction to Machine Learning with Python, by Andreas C. Müller and Sarah Guido.

Web Resources

1. http://www.ru.ac.bd/wpcontent/uploads/sites/25/2019/03/207_05_01_Rajchka_Using-Python-for-machinelearning-2015.pdf
2. Foundation of Machine Learning: https://cs.nyu.edu/~mohri/mlbook/
3. Dive into Deep Learning: http://d2l.ai/
4. A brief introduction to machine learning for Engineers: https://arxiv.org/pdf/1709.02840.pdf
5. Feature selection: https://dl.acm.org/doi/pdf/10.5555/944919.944968
6. Introductory Machine Learning Nodes : http://lcsl.mit.edu/courses/ml/1718/MLNotes.pdf

Massive Open Online Courses (MOOCs)

1. Introduction to Machine Learning: https://nptel.ac.in/courses/106105152

Artificial Intelligence & Data Science
KSE
Pune

Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Generative Artificial Intelligence (PCC-26504-MTAD)		
Teaching Scheme: **Theory: 04 Hrs/Week** **Tutorial:**	**Credit Scheme:** **TH: 04**	**Examination Scheme:** **Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks** **Total Marks: 100 Marks**

Course Introduction:

Generative Artificial Intelligence introduces postgraduate students to the state-of-the-art paradigms of content generation, mapping the technical shift from classical machine learning to modern deep generative architectures. Through a comprehensive progression across five specialized units, the curriculum details the mathematical and structural frameworks of autoregressive models, Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), Diffusion Models, and Transformer self-attention mechanisms. Students explore both single-modality and multimodal architectures (such as GPT, BERT, LLaMA, DALL·E, and Stable Diffusion) while mastering practical workflows for text, image, and audio generation. The course emphasizes hands-on development using industry-standard libraries like Hugging Face, OpenAI APIs, and TensorFlow/Keras to build custom code copilots and creative AI content. Finally, it investigates industrial use cases in healthcare and finance, synthetic data simulation methods, and emerging global governance regulations, preparing students to lead responsibly in future frontiers like Explainable AI and Edge AI.

Prerequisites:

- Artificial Intelligence

Course Objectives: The objective of this course is:

1. To understand the fundamentals of Generative AI models and their applications.
2. To explore various architectures like GANs, VAEs, and Transformers.
3. To apply Generative AI techniques for text, image, and audio generation.
4. To examine real-world applications and emerging trends in Generative AI.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Explain the evolution, types, and ethical considerations of Generative AI models.

CO2: Analyse the architecture and working principles of Transformers, Large Language Models (LLMs), and multimodal generative models.

CO3: Implement AI-based text, image, and audio generation techniques using state-of-the-art models.

CO4: Utilize industry-standard frameworks (Hugging Face, OpenAI APIs, TensorFlow/Keras) to build and fine-tune generative AI models.

CO5: Evaluate real-world applications, challenges, and emerging trends in Generative AI across different domains.

Artificial Intelligence & Data Science ★ KSE Pune

Course Contents		
Unit I	**Introduction to Generative AI**	**12 Hours**
Introduction to AI-driven content generation, Evolution of generative models: From rule based AI to deep generative models, Types of Generative Models: Autoregressive models, Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), Diffusion Models, Ethical Considerations in Generative AI: Bias, misinformation, deepfakes, copyright issues.		
Unit II	**Foundation Models & Architectures**	**12 Hours**
Transformers and Self-Attention Mechanism, Large Language Models (LLMs): GPT, BERT, T5, LLaMA, Pre-training vs. Fine-tuning approaches, Multimodal Generative AI: DALL·E, CLIP, Stable Diffusion.		
Unit III	**Generative AI for Text, Image & Audio Generation**	**12 Hours**
Text Generation: Autoregressive models, ChatGPT, summarization, creative writing, Image Generation: StyleGAN, Stable Diffusion, DeepDream, AI Art, Audio Generation: Text to-Speech (TTS), AI Music Composition, Speech Synthesis, Evaluation Metrics: Perplexity, BLEU, FID score, Inception Score.		
Unit IV	**Tools, Frameworks, and Applications**	**12 Hours**
Generative AI in Software Development: GitHub Copilot, Code Llama, Frameworks & Libraries: Hugging Face, OpenAI APIs, TensorFlow/Keras for Generative AI, Fine-tuning & Custom Model Development, AI for Content Creation: Deepfake detection, AI-generated movies, virtual influencers.		
Unit V	**Industry Use Cases & Future Trends**	**12 Hours**
Generative AI in Healthcare, Finance, and Creative Industries, Synthetic Data Generation & AI for Simulations, Regulations & AI Governance: Transparency, accountability, and ethical AI, Future Trends in Generative AI: Explainable AI, Federated Learning, Edge AI.		

Learning Resources

Textbooks

T1. Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play – David Foster, O'Reilly Media, 2nd Edition, 2023.

T2. Hands-On Generative AI with Transformers – Utkarsh Sinha, Packt Publishing, 2023.

Reference Books

R1. Deep Learning for Natural Language Processing – Palash Goyal, Sumit Pandey, Karan Jain, A press, 2018.

R2. GANs in Action: Deep Learning with Generative Adversarial Networks – Jakub Langr, Vladimir Bok, Manning Publications, 2019.

R3. Building Machine Learning Powered Applications – Emmanuel Ameisen, O'Reilly Media, 2020.

Artificial Intelligence & Data Science ★ KSE Pune

R4. Transformers for Natural Language Processing – Denis Rothman, Packt Publishing, 2022

Web Resources

1. Prompt Engineering for ChatGPT: www.coursera.org/learn/prompt-engineering?utm_source=chatgpt.com
2. Generative AI Full Course 2025: www.youtube.com/watch?v=-v9PiM6cqLM

Massive Open Online Courses (MOOCs)

1. onlinecourses.swayam2.ac.in/nou25_ma05/preview - Generative AI for Everyday
2. https://onlinecourses.swayam2.ac.in/e-learning/preview/imb25_mg46 - Generative AI and Large Language Models
3. www.coursera.org/learn/prompt-engineering?utm_source=chatgpt.com - Prompt Engineering for ChatGPT.
4. www.youtube.com/watch?v=-v9PiM6cqLM - Generative AI Full Course 2025

Artificial Intelligence & Data Science
KSE
Pune

Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Computational Laboratory-I (LC-26505-MTAD)		
Teaching Scheme: **Practical:** 04 Hrs/Week	**Credit Scheme:** **PR/OR:** 02	**Examination Scheme:** **Term work:** 25 Marks **Oral:** 25 Marks **Total Marks:** 50 Marks

Course Introduction:

Computational Laboratory-I serves as the primary practical core course designed to provide hands-on experience and experiential learning by implementing foundational algorithms across the four program core courses of Semester I: Probability & Statistics for Data Science, Advanced Natural Language Processing , Learning Methods for AI , and Generative Artificial Intelligence. The laboratory curriculum enables postgraduate students to transition from mathematical theory to execution by designing Python-based programs for dimensionality reduction, probability simulation, named entity extraction using spaCy, and transformer fine-tuning for intent classification. Furthermore, students develop advanced skills in predictive modeling through multi-class classification and clustering, explore probabilistic coordination and election protocols in distributed processes, and implement deep generative models such as DCGANs for image generation and pre-trained GPT architectures for text synthesis. The practical coursework culminates in a comprehensive mini-project, preparing students to architect and deploy production-ready AI pipelines to solve complex, real-world engineering challenges.

Prerequisites:

- Probability & Statistics for Data Science
- Advanced Natural Language Processing
- Learning Methods for AI
- Generative Artificial Intelligence

Course Objectives: The objective of this course is:

1. Provide knowledge of Probability & Statistics for Data Science for solving engineering problems.
2. Explain various algorithms for solving engineering problems.
3. To apply various Learning Methods for AI for solving engineering problems.
4. Describe Generative Artificial Intelligence for solving engineering problems.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Apply knowledge of Probability & Statistics for Data Science for solving engineering problems.

CO2: Choose appropriate algorithms for solving engineering problems.

CO3: Develop Learning Methods for AI for improving engineering problem solutions.

CO4: Develop Generative Artificial Intelligence environment while solving engineering problems.

Artificial Intelligence & Data Science ★ KSE Pune

Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Guidelines:

- Computational Laboratory-I assignments are based on 4 Programme Core courses (Probability & Statistics for Data Science, Advanced Natural Language Processing , Learning Methods for AI , Generative Artificial Intelligence)
- Respective course faculty assigned will frame course objectives, course outcomes, list of assignments, mini projects, if any
- The list of sample assignments is given for preparation of assignments which will be approved by the respective heads of Department before conduction of laboratory.

List of Assignments

Part A: Probability & Statistics for Data Science (Any One)

1. Write a program for data transformation and reduce data dimensions in linear algebra. Implement matrix operations - addition, multiplication, inverse, eigenvalues, using Python NumPy.
2. Write a program for probability distributions. Simulate and visualize - normal, binomial, and Poisson distributions and calculate expected value and variance using Python, Matplotlib, SciPy.

Part B - Advanced Natural Language Processing (Any One)

1. Write a program for named entity recognition. Apply NER tagging to extract entities for custom text dataset using Python, spaCy.
2. Write a program for Intent Classification using BERT. Develop and fine-tune a pre-trained BERT model for an intent classification dataset using tools like HuggingFace Transformers, PyTorch/TensorFlow.

Part C - Learning Methods for AI (Any TWO)

1. Support vector machine for multiclass classification using Python: Implement a classification model using Support vector machine algorithm for news article topic classification like science, technology, politics, sports, etc. Use sklearn dataset and/or fetch dataset using APIs.
2. Write a program for classification using Decision Trees and Random Forest. Train models on the Titanic dataset and evaluate using accuracy, precision, recall using tools Scikit-learn, Pandas.
3. Convolutional Neural Network for face recognition using Python: Implement a real-world application like AI Face Recognition Entry Agent using CNN. It detects and recognizes faces in real-time environment. It should grant or deny the entry based on known faces of pre-trained databases.
4. Write a program for K-Means Clustering for Customer Segmentation. Apply K-Means on customer data to identify distinct customer groups based on features like spending, income, etc. using Python, Scikit-learn, Seaborn.

Part D - Generative Artificial Intelligence (Any TWO)

1. Coordination Problem. Two processes, A and B, communicate by sending and receiving messages on a bidirectional channel. Neither process can fail. However, the channel can



experience transient failures, resulting in the loss of a subset of the messages that have been sent. Devise a protocol where either of two actions α and β are possible, but (i) both processes take the same action and (ii) neither take both actions.

2. Write a program for Image Generation using DCGAN. Develop a Deep Convolutional GAN for MNIST dataset to generate handwritten digit images using TensorFlow / PyTorch.
3. Election Problem. A set of processes P1, P2, ..., Pn must select a leader. Each process Pi has a unique identifier uid(i). Devise a protocol so that all of the processes learn the identity of the leader. Assume all processes start executing at the same time and that all communicate using broadcasts that are reliable.
4. Write a program for Text Generation using GPT pre-trained model to generate coherent text given a seed prompt. Explore text temperature and length effects using tools like HuggingFace Transformers, Colab.

Mini-Projects (Any ONE)

1. Students are expected to develop a mini project based on above topics (Any ONE).

Learning Resources

Reference Books:

R1. Sebastien Roch, "Mathematical Methods in Data Science with Python".

R2. Steven Bird, Ewan Klein, Edward Loper, "Security Essentials, Natural Language Processing with Python", O'Reilly, 2009.

R3. Introduction to Machine Learning with Python by Andreas C. Müller and Sarah Guido.

R4. Hands-On Generative AI with Transformers – Utkarsh Sinha, Packt Publishing, 2023.

Artificial Intelligence & Data Science
KSE
Pune

Elective I: Cloud Technology (PEC-26506A-MTAD)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory: 03 Hrs/Week	TH: 03	Comprehensive Continuous Evaluation (CCE): 50 Marks End Semester: 50 Marks Total Marks: 100 Marks

Course Introduction:

Cloud Technology offers postgraduate students a comprehensive, multi-dimensional overview of the architectural paradigms, deployment models, and computational environments that underpin modern distributed computing systems. Over a five-unit trajectory, the curriculum transitions students from historical centralized and distributed frameworks (such as cluster, grid, and network-based distributed computing models) to modern virtualization structures and cloud deployment services (IaaS, PaaS, SaaS, and XaaS). Students are equipped with practical capabilities in cloud programming using major commercial cloud ecosystems, including Amazon AWS, Microsoft Azure, and Google App Engine. By analyzing the entire course, it balances this hands-on engineering focus with essential governance and compliance guidelines, including multi-tenant access authentication, reliability and fault tolerance, privacy policies, and interoperability standards, preparing graduates to confidently architect and secure high-performance, cloud-native solutions.

Prerequisites:

- Computer Network

Course Objectives: The objective of this course is:

1. To introduce concepts of Cloud Computing.
2. To shed light on the Security issues in Cloud Computing.
3. To introduce the Cloud Standards.
4. To Implement various cloud computing environments.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Articulate the main concepts, key technologies, strengths, and limitations of cloud computing and the possible applications for state-of-the-art cloud computing.

CO2: Identify the architecture and infrastructure of cloud computing, including SaaS, PaaS, IaaS, public cloud, private cloud, hybrid cloud, etc.

CO3: Explain the core issues of cloud computing such as security, privacy, and interoperability.

CO4: Provide the appropriate cloud computing solutions and recommendations according to the applications used.

Course Contents

Unit I	Introduction to Cloud Computing	08 Hours

History of Centralized and Distributed Computing - Overview of Distributed Computing, Cluster computing, Grid computing. Technologies for Network based systems- System models for Distributed and cloud computing- Software environments for distributed systems and clouds.

Artificial Intelligence & Data Science
KSE
Pune

Unit II	Cloud Computing Models	08 Hours
Introduction to Cloud Computing- Cloud issues and challenges - Properties - Characteristics - Service models, Deployment models. Cloud resources: Network and API - Virtual and Physical computational resources - Data-storage. Virtualization concepts - Types of Virtualizations-Introduction to Various Hypervisors - High Availability (HA)/Disaster Recovery (DR) using Virtualization, Moving VMs.		
Unit III	**Service Models**	**08 Hours**
Service models - Infrastructure as a Service (IaaS) - Resource Virtualization: Server, Storage, Network - Case studies. Platform as a Service (PaaS) - Cloud platform & Management: Computation, Storage - Case studies. Software as a Service (SaaS) - Web services - Web 2.0 - Web OS - Case studies – Anything as a service (XaaS).		
Unit IV	**Cloud Programming**	**08 Hours**
Cloud Programming and Software Environments – Parallel and Distributed Programming paradigms – Programming on Amazon AWS and Microsoft Azure – Programming support of Google App Engine – Emerging Cloud software Environment.		
Unit V	**Cloud Access and Security**	**08 Hours**
Cloud Access: authentication, authorization and accounting - Cloud Provenance and meta-data - Cloud Reliability and fault-tolerance - Cloud Security, privacy, policy and compliance- Cloud federation, interoperability and standards.		

Learning Resources

Textbooks

T1. Kai Hwang, Geoffrey C. Fox and Jack J. Dongarra, "Distributed and cloud computing from Parallel Processing to the Internet of Things", Morgan Kaufmann, Elsevier – 2012.

T2. Toby Velte, Anthony Velte, Robert Elsenpeter, Robert C. Elsenpeter, "Cloud Computing, A Practical Approach", McGraw Hill LLC, ISBN: 9780071626958, 0071626956, 2009.

Reference Books

R1. Barrie Sosinsky, "Cloud Computing Bible" John Wiley & Sons, 2010.

R2. Tim Mather, Subra Kumaraswamy, and Shahed Latif, "Cloud Security and Privacy an Enterprise Perspective on Risks and Compliance", O'Reilly 2009.

Artificial Intelligence & Data Science ★ KSE Pune

Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Elective I: Deep Learning (PEC-26506B-MTAD)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory: 03 Hrs/Week Tutorial:	TH: 03	Comprehensive Continuous Evaluation (CCE): 50 Marks End Semester: 50 Marks Total Marks: 100 Marks

Course Introduction:

Deep Learning offers a comprehensive exploration of neural network architectures, optimization strategies, and deployment methodologies that define modern artificial intelligence systems. Over a five-unit progression, students build a rigorous theoretical and practical foundation, transitioning from classic perceptrons and multi-layer perceptrons to deep convolutional networks (such as LeNet, AlexNet, ResNet, and Inception) and recurrent architectures (Vanilla RNNs, LSTMs, and GRUs) for processing spatial, sequential, and temporal datasets. The curriculum addresses critical training and engineering workflows, including data preprocessing, regularization, hyperparameter tuning, and imbalanced dataset management, before introducing advanced generative paradigms (such as GANs and Variational Autoencoders), Deep Reinforcement Learning, and Explainable AI (XAI). By directly integrating these deep learning concepts with practical case studies in computer vision, speech recognition, and recommender systems across domains like healthcare and robotics, the course prepares postgraduate students to design, implement, and deploy scalable, high-performance, and ethically grounded deep models.

Prerequisites:

- Linear Algebra
- Statistics and Probability
- Advanced Machine Learning

Course Objectives: The objective of this course is:

1. Understand fundamental concepts of deep learning.
2. Apply deep learning algorithms to solve real-world problems.
3. Develop and evaluate deep learning models.
4. Critically analyze deep learning methodologies.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Fundamentals: Demonstrate mastery of fundamental deep learning concepts.
CO2: Application: Apply deep learning algorithms to solve practical engineering problems.
CO3: Model Development: Design, implement, and evaluate deep learning models.
CO4: Critical Analysis: Critically evaluate deep learning methodologies and their limitations.
CO5: Problem Solving: Develop and implement solutions using deep learning techniques appropriate to given contexts.

Course Contents

Unit I	**Foundations of Deep Learning**	**08 Hours**



Introduction to Artificial Neural Networks (ANNs), Perceptrons and Multilayer Perceptrons (MLPs), Activation Functions (Sigmoid, Tanh, ReLU, Leaky ReLU), Backpropagation Algorithm, Optimization Algorithms (Gradient Descent, Stochastic Gradient Descent, Adam), Regularization Techniques (Dropout, L1/L2 Regularization), Introduction to Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) – basic concepts only. Assessment: Homework assignments focused on implementing basic neural networks using Python libraries (e.g., TensorFlow/Keras).		
Unit II	**Deep Learning Architectures and Applications**	**08 Hours**
Convolutional Neural Networks (CNNs): Architectures (LeNet, AlexNet, ResNet, Inception), Applications (Image Classification, Object Detection, Image Segmentation), Recurrent Neural Networks (RNNs): Architectures (Vanilla RNN, LSTM, GRU), Applications (Natural Language Processing, Time Series Analysis), Long Short-Term Memory (LSTM) Networks and Gated Recurrent Units (GRUs), Generative Adversarial Networks (GANs). Assessment: Practical project on image classification or a simple NLP task using a chosen deep learning framework.		
Unit III	**Deep Learning Model Development**	**08 Hours**
Data Preprocessing for Deep Learning, Model Selection, Training Strategies, and Hyperparameter Tuning, Evaluation Metrics for Deep Learning Models (Accuracy, Precision, Recall, F1-Score, AUC-ROC), Model Deployment and Monitoring, Techniques for handling imbalanced datasets in deep learning, Introduction to transfer learning. Assessment: A medium-sized project requiring the design, implementation, and evaluation of a deep learning model on a given dataset (e.g., a complex image classification problem or a text classification task).		
Unit IV	**Advanced Deep Learning Techniques**	**08 Hours**
Autoencoders: Undercomplete Autoencoders, Regularized Autoencoders-Sparse Autoencoders, Stochastic Encoders and Decoders, Denoising Autoencoders, Contractive Autoencoders, Applications of Autoencoders. Ensemble Methods in Deep Learning, Autoencoders and Variational Autoencoders (VAEs), Deep Reinforcement Learning (basics), Explainable AI (XAI) and its role in DL, Ethical considerations in deep learning. Assessment: A presentation on advanced deep learning techniques along with a critical analysis of their strengths and limitations.		
Unit V	**Deep Learning Applications in AI and DS**	**08 Hours**



Deep Learning applications in Computer Vision, Data Visualization Techniques, and other relevant fields, Image Classification, Social N/w/ analysis, Speech Recognition, Recommender system, Case studies on the application of Deep Learning in areas such as robotics, signal processing, and healthcare. Assessment: A final project focusing on a deep learning application that solves a specific problem in Artificial Intelligence and Data Science (e.g., anomaly detection using sensor data, image-based defect detection).

Learning Resources

Textbooks

T1. Deep Learning by Ian Goodfellow, Yoshua Bengio, and Aaron Courville.

Reference Books

1. Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow by Aurélien Géron
2. Neural Networks and Deep Learning by Michael Nielsen
3. Deep Learning with Python by Francois Chollet

Web Resources

1. https://www.deeplearningbook.org/
2. https://www.youtube.com/channel/UCF9O8Vj-FEbRDA5DcDGz-Pg/videos

Massive Open Online Courses (MOOCs)

1. https://onlinecourses.nptel.ac.in/noc20_cs62/preview
2. https://swayam.gov.in/search_courses?searchText=deep%20learning

Artificial Intelligence & Data Science
KSE
Pune

Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Elective I: Blockchain Technology (PEC-26506C-MTAD)		
Teaching Scheme: **Theory: 03 Hrs/Week**	**Credit Scheme:** **TH: 03**	**Examination Scheme:** **Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks** **Total Marks: 100 Marks**

Course Introduction:

Blockchain Technology is a specialized elective designed to introduce postgraduate students to the architectural paradigms, security mechanisms, and scalability layers of decentralized ledger systems. Over a five-unit progression, the curriculum guides students through cryptographic foundations, encryption, key management, and financial applications such as decentralized finance (DeFi) and non-fungible tokens (NFTs). The course addresses critical system challenges including blockchain scalability (sharding, layer 2 scaling solutions) and multi-chain interoperability protocols. It explores the integration of artificial intelligence with decentralized systems (such as AI-powered smart contracts, decentralized AI models, and AI marketplaces) and the convergence of blockchain with cloud computing environments (BaaS). By analyzing the entire syllabus, this course prepares students to design, develop, and deploy scalable cloud-native blockchain architectures across various high-impact fields such as banking, healthcare, judiciary, IoT, real estate, and government sectors.

Prerequisites:

- Cryptography and Network Security

Course Objectives: The objective of this course is:

1. To understand security threats with application of blockchain.
2. To gain knowledge of scalability with understanding operability.
3. To understand Blockchain and AI integration.
4. To analyze and explore the real-world applications of Blockchain technology.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Get acquainted with the concept of Distributed ledger system and Blockchain.
CO2: Understanding scalability challenges with interoperability
CO3: Apply blockchain-AI integration with AI powered solutions.
CO4: Understanding, Designing and Deploying cloud based blockchain solutions.
CO5: Analyze real world applications of Blockchain Technology.

Course Contents

Unit I	Blockchain Security and Applications	08 Hours

Security in Blockchain: Understanding encryption, key management, and security challenges in blockchain networks. Blockchain in Various Industries: Studying applications of blockchain in finance, healthcare, supply chain management, and government services. Decentralized Finance

Artificial Intelligence & Data Science KSE Pune

(DeFi) and NFTs: Learning about DeFi, liquidity pools, yield farming, and creating, minting, and trading NFTs.		
Unit II	**Blockchain Scalability and Interoperability**	**08 Hours**
Blockchain Scalability and Interoperability. Blockchain Scalability: Sharding, Layer 2 scaling solutions, Consensus mechanism optimization, Block size increases, transaction processing optimization. Blockchain Interoperability: Cross-chain bridges, interoperability protocols, Atomic swaps, Ulti-chain applications, Blockchain agnostic solutions, Use cases.		
Unit III	**Blockchain and Artificial Intelligence**	**08 Hours**
Blockchain and AI Integration: Decentralized AI models, AI-powered smart contracts, Blockchain-based AI marketplaces. Applications of Blockchain and AI: Predictive maintenance, Supply chain optimization, Identity verification, Healthcare, Use cases.		
Unit IV	**Blockchain and Cloud Computing**	**08 Hours**
Blockchain and Cloud Computing. Cloud-based Blockchain Solutions: Blockchain-as-a-Service (BaaS), Cloud-based node infrastructure, Scalable blockchain solutions, Challenges and Opportunities in Blockchain and cloud computing.		
Unit V	**Blockchain for Real World Applications**	**08 Hours**
Crypto-currencies, Banking and Financial Services, Supply Chain, Healthcare, Real-Estate, Judiciary, IoT, Insurance, Government Sector, Energy and Utilities, Blockchain Integration with other Domains etc.		

Learning Resources

Textbooks

T1. Martin Quest, "Blockchain Dynamics: A Quick Beginner's Guide on Understanding the Foundations of Bit coin and Other Crypto currencies", Create Space Independent Publishing Platform, 15-May-2018.

T2. Imran Bashir, "Mastering Blockchain: Distributed Ledger Technology, Decentralization and Smart Contracts Explained", Second Edition, Packt Publishing, 2018.

T3. Alex Leverington, "Ethereum Programming", Packt Publishing, 2017.

T4. Elaine Shi, "Foundations of Distributed Consensus and Blockchains", http://elaineshi.com/docs/blockchain-book.pdf, 2020.

T5. Alan T. Norman, "Blockchain Technology Explained: the Ultimate Beginner's Guide About Blockchain Wallet, Mining, Bitcoin, Ethereum, Litecoin, Zcash, Monero, Ripple, Dash, IOTA and Smart Contracts", Amazon Digital Services, 2017.

Reference Books

R1. Bikramaditya Singhal, Gautam Dhameja, Priyansu Sekhar Panda, "Beginning Blockchain: A Beginner's Guide to Building Blockchain Solutions", 2018.

Artificial Intelligence & Data Science ★ KSE Pune

R2. Chris Dannen, "Introducing Ethereum and Solidity: Foundations of Cryptocurrency and Blockchain Programming for Beginners", Apress.
R3. Ritesh Modi, "Solidity Programming Essentials", Packt Publishing, 2018. R4. Chandramouli Subramanian, Asha A George, Abhilash K A and Meena Karthikeyan, "Blockchain Technology", Universities Press, ISBN-9789389211634.
R5. Bahga, Arshdeep, and Vijay Madisetti, "Blockchain applications: a hands-on approach", VPT, 2017.

Web Resources

1. Elaine Shi, "Foundations of Distributed Consensus and Blockchains", http://elaineshi.com/docs/blockchain-book.pdf , 2020.

Artificial Intelligence & Data Science
KSE
Pune

Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Elective I: Recommendation Systems (PEC-26506D-MTAD)		
Teaching Scheme: **Theory: 03 Hrs/Week**	**Credit Scheme:** **TH: 03**	**Examination Scheme:** **Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks** **Total Marks: 100 Marks**

Course Introduction:

Recommendation Systems is a specialized elective course designed to provide postgraduate students with a comprehensive understanding of the algorithmic frameworks, design paradigms, and evaluation methodologies used to build modern personalization engines. Over a five-unit structured path, the course details the foundational goals and domain-specific challenges of recommender systems, proceeding to neighborhood-based and model-based collaborative filtering techniques (including clustering, dimensionality reduction, SVD, latent factor models, and regression-based modeling). Students explore content-based systems, knowledge-based personalization (constraint and case-based architectures), and the design of hybrid models using ensemble-based classification perspectives. Finally, the curriculum covers experimental evaluation design, accuracy metrics in offline testing, and operational limitations, preparing graduates to engineer and optimize scalable recommender platforms for real-world industries.

Prerequisites:

- Learning Methods for AI

Course Objectives: The objective of this course is:

1. Provide knowledge of Basic Models of Recommender Systems.
2. Elaborate Model-Based Collaborative Filtering.
3. Explain Content-Based and Knowledge-Based Recommender Systems.
4. Describe for building a recommender system for Movies.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Understand Basic Models of Recommender Systems
CO2: Explain Model-Based Collaborative Filtering.
CO3: Analyze Content-Based and Knowledge-Based Recommender Systems.
CO4: Build a Recommender Systems for Movies.

Course Contents

Unit I	Introduction to Recommender Systems	08 Hours

Introduction, Goals of Recommender, Basic Models of Recommender Systems, Domain-Specific Challenges in Recommender Systems, Advanced Topics and Applications.

Artificial Intelligence & Data Science ★ KSE Pune

Unit II	IoT: Neighborhood-Based and Model-Based Collaborative Filtering	08 Hours
Introduction, Key Properties of Ratings, Predicting Ratings with Neighborhood-Based Methods, Clustering and Neighborhood-Based Methods, Dimensionality Reduction and Neighborhood, A Regression Modeling View of Neighborhood, Graph Models for Neighborhood-Based Methods, Decision and Regression Trees, Rule-Based Collaborative Filtering, Naive Bayes Collaborative Filtering, Using an Arbitrary Classification Model as a Black-Box, Latent Factor Models, Integrating Factorization and Neighborhood Models.		
Unit III	**Content-Based and Knowledge-Based Recommender Systems**	**08 Hours**
Content-Based and Knowledge-Based Recommender Systems, Basic Components of Content-Based Systems, Preprocessing and Feature Extraction, Learning User Profiles and Filtering, Content-Based Versus Collaborative, Using Content-Based Models for Collaborative Filtering. Constraint-Based Recommender, Case-Based Recommenders, Persistent Personalization in Knowledge-Based Systems.		
Unit IV	**Ensemble-Based and Hybrid Recommender Systems**	**08 Hours**
Ensemble Methods from the Classification Perspective, Weighted Hybrids, Switching Hybrids, Cascade Hybrids, Feature Augmentation Hybrids, Meta-Level Hybrids, Feature Combination Hybrids, Mixed Hybrids.		
Unit V	**Evaluating Recommender Systems**	**08 Hours**
Evaluation Paradigms, General Goals of Evaluation Design, Accuracy Metrics in Offline Evaluation, Limitations of Evaluation Measures.		

Learning Resources

Textbooks

T1. Charu C. Aggarwal, "Recommender Systems", Springer, 2016, ISBN 978-3-3-319-29657-9, ISBN 978-3-319-29659-3 (eBook), DOI 10.1007/978-3-319-29659-3.

T2. Dietmar Jannach, Markus Zanker, Alexander Felfernig and Gerhard Friedrich, "Recommender Systems: An Introduction", Infinity Science Press, 2008 ISBN: 978-1-934015-16-2.

Reference Books

R1. Francesco Ricci, Lior Rokach, Bracha Shapira and Paul B. Kantor, "Recommender Systems Handbook", Springer, ISBN 978-0-387-85819-7, e-ISBN 978-0-387-85820-3, DOI 10.1007/978-0-387-85820-3.

Artificial Intelligence & Data Science ★ KSE Pune

Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Skill Based Laboratory - I (LC-26507-MTAD)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Practical: 02 Hrs/Week	**PR/OR:** 1	**Term work:** 25 Marks **Oral:** 25 Marks **Total Marks:** 50 Marks

Course Introduction:

Skill Based Laboratory are based on the electives chosen by the students (Cloud Computing, Deep Learning, Blockchain Technology, Recommendation Systems)

Prerequisites:

- Cloud Technology
- Deep Learning
- Blockchain Technology
- Recommendation Systems

Course Objectives: The objective of this course is:

1. To introduce hands-on techniques for provisioning, managing, and auto-scaling resources in commercial cloud platforms (AWS, GCP, and Kubernetes).
2. To build deep learning models for classification and sequence analysis using advanced Python libraries and transfer learning.
3. To develop, compile, and interact with smart contracts on both local and distributed test network blockchain architectures.
4. To implement neighborhood-based, content-based, and matrix factorization recommendation pipelines using vector space similarity metrics.
5. To execute an end-to-end engineering mini-project that models and solves a real-world enterprise problem.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Build, orchestrate, and deploy robust, containerized, and auto-scaled microservices across cloud environments (AWS, GKE, Docker).

CO2: Design, evaluate, and optimize deep neural networks (CNNs, LSTMs) utilizing transfer learning for image and text classification datasets.

CO3: Program and audit functional smart contracts and decentralized applications (DApps) with immutable record-keeping.

CO4: Implement personalizing recommender models using collaborative filtering and Singular Value Decomposition (SVD) matrix factorization.

CO5: Author clean, modular codebase for a functional, elective-based mini-project and defend technical design choices in a professional peer review.

Guidelines:

- Skill Based Laboratory are based on the electives chosen by the students (Cloud Computing, Deep Learning, Blockchain Technology, Recommendation Systems)
- Respective course faculty assigned will frame course objectives, course outcomes, list of assignments, mini projects and references
- The list of assignments should be prepared and approved by the respective heads of Department before conduction of laboratory



Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

List of Assignments

Part A: Cloud Computing

1. Create a virtual machine on AWS EC2, launch an EC2 instance and install a web server Apache/Nginx to deploy a static website using tools - AWS Free Tier, EC2, SSH.
2. Manage a cloud storage with Amazon S3 / Google Cloud Storage and upload, retrieve, manage files using CLI and SDKs in S3 buckets using tools - AWS CLI / Google Cloud CLI, Python SDK.
3. Deploy dockerized App on Google Kubernetes Engine to containerize a Python/Flask app and deploy it on GKE with auto-scaling enabled using tools like - Docker, Kubernetes, GCP

Part A - Deep Learning

1. Write a program to build CNN model for Image Classification on the CIFAR-10/MNIST datasets using tools - TensorFlow/Keras or PyTorch.
2. Write a program to build LSTM model to classify movie reviews - positive/negative for sentiment analysis using LSTM using tools - Keras, NLTK, IMDB datasets.
3. Write a program to build and fine-tune a pretrained CNN model for a custom image classification datasets for transfer learning with Pretrained ResNet/VGG using tools - PyTorch/TensorFlow.

Part C - Learning Methods for AI (Any TWO)

1. Support vector machine for multiclass classification using Python: Implement a classification model using Support vector machine algorithm for news article topic classification like science, technology, politics, sports, etc. Use sklearn dataset and/or fetch dataset using APIs.
2. Write a program for classification using Decision Trees and Random Forest. Train models on the Titanic dataset and evaluate using accuracy, precision, recall using tools Scikit-learn, Pandas.
3. Convolutional Neural Network for face recognition using Python: Implement a real-world application like AI Face Recognition Entry Agent using CNN. It detects and recognizes faces in real-time environment. It should grant or deny the entry based on known faces of pre-trained databases.
4. Write a program for K-Means Clustering for Customer Segmentation. Apply K-Means on customer data to identify distinct customer groups based on features like spending, income, etc. using Python, Scikit-learn, Seaborn.

Part A - Blockchain Technology

1. Create a basic blockchain with blocks, proof-of-work, and hashing mechanism using Python.
2. Develop a smart contract to write, deploy and interact with smart contract - voting / wallet using tools Solidity, Remix IDE, MetaMask.
3. Set up a private Ethereum blockchain and deploy a smart contract using Truffle using tools Ganache, Truffle Suite.



Part A - Recommendation Systems
1. Write a program for movie recommendation system using collaborative filtering and cosine similarity. Use user rating matrix and similarity measures and tools - Python, Pandas, Scikit-learn. 2. Write a program for articles / movies recommendation using content-based filtering and TF-IDF similarity matrix. Use tools - Scikit-learn, Pandas. 3. Write a program for matrix factorization using SVD to reduce dimensionality and improve recommendation quality using tools- Python, Surprise library.
Part B - Mini-Projects
1. **Cloud Computing:** Multi-Tier Web Application Deployment on AWS Cloud: Design a full-stack cloud-native app with real-world deployment and deploy a secure, scalable cloud-based web app using tools - EC2, RDS, S3. 2. **Deep Learning:** AI-based Face Mask Detection System using Deep Learning: Design a deep learning model - CNN to detect face masks in real-time video inference using tools - PyTorch/TensorFlow. (Instructor can select any DL models for any real-world problems). 3. **Blockchain Technology:** Blockchain-based Certificate Issuing System: Design DApp architecture with immutability and trust features. Build a DApp that stores and verifies academic certificates securely on blockchain. 4. **Recommendation Systems:** Personalized Book Recommendation System: Develop a hybrid recommender system to handle data preprocessing, model evaluation and hybrid model integration that combines content and collaborative filtering for personalized book suggestions.
Learning Resources
Textbooks: T1. Toby Velte, Anthony Velte, Robert Elsenpeter, Robert C. Elsenpeter, "Cloud Computing, A Practical Approach", McGraw Hill LLC, ISBN: 9780071626958, 2009. T2. François Chollet, "Deep Learning with Python", Manning Publications, Second Edition, 2021. T3. Chris Dannen, "Introducing Ethereum and Solidity: Foundations of Cryptocurrency and Blockchain Programming for Beginners", Apress, 2017. T4. Charu C. Aggarwal, "Recommender Systems", Springer, 2016, ISBN 978-3-3-319-29657-9.



Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

First Year M.Tech. Artificial Intelligence and Data Science (AI&DS)
Curriculum with effect from A.Y. 2026-27
Semester – II

Explainable Artificial Intelligence (PCC-26508-MTAD)		
Teaching Scheme: **Theory:** 04 Hours/Week	**Credit Scheme:** **TH:** 04	**Examination Scheme:** **Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks** **Total Marks: 100 Marks**

Course Introduction:

Explainable Artificial Intelligence (XAI) focuses on the principles, techniques, and tools for making AI and machine learning models transparent, interpretable, and trustworthy. The course covers interpretable machine learning, model-specific and model-agnostic methods, explainability in deep learning and NLP, and visual explanation techniques such as LIME, SHAP, Grad-CAM, PDP, and saliency maps. It also addresses the ethical, legal, and practical aspects of XAI and the evaluation of explanations.

Prerequisites:

- Machine Learning

Course Objectives: The objective of this course is:

1. Understand the importance and need for interpretability in AI systems.
2. Explore techniques for explainability in classical ML and deep learning models.
3. Explore explainability in deep learning and NLP models, including attention-based and saliency based techniques.
4. Gain hands-on experience with XAI tools and libraries.
5. Examine case studies and applications in healthcare, finance, and autonomous systems.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Justify the need for explainability in AI systems.

CO2: Implement and compare various XAI methods.

CO3: Analyze deep learning models using explainability tools such as saliency maps and attention visualizations.

CO4: Apply XAI in real-world domains like healthcare, finance, legal tech, etc.

CO5: Apply visual explanations techniques such as PDP, ICE, Surrogate Models, and Feature Importance.

Course Contents

Unit I	**Introduction to XAI**	**12 Hours**

Artificial Intelligence & Data Science ★ KSE Pune

Introduction to the multidisciplinary topics of explainable AI, what is XAI, why is it important, XAI terminologies Dimensions of interpretability, Ethical and legal implications (GDPR, etc.), Types of explainability and Taxonomy, Explainability in the model development process.		
Unit II	**Interpretable Machine Learning**	**12 Hours**
Broad taxonomy of XAI methods including Intrinsic vs post hoc, model-specific vs. model-agnostic, and local vs. global. Perturbation-based feature attribution (LIME, Anchors, SHAP, etc.) Propagation-based feature attribution methods (e.g., Relevance Propagation, Saliency methods, etc.) Tools: ELI5, LIME		
Unit III	**Explainability in Deep Learning**	**12 Hours**
Feature visualization, Saliency maps, Grad-CAM, and LRP, Attention mechanisms, Concept activation vectors (TCAV) XAI for Natural Language Processing: Attention and explainability in Transformers, Attribution methods for NLP, Challenges in interpreting language models		
Unit IV	**Intrinsically Explainable Models**	**12 Hours**
Intrinsically explainable models including Linear Regression, Logistic Regression, Generalized Linear Model (GLM), Generalized Additive Model (GAM), and Decision Tree.		
Unit V	**Visual Explanations**	**12 Hours**
XAI methods include Partial Dependence Plot (PDP), Conformal Prediction, Individual Conditional Expectation (ICE), Feature Importance, and Saliency Maps. Local Interpretable Model-Agnostic Explanations (LIME), SHAP, Integrated Gradient (IG) Evaluation of explainability		

Learning Resources

Textbooks

T1. Interpretable Machine Learning" by Christoph Molnar (open access)
T2. Explainable AI: Interpreting, Explaining and Visualizing Deep Learning" edited by Samek et al.

Reference Books

R1. Explainable AI for Practitioners", by Michael Munn, David Pitman. Released October 2022 Publisher(s): O'Reilly Media, Inc.ISBN: 9781098119133.
R2. Explainable AI: Interpreting, Explaining and Visualizing Deep Learning, Lecture Notes in Artificial Intelligence, Springer Nature, 2019.
R3. Denis Rothman "Hands-On Explainable AI (XAI with Python", Packt Publishing, 2020.
R4. Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way" By Virginia Dignum.
R5. "Explainable AI: Foundations, Developments, Prospects and Challenges" by Wilfried Grossmann.



Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

AI in Cyber Security (PCC-26509-MTAD)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory: 04 Hours/Week	**TH:** 04	**Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks** **Total Marks: 100 Marks**

Course Introduction:

Artificial Intelligence in Cyber Security focuses on the application of AI, machine learning, deep learning, and NLP techniques for detecting and preventing cyber threats. The course covers cyber threat detection, intrusion detection, malware and phishing analysis, security log analysis, AI-based threat intelligence, adversarial machine learning, privacy-preserving AI, and AI-driven security systems such as SIEM and UEBA.

Prerequisites:

- Machine Learning

Course Objectives: The objective of this course is:

1. To introduce AI techniques and their applicability in cyber security systems.
2. To enable the use of machine learning and deep learning for real-time threat detection.
3. To equip students with the skills to design intelligent, secure systems using data science.
4. To understand adversarial attacks, defense strategies, and AI model robustness.
5. To explore the ethics, privacy, and compliance aspects of AI in security.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Analyze cyber threats using AI/ML models.
CO2: Apply supervised and unsupervised learning for anomaly and intrusion detection.
CO3: Develop AI-based models for malware, phishing, and fraud detection.
CO4: Evaluate adversarial machine learning challenges and secure model training.
CO5: Demonstrate understanding of ethical and privacy issues in AI-driven cyber security system.

Course Contents

Unit I	Foundations of AI in Cybersecurity	12 Hours

Introduction to Cybersecurity, threats, vulnerabilities, attacks, role of AI in cyber privacy and security, cybersecurity data, Logs, events, network flows, user behavior, Introduction to AI/ML/DL for Cybersecurity, Cyber threat detection, imbalanced data, concept drift, real-time requirements.

Unit II	Machine Learning for Cyber Threat Detection	12 Hours



Supervised Learning, SVM, Decision Trees, Naive Bayes, Random Forest, unsupervised learning K-Means, DBSCAN, autoencoders, semi-supervised learning, reinforcement learning in Security, feature Engineering for security datasets. intrusion detection and Prevention using ML		
Unit III	**Deep Learning and NLP in Cybersecurity**	**12 Hours**
Overview of DL and NLP, CNN & RNN for packet analysis and malware detection, LSTM for Log sequence analysis, transformer models like BERT, GPT for threat intelligence, NLP for phishing email, URL detection, dark web monitoring, applications for spam filtering, threat detection, limitations of deep models in Security.		
Unit IV	**Adversarial AI & Model Security**	**12 Hours**
Adversarial machine learning, attack surface in AI models, attack types - evasion, poisoning, inference attacks, Adversarial training, robust ML, privacy-preserving AI, Differential Privacy, security in AI Pipelines - data integrity, model hardening, Explainability in AI security models.		
Unit V	**AI for Cyber Security**	**12 Hours**
Adversarial attacks on ML systems, model poisoning, Security Information and Event Management(SIEM), Phishing Detection, User and Entity Behavior Analytics(UEBA) Behavior Analytics, Malware Analysis, Security Log Analysis, End Point Security, Benefits and Challenges of using AI in cyber security		

Learning Resources

Textbooks

T1. Artificial Intelligence for Cybersecurity: Techniques, Applications, and Challenges, Springer.

T2. Hands-On Machine Learning for Cybersecurity, Soma Halder & Sinan Ozdemir.

Reference Books

R1. Machine Learning and Security, Clarence Chio, David Freeman, O'Reilly.

R2. Cyber Security Essentials, James Graham, Richard Howard and Ryan Otson, CRCPress.

R3. Introduction to Cyber Security, Chwan-Hwa(john) Wu,J. David Irwin, CRC Press T&FGroup.

R4. Introduction to Cyber Security: Jatindra Pandey.

Web Resources

1. https://www.nist.gov/cyberframework
2. Introduction to Cyber Security. http://uou.ac.in/foundation-course
3. Fundamentals of Information Security http://uou.ac.in/progdetail?pid=CEGCS-17
4. Cyber Security Techniques. http://uou.ac.in/progdetail?pid=CEGCS-17
5. Cyber Attacks and Counter Measures: User Perspective http://uou.ac.in/progdetail?pid=CEGCS17
6. Information System. http://uou.ac.in/progdetail?pid=CEGCS-17

Artificial Intelligence & Data Science
KSE
Pune

Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Data Visualization Techniques (PCC-26510-MTAD)		
Teaching Scheme: **Theory:** 04 Hours/Week	**Credit Scheme:** **TH:** 04	**Examination Scheme:** **Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks** **Total Marks: 100 Marks**

Course Introduction:

Data Visualization Techniques focuses on the principles, methods, and tools used to represent and analyze numerical, non-numerical, multidimensional, and graph-based data through effective visualizations. The course covers data mapping, charts, glyphs, parallel coordinates, graph and network visualization, dimensionality reduction, dashboards, visualization system design, and case studies for analytical decision-making

Prerequisites:

- Data Science

Course Objectives: The objective of this course is:

1. Know the art and science of turning data into readable graphics.
2. Design and create data visualizations based on data available and tasks to be achieved,
3. Know the processes include data modeling, data processing, mapping data attributes to graphical attributes, and strategic visual encoding based,
4. Learn and evaluate the effectiveness of visualization designs, and think critically about each design decision, such as choice of color and choice of visual encoding.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Familiarize with the fundamentals of Data Visualization Techniques.
CO2: Perform data visualizations based on data available and tasks to be achieved.
CO3: Apply data modeling, data processing, mapping data attributes and strategic visual encoding.
CO4: Distinguish color and choice of visual encoding.
CO5: Evaluate the effectiveness of visualization designs in real life problem.

Course Contents

Unit I	Fundamentals of Data Visualization Techniques	12 Hours

Foundation: Importance of analytics and visualization in the era of data abundance, 2-D Graphics, 2-D Drawing, 3-D Graphics, Photorealism, Non-Photorealism, The Human Retina, Perceiving Two Dimensions, Perceiving Perspective.

Unit II	Visualization of Numerical Data	12 Hours

Visualization of Numerical Data: Data Mapping, Charts, Glyphs, Parallel Coordinates, Stacked Graphs, Tufte's Design Rules, Using Colors.



Unit III	Visualization of Non-Numerical Data	12 Hours
Visualization of Non-Numerical Data: Graphs and Networks, Embedding Planar Graphs, Graph Visualization, Tree Maps, Principal Component Analysis, Multidimensional Scaling, Packing.		
Unit IV	**Visualization Dashboard**	**12 Hours**
Visualization Dashboard: Visualization Systems, Database Visualization, Visualization System Design.		
Unit V	**Analysis Case Studies**	**12 Hours**
Graph-Theoretic Scagnostics, VisDB, Hierarchical Clustering Explorer, PivotGraph, InterRing Constellation Application of models in Data Visualization Techniques.		

Learning Resources

Textbooks

T1. T. Munzner, Visualization Analysis and Design, CRC Press, 2015.
T2. Edward Tufte, The Visual Display of Quantitative Information (2nd edition), Graphics Press
T3. Colin Ware, Information Visualization: Perception for Design (2nd edition), Morgan Kaufmann
T4. Alberto Cairo, The Functional Art: An Introduction to Information Graphics and Visualization, New Riders, Pearson Education.
T5. Nathan Yau, Data Points: Visualization That Means Something, Wiley
T6. Charles D. Hansen and Chris R. Johnson, Visualization Handbook, Academic Press.
T7. Will Schroeder, Ken Martin, and Bill Lorensen, The Visualization Toolkit: An Object-Oriented Approach to 3D Graphics, Kitware Inc. Publisher

Reference Books

R1. Nathan Yau, Data Points: Visualization That Means Something, Wiley.
R2. Charles D. Hansen and Chris R. Johnson, Visualization Handbook, Academic Press.
R3. Will Schroeder, Ken Martin, and Bill Lorensen, The Visualization Toolkit: An Object-Oriented Approach to 3D Graphics, Kitware Inc. Publisher



Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Computational Laboratory-II (LC-26511-MTAD)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Practical: 04 Hrs/Week	PR/OR: 2	Term work: 25 Marks Oral: 25 Marks Total Marks: 50 Marks

Course Objectives: The objective of this course is:

1. Understand and apply ML techniques to improve transparency and trust in AI systems.
2. Apply explainable AI techniques to understand decision making models.
3. Visually explore and interpret data using various visualization tools and techniques.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Understand feature importance and model behavior on real-world datasets using tools.
CO2: Develop AI-based systems to identify suspicious activity in real-world scenarios.
CO3: Create charts and visual summaries using tools like Matplotlib, Plotly, or Excel

Guidelines

- Computational Laboratory-I assignments are based on 3 Programme Core courses.
- Respective course faculty assigned will frame course objectives, course outcomes, list of assignments, mini projects, if any.
- The list of sample assignments is given for preparation of assignments which will be approved by the respective heads of Department before conduction of laboratory.

List of Assignments

Part A: Explainable Artificial Intelligence (Any One)

1. Write a program to generate and interpret the feature importance charts to explain model behavior using global explainability for health datasets
2. Write a program to Interpret House Price Predictions using regression mode / XGBoost. Use SHAP values to explain price predictions and housing datasets.

Part B: AI for Cyber Security (Any Two)

1. Write a program to distinguish between spam and non-spam emails using Naive Bayes / Logistic Regression model spam detection dataset - SMS Spam Collection.
2. Write a program to detecting unusual login behavior of a model and secure systems from unauthorized access using dataset with login times, user IDs, and IP addresses.

Part C: Data Visualization Techniques (Any Two)

1. Loading a Sample Dataset and Calculating the Mean using NumPy
   a) to compute the Mean, Median, Variance, and Standard Deviation of a Dataset
   b) Use the features to index, slice, split, and iterate.
2. Loading a Sample Dataset and Calculating the Mean using Pandas
   c) to Compute the Mean, Median, and Variance of a Dataset
   d) use the features to index, slice, split, and iterate



3. Visualizing Stock Trends by Using a line plot, basic plots
4. Analyzing Visualizations for choosing the best suitable plot for a scenario.



Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Elective II: Social Media Analytics (PEC-26512A-MTAD)		
Teaching Scheme: **Theory:** 03 Hours/Week	**Credit Scheme:** **TH:** 03	**Examination Scheme:** **Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks** **Total Marks: 100 Marks**

Course Introduction:

Social Media Analytics focuses on the techniques used to collect, preprocess, analyze, and interpret social media data from different platforms. The course covers social media data acquisition, text analytics and NLP, sentiment and topic analysis, social network analysis, community detection, influence and virality analysis, visualization, and applications in areas such as healthcare, politics, disaster management, and agriculture, along with ethical and privacy considerations.

Prerequisites:

- Data Science

Course Objectives: The objective of this course is:

1. To understand the structure, sources, and scale of social media data.
2. To explore methods for collecting, processing, and analysing textual and multimedia social data.
3. To study the use of machine learning and NLP techniques in sentiment and opinion mining.
4. To analyse user interactions and social networks using graph theory concepts.
5. To apply analytics techniques to real-world social media case studies for business, healthcare, disaster management, etc.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Explain social media ecosystems and the characteristics of social data.
CO2: Describe data collection mechanisms and pre-processing techniques.
CO3: Apply theoretical models for sentiment analysis, topic modelling, and opinion mining. •
CO4: Analyse social networks and user behaviours using graph-based metrics.
CO5: Evaluate the impact of social media analytics in various domains with ethical awareness

Course Contents

Unit I	Foundations of Social Media Analytics	08 Hours

Introduction to Social Media, Web platforms and world wide websites (www), types of social media data, Text, Image, Video, Metadata, Social media data characteristics, Volume, Velocity, Variety, Veracity, challenges in social media data analysis, SMA applications, marketing, politics, healthcare, disaster management, SMA architecture, ingestion, storage, analysis, visualization, data ecosystem, analytics opportunities across platforms, domain-specific applications of SMA

Artificial Intelligence & Data Science ★ KSE Pune

Unit II	Data Acquisition and Preprocessing	08 Hours
Overview of social media APIs, Facebook Graph, Reddit, X-Twitter, Instagram, structure and access policies, web scraping, architecture, ethics, and methods, text preprocessing, tokenization, case folding, stopword removal, stemming, lemmatization, handling multilingual text, hashtags, mentions, emojis, metadata, data challenges, spam, bots, noise, redundancy, techniques for extracting and preparing social media data, common challenges in raw social data.		
Unit III	**Text Analytics and Natural Language Processing**	**08 Hours**
Text representation, Bag-of-Words, TF-IDF, Word Embeddings, Sentiment Analysis, Rule-based approaches, ML-based approaches, topic modeling, Latent Dirichlet Allocation (LDA), Non-negative Matrix Factorization (NMF), Named Entity Recognition (NER), Overview of deep learning approaches, BERT, transformers, NLP techniques to sentiment and topic analysis, extracting insight from text data and evaluation.		
Unit IV	**Social Network and User Interaction Analysis**	**08 Hours**
Introduction to graph theory, nodes, edges, degree, centrality, social graphs, user mentions, replies, retweets, network metrics, Betweenness, Closeness, PageRank, community detection: modularity, Louvain algorithm, Influence and virality analysis, Behavioural analytics: likes, shares, content engagement patterns, graph-based representation of social networks		
Unit V	**Applications of social media**	**08 Hours**
Applications of Politics - Election sentiment, propaganda detection, Healthcare - Vaccine hesitancy, pandemic trend tracking, Disaster response - Earthquake, flood, or COVID alerting via social media, Agriculture - Farmer sentiment, feedback analytics, Visual analytics: Word clouds, sentiment timelines, conceptual network diagrams, Ethics - data privacy, fake news, bots, bias, misinformation, real-world applications of SMA, ethical and responsible use of social media data		

Learning Resources

Textbooks

T1. Gupta, Mukul, Gupta, Deepa, Gupta, Parth Mukul, "Social Media and Web Analytics: Turning Insights into Action in a Digital World", PHI publisher, ISBN: 9789354439605, 2024.

T2. Matthew Ganis, Avinash Kohirkar, "Social Media Analytics: Techniques and Insights for Extracting Business Value Out of Social Media", IBM Press

Reference Books

R1. Subodha Kumar, Liangfei Qiu, "Social Media Analytics and Practical Applications: The Change to the Competition Landscape", 1st Edition, CRC Press,

R2. Marshall Sponder, "Social Media Analytics", McGraw-Hill Publisher, ISBN: 2800071768292, 2011.

Web Resources

1. https://nibmehub.com/opac-service/pdf/read/social%20media%20analytics%20strategy%20%20usi

Artificial Intelligence & Data Science
KSE Pune

Massive Open Online Courses (MOOCs)

1. Alex Gonçalves, "Social Media Analytics Strategy Using Data To Optimize Business Performance", Apress publisher, ISBN: 978-1-4842-3102-9.
2. https://nibmehub.com/opac-service/pdf/read/social%20media%20analytics%20strategy%20%20usi



Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Elective II: Quantum Computing (PEC-26512B-MTAD)		
Teaching Scheme: **Theory:** 03 Hours/Week	**Credit Scheme:** **TH:** 03	**Examination Scheme:** **Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks** **Total Marks: 100 Marks**

Course Introduction:

Quantum Computing focuses on the fundamental principles of quantum computation and its transition from classical computing to quantum systems. The course covers qubits, quantum states, quantum gates, quantum circuits, quantum algorithms, quantum programming, quantum cryptography, quantum information theory, and quantum computing hardware, along with emerging applications and challenges.

Prerequisites:

- Computer Organization and Architecture

Course Objectives: The objective of this course is:

1. Being able to analyze simple quantum algorithms and argue optimality.
2. Familiarity with 1-qubit / 2-qubit gate operators and ability to design simple quantum circuits.
3. Ability to read and understand recent results as well as research papers on quantum algorithms.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Explain how the physics of quantum computation is different from classical computational models.

CO2: Describe the theoretical performance improvements that quantum algorithms offer compared to classical algorithms.

CO3: Analyze the life cycle of hybrid applications and decompose their execution on a hybrid quantum-classical computational continuum.

CO4: Develop their own (hybrid) quantum algorithms and implement them using (real or simulated) quantum computers using quantum toolkits such as Qiskit.

Course Contents

Unit I	**Introduction and The Leap from Classical to Quantum**	**08 Hours**

Introduction: cbits and qubits, Reversible and Manipulating operations on cbits and qubits, Circuit diagrams, Measurement gates and state preparation, Constructing arbitrary 1- and 2-Qbit states The Leap from Classical to Quantum: Classical Deterministic Systems, Probabilistic Systems, Quantum Systems, Assembling Systems

Unit II	**Basic Quantum Theory & Architecture**	**08 Hours**

Artificial Intelligence & Data Science ★ KSE Pune

Basic Quantum Theory: Quantum States, Observables, Measuring, Dynamics, Assembling Quantum Systems Architecture: Bits and Qubits, Classical Gates, Reversible Gates, Quantum Gates		
Unit III	**Algorithms & Programming Languages**	**08 Hours**
Algorithms: Deutsch's Algorithm, The Deutsch–Jozsa Algorithm, Simon's Periodicity Algorithm, Grover's Search Algorithm, Shor's Factoring Algorithm, Programming Languages: Programming in a Quantum World, Quantum Assembly Programming, Toward Higher-Level Quantum Programming, Quantum Computation Before Quantum Computers.		
Unit IV	**Theoretical Computer Science and Cryptography**	**08 Hours**
Theoretical Computer Science: Deterministic and Nondeterministic Computations, Probabilistic Computations, Quantum Computations Cryptography: Classical Cryptography, Quantum Key Exchange I: The BB84 Protocol, Quantum Key Exchange II: The B92 Protocol, Quantum Key Exchange III: The EPR Protocol, Quantum Teleportation.		
Unit V	**Information Theory & Hardware**	**08 Hours**
Information Theory: Classical Information and Shannon Entropy, Quantum Information and von Neumann Entropy, Classical and Quantum Data Compression, Error-Correcting Codes Hardware: Quantum Hardware: Goals and Challenges, Implementing a Quantum Computer I: Ion Traps, Implementing a Quantum Computer II: Linear Optics, Implementing a Quantum Computer III: NMR and Superconductors, Future of Quantum Ware.		

Learning Resources

Textbooks

T1 Eleanor Rieffel and Wolfgang Polak, "Quantum Computing: A Gentle Introduction," The MIT Press, 2011, ISBN 978-0-262-01506-6

T2. Quantum Computation and Quantum Information by Nielsen and Chuang (NC), Classical and Quantum Computation by Kitaev, Shen, and Vyalyi (KSV)

T3. Scott Aaronson, "Quantum Computing Since Democritus," Cambridge University Press, March 2013, ISBN: 9780521199568

Reference Books

R1. N. David Mermin, "Quantum Computer Science: An Introduction," Cambridge University Press, August 2007, ISBN: 9780521876582.

R2. Noson S. Yanofsky and Mirco A. Mannucci, "Quantum Computing for Computer Scientists," Cambridge University Press, August 2008, ISBN: 9780521879965.

R3. Denis Rothman "Hands-On Explainable AI (XAI with Python", Packt Publishing, 2020.

R4. IBM Quantum Computing- Qiskit - ibm.com," https://www.ibm.com/quantum/qiskit

Artificial Intelligence & Data Science ★ KSE Pune

Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Elective II: Business Intelligence (PEC-26512C-MTAD)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory: 03 Hours/Week	**TH:** 03	**Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks** **Total Marks: 100 Marks**

Course Introduction:

Business Intelligence focuses on the methods and technologies used to collect, integrate, manage, analyze, and exploit business information for effective decision-making. The course covers business intelligence processes, business models, data warehouses, OLAP, metadata, data profiling, data quality, ETL, information integration, parallel processing, and the use of diverse information sources for business analysis.

Prerequisites:

- Linear algebra and Probability

Course Objectives: The objective of this course is:

1. To understand the fundamentals of Business Intelligence and its role in decision-making.
2. To analyze data warehouses, OLAP, data quality, and information integration techniques.
3. To apply data mining and knowledge discovery techniques for extracting useful business insights.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Critically appraise current theory and practice in Big Data Analytics, Decision Support Systems and Business Intelligence

CO2: Appraise the role of BI strategy in driving companies' insight.

CO3: Apply data analysis and visualization on a sample dataset, appropriate to the level of study, and to provide recommendations on the driven insights to specialist and non-Specialist audiences.

CO4: Identify and appraise emerging trends within the field and evaluate the social and ethical aspects.

CO5: Communicate complex topics and concepts effectively using different communication means such as report writing, use of ICT and/or presentation to specialist and non-specialist audience.

Course Contents

Unit I	Business Intelligence and Information Exploitation	08 Hours

Why Business Intelligence? The Information Asset, Exploiting Information, Business Intelligence and Program Success, what is Business Intelligence? Actionable Knowledge Business Case, Business Intelligence Process, System Infrastructure, Information Access, Delivery, and Analysis Services, Management Issues

Unit II	Business Models	08 Hours



Business Models and Information Flow: The Business Case, Information Processing and Information Flow, the Information Flow Model Usage in Practice Modeling Frameworks Manageme Data Warehouses, Online Analytical Processing, and Metadata: The Business Case, Data Models, The Data Warehouse, The Data Mart, Online Analytical Processing, Metadata, Management Issuesnt Issues.

Unit III	Data Quality and Information Compliance	08 Hours

Business Rules: Business Case, Business Rules Approach, What Is a Business Rule? What Is a Business Rule System? Sources of Business Rules, Management Issues Data Profiling: The Business Case, Data Profiling Activities, Data Model Inference, Attribute Analysis, Relationship Analysis, Management Issues Data Quality and Information Compliance: Business Case, More Than Just Names and Addresses, Types of Errors, Data Cleansing, Business Rule-Based Information Compliance, Management Issues.

Unit IV	Information Integration	08 Hours

Information Integration: The Business Case, ETL: Extract, Transform, Load, Enterprise Application Integration and Web Services, Record Linkage and Consolidation, Management Issues The Value of Parallelism: Business Case, Parallelism and Granularity, Parallel Processing Systems, Dependence, Parallelism and Business Intelligence, Management Issues Alternate Information Contexts: Business Case, Psychographics and Demographics, Geographic Data , Web Behavior Intelligence, Management Issue.

Unit V	Data Enhancement, Knowledge Discovery and Data Mining	08 Hours

Data Enhancement: The Business Case, Types of Data Enhancement, Incremental Enhancements, Batch Enhancements, Standardization Example: Address Standardization, Enhancement Methodologies Management Issues Knowledge Discovery and Data Mining: The Business Case, Data Mining and the Data Warehouse, The Virtuous Cycle, Directed versus Undirected Knowledge Discovery, Six Basic Tasks of Data Mining, Data Mining Techniques, Management Issues Learning Resources Using Publicly Available Data: Business Case, Management Issues, Public Data, Data Resources Semi Structured Data, The Myth of Privacy.

Learning Resources

Text Books

T1. Douglas K. Barry and David Loshin, Business Intelligence: The Savvy Manager's Guide, Morgan Kaufmann Publishers,2003, ISBN-13:978-1-55860-916-7

T2. Ahmed Sherif, Practical Business Intelligence, Packt Publishing Ltd.,2016, ISBN 978-1-78588 543-

Reference Books

R1. Sartaj Singh, Business Intelligence, Excel Books Private Limited, 2011



Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Elective II: Spatial Data Analytics (PEC-26512D-MTAD)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory: 03 Hours/Week	**TH:** 03	**Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks** **Total Marks: 100 Marks**

Course Introduction:

Spatial Data Analytics focuses on the principles and techniques used to collect, manage, analyze, and interpret spatial data using Geographic Information Systems (GIS). The course covers spatial data models, GIS analysis and modeling, raster and vector operations, terrain analysis, spatial interpolation, geostatistics, and spatial decision support systems using techniques such as AHP and multi-criteria decision making.

Prerequisites:

- Data Science

Course Objectives: The objective of this course is:

1. Explain Geographic Information Systems components.
2. Describe Geographic Information Systems Analysis and Modeling.
3. Elaborate Spatial Data Analysis Techniques.
4. Discuss Geostatistical Analysis Techniques.
5. Explore decision making process and decision support systems

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Understand Geographic Information Systems components.
CO2: Apply Geographic Information Systems Analysis and Modeling for real-time example.
CO3: Implement Spatial Data Analysis Techniques for chosen real-time example.
CO4: Analyse Geostatistical Analysis Techniques for chosen real-time example
CO5: Use decision making process to create real-time decision support system.

Course Contents

Unit I	Introduction to Spatial Data	08 Hours

Basic concepts: Definition and history, Components of GIS, Data structure and formats, Spatial data models – Raster and Vector, Data base design - editing and topology creation in GIS, Linkage between spatial and non-spatial data, Data inputting in GIS

Unit II	Introduction to GIS Analysis and Modeling	08 Hours

Spatial Data: Definition, Analysis, Processes & Steps, Software and Tools Raster-Based and VectorBased GIS Modeling, Binary Models, Index Models, Regression Models, Process Models Geodatabase Model, Role of Databases in GIS, Creating, Editing and Managing.



Unit III	Spatial Data Analysis Techniques	08 Hours
Classification Scheme of Vector-Based and Raster-Based GIS Operations Raster–Based Techniques: Methods of Reclassification, Overlay Analysis, Slope and Aspects, Buffering, Cost-Distance Calculation Vector-Based Techniques: Map Manipulation Techniques, Buffering, Overlay Analysis, Network Analysis Digital Terrain Analysis and Modeling: TIN and DEM, Surface Representation & Analysis		
Unit IV	**Geostatistical Analysis Techniques**	**08 Hours**
Introduction to Spatial Interpolation: Control Points Global Methods: Trend Surface Analysis, Regression Models Local Methods: Thiessen Polygons, Density Estimation, Inverse Distance Weighted Interpolation Kriging: Ordinary Kriging (Semivariance, Semivariogram), Universal Kriging.		
Unit V	**Decision Support System**	**08 Hours**
GIS and decision support system, Introduction to decision making process and decision support systems, Introduction of a frame work for planning and decision making, Spatial Decision Making, Development of DSS, DSS Architecture, Principles and components of multiplecriteria decision making, Main multiple-criteria evaluation methods/techniques, Spatial multiple criteria decision making, Multiple-criteria decision making in spatial data analysis Introduction to AHP, Basic Principles of AHP, Effect Table, Pair Wise comparison, Standardization, Consistency, Weightage, performance score, Different method in PWC.		

Learning Resources

Textbooks

T1. Burrough, Peter A. and Rachael McDonnell., (1998), Principles of Geographical Information Systems. Oxford University Press, New York.

T2. Laurini, Robert and Derek Thompson. , (1992), Fundamentals of Spatial Information Systems. Academic Pr., London.

T3. Kluwer Fotheringham A S, O'Kelly M E., (1998), Spatial Interaction Models: Formulations and Applications

T4.Paul Longley, Michael Goodchild, David Maguire and David Rhind:, (2005), Geographical Information Systems: Principles, Techniques, Applications and Management. John Wiley & Sons.

Reference Books

R1. Bonczek, R.H., C.W. Holsapple, and A.B. Whinston, (1981), Foundations of Decision Support Systems, Academic Press, New York.

R2. House, W.C. (ed.), (1983). Decision Support Systems, Petrocelli, New York.

R3. Sprague, R.H., and Carlson, E.D., (1982). Building Effective Decision Support Systems, PrenticeHall, Englewood Cliffs NJ.

Artificial Intelligence & Data Science
KSE
Pune

Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Elective III: Federated Learning (PEC-26513A-MTAD)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory: 03 Hours/Week	**TH:** 03	**Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks Total Marks: 100 Marks**

Course Introduction:

Federated Learning focuses on decentralized machine learning, where models are trained collaboratively across multiple devices or organizations without directly sharing their local data. The course covers federated learning architectures and algorithms, model aggregation, privacy and security techniques, federated learning frameworks, communication and optimization challenges, and real-world applications in healthcare, finance, smart devices, and agriculture.

Prerequisites:

- Machine Learning

Course Objectives: The objective of this course is:

1. To understand the principles and motivation behind federated learning.
2. To explore different architectures, algorithms, and challenges in federated learning.
3. To learn the role of federated learning in privacy-preserving AI and edge computing.
4. To analyze federated learning use cases in industry sectors like healthcare, IoT, and finance.
5. To discuss security threats, mitigation techniques, and ethical issues in federated learning.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Describe the fundamentals, types, and advantages of federated learning.

CO2: Analyze and compare federated learning architectures and optimization algorithms CO3: Identify privacy, security, and communication challenges in federated learning.

CO4: Apply federated learning concepts to real-world use cases and frameworks.

CO5: Evaluate federated learning performance under non- independent and identically distributed (non-IID) data and system heterogeneity

Course Contents

Unit I	Introduction to Federated Learning	08 Hours

Introduction to federated learning, need for federated learning, data privacy, decentralization, edge AI. Centralized vs decentralized vs federated paradigms, clients, server, rounds, model aggregation, Benefits, privacy-preserving, bandwidth-efficient, collaborative learning, federated learning life cycle. concept of drivers of federated learning, centralized ML and federated ML paradigms.

Unit II	Federated Learning Architectures and Algorithms	08 Hours

Architecture of federated learning, horizontal federated learning, vertical federated learning, federated transfer learning federated learning, federated averaging, algorithm and work flow,



federated SGD, advanced aggregation techniques, synchronous and asynchronous federated learning, optimization challenges with non- independent and identically distributed (non-IID) data, compression and quantization for efficient communication in federated learning, performance evaluation under various data and system conditions.		
Unit III	**Privacy and Security in Federated Learning**	**08 Hours**
Threats in federated learning, model inversion, poisoning, membership inference Differential Privacy (DP) and differentially private stochastic gradient descent (DP-SGD), Secure Aggregation: encryption and secure multiparty computation (SMC), homomorphic encryption and federated cryptographic techniques, anonymity and trust in decentralized settings, privacy-preserving techniques in federated models, security vulnerabilities and defenses in federated learning systems.		
Unit IV	**Federated Learning Systems and Frameworks**	**08 Hours**
Architecture of federated system: clients, coordinator, data flow. Open-source frameworks, TensorFlow Federated, PySyft, Flower, NVIDIA Clara. Resource constraints: computation, memory, communication. System-level challenges: stragglers, client selection, dropouts, Bench-marking and evaluation of federated system models, federated system development using tools and frameworks, system-level considerations in real-world deployment		
Unit V	**Applications of Federated Learning**	**08 Hours**
Real-world applications of federated system, Healthcare: cross-institutional model training, Finance: fraud detection. Smart devices: predictive keyboard, face unlock. Agriculture: collaborative crop disease detection. Ethical aspects: bias, fairness, transparency. federated reinforcement learning, personalized federated system, swarm learning, ethical and legal implications of deploying federated system.		

Learning Resources

Textbooks

T1. Heiko Ludwig, Nathalie Baracaldo, "Federated Learning: A Comprehensive Overview of Methods and Applications", Springer, 2022.

T2. Jayakrushna Sahoo, Mariya Ouaissa, Akarsh K. Nair, "Federated Learning: Principles, Paradigms, and Applications", 1st Edition,9781003497196, Apple Academic Press, 2024.

Reference Books

R1. George Jeno, "Federated Learning with Python: Design and implement a federated learning system and develop applications using existing frameworks", Packt Publishing, ISBN-13978-1803247106, 2022.

R2. Shelly Gupta, Puneet Garg, Jyoti Agarwal, Hardeo Kumar Thakur, Satya Prakash Yadav, "Federated Learning Based Intelligent Systems to Handle Issues and Challenges in Iovs- Part 1", ISBN: 978-981-5313-03-1 (Print), ISBN: 978-981-5313-02-4 (Online), 2024.

Artificial Intelligence & Data Science
KSE Pune

Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Massive Open Online Courses (MOOCs)

1. George Drosatos, Pavlos S. Efraimidis, Pavlos S. Efraimidis, Avi Arampatzis, "Federated and Transfer Learning Applications", ISBN 978-3-7258-0076-6, 2024. https://www.mdpi.com/books/reprinfederated-and-transfer-learning-applications



Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Elective III: Big Data Analytics (PEC-26513B-MTAD)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory: 03 Hours/Week	**TH:** 03	**Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks** **Total Marks: 100 Marks**

Course Introduction:

Big Data Analytics focuses on the concepts, technologies, and tools used to store, process, manage, and analyze large and complex datasets. The course covers Big Data fundamentals, Hadoop and HDFS, MapReduce, the Hadoop ecosystem including Pig, Hive, and HBase, and data analytics using R and Learning Methods for AI.

Prerequisites:

- Data Science

Course Objectives: The objective of this course is:

1. Understand the Big Data Platform and its Use cases.
2. Provide an overview of Apache Hadoop.
3. Provide HDFS Concepts and Interfacing with HDFS.
4. Understand Map Reduce Jobs and Provide hands on Hadoop Ecosystem.
5. Apply analytics on Structured, Unstructured Data.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Identify Big Data and its Business Implications.
CO2: List the components of Hadoop and Manage Job Execution in Hadoop Environment.
CO3: Access and Process Data on Distributed File System.
CO4: Develop Big Data Solutions and analyze Infosphere Big Insights.
CO5: Apply Learning Methods for AI using R.

Course Contents

Unit I	INTRODUCTION TO BIG DATA AND HADOOP	08 Hours

Types of Digital Data, Introduction to Big Data, Big Data Analytics, History of Hadoop, Apache Hadoop, Analysing Data with Unix tools, Analysing Data with Hadoop, Hadoop Streaming, Hadoop Echo System, IBM Big Data Strategy, Introduction to Infosphere Big Insights and Big Sheets.

Unit II	HDFS (Hadoop Distributed File System)	08 Hours

The Design of HDFS, HDFS Concepts, Command Line Interface, Hadoop file system interfaces, Data flow, Data Ingest with Flume and Scoop and Hadoop archives, Hadoop I/O: Compression, Serialization, Avro and File-Based Data structures.



Unit III	Map Reduce	08 Hours
Anatomy of a Map Reduce Job Run, Failures, Job Scheduling, Shuffle and Sort, Task Execution, Map Reduce Types and Formats, Map Reduce Features.		
Unit IV	**Hadoop Eco System**	**08 Hours**
Pig : Introduction to PIG, Execution Modes of Pig, Comparison of Pig with Databases, Grunt, Pig Latin, User Defined Functions, Data Processing operators. 47 Hive : Hive Shell, Hive Services, Hive Metastore, Comparison with Traditional Databases, HiveQL, Tables, Querying Data and User Defined Functions. Hbase : HBasics, Concepts, Clients, Example, Hbase Versus RDBMS. Big SQL		
Unit V	**Data Analytics with R**	**08 Hours**
Machine Learning: Introduction, Supervised Learning, Unsupervised Learning, Collaborative Filtering. Big Data Analytics with BigR.		

Learning Resources

Textbooks

T1. Tom White "Hadoop: The Definitive Guide" Third Edition, O'Reilly Media, 2012.
T2. Seema Acharya, Subhasini Chellappan, "Big Data Analytics" Wiley 2015.

Reference Books

R1. Michael Berthold, David J. Hand, "Intelligent Data Analysis", Springer, 2007
R2. Jay Liebowitz, "Big Data and Business Analytics" Auerbach Publications, CRC press (2013).
R3. Tom Plunkett, Mark Hornick, "Using R to Unlock the Value of Big Data: Big Data Analytics with Oracle R Enterprise and Oracle R Connector for Hadoop", McGraw-Hill/Osborne Media (2013), Oracle press.
R4. Anand Rajaraman and Jeffrey David Ulman, "Mining of Massive Datasets", Cambridge University Press, 2012.
R5. Bill Franks, "Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics", John Wiley & sons, 2012.
R6. Glen J. Myat, "Making Sense of Data", John Wiley & Sons, 2007 R7.
Pete Warden, "Big Data Glossary", O'Reily, 2011.
R8. Michael Mineli, Michele Chambers, Ambiga Dhiraj, "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses", Wiley Publications, 2013.
R9. ArvindSathi, "Big Data Analytics: Disruptive Technologies for Changing the Game", MC Press, 2012
R10. Paul Zikopoulos ,Dirk DeRoos , Krishnan Parasuraman , Thomas Deutsch , James Giles , David Corigan , "Harness the Power of Big Data The IBM Big Data Platform ", Tata McGraw Hill Publications, 2012.

Artificial Intelligence & Data Science
KSE
Pune

Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Elective III: Reinforcement Learning (PEC-26513C-MTAD)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory: 03 Hours/Week	**TH:** 03	**Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks** **Total Marks: 100 Marks**

Course Introduction:

Reinforcement Learning focuses on learning optimal decision-making through interaction between an agent and its environment. The course covers probability concepts, Markov Decision Processes, Bellman equations, dynamic programming, Monte Carlo and temporal-difference methods, SARSA and Q-Learning, function approximation, policy gradients, and actor-critic methods.

Prerequisites:

- Machine Learning

Course Objectives: The objective of this course is:

1. Explain dynamic programming-based methods.
2. Describe iteration and value iteration algorithms for control of MDP.
3. Elaborate incremental computation of Monte Carlo Methods, Temporal Difference (TD) methods.
4. Provide knowledge of different on and off policy evaluation and control methods with function approximation
5. Explain naive REINFORCE algorithm, different variance reduction techniques and Actor-Critic methods.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Evaluate policies using dynamic programing-based methods.
CO2: Understand policy iteration and value iteration algorithms for control of MDP.
CO3: Understand incremental computation of Monte Carlo Methods, different Temporal Difference (TD) methods.
CO4: Get familiar with different on and off policy evaluation and control methods with function approximation.
CO5: Get familiar with naive REINFORCE algorithm, different variance reduction techniques and Actor-Critic methods.

Course Contents

Unit I	Introduction of Reinforcement Learning	08 Hours

Origin and history of Reinforcement Learning research. Its connections with other related fields and with different branches of machine learning. Brush up of Probability concepts - Axioms of probability, concepts of random variables, PMF, PDFs, CDFs, Expectation. Concepts of joint and multiple random variables, joint, conditional and marginal distributions. Correlation and independence.



Unit II	Markov Decision Process	08 Hours
Introduction to RL terminology, Markov property, Markov chains, Markov reward process (MRP). Introduction to and proof of Bellman equations for MRPs along with proof of existence of solution to Bellman equations in MRP. Introduction to Markov decision process (MDP), state and action value functions, Bellman expectation equations, optimality of value functions and policies, Bellman optimality equations		
Unit III	**Prediction and Control by Dynamic Programing**	**08 Hours**
Overview of dynamic programing for MDP, definition and formulation of planning in MDPs, principle of optimality, iterative policy evaluation, policy iteration, value iteration, Banach fixed point theorem, proof of contraction mapping property of Bellman expectation and optimality operators, proof of convergence of policy evaluation and value iteration algorithms, DP extensions		
Unit IV	**Temporal Difference Methods for Model Free Prediction and Control**	**08 Hours**
Overview of Monte Carlo methods for model free RL, First visit and every visit Monte Carlo, Monte Carlo control, On policy and off policy learning, Importance sampling. Incremental Monte Carlo Methods for Model Free Prediction, Overview TD(0), TD(1) and TD(λ), k-step estimators, unified view of DP, MC and TD evaluation methods, TD Control methods - SARSA, Q-Learning and their variants.		
Unit V	**Function Approximation Methods**	**08 Hours**
Getting started with the function approximation methods, Revisiting risk minimization, gradient descent from Machine Learning, Gradient MC and Semi-gradient TD(0) algorithms, Eligibility trace for function approximation, Afterstates, Control with function approximation, Least squares, Experience replay in deep Q-Networks. Log-derivative trick, Naive REINFORCE algorithm, bias and variance in Reinforcement Learning, Reducing variance in policy gradient estimates, baselines, advantage function, actor-critic methods.		

Learning Resources

Textbooks

T1. Reinforcement Learning: An Introduction", Richard S. Sutton and Andrew G. Barto, 2nd Edition, The MIT Press, 2018.
T2. An Introduction to Deep Reinforcement Learning", Vincent François-Lavet, Peter Henderson, Riashat Islam, Marc G. Bellemare and Joelle Pineau,
T3. "Algorithms for Reinforcement Learning", Csaba Szepesvari

Reference Books

R1. "Probability, Statistics, and Random Processes for Electrical Engineering", 3rd Edition
R2. Alberto Leon-Garcia, Lecture Notes on Introduction to Probability, Dimitri P. Bertsekas and John N. Tsitsiklis

Artificial Intelligence & Data Science
KSE Pune

Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Elective III: Game Theory and Applications (PEC-26513D-MTAD)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory: 03 Hours/Week	**TH:** 03	**Comprehensive Continuous Evaluation (CCE): 50 Marks** **End Semester: 50 Marks** **Total Marks: 100 Marks**

Course Introduction:

Game Design and Artificial Intelligence focuses on the principles, processes, tools, and techniques used to design and develop interactive digital games. The course covers game architecture, game programming, user interface design, game design processes, 2D and 3D modelling, animation, and AI techniques for game development. It also introduces pathfinding, decision-making, decision trees, and strategy-based AI for creating intelligent and interactive game environments using tools such as Unity.

Prerequisites:

2. Artificial Intelligence

Course Objectives: The objective of this course is:

1. To understand the game concept and visualization.
2. To study user interface and various elements of game.
3. To understand game design process and apply to develop game applications.
4. To be able to design 2D and 3D Models using tools.
5. To study AI principles to problem solving, knowledge representation, learning.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1: Understand game logic and visualization
CO2: Understand user interface and various elements of game.
CO3: Understand the game design process and decision making.
CO4: Students will be able to design various 2D and 3D game characters.
CO5: Apply the principles of AI to problem solving, knowledge representation, learning.

Course Contents

Unit I	Fundamentals of Game Theory	08 Hours

Introduction to Game, Game History, Game Architecture, Game Logic, Game View for the Human Player, Game Views for AI Agents, game idea visualization and storytelling, and game documentation. Game Programming, Software Tools, Building the Game, creating a Project, Creating Build Scripts, Introduction to Unity, prototyping using Unity, Photoshop and flash.
Case study: Write a game architecture for target shooting game.

Unit II	Game UI Design Elements	08 Hours



Introduction to User Interface, Web and Graphic User Interfaces, Visual Design Basics, Usability, Visual and Interactive elements, buttons, icons, spacing, sliders and scrollbars, typography, color schemes, text style, shapes and layers to create screens, fundamental principles of visual design,pixel precision to UI elements. Case study:Design the UI for "candy catch game "and represent essential components.		
Unit III	**Game Design Process**	**08 Hours**
Pre-Production, Production, Post-Production, MDA Framework, Planning and Iteration: Under planning and Over planning, Iteration, Knowledge Creation: Knowledge Creation Methods, Rumination, Research, artistic methods, Brainstorming, written analysis, Debate, testing,metrics, invented methods, Organic Process, Dependencies:Dependency Stack, Cascading uncertainty, Design Backlog, Authority, Motivation:Extrinsic Rewards, meaningful work, expectations-Driven motivation, Complex Decisions: Decision Effects, Decision Effects Case Study, Values: Openness, Candor, Humility, Hunger.		
Unit IV	**2D & 3D Modeling for Game**	**08 Hours**
Definition of Computer-based Animation, Basic Types of Animation: Real Time, Non-real-time, Definition of Modelling. Understanding 2D Splines & shape using unity, Extrude & Bevel 2D object to 3D using Unity SNAPS, Understanding Loft & terrain, Creation of 3D objects. Modelling with Polygons, using Unity, Creating and Destroying Game Objects , the scene view, building simple and complex games in 2D and 3D using Unity. Case study: Design 2D and 3D face models.		
Unit V	**Game Applications**	**08 Hours**
Introduction to AI, Model of Game AI, Movement, Decision Making, Strategy, Infrastructure, Game AI, The Pathfinding Graphs, Weighted Graphs, Directed Weighted Graphs, Dijkstra Algorithm, Data Structures and Interfaces, Performance of Dijkstra, Decision Making, Decision Tree, Knowledge Representation, Implementation of Nodes, Performance of Decision Trees, Balancing the Tree, Random Decision Trees. Case study: AI in Video Games.		

Learning Resources

Textbooks

T1. Florian Bartholomae, Marcus Wiens, "Game Theory and Applications: A Guide for Students and Researchers", 1st edition, 2024.

T2. Dario Calonaci, Designing User Interfaces: Exploring User Interfaces, UI Elements, Design Prototypes and the Figma UI Design Tool, BPB Publisher, ISBN-13-978-9389898743, 2021.

Reference Books

R1. Michael E. Mortenson , 3D Modeling, Animation, and Rendering: An Illustrated Lexicon, Color Edition, Create Space Independent Publishing Platform Publisher, ISBN:1453728481, 2010

R2. Tatsuro Ichiishi, Abraham Neyman and Yair Tauman, "Game Theory and Applications ", Elsevier.

Artificial Intelligence & Data Science ★ KSE Pune

R3. Tom Plunkett, Mark Hornick, "Using R to Unlock the Value of Big Data: Big Data Analytics with Oracle R Enterprise and Oracle R Connector for Hadoop", McGraw-Hill/Osborne Media (2013), Oracle press.

Massive Open Online Courses (MOOCs)

1. Mike MrMike McShaffry , David Rez Graham, "Game Coding Complete", Fourth Edition, Course Technology, a part of Cengage Learning Publisher, ISBN-13: 978-1-133-77657-4. https://ms.zlibrary.sk/book/2072433/3f7bf2/game-coding-complete-fourth-edition.html
2. Gabe Zechermann, Christopher Cunningham, Gamification by Design, Oreilly, ISBN: 978-1-449- 39767-8, 2011. http://storage.libre.life/Gamification_by_Design.pdf
3. Jun Tanimoto, Fundamentals of Game Theory and its Applications, Volume 6, Springer, ISBN: 978-4-431-54961-1, 2015. https://link.springer.com/book/10.1007/978-4-431-54962-8



Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Technical Seminar - I (TS-26514-MTAD)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Practical: 04 Hrs/Week	**PR/OR:** 2	**Term work:** 25 Marks **Oral:** 25 Marks **Total** **Marks:** 50 Marks

Course Introduction:

The seminar aims to enhance students' research, presentation, and critical thinking skills, preparing them for advanced academic pursuits and professional careers. Technical Seminars will provide students with the opportunity and support to improve their self-study skills using modern information technologies and apply new knowledge and skills in practice, including in new areas.

Course Objectives: The objective of this course is:

1. Deepen Technical Knowledge: To enable students to explore a specialized topic within Artificial Intelligence and Data Science beyond the regular curriculum, fostering in-depth understanding.
2. Develop Research Skills: To provide practical experience in identifying, acquiring, evaluating, and synthesizing information from various technical sources (research papers, standards, technical reports).
3. Enhance Communication Skills: To cultivate effective oral and visual presentation skills, enabling students to articulate complex technical concepts clearly and concisely to a knowledgeable audience.
4. Foster Critical Thinking: To encourage students to critically analyze existing research, identify challenges, propose solutions, and engage in constructive discussions.
5. Promote Independent Learning: To encourage self-directed learning and the ability to stay updated with emerging technologies and research trends.
6. Prepare for Thesis/Dissertation: To serve as a foundational step for the Master's thesis/dissertation, allowing students to explore potential research areas.

Course Outcomes: After successful completion of the course, the learner will be able to:

CO1 : Formulate the goals and objectives of scientific research;

CO2 : Search, evaluate and analyze information about the achievements of science and technology in the target area and beyond;

CO3 : Interpret data from different fields of science and technology;

CO4 : Build the logic of reasoning and statements;

CO5 : Create, design and edit text documents in accordance with the requirements of the organization or publisher.

Artificial Intelligence & Data Science KSE Pune

Shalaka Foundation's

KEYSTONE SCHOOL OF ENGINEERING

Guidelines for Seminar:

- **Responsibility of the students:**

The Seminar should be carried out individually by each student.

– A student should identify the area or topics in recent trends and developments in consultation with the guide.

– A student should report to his/her respective guide regularly (at least once in a week) and report the progress of the seminar work.

– A student should follow the timelines and deadlines and inform the supervisor in case of any difficulty/delay.

– Students should maintain the record of all the meetings and remarks given by guide/reviewers and progress of the work in the diary. The diary must be presented during each review presentation to the reviewers.

– A student should conduct the research ethically, adhere to the academic integrity standards, and cite sources whenever using any existing results.

– A student should incorporate constructive feedback to improve the quality and rigor of the research work towards the seminar.

– For final examination, students should complete the Seminar Report in all aspects, including formatting and citation.

– Each student should prepare the report, get it approved by their guide and submit the duly signed copy within the deadline.

– A student should invest time and effort in preparing for seminar presentations and the oral presentation of the seminar.

- **Topic Selection—Relevance:**

Topics must be directly related to Computer Engineering, encompassing current research trends, emerging technologies, advanced concepts, or interdisciplinary applications.

– Scope: The topic should be sufficiently focused to allow for in-depth exploration within the seminar timeframe, yet broad enough to demonstrate a comprehensive understanding. Avoid overly narrow or excessively broad topics.

– Novelty (Desired): While not strictly a research paper, students are encouraged to explore topics that have recent advancements, open problems, or areas where their unique insights can be presented. Avoid merely summarizing introductory textbook material.

– Guide / Supervisor Approval: Each student must select a seminar topic in consultation with and obtain approval from an assigned faculty supervisor. The supervisor will guide the student in refining the topic and identifying relevant resources.

– Examples of Broad Areas: Artificial Intelligence/Machine Learning, Data Science & Big Data, Cybersecurity, Cloud Computing, Internet of Things (IoT), Computer Networks, Software Engineering, High-Performance Computing, Embedded Systems, Computer Vision, Natural Language Processing, Blockchain, and Quantum Computing.

- **Seminar Structure and Deliverable:**

The technical seminar typically involves the following stages and deliverables:

– Topic Proposal (2-3 weeks after topic approval)

– A concise document (1-2 pages) outlining:

* Proposed Seminar Title
* Brief Description/Abstract of the Topic



* Motivation and Relevance to Computer Engineering
* Preliminary List of Key References (at least 5-7 reputable sources)
* Tentative Scope and Outline of the Presentation
* Submission: To the faculty supervisor for approval.
* Literature Review and Research (Ongoing): Sources: Students must primarily rely on peer-reviewed academic sources (IEEE Xplore, ACM Digital Library, SpringerLink, arXiv, Google Scholar), reputable conference proceedings, and established industry standards. Wikipedia and unverified blogs are generally not acceptable as primary sources.
* Critical Analysis: Beyond mere summarization, students are expected to critically analyze the literature, identifying different approaches, their advantages/disadvantages, open issues, and potential future directions.
* Note-Taking & Organization: Maintain systematic notes and organize research material effectively.

• **Seminar Report (Due 2-3 weeks before presentation):**
– A written report (typically 20-25 pages, excluding references and appendices) detailing the seminar content.
– Format: Follow a professional academic paper format (e.g., IEEE Transaction style).
– Sections:
* Abstract: A concise summary of the seminar topic and key findings.
* Introduction: Background, motivation, problem statement (if applicable), and outline of the report.
* Literature Review/Background: Detailed discussion of relevant concepts, theories, and existing work.
* Core Content: In-depth exploration of the chosen topic, presenting different methodologies, architectures, algorithms, or challenges as relevant.
* Analysis/Discussion: Critical evaluation of the presented material, comparing different approaches, discussing implications, and identifying gaps.
* Future Trends/Conclusion: Summarization of key takeaways, potential future directions, and concluding remarks.
* References: A comprehensive list of all cited sources, properly formatted.
* Appendices (Optional): Supplementary material if necessary.

• **Oral Presentation:**
– Duration: Typically 25-30 minutes for presentation, followed by 10-15 minutes for Q&A.
– Audience: Faculty members, peers, and potentially other interested individuals.
– Content: The presentation should effectively convey the key aspects of the seminar topic. It should not simply be a reading of the report.
– Visual Aids: High-quality presentation slides (e.g., PowerPoint, Google Slides, LaTeX Beamer) are mandatory. Slides should be clear, concise, and visually appealing and support the oral delivery. Avoid excessive text on slides.
– Delivery: Clear articulation, confident posture, good eye contact, and appropriate pace. Practice the presentation thoroughly.
– Q&A Session: Be prepared to answer questions from the audience on all aspects of the seminar.



topic. Demonstrate a strong understanding and ability to defend your perspectives.

• **Evaluation Criteria:**

The technical seminar will be evaluated based on the following criteria: – Topic Selection and Scope (10%): Relevance, timeliness, and appropriate depth of the chosen topic. Clarity and focus of the topic proposal.

– Literature Review and Research (25%): Breadth and depth of literature surveyed. Quality and credibility of sources used. Critical analysis and synthesis of information.

– Seminar Report/Paper (30%): Clarity, organization, and logical flow of content. Technical accuracy and depth of discussion. Adherence to academic writing standards (grammar, spelling, formatting, referencing). Originality in synthesis and critical insights. Absence of plagiarism.

– Oral Presentation (35%): Content: Clarity, completeness, and accuracy of the presented material. Organization: Logical flow, effective use of time. Visual Aids: Quality, clarity, and effectiveness of slides. Delivery: Confidence, clarity of speech, enthusiasm, engagement with the audience. Q&A: Ability to answer questions accurately, comprehensively, and confidently.

Learning Resources

Textbooks

T1. Engineering Communication" by Charles W. Knisely & Karin I. Knisely

T2. Technical Communication: Principles and Practice" by Meenakshi Raman & Sangeeta Sharma

T3. The Craft of Scientific Presentations" by Michael Alley

Reference Books

1. https://nptel.ac.in/courses/109/106/109106180/
2. https://www.udemy.com/course/technical-writing/
3. https://www.edx.org/course/writing-in-the-sciences

Approved By

BOS Chairperson
Department Of Artificial Intelligence & Data Science

Chairperson
Board of Studies
Artificial Intelligence & Data Science
Department
Keystone School of Engineering, Pune.

Artificial Intelligence & Data Science
KSE
Pune

Chairperson
Academic Council
Keystone School of Engineering,
Pune.