TULSIRAMJI GAIKWAD-PATIL
College of Engineering & Technology
Mohgaon, Wardha Road, Nagpur - 441 108

-- An Autonomous Institute --

Department of
Artificial Intelligence and Machine Learning

B.Tech.
Artificial Intelligence and Machine Learning

Syllabus of
Honors in Data Science

Considering

National Education Policy 2020

From

Academic Year 2025-26

Scheme of
Honors in Data Science

Sr. No.	Course Code	Course Title	T/P	Contact Hours			Credits	Total Marks
				L	P	Hrs.		
1	BAI12308	Introduction to Explainable AI (XAI)	T	3	-	3	3	100
2	BAI12408	Fundamentals of Exploratory Data Analysis	T	3	-	3	3	100
3	BAI13511	Introduction to Computational Complexity	T	3	-	3	3	100
4	BAI13613	Data Analytics with Python	T	3	-	3	3	100
5	BAI14707	Advance machine learning	T	3	-	3	3	100
6	BAI14809	Capstone Project in Data Science	T	3	-	3	3	100
		Total	**-**	**18**	**8**	**18**	**18**	**600**

	Tulsiramji Gaikwad-Patil College of Engineering and Technology Wardha Road, Nagpur- 441 108 **NAAC Accredited (A+ Grade)** **(An Autonomous Institution Affiliated to RTM Nagpur University, Nagpur)**	

Second Year (Semester-III) B. Tech. Artificial Intelligence and Machine Learning

Teaching Scheme		Course code: BAI12308 Course Name: - Introduction to Explainable AI (XAI)	Examination Scheme	
Theory	3 Hrs./wk.		CT-I	15 Marks
Tutorial	-		CT-II	15 Marks
Total Credits	3		CA	10 Marks
Duration of ESE: 3 Hrs.			ESE	60 Marks
			Total	100 Marks

Course Objectives:

1	**Understand** the importance of explain ability in AI and its impact on stakeholders.
2	**Explore** different techniques and methods for making AI systems explainable.
3	**Analyze** the trade-offs between model complexity and interpretability.
4	**Examine** the ethical and societal implications of XAI.
5	**Apply** XAI techniques to real-world datasets and scenarios.

Course Contents

Unit I	**Introduction to Explainable AI (XAI)**: Motivations for XAI Importance of interpretability and transparency Techniques for XAI, Model-specific interpretability methods (e.g., decision trees, rule- based systems) Model-agnostic interpretability methods (e.g., LIME, SHAP) Post-hoc explanation techniques (e.g., feature importance, counterfactual explanation
Unit II	**Interpretable Models**: Linear models, Decision trees and rule-based systems Symbolic AI approaches, Interpretable Neural Networks, Sparse neural networks, Attention mechanisms, Layer-wise relevance propagation (LRP).
Unit III	**Evaluation of XAI Methods:** Quantitative metrics for interpretability, Human-centric evaluation methods, Ethical and Societal Implications of XAIB, is and fairness in interpretable AI, Trust and accountability in AI systems, Regulatory considerations.
Unit IV	**Evaluation of XAI Methods**: Quantitative metrics for interpretability Human-centric evaluation methods, Ethical and Societal Implications of XAI Bias and fairness in interpretable AI, Trust and accountability in AI systems Regulatory considerations.
Unit V	**Applications of XAI:** Healthcare (e.g., medical diagnosis, personalized treatment) Finance (e.g., credit scoring, fraud detection), Autonomous systems (e.g., self-driving cars, drones).

Text Books

T.1	"Interpretable Machine Learning" by Christoph Molnar
T.2	"Explainable AI: Interpreting, Explaining and Visualizing Deep Learning" by L. Liu and G. Hu
T.3	Research papers and articles from relevant conferences and journals (e.g., NeurIPS, ICML, AAAI)
Reference Books	
R.1	"Interpretable Machine Learning: A Guide for Making Black Box Models Explainable" by Christoph Molnar
R.2	"Explainable AI: Interpreting, Explaining and Visualizing Deep Learning" by L. Liu and G. Hu
R.3	"Explainable AI in Healthcare: Exploring Interpretable Models and Learning from Patient Data" edited by F. E. Elsayed and B. G. Stoecklin
Useful Links	
1	https://christophm.github.io/interpretable-ml-book
2	https://royalsocietypublishing.org/doi/full/10.1098/rsta.2018.0085
3	https://arxiv.org/abs/1812.02953

	Course Outcomes	**CL**	**Class Session**
CO1	**Understand** the core motivations behind the need for explainability in AI systems	2	9
CO2	**Evaluate** trade-offs between model complexity and interpretability**.**	5	9
CO3	**Design** human-centric evaluations (e.g., user studies) to test how understandable a model’s output is to humans.	6	9
CO4	**Integrate** regulatory and legal considerations into the development and deployment of AI systems.	3	9
CO5	**Apply** XAI methods to real-world domains like healthcare**,** finance**,** and autonomous systems**.**	3	9

	Tulsiramji Gaikwad-Patil College of Engineering and Technology Wardha Road, Nagpur- 441 108 NAAC Accredited (A+ Grade) (An Autonomous Institution Affiliated to RTM Nagpur University, Nagpur)	

Second Year (Semester – IV) B.Tech. Artificial Intelligence and Machine Learning

Teaching Scheme		Course Code: BAI12408 Course Name: Fundamentals of Exploratory Data Analysis	Examination Scheme	
Theory	3 Hrs./wk.		**CT-1**	15 Marks
Tutorial	-		**CT-2**	15 Marks
Total Credits	3		**CA**	10 Marks
Duration of ESE: 3 Hrs.			**ESE**	60 Marks
			Total	100 Marks

Course Objective:

1	To outline an overview of exploratory data analysis.
2	To implement data visualization using Matplotlib.
3	To perform univariate data exploration and analysis
4	To apply bivariate data exploration and analysis
5	To use Data exploration and visualization techniques for multivariate and time series data.

Course Contents

Unit I	**Exploratory Data Analysis**: EDA fundamentals – Understanding data science – Significance of EDA – Making sense of data – Comparing EDA with classical and Bayesian analysis – Software tools for EDA - Visual Aids for EDA- Data transformation techniques-merging database, reshaping and pivoting, Transformation techniques.
Unit II	**EDA Using Python:** Data Manipulation using Pandas – Pandas Objects – Data Indexing and Selection – Operating on Data – Handling Missing Data – Hierarchical Indexing – Combining datasets – Concat, Append, Merge and Join – Aggregation and grouping – Pivot Tables – Vectorized String Operations.
Unit III	**Univariate Analysis:** Introduction to Single variable: Distribution Variables - Numerical Summaries of Level and Spread - Scaling and Standardizing – Inequality.
Unit IV	**Bivariate Analysis**: Relationships between Two Variables - Percentage Tables - Analysing Contingency Tables - Handling Several Batches - Scatterplots and Resistant Lines.
Unit V	Case study 1: Customer Churn Prediction Using Multivariate Analysis. Case study 2: Impact of Advertising Spend on Sales Across Regions. Case study 3: Time Series Forecasting for Electricity Demand. Case study 4: Retail Sales Analysis Using Grouping & Resampling. Case study 5: Stock Market Trend Analysis Using Time Series.

Text Books

T.1	Suresh Kumar Mukhiya, Usman Ahmed, “Hands - On Exploratory Data Analysis with Python”, Packt Publishing, 2020.
T.2	Jake Vander Plas, "Python Data Science Handbook: Essential Tools for Working with Data", First Edition, O Reilly, 2017.
T.3	Catherine Marsh, Jane Elliott, “Exploring Data: An Introduction to Data Analysis for Social Scientists”, Wiley Publications, 2nd Edition, 2008.

Reference Books

R.1	Eric Pimpler, Data Visualization and Exploration with R, GeoSpatial Training service, 2017.
R.2	Claus O. Wilke, “Fundamentals of Data Visualization”, O’reilly publications, 2019
R.3	Matthew O. Ward, Georges Grinstein, Daniel Keim, “Interactive Data Visualization: Foundations, Techniques, and Applications”, 2nd Edition, CRC press, 2015.

Useful Links

1	https://www.bing.com/videos/riverview/relatedvideo?&q=Exploratory+Data+Analysis+NPTEL+Course&&mid=8917A8B3C2F3429EFA538917A8B3C2F3429EFA53&&FORM=VRDGAR
2	https://www.bing.com/videos/riverview/relatedvideo?q=Exploratory%20Data%20Analysis%20NPTEL%20Course&mid=62EB57210CB121124C3B62EB57210CB121124C3B&ajaxhist=0

	Course Outcomes	**CL**	**Class Session**
1	**Understand** the fundamentals of exploratory data analysis.	2	9
2	**Implement** the data visualization using Matplotlib.	3	9
3	**Perform** univariate data exploration and analysis.	3	9
4	**Apply** bivariate data exploration and analysis.	3	9
5	**Use** Data exploration and visualization techniques for multivariate and time series data.	3	9

Tulsiramji Gaikwad-Patil College of Engineering and Technology
Wardha Road, Nagpur- 441 108
NAAC Accredited (A+ Grade)
(An Autonomous Institution Affiliated to RTM Nagpur University, Nagpur)

Third Year (Semester – V) B.Tech. Artificial Intelligence and Machine Learning

<table>
<tr><th colspan="2">Teaching Scheme</th><td rowspan="7">Course Code: BAI13511
Course Name: Introduction to Computational Complexity</td><th colspan="2">Examination Scheme</th></tr>
<tr><td>Theory</td><td>3 Hrs./wk.</td><td>CT-1</td><td>15 Marks</td></tr>
<tr><td>Tutorial</td><td>-</td><td>CT-2</td><td>15 Marks</td></tr>
<tr><td>Total Credits</td><td>3</td><td>CA</td><td>10 Marks</td></tr>
<tr><td colspan="2" rowspan="2">Duration of ESE: 3 Hrs.</td><td>ESE</td><td>60 Marks</td></tr>
<tr><td>Total</td><td>100 Marks</td></tr>
</table>

Course Objective:

1	Understand key concepts in computational complexity, including NP-completeness, time and space hierarchy, and the P vs NP problem.
2	Explain advanced space complexity classes and key theorems in computational complexity theory.
3	Introduce concepts of randomized and nonuniform computation, including BPP and circuit complexity.
4	Explore advanced complexity theory topics including circuit complexity, randomized algorithms, and #P-completeness.
5	Recognize key theorems and concepts in computational complexity, including Valiant-Vazirani, Toda's Theorem, communication complexity, and interactive proofs.

Course Contents

Unit I	**Unit I: Foundations of Complexity Theory**-Review of NP-Completeness, P vs NP, Cook–Levin Theorem, Time Hierarchy and Polynomial Hierarchy, Formal models of computation: Turing Machines, RAM, Circuit Models, Complexity measures: time, space, nondeterminism, Decision problems, Karp/Turing reductions, completeness, NP structure: verifiers, witnesses, certificate complexity, Origins of NP-completeness and hardness notions, Diagonalization foundations for hierarchy theorems, Introduction to Space Complexity; basic class inclusions ($L \subseteq NL \subseteq P \subseteq NP$)
Unit II	**Unit II: Space Complexity and Relativization**-Savitch's Theorem; deterministic vs nondeterministic space, Space constructibility and SPACE(f(n)), NL machines: log-space TMs, configurations, log-space reductions, NL-completeness; NL = coNL, PSPACE and PSPACE-Complete problems (QBF, games), Space Hierarchy Theorem, Oracle Turing Machines and relativization, Baker–Gill–Solovay Theorem and its implications.
Unit III	**Randomization and Nonuniform Computation**-Randomized complexity classes: RP, coRP, BPP, ZPP, Probabilistic Turing Machines, error reduction, amplification, Derandomization, pseudorandom generators, BPP in the Polynomial Hierarchy (Sipser–Gács–Lautemann Theorem), Nonuniform computation: P/poly, advice strings, cryptographic relevance, Circuit Complexity: AC^0, AC^k, NC, TC classes, Circuit size/depth, uniformity vs nonuniformity, Lower bound techniques: combinatorial, diagonalization, communication-based
Unit IV	**Lower Bounds, Algebraic Complexity & Counting Classes**-Parity not in AC^0; Switching Lemma and random restrictions, Karp–Lipton Theorem and PH collapse implications, Adleman's Theorem: randomness within polynomial time, Polynomial Identity Testing (PIT): arithmetic

	circuits, Schwarz–Zippel lemma, Isolation Lemma and randomized algorithms, Perfect Matching in RNC^2; NC vs RNC, Counting complexity: definition of #P, reductions, relationships, #P-Completeness; Valiant's Theorem—Permanent is #P-Complete
Unit V	**Advanced Theorems, Communication & Interactive Proofs-**Valiant–Vazirani Theorem (SAT → Unique-SAT), Toda's Theorem: PH ⊆ P^#P, Communication Complexity: deterministic, randomized, nondeterministic models; discrepancy, rank, fooling sets, Monotone computation: monotone circuits, lower bounds (matching, clique); Razborov's method, Interactive Proofs: IP vs PSPACE, verifier–prover model, relevance to cryptography (zero-knowledge, PCP).
Text Books	
T.1	Computational Complexity, by Christos Papadimitriou
T.2	Computational Complexity: A Modern Approach, by Sanjeev Arora and Boaz Barak.
T.3	Introduction to the Theory of Computation by Michael Sipser
Reference Books	
R.1	Sanjeev Arora and Boaz Barak, Computational Complexity: A ModernApproach, Cambridge University Press, Edition I, 2009
R.2	O. Goldreich. Computational complexity: a conceptual perspective. Cambridge University Press, 2008
R.3	O. Goldreich. P, NP, and NP-completeness. Cambridge University Press, 2010.
Useful Links	
1	Computational Complexity - Course
2	NOC:Computational Complexity Theory NPTEL Course - Free Practice Questions & Materials | NPTELPrep

	Course Outcomes	CL	Class Session
1	**Interpret** key concepts in computational complexity, including NP-completeness, P vs NP, and space complexity.	2	9
2	**Describe** key results and theorems in space complexity, including NL-completeness, PSPACE-completeness, and space hierarchy.	2	9
3	**Apply** the concepts of randomized and nonuniform computation, including BPP and circuit complexity.	3	9
4	**Recognize** key results in computational complexity, including circuit lower bounds, randomized algorithms, and #P-completeness.	2	9
5	**Demonstrate** key complexity theorems and models, including Valiant-Vazirani, Toda's Theorem, communication complexity, and interactive proofs.	3	9

	Tulsiramji Gaikwad-Patil College of Engineering and Technology Wardha Road, Nagpur- 441 108 **NAAC Accredited (A+ Grade)** **(An Autonomous Institution Affiliated to RTM Nagpur University, Nagpur)**	

Third Year (Semester – VI) B.Tech. Artificial Intelligence and Machine Learning

Teaching Scheme		**Course Code: BAI13613** **Course Name: Data Analytics with Python**	**Examination Scheme**	
Theory	3 Hrs./wk.		**CT-1**	15 Marks
Tutorial	-		**CT-2**	15 Marks
Total Credits	3		**CA**	10 Marks
Duration of ESE : 3 Hrs.			**ESE**	60 Marks
			Total	100 Marks

Course Objective:

1	**Understand** foundational concepts of data analytics using Python, including probability, sampling, and distributions.
2	**Apply** hypothesis testing, two-sample comparisons, and ANOVA for statistical decision-making.
3	**Analyse** regression techniques and statistical inference methods, including MLE, linear, and logistic regression.
4	**Evaluate** regression models using ROC curves, chi-square tests, and clustering fundamentals.
5	**Develop** practical skills in clustering and decision tree techniques for classification and regression tasks.

Course Contents

Unit I	Introduction to data analytics and Python fundamentals, Introduction to probability, Sampling and sampling distributions
Unit II	Hypothesis testing, two sample testing and introduction to ANOVA
Unit III	Two-way ANOVA and linear regression, Linear regression and multiple regression, Concepts of MLE and Logistic regression

Unit IV	ROC and Regression Analysis Model Building, c^2 Test and introduction to cluster analysis
Unit V	Clustering analysis, Classification and Regression Trees (CART)

Text Books

T.1	McKinney, W. (2012). Python for data analysis: Data wrangling with Pandas, NumPy, and IPython. " O'Reilly Media, Inc.".
T.2	Swaroop, C. H. (2003). A Byte of Python. Python Tutorial.
T.3	Ken Black, sixth Editing. Business Statistics for Contemporary Decision Making. "John Wiley & Sons, Inc".

Reference Books

R.1	Anderson Sweeney Williams (2011). Statistics for Business and Economics. "Cengage Learning".
R.2	Douglas C. Montgomery, George C. Runger (2002). Applied Statistics & Probability for Engineering. "John Wiley & Sons, Inc"
R.3	Jay L. Devore (2011). Probability and Statistics for Engineering and the Sciences. "Cengage Learning".

Useful Links

1	Data Analytics with Python - Course
2	https://www.bing.com/videos/riverview/relatedvideo?q=data+analytics+with+python+nptel+course+syllabus&mid=FA6630CBDBA601BAE8BCFA6630CBDBA601BAE8BC&FORM=VIRE

	Course Outcomes	CL	Class Session
1	**Understand** the fundamentals of data analytics using Python, along with basic concepts of probability, sampling, and sampling distributions.	2	9
2	**Apply** hypothesis testing, compare two sample data sets, and interpret results using ANOVA techniques.	3	9
3	**Implement** statistical modeling techniques such as ANOVA, regression (linear, multiple, logistic), and Maximum Likelihood Estimation (MLE) for data analysis and inference.	3	9
4	**Execute** ROC analysis, regression modeling, Chi-square tests, and basic clustering techniques for data-driven decision-making.	3	9
5	**Employ** clustering and decision tree techniques such as CART for effective data classification and regression analysis.	3	9

	Tulsiramji Gaikwad-Patil College of Engineering and Technology Wardha Road, Nagpur- 441 108 **NAAC Accredited (A+ Grade)** **(An Autonomous Institution Affiliated to RTM Nagpur University, Nagpur)**	

Fourth Year (Semester – VII) B.Tech. Artificial Intelligence and Machine Learning

Teaching Scheme		Course Code / Name	Examination Scheme	
Theory	3 Hrs./wk.	**Course Code: BAI14707** **Course Name: Advance machine learning**	**CT-1**	15 Marks
Tutorial	-		**CT-2**	15 Marks
Total Credits	3		**CA**	10 Marks
Duration of ESE: 3 Hrs.			**ESE**	60 Marks
			Total	100 Marks

Course Objective:

1	**To Provide** advanced theoretical understanding of modern machine learning paradigms, including representation learning, probabilistic modeling, and optimization strategies.
2	**To Develop** the ability to implement and experiment with cutting-edge ML techniques such as self-supervised learning, graph learning, meta-learning, and generative models.
3	**To Build** competency in designing scalable and efficient ML systems using distributed training, model compression, and hardware-aware optimization.
4	**To Foster** analytical skills to assess robustness, fairness, interpretability, and security of ML models through contemporary evaluation frameworks.
5	**To Promote** responsible AI development by integrating ethical principles, governance frameworks, and lifecycle management of deployed ML systems.

Course Contents

Unit I	Advanced Optimization Techniques in ML-Convex vs Non-Convex Optimization, Advanced Gradient Methods: Adam, RMSProp, AdaGrad, AMSGrad, Second-order & Quasi-Newton Methods: L-BFGS, Natural Gradient, Optimization on Manifolds; Trust-Region Methods, Training Stability, Gradient Clipping, Learning Rate Scheduling, Large-scale/Distributed Optimization: Data Parallelism, Model Parallelism.
Unit II	Representation Learning & Self-Supervised Methods,Representation Learning Fundamentals,Autoencoders: Variants (Denoising, Sparse, Contractive),Self-Supervised Pretext Tasks (contrastive learning, masked modeling), Contrastive Learning Methods: SimCLR, MoCo, BYOL, Metric Learning: Triplet Loss, Siamese Networks, Foundation Models & Pretraining Paradigms
Unit III	Probabilistic Deep Learning & Generative Models,Variational Inference & Bayesian Deep Learning, Variational Autoencoders (VAE) – Theory & Applications, Normalizing Flows (RealNVP, Glow), Generative Adversarial Networks (GANs): DCGAN, WGAN, CycleGAN, Diffusion Models: Score-based generative modeling, Uncertainty Quantification in Deep Models
Unit IV	Advanced Sequence & Graph-based Learning, Transformer Architecture: Self-Attention, Multi-Head Attention, Large Language Models (LLMs): Training Pipeline & Tokenization, Graph Machine Learning: GNNs, GCN, GAT, Message-Passing Networks, Temporal Models: Temporal Convolution, Neural ODEs, Multimodal Learning: Vision–Language Models (CLIP, ViLBERT)
Unit V	Causal, Ethical & Trustworthy Machine Learning- Causal ML: Structural Causal Models, Counterfactuals, Do-Calculus, Causal Inference Methods: Propensity Scoring, IV Methods, CATE, Uplift Models, Explainability: SHAP, LIME, Integrated Gradients, Counterfactual Explanations, Fairness in ML: Bias Metrics, Mitigation Techniques, Adversarial ML: FGSM, PGD, Robust Training, ML Security, Privacy, Federated Learning & Differential Privacy

Text Books	
T.1	Satyajit Das, Bharatkumar Sharma, “Applied Accelerated Artificial Intelligence” (NPTEL video series & lecture notes), IIT Palakkad/IIT Goa, 2022.
T.2	Vini Madhavan, John Owens, “Programming Massively Parallel Processors: A Hands-on Approach”, 4th Ed., Morgan Kaufmann, 2022.
T.3	RAPIDS & Nvidia Docs, “RAPIDS AI User Documentation”, Nvidia, latest online edition.

Reference Books	
R.1	R.1 Aurélien Géron, “Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow”, 3rd Ed., O'Reilly, 2023.
R.2	R.2 Research and whitepapers from ACM Digital Library, NeurIPS, NVIDIA, and OpenACC.

Useful Links	
1	https://nptel.ac.in/courses/106106238
2	https://iitgoa.ac.in/aishikshaai/

	Course Outcomes	**CL**	**Class Session**
1	**Analyze** advanced machine learning architectures and theoretical foundations including representation learning, probabilistic models, and optimization techniques.	4	9
2	**Apply** state-of-the-art ML methods such as self-supervised learning, contrastive learning, GNNs, and meta-learning to solve complex data-driven problems.	3	9
3	**Design** scalable ML pipelines using distributed training, efficient model compression, and hardware-aware optimization for real-world deployment.	6	9
4	**Evaluate** model robustness, fairness, interpretability, and security through modern frameworks such as adversarial evaluation and explainability metrics.	5	9
5	**Create** responsible and high-performance ML solutions that integrate ethical AI considerations, continuous monitoring, and compliance with modern AI governance frameworks.	6	9

Fourth Year (Semester – VIII) B.Tech. Artificial Intelligence and Machine Learning

Teaching Scheme		Course	Examination Scheme	
Theory	3 Hrs./wk.	**Course Code: BAI14809**	Seminar 1	20 Marks
Tutorial	-	**Course Name: Capstone Project in Data Science**	Seminar 2	20 Marks
Total Credits	3		Seminar 3	20 Marks
Duration of ESE: 3 Hrs.			Seminar 4	20 Marks
			Seminar 5	20 Marks
			Total	100 Marks

Course Objective:

1	**Enable** students to identify, define, and scope real-life data science problems.
2	**Provide** hands-on experience with the complete data science project lifecycle.
3	**Strengthen** team-based and independent project execution, communication, and analytical skills.
4	**Encourage** the application of advanced data science and AI techniques for real-world datasets.
5	**Instill best practices** in data handling, model validation, presentation, and reproducibility.

Course Contents

Seminar -1	**Problem Formulation and Planning**: Identification of a real-world problem (industry, research, societal), Literature survey and baseline analysis, Defining specific objectives, outcomes, and deliverables, Project planning: timelines, roles, and milestones
Seminar -2	**Data Acquisition, Preparation, and Exploration**: Raw data acquisition: APIs, web scraping, industry/research datasets Data cleansing, handling missing data, transformations Exploratory data analysis (EDA): visualization and statistical summaries Feature engineering and data partitioning
Seminar -3	**Model Building and Validation**: Selection and implementation of appropriate algorithms (supervised/unsupervised) Model training, parameter tuning, cross-validation Performance metrics, error analysis, iterative improvement Advanced topics: ensemble methods, deep learning, if appropriate
Seminar -4	**Solution Deployment and Communication:** Deployment strategies: dashboards, REST APIs, cloud, portable notebooks Communication of findings: technical report writing, presentations, data visualization Ethical, privacy, societal consideration in deployment Collaboration and version control best practices
Seminar -5	**Final Demonstration and Documentation:** Final product demonstration (proof-of concept/prototype/tool) Documentation of process, results, and reproducibility resources Viva-voce and Q&A Reflection, lessons learned, and future scope

Text Books

T.1	Cathy O'Neil and Rachel Schutt, "Doing Data Science," O'Reilly, 2014.
T.2	Aurélien Géron, "Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow," 3rd Ed., O'Reilly, 2023.
T.3	Joel Grus, "Data Science from Scratch," 2nd Ed., O'Reilly, 2019.

Reference Books

R.1	Hadley Wickham & Garrett Grolemund, "R for Data Science," O'Reilly, 2017.

R.2	Research articles, datasets, and documentation from Kaggle/Data.gov, and open repositories.

Useful Links

1	https://nptel.ac.in/courses/106106212 (NPTEL Data Science for Engineers)
2	https://www.kaggle.com/learn/overview

	Course Outcomes	CL	Class Session
1	**Understand** problem identification and precise project definition in data science.	2	9
2	**Apply** end-to-end data science practices on real-world or research datasets.	4	9
3	**Analyze** and report results, using state-of-the-art tools, models, and evaluation metrics.	3	9
4	**Evaluate** and communicate outcomes via presentations, technical documentation, and live demonstrations.	5	9
5	**Design** and deliver a functional prototype/solution addressing a significant data science challenge.	4	9