Anekant Education Society's

Tuljaram Chaturchand College

of Arts, Science and Commerce, Baramati

(Autonomous)

M.Sc. Degree Program in Data Science

(Faculty of Science & Technology)

CBCS Syllabus

M.Sc.(Data Science) Part – II Semester – IV

For Department of Statistics

Tuljaram Chaturchand College of Arts, Science and Commerce, Baramati

Choice Based Credit System Syllabus (2023 Pattern)

(As Per NEP 2020)

To be implemented from Academic Year 2024-2025

Program Outcomes for M.Sc.

PO1. Comprehensive Knowledge and Understanding:

Postgraduates will possess a profound understanding of their field, encompassing foundational theories, methodologies, and key concepts within a multidisciplinary context.

PO2. Practical, Professional, and Procedural Knowledge:

Postgraduates will acquire practical skills and expertise necessary for professional tasks, including industry standards, regulations, and ethical considerations, with effective application in real-world scenarios.

PO3. Entrepreneurial Mindset, Innovation, and Business Understanding:

Postgraduates will cultivate an entrepreneurial mindset, identify opportunities, foster innovation, and understand business principles, market dynamics, and risk management strategies.

PO4. Specialized Skills, Critical Thinking, and Problem-Solving:

Postgraduates will demonstrate proficiency in technical skills, analytical abilities, effective communication, and leadership, adapting and innovating in response to changing circumstances.

PO5. Research, Analytical Reasoning, and Ethical Conduct:

Postgraduates will exhibit observational and inquiry skills, formulate research questions, utilize appropriate methodologies for data analysis, and adhere to research ethics while effectively reporting findings.

PO6. Communication, Collaboration, and Leadership:

Postgraduates will effectively communicate complex information, collaborate in diverse teams, demonstrate leadership qualities, and facilitate cooperative efforts toward common goals.

PO7. Digital Proficiency and Technological Skills:

Postgraduates will demonstrate proficiency in using ICT, accessing information sources, analyzing data using appropriate software, and adapting to technological advancements.

PO8. Multicultural Competence, Inclusive Spirit, and Empathy:

Postgraduates will engage effectively in multicultural settings, respect diverse perspectives, lead diverse teams, and demonstrate empathy and understanding of others' perspectives and emotions.

PO9. Value Inculcation, Environmental Awareness, and Ethical Practices:

Postgraduates will embrace ethical and moral values, practice responsible citizenship, recognize and address ethical issues, and promote sustainability and environmental conservation.

PO10. Autonomy, Responsibility, and Accountability:

Postgraduates will apply knowledge and skills independently, manage projects effectively, and demonstrate responsibility and accountability in work and learning contexts, contributing to societal

well-being.

Anekant Education Society's

Tuljaram Chaturchand College,Baramati

(Autonomous)

Board of Studies (BOS) in Statistics

From 2022-23 to 2024-25

Sr.No.	Name	Designation
1.	Prof. Dr. Vikas C. Kakade	Chairman
2.	Prin. Dr. Avinash S. Jagtap	Member
3.	Dr. Neeta K. Dhane	Member
4.	Dr.Vaishali V. Patil	Member
5.	Mrs. Sarita D. Wadkar	Member(Ad hoc)
6.	Mr. Chandrashekhar P. Swami	Member
7.	Ms. Priti M. Mohite	Member(Ad hoc)
8.	Ms. Nilambari A. Jagtap	Member (Ad hoc)
9.	Miss.Kalyani C. Kale	Member (Ad hoc)
10.	Ms.Pooja S. Zanjurne	Member (Ad hoc)
11.	Dr. Akanksha S. Kashikar	Vice-Chancellor Nominee
12.	Prin. Dr. Rajendra G. Gurao	Expert from other University
13.	Mr. Rohan Koshti	Expert from other University
14.	Mr. Saurabh Kadam	Industry Expert
15.	Dr. Jaya L. Limbore	Meritorious Alumni

16.	Miss. Priya N. Rakate	Invitee Member
17.	Ms. Ankita G. Deshmukh	Invitee Member
18.	Ms. Shital B. Choudhar	Invitee Member
19.	Miss. Kiran Banda (M.Sc. II)	Student Representative
20.	Mr. Rushikesh Pandhare (M.Sc. II)	Student Representative
21.	Mr. Bharat Jambhulkar (TYBSc)	Student Representative
22.	Miss. Prapti Mane (TYBSc)	Student Representative

Credit Distribution Structure for M.Sc.(Data Science)Part-I

Level	Semester	Major		Research Methodology (RM)	OJT/FP	RP	Cum. Cr.	Degree
		Mandatory	Electives					
6.0	Sem-I	DSC-501-MJM: Probability and Statistics for Data Science(Credit 04) DSC-502-MJM: Data Analytics Using R (Credit 04) DSC -503-MJM: Data Science Practical – I (Credit 02) DSC-504-MJM: Data Science Practical – II (Credit 02)	DSC-511-MJE (A): Data Base Management System DSC -511-MJE (B): Stochastic Models and Applications (Credit 04)	DSC -521-RM: Research Methodology (Credit 04)	--	--	**20**	PG Diploma (after 3 Year Degree)
	Sem-II	DSC -551-MJM: Machine Learning and Artificial intelligence(Credit 04) DSC-552-MJM: Regression Analysis and Predictive Models (Credit 04) DSC-553-MJM: Data Science Practical – III (Credit 02) DSC -554-MJM: Data Science Practical – IV (Credit 02)	DSC -561-MJE (A): Bayesian Inference DSC -561-MJE (B): Computational Statistics (Credit 04)	--	DSC-581-OJT/FP: On Job Training/ Field Project	--	**20**	

Credit Distribution Structure for M.Sc.(Data Science) Part-II

Level	Semester	Major		Research Methodology (RM)	OJT /FP	RP	Cum. Cr.	Degree
		Mandatory	Electives					
6.5	Sem-III	DSC-601-MJM: Multivariate Analysis (Credit 04)	DSC-611-MJE(A): Business and Project Management (Credit 02)	--	--	DSC-621-RP: Research Project (Credit 04)	**20**	PG Diploma (after 3 Year Degree)
		DSC-602-MJM: Time Series Analysis and Forecasting (Credit 04)	DSC-611-MJE(B): Text Mining and Natural Language Processing (Credit 02)					
		DSC-503-MJM: Statistics Practical – V (Credit 02)	DSC-612-MJE(A):Practical Based on Business and Project Management (Credit 02)					
		DSC-504-MJM: Statistics Practical – VI (Credit 02)	DSC-612-MJE(B): Practical Based on Text Mining and NLP (Credit 02)					
	Sem-IV	DSC-651-MJM: Introduction to Generative AI and Large Language Models (Credit 04)	DSC-661-MJE (A): Supply Chain and Logistics Analytics (Credit 02)	--	--	DSC-621-RP: Research Project (Credit 06)	**20**	
		DSC-652-MJM: Advanced Machine Learning and Deep Learning (Credit 04)	DSC-661-MJE(B): Image Processing and Computer Vision (Credit 02)					
		DSC-653-MJM: Data Science Practical – VI (Credit 02)	DSC-662-MJE (A): Introduction to Hadoop (Credit 02)					
			DSC-662-MJE (B): Web Application Development (Credit 02)					

Course Structure for M.Sc. Part-I (Data Science)(2023 Pattern)

Sem	Course Type	Course Code	Course Title	Theory/ Practical	No. of Credits
I	Major (Mandatory)	DSC-501-MJM	Probability and Statistics for Data Science	Theory	04
	Major (Mandatory)	DSC -502-MJM	Data Analytics Using R	Theory	04
	Major (Mandatory)	DSC -503-MJM	Data Science Practical – I	Practical	02
	Major (Mandatory)	DSC -504-MJM	Data Science Practical – II	Practical	02
	Major (Elective)	DSC-511-MJE (A)	Data Base Management System	Theory	04
		DSC -511-MJE (B)	Stochastic Models and Applications	Theory	
	Research Methodology (RM)	DSC -521-RM	Research Methodology	Theory	04
	Total Credits Semester I				**20**
II	Major (Mandatory)	DSC -551-MJM	Machine Learning and Artificial intelligence	Theory	04
	Major (Mandatory)	DSC-552-MJM	Regression Analysis and Predictive Models	Theory	04
	Major (Mandatory)	DSC-553-MJM	Data Science Practical – III	Practical	02
	Major (Mandatory)	DSC-554-MJM	Data Science Practical – IV	Practical	02
	Major (Elective)	DSC -561-MJE (A)	Bayesian Inference	Theory	04
		DSC -561-MJE (B)	Computational Statistics	Theory	
	On Job Training (OJT)/Field Project (FP)	DSC -581-OJT/FP	On Job Training Field Project	Training/Project	04
	Total Credits Semester-II				**20**
	Cumulative Credits Semester I and II				**40**

Course Structure for M.Sc. Part-II (Data Science)(2023 Pattern)

Sem	Course Type	Course Code	Course Title	Theory/ Practical	No. of Credits
III	Major (Mandatory)	DSC-601-MJM	Multivariate Analysis	Theory	04
	Major (Mandatory)	DSC -602-MJM	Time Series Analysis and Forecasting	Theory	04
	Major (Mandatory)	DSC -603-MJM	Data Science Practical – V	Practical	02
	Major (Mandatory)	DSC -604-MJM	Data Science Practical – VI	Practical	02
	Major (Elective)	DSC-611-MJE(A)	Business and Project Management	Theory	02
		DSC -611-MJE(B)	Text Mining and Natural Language Processing	Theory	
		DSC-612-MJE (A)	Practical Based on Business and Project	Practical	02
		DSC -612-MJE (B)	Practical Based on Text Mining and NLP	Practical	
	Research Project (RP)	DSC -621-RP	Research Project	Project	04
				Total Credits Semester III	**20**
IV	Major (Mandatory)	DSC -651-MJM	Introduction to Generative AI and Large Language Models	Theory	04
	Major (Mandatory)	DSC-652-MJM	Advance machine Learning and Deep Learning	Theory	04
	Major (Mandatory)	DSC-653-MJM	Data Science Practical – VI	Practical	02
	Major (Elective)	DSC -661-MJE (A)	Supply Chain and Logistics Analytics	Theory	02
		DSC -661-MJE (B)	Image Processing and Computer Vision	Theory	
		DSC -662-MJE (A)	Introduction to Hadoop	Practical	02
		DSC -662-MJE (B)	Web Application Development	Practical	
	Research Project (RP)	DSC -681-RP	Research Project	Project	06
				Total Credits Semester-IV	**20**
				Cumulative Credits Semester III and IV	**40**

CBCS Syllabus as per NEP 2020 for M.Sc. Part-II Data Science (2023 Pattern)

Name of the Programme	: M.Sc. Data Science
Program Code	: PSDSC
Class	: M.Sc. Part – II
Semester	: IV
Course Type	:Major Mandatory Theory
Course Name	: Introduction to Generative AI and Large Language Models
Course Code	: DSC-651-MJM
No. of Credits	: 4
No. of Teaching Hours	: 60

Course Objectives:

1. Explain the key concepts, evolution, and applications of Generative AI and Large Language Models (LLMs).
2. Apply probability, embeddings, and matrix operations in language modeling.
3. Examine the structure of deep learning models, loss functions, optimizers, and the Transformer architecture.
4. Explain various tokenization methods, sampling techniques, and controlled text generation approaches.
5. Differentiate between fine-tuning and in-context learning, and apply prompt engineering techniques.
6. Utilize LLMs for statistical queries, report summarization, code assistance, and data analysis.
7. Assess ethical considerations, AI biases, security challenges, and emerging trends in generative AI.

Course Outcomes:

By the end of the course, students will be able to:

CO1. Demonstrate knowledge of generative models, their evolution, and real-world applications.

CO2. Utilize probability, embeddings, and evaluation metrics in language modeling tasks.

CO3. Build and optimize neural network models, understanding backpropagation and optimization strategies.

CO4. Execute different tokenization and text generation methods using advanced sampling strategies.

CO5. Develop structured prompts for LLMs and apply fine-tuning techniques for specific applications.

CO6. Implement LLMs in programming, data analysis, and retrieval-augmented generation (RAG) tasks.

CO7. Critically evaluate ethical challenges, regulatory frameworks, and future trends in AI and LLMs.

Topics and Learning Points

Unit 1 **(10L)**

Introduction to GenAI and LLMs:

- Definition and Scope of Generative AI
- Evolution from rule-based to generative systems
- Generative Models vs Discriminative Models
- Real-world Applications: Chatbots, Content Generation, Coding, Healthcare
- Milestone Models: GPT-1 to GPT-4, BERT, LLaMA, Claude, Gemini

Mathematical & Statistical Foundations:

- Probability in Language Modeling: Chain Rule, Conditional Probabilities
- Introduction to Embeddings: Vector Representations of Words
- Matrix Operations: Dot Product, Softmax, Attention
- Statistical Evaluation Metrics: Perplexity, BLEU, ROUGE, F1 Score

Unit 2: **(15 L)**

Neural Networks and Deep Learning Basics

- Neuron, Layers, Forward Pass, Back propagation
- Loss Functions: Cross Entropy, MSE
- Optimizers: SGD, Adam
- Introduction to Language Modeling: From N-grams to Neural Networks

The Transformer Architecture

- Self-Attention: Queries, Keys, Values
- Positional Encoding

- Encoder-Decoder Structure
- Variants: Encoder-only (BERT), Decoder-only (GPT), Encoder-Decoder (T5, BART)

Unit 3: (15 L)

Tokenization and Text Generation

- What is a Token? Byte Pair Encoding, WordPiece
- Tokenization Process in LLMs
- Sampling Methods: Greedy, Beam Search, Top-k, Top-p
- Controlled Text Generation (Temperature, Repetition Penalty)

Module 6: Prompt Engineering & Fine-Tuning

- Prompt Design: Basic, Chain-of-Thought, Instruction Tuning
- Zero-shot, Few-shot, and Multi-turn Prompts
- Fine-Tuning vs. In-Context Learning
- Parameter-Efficient Fine-Tuning (LoRA, QLoRA, PEFT)

Unit 4: Practical Implementation and Applications (20 L)

Applications of LLMs in CS & Statistics

- Natural Language Interfaces for Statistical Queries
- Summarization of Statistical Reports
- LLMs for Code Assistance (Python, R)
- Conversational Agents for Data Analysis
- Introduction to RAG (Retrieval-Augmented Generation),FAISS,Pinecone,LangChain,LlamaIndex

Responsible GenAI & Future Trends

- Ethics in AI: Bias, Fairness, Privacy
- AI Hallucination and Misuse
- Copyright & Licensing Issues with Generated Content
- Open-source Models and Democratization of LLMs
- Future Directions: Multimodal AI, AGI, Agentic Workflows
- Prompt Engineering & AI Agents – Chain-of-thought, AutoGPT, OpenAI Assistants API.
- Multimodal AI – Text-to-image (Stable Diffusion, DALL·E), Text-to-video (Sora, RunwayML).
- LLM Deployment – AWS SageMaker, Vertex AI, On-prem Llama2, Edge AI models.
- Reinforcement Learning & AI Alignment – RLHF, PPO, Self-improving AI.
- AI Security & Ethics – Jailbreak attacks, bias detection, GDPR, AI Act compliance.

- Code Generation & AI Coding Assistants – Code Llama, OpenAI Codex, GitHub Copilot.
- Generative AI for Business – AI-powered content creation, marketing automation.
- Self-Improving Autonomous AI – Meta AI agents, AI-driven decision-making.

References:

1. Deep Learning" by Ian Good fellow, Yoshua Bengio, and Aaron Courville: 2016, MIT Press,Volume: 1
2. Transformers for Natural Language Processing by Denis Rothman: 2021, Packt Publishing, Volume: 1
3. GANs in Action: Deep Learning with Generative Adversarial Networks"by Jakub Langr and Vladimir Bok:: 2018, : Manning Publications, Volume: 1
4. Natural Language Processing with Transformers" by Lewis Tunstall, Leandro von Werra, and Thomas Wolf: 2022, O'Reilly Media, Volume: 1
5. Attention is All You Need" by Vaswani et al. (2017)
6. Language Models are Few-Shot Learners" by Radford et al. (2020)
7. Generative Adversarial Networks" by Goodfellow et al. (2014)
8. Auto-Encoding Variational Bayes" by Kingma and Welling (2013)
9. [Chris Olah's Blog](http://colah.github.io/): Provides in-depth articles on neural networks, transformers, and generative models.
10. Kaggle, [OpenAI GPT-2](https://github.com/openai/gpt-2): The official repository for GPT-2, useful for exploring text generation tasks., [TensorFlow GAN]
11. - MIT Press (https://www.deeplearningbook.org)
12. 2. "Natural Language Processing with Transformers" by Lewis Tunstall et al. (O'Reilly, 2022)
13. "The Hundred-Page Machine Learning Book" by Andriy Burkov
14. - https://themlbook.com/wiki
15. "Python Machine Learning" by Sebastian Raschka

Programme Outcomes and Course Outcomes Mapping:

Course Outcomes	Programme Outcomes (POs)								
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**	**PO9**
CO1	3	2	2	3	2	2	3	1	2
CO2	3	3	2	3	3	2	3	1	2
CO3	3	3	2	3	3	2	3	1	2
CO4	3	3	2	3	3	2	3	1	2
CO5	3	3	3	3	3	2	3	1	2
CO6	3	3	3	3	3	2	3	1	2
CO7	3	3	3	3	3				

Weight: 1 - Partially related 2 - Moderately Related 3 - Strongly related

Justification for CO-PO Mapping

CO1, CO2 → PO1, PO2, PO4 (Comprehensive Knowledge, Practical Expertise, and Problem-Solving)

Students gain a strong foundation in generative models, embeddings, and evaluation metrics for LLMs.

CO3, CO4 → PO3, PO4, PO7 (Technical Skills, Innovation, and Digital Proficiency)

Covers neural networks, optimization, tokenization, and text generation techniques essential for AI innovation

CO5, CO6 → PO2, PO5, PO7 (Practical Implementation and Research Skills)

Focuses on LLM fine-tuning, prompt engineering, and AI-driven applications in real-world domains

CO7 → PO5, PO6, PO8, PO9 (Ethics, Collaboration, and Responsibility in AI Development)

CBCS Syllabus as per NEP 2020 for M.Sc. Part-II Data Science (2023 Pattern)

Name of the Programme	: M.Sc. Data Science
Program Code	: PSDSC
Class	: M.Sc. Part – II
Semester	: IV
Course Type	:Major Mandatory Theory
Course Name	: Advanced Machine Learning and Deep Learning
Course Code	: DSC-652-MJM
No. of Credits	: 4
No. of Teaching Hours	: 60

Course Objectives:

Students successfully completing this course will be able to:

1. Understand the fundamentals of deep learning, including its history, capabilities, and the basics of neural networks.
2. Gain proficiency in implementing and training neural networks using back propagation, auto encoders, and batch normalization.
3. Learn to identify and address common issues like over fitting and generalization in deep learning models.
4. Gain advanced knowledge of deep architectures, including recurrent neural networks (RNNs), LSTM, GRU, and Generative Adversarial Networks (GANs).

Course Outcomes:

By the end of the course, students will be able to:

CO1. Evaluate, in the context of a case study, the advantages and disadvantages of deep learning neural network architectures and other approaches.

CO2. Implement deep learning models in Python using the PyTorch library and train them with real-world datasets.

CO3. Design convolution networks for handwriting and object classification from images or video.

CO4. Design recurrent neural networks with attention mechanisms for natural language classification, generation, and translation.

CO5. Explore ethical challenges and considerations in the application of deep learning.

CO6. Evaluate the performance of different deep learning models (e.g., with respect to the bias-variance trade-off, overfitting and underfitting, estimation of test error).

CO7. Perform regularization, training optimization, and hyperparameter selection on deep models.

Topics and Learning Points

Unit1: (15 L)

Introduction to Reinforcement Learning, Markov Decision Processes (MDP), Dynamic Programming, Monte Carlo Methods, Temporal Difference Learning (TD), Exploration vs. Exploitation, Deep Reinforcement Learning.

Unit 2: (15 L)

Policy Gradient Methods: Understanding Policy-Based Methods, REINFORCE Algorithm, Actor-Critic Methods, Advanced Topics in RL: Advantage Actor-Critic (A2C), Proximal Policy Optimization (PPO) ,Trust Region Policy Optimization (TRPO), Evaluation and Fine-Tuning

Unit 3: (15 L)

Introduction to Deep Learning: Basics of Artificial Intelligence, Machine Learning, and Deep Learning, Neural Networks: Biological vs. Artificial Neural Networks, Perceptron Model and Multilayer Perceptrons (MLP), Activation Functions and Loss Functions, Training Neural Networks:Forward and Backward Propagation, Gradient Descent and Optimization Algorithms (SGD, Adam, RMSprop), Hyperparameter Tuning (Learning Rate, Batch Size, Epochs), Regularization Techniques (L1, L2, Dropout),Deep Neural Networks.

Unit 4: (15L)

Convolutional Neural Networks (CNNs):Understanding Image Data, Convolutional Layers, Pooling Layers, and Fully Connected Layers, Popular CNN Architectures (LeNet, AlexNet, VGG, ResNet), Transfer Learning and Fine-Tuning

Recurrent Neural Networks (RNNs) and LSTMs: Sequential Data and Time Series Data, Basics of RNNs and Gated Recurrent Units (GRUs), Long Short-Term Memory Networks (LSTMs), Applications of RNNs in Text and Speech Processing

Advanced Topics: Autoencoders and Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), Attention Mechanism and Transformers, Real-world Applications of Deep Learning

References:

1. François Chollet,Deep Learning with Python, Manning Publications Co, ISBN: 9781617294433
2. Rajiv Chopra,Deep Learning - A Practical Approach, Khanna Publications, ISBN: 9789386173416
3. Ian Good fellow, Yoshua Bengio, Aaron Courville,Deep Learning, An MIT Press book.

Programme Outcomes and Course Outcomes Mapping:

Course Outcomes	Programme Outcomes (POs)									
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10
CO1	3	3	-	3	-	-	-	-	-	-
CO2	3	3	-	3	-	-	3	-	-	-
CO3	2	2	-	-	-	-	-	-	-	1
CO4	2	-	-	3	3	3	-	-	-	-
CO5	2	-	-	3	3	-	-	-	-	-
CO6	2	2	-	-	3	-	3	-	-	2
CO7	2	2	-	-	-	3	3	-	-	-

Weight:1 - Partially Related2 - Moderately Related3 - Strongly Related

PO1. Comprehensive Knowledge and Understanding:

CO1: (Weightage: 3)

CO2: (Weightage: 3)

CO3: (Weightage: 2)

CO4: (Weightage: 3)

CO5: (Weightage: 2)

CO6: (Weightage: 2)

CO7: (Weightage: 2)

Justification: These outcomes ensure a solid understanding of deep learning concepts, including neural networks, CNNs, RNNs, and applications in various domains.

PO2. Practical, Professional, and Procedural Knowledge:

CO1: (Weightage: 3)

CO2: (Weightage: 3)

CO3: (Weightage: 2)

CO6: (Weightage: 3)

CO7: (Weightage: 2)

Justification: Learners apply deep learning techniques using libraries like TensorFlow and PyTorch, gaining hands-on experience in real-world problem-solving.

PO4. Specialized Skills, Critical Thinking, and Problem-Solving:

CO1: (Weightage: 3)

CO2: (Weightage: 3)

CO4: (Weightage: 3)

CO5: (Weightage: 3)

Justification: Problem-solving in tasks such as image classification, NLP, and time series analysis using specialized deep learning models.

PO5. Research, Analytical Reasoning, and Ethical Conduct:

CO4: (Weightage: 3)

CO5: (Weightage: 3)

CO7: (Weightage: 3)

Justification: Learners will interpret deep learning research papers, analyze model performance, and ensure ethical AI practices.

PO6. Communication, Collaboration, and Leadership:

CO4: (Weightage: 3)

CO7: (Weightage: 3)

Justification: Involves collaborative projects, presentations, and discussions to communicate findings effectively.

PO7. Digital Proficiency and Technological Skills:

CO2: (Weightage: 3)

CO6: (Weightage: 3)

CO7: (Weightage: 3)

Justification: Strong hands-on skills in implementing and fine-tuning models using advanced tools and technologies.

PO8. Multicultural Competence, Inclusive Spirit, and Empathy:

Not Directly Addressed

Justification: This course focuses on technical concepts without directly addressing multicultural competence.

PO9. Value Inculcation, Environmental Awareness, and Ethical Practices:

Not Directly Addressed
Justification: Ethical considerations will be implied through responsible AI development, though no explicit focus is given.

PO10. Autonomy, Responsibility, and Accountability:

CO3: (Weightage: 1)

CO6: (Weightage: 2)

Justification: Independent project work and case studies encourage self-directed learning and accountability.

CBCS Syllabus as per NEP 2020 for M.Sc. Part-II Data Science (2023 Pattern)

Name of the Programme	: M.Sc. Data Science
Program Code	: PSDSC
Class	: M.Sc. Part – II
Semester	: IV
Course Type	: Major Mandatory Practical
Course Name	: Data Science Practical – VI
Course Code	: DSC-653-MJM
No. of Credits	: 2
No. of Teaching Hours	: 60

Course Objectives:

1. To develop proficiency in implementing state space search algorithms (BFS, DFS, IDS) for problem-solving tasks.
2. To apply uninformed search strategies to solve real-world problems, analysing their effectiveness.
3. To acquire knowledge and skills in knowledge representation using propositional and first-order logic.
4. To gain practical experience in logic programming with Prolog, including building expert systems through case studies.
5. To explore and understand various deep learning techniques, including CNNs, RNNs, GANs, and reinforcement learning, and apply them to solve complex tasks in computer vision and sequence modelling.

Course Outcomes:

By the end of the course, students will be able to:

CO1. Demonstrate proficiency in implementing state space search algorithms such as BFS, DFS, and IDS, and apply them to solve a variety of problems.

CO2. Apply uninformed search strategies to real-world problems, analyzing their efficiency and effectiveness in finding solutions.

CO3. Understand and employ knowledge representation techniques in both propositional and first-order logic to model complex domains.

CO4. Develop practical skills in logic programming using PROLOG and apply it to solve logic-based problems.

CO5. Design and develop expert systems using PROLOG, applying them to real-world scenarios for decision-making and problem-solving.

CO6. Explore and implement planning techniques like Blocks World and STRIPS, assessing their suitability for various planning problems.

CO7. Gain hands-on experience with perceptron learning rule and activation functions, understanding their role in neural network training.

Topics and Learning Points

Sr. No.	Title of Experiments
1.	Setup & Intro: Python, Jupyter, HuggingFace
2.	Bias Detection in LLMs
3.	Training a Simple GAN for Image Generation
4.	Fine-Tuning GPT for Custom Datasets
5.	Implementing a VAE for Latent Space Exploration
6.	Masked Language Modeling (MLM) with Statistical Evaluation
7.	Data Augmentation for Fine-Tuning LLMs
8.	Implementing State Space Search Algorithms: BFS, DFS, and IDS
9.	Practical Implementation of Logic Programming with PROLOG
10.	Introduction to Deep Learning Basics with TensorFlow
11.	Text Generation with Sampling Methods
12.	Exploring Recurrent Neural Networks with PyTorch

Programme Outcomes and Course Outcomes Mapping:

Course Outcomes	Programme Outcomes (POs)								
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9
CO1	3	3	3	3					3
CO2	2	3							
CO3									
CO4				3	2				
CO5				1			3		
CO6						2			
CO7									

Weight: 1 - Partially related 2 - Moderately Related 3 - Strongly related

PO1. Disciplinary Knowledge

CO1: Demonstrate proficiency in implementing state space search algorithms such as BFS, DFS, and IDS, and apply them to solve a variety of problems.

Justification: Strongly Related (Weightage: 3)

Proficiency in implementing state space search algorithms directly contributes to disciplinary knowledge in artificial intelligence and problem-solving techniques.

CO3: Understand and employ knowledge representation techniques in both propositional and first-order logic to model complex domains.

Justification: Moderately Related (Weightage: 2)

Understanding knowledge representation techniques is moderately related to disciplinary knowledge as it provides a foundational understanding of how complex domains are represented and manipulated in AI systems.

PO2. Critical Thinking and Problem Solving

CO1: Demonstrate proficiency in implementing state space search algorithms such as BFS, DFS, and IDS, and apply them to solve a variety of problems.

Justification: Strongly Related (Weightage: 3)

Proficiency in implementing search algorithms requires critical thinking and problem-solving skills to analyze problem characteristics and select appropriate algorithms.

CO2: Apply uninformed search strategies to real-world problems, analyzing their efficiency and effectiveness in finding solutions.

Justification: Strongly Related (Weightage: 3)

Applying search strategies involves critical thinking to evaluate their efficiency and effectiveness in solving real-world problems, contributing directly to critical thinking and problem-solving skills.

PO3. Social Competence

CO5: Design and develop expert systems using PROLOG, applying them to real-world scenarios for decision-making and problem-solving.

Justification: Partially Related (Weightage: 1)

Designing and developing expert systems involves considerations of ethical implications and societal impacts, contributing partially to social competence.

PO4. Research-related Skills and Scientific Temper

CO1: Demonstrate proficiency in implementing state space search algorithms such as BFS, DFS, and IDS, and apply them to solve a variety of problems.

Justification: Strongly Related (Weightage: 3)

Proficiency in implementing search algorithms fosters research-related skills and scientific temper by experimenting with different algorithms and analyzing their performance.

CO4: Develop practical skills in logic programming using PROLOG and apply it to solve logic-

based problems.

Justification: Strongly Related (Weightage: 3)

Developing practical skills in logic programming involves research-related skills and fosters a scientific temper by exploring different problem-solving approaches using PROLOG.

PO5. Trans-disciplinary Knowledge

CO6: Explore and implement planning techniques like Blocks World and STRIPS, assessing their suitability for various planning problems.

Justification: Moderately Related (Weightage: 2)

Exploring planning techniques involves elements of trans-disciplinary knowledge as it requires understanding various domains and their planning requirements.

PO6. Personal and Professional Competence

CO7: Gain hands-on experience with perceptron learning rule and activation functions, understanding their role in neural network training.

Justification: Partially Related (Weightage: 1)

Gaining hands-on experience with neural network training techniques contributes partially to personal and professional competence, focusing more on technical skills development.

PO7. Effective Citizenship and Ethics

CO5: Partially Related (Weightage: 1)

Justification:Designing expert systems involves considerations of ethical implications and effective citizenship in decision-making processes, contributing partially to this outcome.

PO9. Self-directed and Life-long Learning

CO1: Demonstrate proficiency in implementing state space search algorithms such as BFS, DFS, and IDS, and apply them to solve a variety of problems.

Justification: Strongly Related (Weightage: 3)

Proficiency in implementing search algorithms aligns closely with self-directed and life-long learning by fostering continuous exploration and improvement in problem-solving techniques.

CBCS Syllabus as per NEP 2020 for M.Sc. Part-II Data Science (2023 Pattern)

Name of the Programme	: M.Sc. Data Science
Program Code	: PSDSC
Class	: M.Sc. Part – II
Semester	: IV
Course Type	:Major Mandatory Theory
Course Name	: Supply Chain and Logistics Analytics
Course Code	: DSC-661-MJE (A)
No. of Credits	: 2
No. of Teaching Hours	: 30

Course Objectives:

1. To develop an understanding of the various supply chain and logistics processes.
2. To develop knowledge on structures, decision phases, measures and tools of supply
3. chains.
4. To develop understanding on the strategic, tactical and operational decision tools of
5. supply chains.
6. To impart knowledge on logistics management and related advanced tools and
7. techniques.
8. To develop and analyse the role of digital transformation of the supply chains and logistics.

Course Outcomes:

By the end of the course, students will be able to:

CO1. understand the various stages of a supply chain, from sourcing raw materials to delivering finished products to end customers.

CO2. understand the structures, decision phases, measures and tools of supply chains.

CO3. understand the strategic, tactical and operational decision tools of supply chains.

CO4. understand knowledge on logistics management and related advanced tools and techniques.

CO5. understand the role of distribution centres, warehouses, and transportation in the supply chain.

CO6. learn about collaborative approaches to planning, forecasting, and replenishing inventory.

CO7. apply supply chain and logistics concepts to real-world case studies.

Topics and Learning Points

Unit 1:

Supply Chain Management and Descriptive Analytics (15 L)

- Introduction to Supply Chain Management
  - Overview of Supply Chain Management (SCM)
  - Key Components of SCM: Sourcing, Production, Distribution, and Logistics
  - Supply Chain Network Design
  - Role of Data Science in SCM
- Data Collection and Visualization
  - Data Sources in Supply Chain
  - Data Cleaning and Preprocessing Techniques
  - Visualization Tools for SCM using Tableau and Power BI
  - Exploratory Data Analysis (EDA) on Supply Chain Data
- Descriptive Analytics in SCM
  - Demand Forecasting Methods (Moving Average, Exponential Smoothing)
  - Inventory Analysis and Management
  - ABC Analysis
  - Descriptive KPI Dashboards for Supply Chains

Unit 2: Predictive and Prescriptive Analytics for Supply Chains (15 L)

- **Predictive Analytics**
  - Time Series Forecasting Techniques (ARIMA, SARIMA)
  - Regression Models for Demand Forecasting
  - Predictive Maintenance using Sensor Data
  - Real-world Case Studies
- Prescriptive Analytics and Optimization
  - Linear and Integer Programming
  - Network Optimization for Logistics and Transportation
  - Simulation for Supply Chain Decision-Making
  - Case Study: Supply Chain Network Optimization
- Risk Management and Sustainability
  - Identifying and Mitigating Supply Chain Risks
  - Simulation and Scenario Analysis
  - Carbon Footprint Calculation and Analysis
  - Circular Economy and Reverse Logistics
  - Sustainable Sourcing with Data Science

References:

1. Neha Tikoo,Logistics and Supply Chain Management
2. Dr. Dawei Lu,Fundamentals of Supply ChainManagement
3. David Simchi – Levi & Philip Kaminsk, Designing and Managing the SupplyChain, McGraw-Hill Companies Inc.
4. David Taylor and David Brunt, Manufacturing Operations and Supply ChainManagement, Vikas Thomson Learning, 2001.
5. Donald J. Bowersox & David J. Closs, Logistical Management, TMH.
6. Jeremy F. Shapiro, Modeling and Supply Chain,. Thomson Learning, 2001.

Programme Outcomes and Course Outcomes Mapping:

Course Outcomes	Programme Outcomes (POs)								
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9
CO1	3	2		3					2
CO2	3	2			2				
CO3	3				2				
CO4	3			3	2				
CO5	3		1	1					
CO6	3	2							
CO7	3	2							

PO1. Disciplinary Knowledge (Weightage: 3 - Strongly Related)

CO1. Understand the various stages of a supply chain, from sourcing raw materials to delivering finished products to end customers.

Justification: Understanding the stages of a supply chain is fundamental knowledge in supply chain management, making it strongly related to disciplinary knowledge.

CO2. Understand the structures, decision phases, measures, and tools of supply chains.

Justification: Understanding supply chain structures and decision phases directly contributes to disciplinary knowledge in supply chain management.

CO3. Understand the strategic, tactical, and operational decision tools of supply chains.

Justification: Understanding decision tools in supply chains enhances disciplinary knowledge by providing insights into strategic, tactical, and operational decision-making processes.

CO4. Understand knowledge on logistics management and related advanced tools and techniques.

Justification: Knowledge of logistics management and advanced tools enriches disciplinary knowledge in supply chain management by incorporating advanced techniques and methodologies.

CO5. Understand the role of distribution centers, warehouses, and transportation in the supply chain.

Justification: Understanding the roles of distribution centers, warehouses, and transportation directly contributes to disciplinary knowledge by providing insights into the operational aspects of supply chain management.

CO6. Learn about collaborative approaches to planning, forecasting, and replenishing inventory.

Justification: Learning about collaborative approaches enhances disciplinary knowledge by providing insights into cooperative planning and inventory management practices in supply chains.

CO7. Apply supply chain and logistics concepts to real-world case studies.

Justification: Applying concepts to real-world case studies reinforces disciplinary knowledge by demonstrating how theoretical concepts are applied in practical supply chain scenarios.

PO2. Critical Thinking and Problem Solving (Weightage: 2 - Moderately Related)

CO1. Understand the various stages of a supply chain, from sourcing raw materials to delivering finished products to end customers.

Justification: Understanding the stages of a supply chain involves critical thinking to identify potential bottlenecks and inefficiencies in the supply chain process.

CO2. Understand the structures, decision phases, measures, and tools of supply chains.

Justification: Understanding supply chain structures and decision phases requires critical thinking to analyze their effectiveness and identify areas for improvement.

CO6. Learn about collaborative approaches to planning, forecasting, and replenishing inventory.

Justification: Learning about collaborative approaches involves critical thinking to assess the benefits and challenges of collaborative planning and forecasting in supply chains.

CO7. Apply supply chain and logistics concepts to real-world case studies.

Justification: Applying concepts to real-world case studies enhances critical thinking by requiring students to analyze complex supply chain problems and propose effective solutions.

PO3. Social Competence (Weightage: 1 - Partially Related)

CO5. Understand the role of distribution centers, warehouses, and transportation in the supply chain.

Justification: Understanding the role of distribution centers, warehouses, and transportation may involve considerations of social impacts and responsibilities, making it partially related to social competence.

PO4. Research-related Skills and Scientific Temper (Weightage: 2 - Moderately Related)

CO2. Understand the structures, decision phases, measures, and tools of supply chains.

Justification: Understanding supply chain structures and decision phases involves research-related skills to analyze and evaluate their effectiveness.

CO3. Understand the strategic, tactical, and operational decision tools of supply chains.

Justification: Understanding decision tools in supply chains requires research-related skills to assess

their effectiveness and applicability in different contexts.

CO4. Understand knowledge on logistics management and related advanced tools and techniques.

Justification: Understanding advanced tools and techniques in logistics management involves research-related skills to evaluate their effectiveness and applicability.

PO9. Self-directed and Life-long Learning (Weightage: 2 - Moderately Related)

CO1. Understand the various stages of a supply chain, from sourcing raw materials to delivering finished products to end customers.

Justification: Understanding the stages of a supply chain involves continuous learning to keep up with industry advancements and adapt to changes, making it moderately related to self-directed and life-long learning.

CBCS Syllabus as per NEP 2020 for M.Sc. Part-II Data Science (2023 Pattern)

Name of the Programme	: M.Sc. Data Science
Program Code	: PSDSC
Class	: M.Sc. Part – II
Semester	: IV
Course Type	: Major Elective Theory
Course Name	: Image Processing and Computer Vision
Course Code	: DSC-661-MJE(B)
No. of Credits	: 2
No. of Teaching Hours	: 30

Course Objectives:

1. Understand the fundamentals of image representation, pixel manipulation, and transformations.
2. Learn and apply image processing techniques such as filtering, edge detection, and feature extraction.
3. Implement object detection and recognition using both traditional computer vision methods and deep learning models.
4. Explore convolutional neural networks (CNNs) and their role in image classification and pattern recognition.
5. Gain hands-on experience with OpenCV, TensorFlow, and PyTorch for image processing and deep learning applications.
6. Apply image processing techniques in real-world domains such as healthcare, security, and autonomous systems.
7. Develop practical skills for project-based learning, including dataset preparation, model training, and performance evaluation.

Course Outcomes:

By the end of the course, students will be able to:

CO1. Explain fundamental concepts of digital image processing, including color models, transformations, and filtering.

CO2. Apply feature extraction techniques such as SIFT, HOG, and edge detection to analyze images.

CO3. Develop object detection and recognition models using OpenCV and deep learning frameworks.

CO4. Design and implement CNN architectures for image classification tasks.

CO5. Integrate machine learning and deep learning approaches for real-world image processing applications.

CO6. Evaluate the performance of image processing models using appropriate metrics and

optimization techniques.

CO7. Demonstrate the ability to build and deploy a functional image processing project, solving a real-world problem in healthcare, security, or autonomous systems.

Topics and Learning Points

Unit 1: Fundamentals of Image Processing and Feature Extraction

Introduction to Digital Images: Image representation, pixel values, color spaces (RGB, Grayscale, HSV), Image Transformations: Scaling, rotation, translation, filtering, thresholding, Edge Detection: Sobel, Prewitt, Canny edge detection, Feature Extraction: SIFT, SURF, ORB, HOG descriptors,Image Segmentation Techniques: Thresholding, Region-based segmentation, Watershed algorithm, Hands-on using Python (OpenCV, scikit-image). **(10L)**

Unit 2: Object Detection and Recognition

Object Detection Basics: Contours, bounding boxes, template matching, Feature Matching Techniques: Keypoint matching using SIFT, ORB, and Brute Force Matcher (BFM), Haar Cascades & HOG Classifiers: Face detection using OpenCV, Introduction to Deep Learning-based Detection: YOLO (You Only Look Once), SSD (Single Shot MultiBox Detector), Handwriting Recognition and OCR (Optical Character Recognition) using Tesseract., Hands-on Projects: Object detection in real-world images and videos. **(10L)**

Unit 3: Convolutional Neural Networks (CNNs) and Applications

Introduction to CNNs: Layers in CNN, convolution operation, pooling, activation functions, Popular CNN Architectures: AlexNet, VGGNet, ResNet, EfficientNet, Transfer Learning for Image Classification,Applications of CNNs:Healthcare: Disease detection (COVID-19 X-ray classification, tumor segmentation)., Security: Facial recognition, surveillance systems., Autonomous Systems: Self-driving car vision, object tracking, Hands-on Projects: Implementing CNNs using TensorFlow/Keras & PyTorch. **(10L)**

References:

1. Digital Image Processing" – Rafael C. Gonzalez & Richard E. Woods
2. "Computer Vision: Algorithms and Applications" – Richard Szeliski
3. "Pattern Recognition and Machine Learning" – Christopher M. Bishop
4. "Deep Learning for Computer Vision" – Adrian Rosebrock
5. "Mastering OpenCV with Python" – Alberto Fernández Villán

6. "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" – Aurélien Géron
7. OpenCV Documentation – https://docs.opencv.org/
8. TensorFlow & Keras Documentation – **https://www.tensorflow.org/**
9. PyTorch Documentation – https://pytorch.org/docs/

Programme Outcomes and Course Outcomes Mapping:

Course Outcomes	Programme Outcomes (POs)								
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9
CO1	3	3	2	3	2	2	3	1	1
CO2	3	3	2	3	2	2	3	1	1
CO3	3	3	2	3	2	2	3	1	1
CO4	3	3	2	3	2	2	3	1	1
CO5	3	3	2	3	2	2	3	1	1
CO6	3	3	2	3	2	2	3	1	1
CO7	3	3	2	3	2	2	3	1	1

Weight: 1 - Partially related 2 - Moderately Related 3 - Strongly related

Justification for CO-PO Mapping

CO1, CO2 → PO1 & PO2 (Comprehensive Knowledge & Practical Expertise)

Covers image representation, transformations, and feature extraction techniques.

Hands-on implementation using OpenCV, TensorFlow, and PyTorch

CO3, CO4, CO5 → PO3 (Entrepreneurial Mindset & Innovation)

Object detection and CNN-based models help students innovate in medical imaging, security, and self-driving cars.

CO2, CO3, CO4 → PO4 (Specialized Skills & Problem-Solving)

Application of image classification, object detection, and feature matching in practical scenarios.

CO5, CO6 → PO5 (Research & Analytical Reasoning)

Students conduct comparative analysis of CNN architectures, optimize models, and evaluate performance metrics.

CO6, CO7 → PO6 (Communication & Collaboration)

Students work on real-world projects, document findings, and present research outcomes.

CO4, CO5 → PO7 (Digital Proficiency & Technological Skills)

Covers advanced AI tools like YOLO, Faster R-CNN, and GANs for high-performance image processing.

CO7 → PO10 (Autonomy & Responsibility)

Students independently develop and deploy a fully functional image processing application as a project.

CBCS Syllabus as per NEP 2020 for M.Sc. Part-II Data Science (2023 Pattern)

Name of the Programme	: M.Sc. Data Science
Program Code	: PSDSC
Class	: M.Sc. Part – II
Semester	: IV
Course Type	: Major Elective Practical
Course Name	: Introduction to Hadoop
Course Code	: DSC-662 -MJE(A)
No. of Credits	: 2
No. of Teaching Hours	: 60

Course Objectives:

1. To introduce students to the foundational concepts of Hadoop and its ecosystem.
2. To enable students to set up and configure Hadoop clusters in pseudo-distributed mode.
3. To familiarize students with Hadoop Distributed File System (HDFS) and basic file management commands.
4. To provide hands-on experience in uploading various types of datasets to HDFS.
5. To introduce students to the MapReduce programming paradigm and its implementation within Hadoop.
6. To facilitate understanding of Hadoop environment variables and core configuration files for customization.

Course Outcomes:

By the end of the course, students will be able to:

CO1. Proficiency in setting up and configuring single-node Hadoop clusters to understand fundamental components.

CO2. Mastery of essential HDFS commands for efficient file management within the Hadoop ecosystem.

CO3. Capability to upload diverse datasets to HDFS using command-line tools and web interfaces for seamless data ingestion.

CO4. Understanding of MapReduce programming through execution of the Word Count example and analysis of job logs.

CO5. Competency in configuring Hadoop environment variables and exploring core configuration files for customized deployments.

CO6. Hands-on experience in writing and executing custom MapReduce jobs, enhancing data

processing skills within Hadoop.

CO7. Familiarity with various Hadoop ecosystem tools such as HBase, Hive, Pig, and Sqoop, and their respective use cases.

Topics and Learning Points

Sr. No.	Title of Experiments
1.	**Setting up a Single-Node Hadoop Cluster**: Install Hadoop on a single machine in pseudo-distributed mode to understand the basic components.
2.	**Exploring HDFS Commands**: Learn basic Hadoop File System commands such as ls, mkdir, put, get, cat, etc., to interact with files in HDFS.
3.	**Uploading Data to HDFS**: Upload sample datasets (text files, CSV files) to HDFS using Hadoop commands or HDFS web interfaces.
4.	**Running Word Count Example**: Execute the classic Word Count example using MapReduce to understand the basic MapReduce programming paradigm.
5.	**Understanding MapReduce Logs**: Analyze MapReduce job logs to understand how MapReduce jobs are executed and debug potential issues.
6.	**Configuring Hadoop Environment Variables**: Learn to set up environment variables such as HADOOP_HOME, JAVA_HOME, and PATH for Hadoop operations.
7.	**Monitoring Hadoop Services**: Use Hadoop web interfaces (like the Resource Manager UI) to monitor Hadoop services and cluster health.
8.	**Using Hadoop Streaming:** Execute MapReduce jobs using Hadoop Streaming with basic scripts written in languages like Python or Bash.
9.	**Introduction to Hadoop Streaming**: Write a basic mapper and reducer script and execute a MapReduce job using Hadoop Streaming.
10.	**Understanding YARN Resource Allocation**: Learn how YARN allocates resources to different applications and jobs running on the cluster.
11.	**Managing Hadoop Services**: Start, stop, and restart Hadoop services (NameNode, DataNode, ResourceManager, etc.) using command-line tools.
12.	**Understanding Hadoop Ecosystem**: Explore basic components of the Hadoop ecosystem like HBase, Hive, Pig, and Sqoop, and understand their use cases.

References:

1. White T. (2015). Hadoop: The Definitive Guide, O'Reilly Media.
2. Sammer E. (2012). Hadoop Operations, O'Reilly Media.
3. Turkington G. (2014). Learning Hadoop 2, Packt Publishing.
4. Lam C. (2010). Hadoop in Action, Manning Publications.
5. Venner J. (2011). Pro Hadoop, Apress.
6. Perera S., Gunarathne T. (2013). Hadoop MapReduce Cookbook, Packt Publishing.

7. Murthy A.C., Vavilapalli V. (2014). Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2, Addison-Wesley Professional.

Programme Outcomes and Course Outcomes Mapping:

Course Outcomes	Programme Outcomes (POs)								
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9
CO1	3								
CO2	3								
CO3	3								
CO4		3							
CO5									2
CO6		3							
CO7				1					

Weight: 1 - Partially related 2 - Moderately Related 3 - Strongly related

PO1. Disciplinary Knowledge (Weightage: 3 - Strongly Related)

CO1. Proficiency in setting up and configuring single-node Hadoop clusters to understand fundamental components.

Justification: Developing proficiency in setting up Hadoop clusters directly contributes to disciplinary knowledge in distributed systems and big data technologies.

CO2. Mastery of essential HDFS commands for efficient file management within the Hadoop ecosystem.

Justification: Mastering HDFS commands enhances disciplinary knowledge in data management within distributed systems, which is essential in the field of big data.

CO3. Capability to upload diverse datasets to HDFS using command-line tools and web interfaces for seamless data ingestion.

Justification: The capability to upload datasets to HDFS contributes to disciplinary knowledge by providing practical skills in data ingestion, a crucial aspect of big data systems.

PO2. Critical Thinking and Problem Solving (Weightage: 3 - Strongly Related)

CO4. Understanding of MapReduce programming through execution of the Word Count example and analysis of job logs.

Justification: Understanding MapReduce programming requires critical thinking skills to analyze job logs and identify potential issues or optimizations.

CO6. Hands-on experience in writing and executing custom MapReduce jobs, enhancing data processing skills within Hadoop.

Justification: Writing and executing custom MapReduce jobs involve problem-solving skills to design efficient data processing algorithms tailored to specific requirements.

PO4. Research-related Skills and Scientific Temper (Weightage: 1 - Partially Related)

CO7. Familiarity with various Hadoop ecosystem tools such as HBase, Hive, Pig, and Sqoop, and their respective use cases.

Justification: Familiarity with Hadoop ecosystem tools enhances research-related skills by providing exposure to various data processing and analysis tools commonly used in research and industry.

PO9. Self-directed and Life-long Learning (Weightage: 2 - Moderately Related)

CO5. Competency in configuring Hadoop environment variables and exploring core configuration files for customized deployments.

Justification: Configuring Hadoop environment variables and exploring core configuration files contribute to self-directed and life-long learning by enabling students to adapt to evolving technologies and customize deployments as needed.

CBCS Syllabus as per NEP 2020 for M.Sc. Part-II Data Science (2023 Pattern)

Name of the Programme	: M.Sc. Data Science
Program Code	: PSDSC
Class	: M.Sc. Part – II
Semester	: IV
Course Type	: Major Elective Practical
Course Name	: Web Application Development
Course Code	: DSC-662-MJE (B)

No. of Credits : 2
No. of Teaching Hours : 60

Course Objectives:

1. To introduce students to **modern web development technologies** for building dynamic web applications.
2. To provide hands-on experience in **front-end and back-end frameworks** such as **React.js/Vue.js** and **Node.js/Express.js**.
3. To enable students to design and implement **responsive user interfaces** for web applications.
4. To teach students how to integrate **databases (SQL & NoSQL)** and perform CRUD operations.
5. To equip students with skills for **building and consuming REST APIs** for web applications.
6. To introduce students to **web application security**, including authentication, authorization, and data protection.
7. To guide students in **deploying web applications** on cloud platforms such as **Heroku, Vercel, or AWS**.

Course Objectives:

By the end of the course, students will be able to:

CO1. Understand and apply **full-stack web development concepts** in a practical setting.

CO2. Develop **interactive web applications** using front-end frameworks like **React.js or Vue.js**.

CO3. Implement **back-end logic and API handling** using **Node.js and Express.js**.

CO4. Integrate web applications with **databases (MySQL/MongoDB)** and perform efficient data management.

CO5. Apply **authentication and authorization techniques** such as JWT and OAuth for secure user access.

CO6. Debug, optimize, and test web applications to improve **performance and user experience**.

CO7. Deploy and host a fully functional **web application** on a cloud platform and present it as a project.

Topics and Learning Points

This course introduces students to full-stack web development using **Python**, focusing on building and deploying a complete web application. It covers Flask/Django for backend development, HTML, CSS, Bootstrap for frontend design, and SQLite/PostgreSQL/MongoDB for database management. Students will learn API development, authentication (JWT/OAuth), data visualization (Dash/Streamlit), and deployment on cloud platforms like Heroku or AWS**.** The course follows a project-based approach**,** where students will develop a functional web app integrating backend logic, user interaction, database connectivity, and security best practices. Assessment includes

hands-on practical work, project implementation, and a final demonstration.

References:

1. Flask Web Development: Developing Web Applications with Python" – Miguel Grinberg
2. "Django for Beginners" – William S. Vincent
3. "Python Crash Course" – Eric Matthes
4. "Web Development with Django" – Ben Shaw & Saurabh Badhwar
5. "Full Stack Python" – Matt Makai

Programme Outcomes and Course Outcomes Mapping:

Course Outcomes	Programme Outcomes (POs)								
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9
CO1	3	3	2	3	2	2	3	1	1
CO2	3	3	2	3	2	2	3	1	1
CO3	3	3	2	3	2	2	3	1	1
CO4	3	3	2	3	2	2	3	1	1
CO5	3	3	2	3	2	2	3	1	1
CO6	3	3	2	3	2	2	3	1	1
CO7	3	3	2	3	2	2	3	1	1

Weight: 1 - Partially related 2 - Moderately Related 3 - Strongly related

Justification for CO-PO Mapping

CO1, CO2, CO3, CO4, CO5, CO6, CO7 → PO1 & PO2 (Comprehensive Knowledge & Practical Expertise)

The course provides students with a strong understanding of full-stack web development, including front-end, back-end, and database integration.

Students gain hands-on experience in HTML, CSS, JavaScript, React.js, Node.js, and database management, essential for real-world applications.

CO1, CO2, CO3, CO4, CO5, CO6, CO7 → PO3 (Entrepreneurial Mindset & Innovation)

Students will develop a working web application, enhancing their ability to identify problems, design solutions, and innovate.

They learn about market trends and business perspectives through project-based learning.

CO1, CO2, CO3, CO4, CO5, CO6, CO7 → PO4 (Specialized Skills & Problem-Solving)

Students will demonstrate technical proficiency in API handling, database integration, and authentication.

Debugging and performance optimization ensure critical thinking and analytical skills.

CO1, CO2, CO3, CO4, CO5, CO6, CO7 → PO5 (Research & Analytical Reasoning)

Developing a web application requires research into best practices, frameworks, security protocols, and database optimizations.

Students apply analytical reasoning while implementing solutions for scalability and security.

CO1, CO2, CO3, CO4, CO5, CO6, CO7 → PO6 (Communication & Collaboration)

Web applications require team collaboration, version control (Git), and effective communication for project management.

CO1, CO2, CO3, CO4, CO5, CO6, CO7 → PO7 (Digital Proficiency & Technological Skills)

Students work with modern tools like React.js, Node.js, MySQL, MongoDB, Git, Docker, and cloud deployment, making them industry-ready.

CO1, CO2, CO3, CO4, CO5, CO6, CO7 → PO8 (Multicultural Competence & Empathy)

By working on user-friendly and accessible web applications, students understand user needs and create inclusive digital experiences.

CO1, CO2, CO3, CO4, CO5, CO6, CO7 → PO9 (Ethical Practices & Environmental Awareness)

Secure coding, data privacy, ethical hacking, and responsible digital practices ensure students recognize the impact of their work.

CO1, CO2, CO3, CO4, CO5, CO6, CO7 → PO10 (Autonomy & Responsibility)

Completing a full-fledged web application independently demonstrates self-learning, time management, and accountability in development tasks.