Shri Vile Parle Kelavani Mandal's

Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

Scheme and Detailed syllabus (DJS23)

Final Year B.Tech

in

Computer Science and Engineering (Data Science)

(Semester VIII)

Shri Vile Parle Kelavani Mandal's
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING
(Autonomous College Affiliated to the University of Mumbai)
NAAC Accredited with "A" Grade (CGPA : 3.18)

Scheme of Semester VIII for Department of Computer Science and Engineering (Data Science)
Academic Year 2026-2027

Sr	Course Code	Course	Teaching Scheme				Semester End Examination (SEE) - A						Continuous Assessment (CA) - B						Aggregate (A+B)	Credits Earned	
			Theory (Hrs.)	Practical (Hrs.)	Tutorial (Hrs.)	Credits	Duration (Hrs)	Theory	Oral	Pract	Oral & Pract	SEE Total (A)	Term Test 1 (TT1)	Term Test 2 (TT2)	Term Test 3 (TT3)	Term Test Total (TT1+TT2+TT3)	Term Work	Total CA (B)			
1#	DJS23DCPE811	Bioinformatics	3	--	--	3	2	60	--	--	--	60	15	15	10	40	--	40	100	3	3
	DJS23DCPE812	Green Data Science																			
	DJS23DCPE813	Causal Inference and Decision Intelligence																			
	DJS23DCPE814	Immersive Systems																			
2	DJS23IPELX14	Internship (OJT) cum Project	--	36	--	18	2	--	--	--	200	200	--	--	--	--	400	400	600	18	18
3	DJS23DCSC801	Agentic AI	1	--	--	1	--	--	--	--	--	--	15	15	10	40	--	40	40	1	1
		Total	**4**	**36**	--	**22**	**4**	**60**	--	--	**200**	**260**	**30**	**30**	**20**	**80**	**400**	**480**	**740**	**22**	**22**

Any one Programme Elective Course (PEC) from 1#

Prepared by Checked by Head of Department Vice Principal Principal

Continuous Assessment (A):

Course	Assessment Tools	Marks	Time (hrs.)
Theory	a. Term Test 1 (based on 40 % syllabus)	15	45
	b. Term Test 2 (on next 40 % syllabus)	15	45
	c. Assignment / course project / group discussion / presentation / quiz/ any other.	10	--
	Total marks (a + b + c)	40	--
Audit course	Performance in the assignments / quiz / power point presentation / poster presentation / group project / any other tool.	--	As applicable
Laboratory	Performance in the laboratory and documentation.	25	
Tutorial	Performance in each tutorial & / assignment.	25	
Laboratory &Tutorial	Performance in the laboratory and tutorial.	50	

The final certification and acceptance of term work will be subject to satisfactory performance upon fulfilling minimum passing criteria in the term work / completion of audit course.

Semester End Assessment (B):

Course	Assessment Tools	Marks	Time (hrs.)
Theory / * Computer based	Written paper based on the entire syllabus.	60	2
	* Computer based assessment in the college premises.		
Oral	Questions based on the entire syllabus.	25	As applicable
Practical	Performance of the practical assigned during the examination and the output / results obtained.	25	2
Oral & Practical	Project based courses - Performance of the practical assigned during the examination and the output / results obtained. Based on the practical performed during the examination and on the entire syllabus.	As per the scheme	2

Prepared by Checked by Head of the Department Principal

Program: Computer Science and Engineering (Data Science) **Semester: VIII**

Course: Bioinformatics (DJS23DCPE811)

Pre-requisite: Linear Algebra, Mathematics for Intelligent Systems and Machine Learning.

Objectives: To develop analytical skills in learners for applying data science techniques to biological datasets, including sequence analysis, statistical methods, and data-driven approaches in genomics and proteomics.

Outcomes: On completion of the course, the learner will be able to:

1. Apply statistical and data pre-processing techniques to analyze biological datasets including genomic, transcriptomic, and proteomic data.
2. Analyze biological sequences and genomic data using sequence alignment methods, feature engineering, and exploratory data analysis techniques.
3. Analyze protein data and bioinformatics applications to interpret biological patterns and solve real-world problems in healthcare and precision medicine.

Bioinformatics (DJS23DCPE811)		
Unit	**Description**	**Duration**
1	**Introduction to Bioinformatics:** Role of bioinformatics in data science, Types of biological datasets: Genomics, Transcriptomics, Proteomics, Data Fromats: FASTA, FASTQ, GenBank, EMBL, UniProt, PDB. Data retrieval, storage, and challenges, Characteristics of biological datasets. Applications in healthcare and precision medicine.	**06**
2	**Statistical Analysis of Biological Data:** Probability distributions in biological data, Hypothesis testing (t-test, chi-square), Correlation analysis, Regression, Variance analysis, Biological vs technical variation.	**06**
3	**Sequence Analysis:** Biological sequences as data, Sequence Analysis Algorithms: Pairwise Sequence Alignment (Needleman-Wunsch Algorithm (global), Smith-Waterman Algorithm(local)), Multiple Sequence Alignment (MSA), Heuristic Search Methods (BLAST, FASTA). Scoring schemes (PAM, BLOSUM).	**08**
4	**Genomics Data Analysis:** Genome and genomics, Biological sequences as data structures, preprocessing of genomic data (quality filtering, handling ambiguous bases, sequence transformation), feature engineering for sequences (one-hot encoding, integer encoding, k-mer representation, GC-content), exploratory data analysis of genomic datasets (nucleotide distribution, k-mer frequency analysis, gene expression matrix analysis, correlation analysis), visualization techniques (sequence composition plots, gene expression heatmaps, boxplots).	**08**
5	**Protein Data Analysis:** Protein data as structured and sequence-based representations, amino acid sequences as data, levels of protein structure (primary, secondary, tertiary, quaternary), structure–function relationship as pattern interpretation, protein data sources (Protein Data Bank	**08**

	(PDB)), preprocessing and transformation of protein structural data, feature extraction (amino acid composition, physicochemical properties, structural descriptors), analysis of protein–protein interactions as network/graph data, visualization of protein structures and interaction networks.	
6	**Application of Bioinformatics in Data Science:** Disease gene identification, biomarker discovery, precision medicine, drug discovery, protein–protein interaction network analysis, cancer genomics analysis, gene expression profiling.	**06**
	Total	**42**

Books Recommended:

***Textbooks*:**

1. David A. Hendrix, Applied Bioinformatics, 2020.
2. Lesk, A. M., Introduction to Bioinformatics, University Press, 5th Edition, 2019.

***Reference Books*:**

1. Shotts, W. E., The Linux Command Line: A Complete Introduction, No Starch Press, 2nd Edition, 2019.
2. Pevsner, J., Bioinformatics and Functional Genomics, Wiley-Blackwell, 3rd Edition, 2015.
3. M. Michael Gromiha, Protein Bioinformatics: From Sequence to Function, Academic Press, 2010.
4. Harisha, S., Fundamentals of Bioinformatics, I K International Publishing House, 2010.

Web Links:

1. Oregon State University: https://open.oregonstate.education/appliedbioinformatics/
2. NPTEL Course: https://onlinecourses.nptel.ac.in/noc26_bt34/preview

Program: B.Tech in Computer Science and Engineering (Data Science) **Semester: VIII**

Course: Green Data Science (DJS23DCPE812)

Pre-requisite: Fundamentals of Data Analysis and Applied Data Science.

Objective: To develop understanding of Green Data Science principles for designing energy-efficient data acquisition, processing, and machine learning systems; To enable analysis of sustainability-driven data science applications considering computational efficiency, infrastructure optimization, and ethical frameworks such as ESG and Responsible AI.

Outcomes: On completion of the course, the learner will be able to:

1. Analyze the role of Green Data Science in promoting sustainability and addressing environmental challenges.
2. Apply data collection and energy-efficient processing methods for sustainable data management.
3. Analyze data science applications in areas like climate, energy, and smart systems.
4. Assess environmental impact, ethics, and policy aspects in data-driven systems.

Green Data Science (DJS23DCPE812)		
Unit	**Description**	**Duration**
1	**Introduction to Green Data Science:** Fundamentals of Green Data Science, sustainability in data-driven systems, environmental impact of large-scale computing, carbon footprint of AI/ML pipelines, energy consumption in training and inference, Green AI vs conventional AI, efficiency–accuracy trade-offs, and role of data science in achieving SDGs.	**06**
2	**Sustainable Data Acquisition and Management:** Data acquisition methods including adaptive sampling, duty cycling, wake-up radio, sensor selection/scheduling, event-triggered sensing, in-sensor processing, sensor fusion, and data compression, along with data quality assessment, missing data handling, bias detection, and ethical data management.	**06**
3	**Energy-Efficient and Sustainable Data Processing:** Energy-efficient processing techniques including structured pruning, mixed-precision quantization, knowledge distillation, structured and dynamic sparsity, hardware-aware NAS, TinyML, early-exit networks, and low-rank adaptation, with emphasis on performance-per-watt and reduced training/inference cost.	**07**
4	**Data Science for Sustainability Applications:** Application methods including graph-based spatio-temporal modeling, digital twin modeling, climate informatics, reinforcement learning for smart-grid optimization, semantic segmentation for remote sensing, and physics-informed machine learning, for sustainability problems in climate, energy, agriculture, and environmental monitoring.	**08**
5	**Green Computing & Infrastructure:** Infrastructure techniques including carbon-aware scheduling, dynamic voltage and frequency scaling (DVFS), energy-proportional computing, VM/container scheduling, edge-cloud collaborative computing, thermal-aware data center optimization, and renewable-aware workload management.	**07**

6	**Sustainability Assessment and Reporting** Introduction to Sustainability Assessment, Sustainability Metrics for Data Science Systems, Energy Consumption Assessment of Machine Learning Models, Carbon Footprint Measurement and Analysis using CodeCarbon and CarbonTracker, Resource Utilization Analysis, Lifecycle Assessment (LCA) of Data Science Solutions, ESG-Based Evaluation of Data-Driven Systems, Sustainability Reporting and Dashboard using Microsoft Sustainability Manager, Case Studies on Sustainable Data Science Applications.	**08**
	Total	**42**

Books Recommended:

Textbooks:

1. Raghavendra Selvan, Sustainable AI: Tools for Moving Toward Green AI, O'Reilly Media, 2024.
2. Wu-Chun Feng (Ed.), Green Computing: Large Scale Energy Efficiency, CRC Press,2014.
3. Bhuvan Unhelkar, Green IT Strategies and Applications – Using Environmental Intelligence, CRC Press, 2014.

Reference Books:

1. Green Artificial Intelligence: Sustainable Techniques and Applications, Springer, 2025.
2. Data Science Applied to Sustainability Analysis, Jennifer Dunn and Prasanna Balaprakash (Eds.), Elsevier, 2021.
3. TinyML: Machine Learning with TensorFlow Lite on Arduino and Ultra-Low-Power Microcontrollers, Pete Warden and Daniel Situnayake, O'Reilly Media, 2019.
4. Building Green Software: A Sustainable Approach to Software Development and Operations, Anne Currie, Sarah Hsu, and Sara Bergman, O'Reilly Media, 2024.
5. Life Cycle Assessment: Theory and Practice, Michael Z. Hauschild, Ralph K. Rosenbaum, and Stig Irving Olsen (Eds.), Springer, 2018.

Weblinks:

1. NPTEL: https://onlinecourses.nptel.ac.in/noc26_hs100/preview
2. MIT Sustainability AI: https://www.mit.edu
3. Cornell AI for Sustainability: https://sustainability.ai.cornell.edu/
4. UN SDG Platform: https://sdgs.un.org/goals
5. UN Data Portal: https://data.un.org

Program: B.Tech in Computer Science and Engineering (Data Science) **Semester: VIII**

Course: Causal Inference & Decision Intelligence (DJS23DCPE813)

Pre-requisite: Statistics for data science and Machine learning

Objectives: To equip students with the principles and tools of causal inference for making reliable, data-driven decisions beyond prediction.

Outcomes: On completion of the course, learner will be able to:

1. Apply and differentiate correlation and causation using causal reasoning frameworks and real-world scenarios.
2. Analyze causal graphs (DAGs) to identify confounders, mediators, and colliders in practical problems.
3. Analyze and estimate causal effects using techniques such as ATE, propensity score methods, and RCTs.
4. Evaluate the decision-making strategies by integrating causal inference methods, experimentation, and policy learning.

Causal Inference & Decision Intelligence (DJS23DCPE813)		
Unit	**Description**	**Duration**
1	**Introduction to Causality & Decision Failures:** Correlation vs causation; limits of predictive ML, Real-world failures (healthcare, policy, recommendation systems), Introduction to causal thinking, Introduction to Judea Pearl's causal reasoning framework, Pearl's Ladder of Causation (association, intervention, counterfactuals), Case-led: misleading dashboards, Simpson's paradox.	**06**
2	**Causal Graphs, DAGs & Structural Models:** Directed Acyclic Graphs, assumptions in causal diagrams, Structural Causal Models, confounders/mediators/colliders, d-separation and conditional independence. Visual problem solving through DAG drawing for education, pricing, and traffic scenarios.	**08**
3	**Identification & Estimation of Causal Effects:** Potential outcomes framework (Rubin model), Counterfactual reasoning & fundamental problem, SUTVA, ignorability, overlap, ATE (Average Treatment Effect: effect of treatment on the overall population. ATT (Average Treatment Effect on the Treated), ATC (Average Treatment Effect on the Control), ATE = E[Y(1) − Y(0)]: Average difference between treated and untreated outcomes, Backdoor & front door criteria (DAG intuition), Do-operator (conceptual understanding), Propensity score: Estimation Matching (nearest neighbor, caliper), Covariate balance checking, Inverse Probability Weighting (IPW),Sensitivity to hidden confounders.	**08**
4	**Experimental Design, RCTs & A/B Testing:** Randomized Controlled Trials (RCTs), A/B testing design (industry practices) Randomization strategies, Bias control & confounders, Metric selection (north-star vs proxy), Power analysis (intuition), Pitfalls: Peeking, novelty effect, interference (network effects) and online experimentation practices used in companies like Google and Netflix.	**06**
5	**Advanced Causal Methods & Causal ML:** Instrumental variables, Regression discontinuity, Difference-in-differences, Heterogeneous treatment effects, Uplift modeling, Causal ML additions: Meta-learners: T-learner, S-learner, X-learner, Intro to Double Machine Learning (DML),	**08**

	Basic idea of causal forests. Case studies: Retail (coupon targeting), Banking (risk intervention), Mobility (pricing & demand)	
6	**Decision Intelligence & Real Applications**: Decision-making under uncertainty, Translating causal insights into policies, Policy learning basics, Connection to Reinforcement Learning: Offline vs. online decision-making, Counterfactual policy evaluation (conceptual), Building decision pipelines: Problem → DAG → Estimation → Decision, Ethics in causal decision-making, Capstone-style case studies: Product analytics, Healthcare decisions, Economic policy.	**06**
	Total	**42**

Books Recommended:

Textbooks:

1. Miguel Hernán & James Robins, "Causal Inference: What If", Oxford University Press, 2025.
2. Causal Inference: The Mixtape – Scott Cunningham, Yale University Press, 2021.

Reference Books:

1. Causal Inference for Data Science – Aleix Ruiz de Villa, Publisher: Miguel Hernán & James Robins, Oxford University Press, 2025.
2. Judea Pearl, "Causality: Models, Reasoning, and Inference". Cambridge Press, 2009.

Web links:

1. Johns Hopkins Engineering for Professionals: Introduction to Causal Inference - 625.691 | Hopkins EP Online : https://ep.jhu.edu/courses/625691-introduction-to-causal-inference/
2. Causal Inference | Coursera: https://www.coursera.org/learn/causal-inference
3. Causal AI: An Introduction : https://www.udemy.com/course/causal-ai-an-introduction/

Program: B.Tech in Computer Science and Engineering (Data Science) **Semester: VIII**

Course: Immersive Systems (DJS23DCPE814)

Pre-requisite: Computer System Fundamentals and Design Thinking.

Objective:

To understand the principles of Human-Computer Interaction (HCI), design and develop AR/VR-based interactive systems, apply 3D visualization and interaction techniques for data-driven applications, and evaluate user experience (UX) and usability in immersive environments.

Outcomes: On completion of the course, the learner will be able to:

1. Analyze principles of Human-Computer Interaction to design user-centered and accessible interactive systems.
2. Examine interaction design principles and usability testing methods to improve user experience in software systems
3. Analyze user requirements using research methods, task analysis, and prototyping techniques to develop effective user interfaces
4. Evaluate AR/VR technologies and develop immersive solutions for practical applications.

Immersive Systems (DJS23DCPE814)		
Unit	**Description**	**Duration**
1	**Foundations of HCI:** Introduction to Human-Computer Interaction, History and evolution of user interfaces, Interaction paradigms (CLI, GUI, touch, voice, immersive), Human factors: cognition, perception, memory; Usability principles (learnability, efficiency, satisfaction).	**06**
2	**User-Centered Design (UCD):** Principles of user-centered design, User research methods (interviews, surveys, personas), Task analysis and user journey mapping, Prototyping (low-fidelity & high-fidelity), Accessibility and inclusive design standards, UX Design Considerations for Immersive Environments, Applications of Immersive Technologies in Human–Computer Interaction.	**08**
3	**Interaction Design & Evaluation:** Interaction design principles (affordances, feedback, constraints), Heuristic evaluation (e.g., Jakob Nielsen's heuristics), Usability testing methods (A/B testing, think-aloud protocol), UX metrics and evaluation techniques, Cognitive models (e.g., GOMS, Fitts' Law).	**07**
4	**Introduction to AR & VR:** Basics of Augmented Reality and Virtual Reality, Differences between AR, VR, and MR (Mixed Reality), Hardware overview (headsets, sensors, controllers), Key platforms (e.g., Unity, Unreal Engine), Applications in gaming, healthcare, education, industry.	**06**
5	**Designing for AR/VR Experiences**: 3D interaction techniques (gaze, gesture, controllers), Spatial UI/UX design principles, Presence, immersion, and user experience in VR, Challenges: motion sickness, latency, ergonomics, Storyboarding and prototyping immersive environments.	**07**

6	**Development & Future Trends in AR/VR:** Basics of AR/VR development workflows, SDKs and tools (ARCore, ARKit, WebXR) Testing and evaluation of immersive systems, Ethics and privacy in immersive tech Future trends: metaverse, haptics, AI integration.	**08**
	Total	**42**

Books Recommended:

Textbooks:

1. Designing the User Interface: Strategies for Effective Human-Computer Interaction – Ben Shneiderman et al., Pearson, 2022.
2. Augmented Reality Principles and Practice, Dieter Schmalstieg, Tobias Höllerer, Addison-Wesley, 2016.
3. Human-Computer Interaction – Alan Dix, Janet Finlay, Gregory D. Abowd, Russell Beale, Pearson, Latest Edition, 2004.

Reference Books:

1. Interaction Design Beyond Human-Computer Interaction, Helen Sharp, Yvonne Rogers, Jenny Preece, Wiley, 2019.
2. The UX Book: Agile UX Design for a Quality User Experience – Rex Hartson, Pardha Pyla, Elsevier, 2019.
3. Designing Immersive 3D Experiences – Focused on AR/VR UX design (modern addition).
4. Augmented Human: How Technology Is Shaping the New Reality – O'Reilly, 2017
5. Virtual Reality, Steven M. LaValle, Cambridge University Press, 2017.

Weblinks:

1. NPTEL: https://nptel.ac.in/courses (HCI & UI/UX Courses)
2. MIT OpenCourseWare (HCI): https://ocw.mit.edu
3. Stanford University: https://hci.stanford.edu
4. Unity Technologies Learn Platform: https://learn.unity.com
5. Google AR & VR Development (ARCore): https://developers.google.com/ar
6. Apple ARKit Resources: https://developer.apple.com/augmented-reality/

Prepared by Checked by Head of the Department Principal