Shri Vile Parle Kelavani Mandal's

Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

Scheme and Detailed Syllabus (DJS22)

Final Year B.Tech

in

Artificial Intelligence (AI) and Data Science

Honors Degree Program

in

Computational Biology

(Semester VIII)

To be Prepared by: - Board of Studies in Artificial Intelligence (AI) and Data Science

To be Recommended by: - Academic Council of Dwarkadas J. Sanghvi College of Engineering

To be Approved by: - Governing Body of Dwarkadas J. Sanghvi College of Engineering

With effect from the Academic Year: 2025-2026

Scheme for Undergraduate Program in Artificial Intelligence (AI) and Data Science with honors in Computational Biology:
DJS22(Autonomous)   (Academic Year 2025-2026)

Sr. No	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)	Total (TT1 +TT2)	Term work	CA Total(B)			
	SEM V																			
1	DJS22ADHN1C1	Introduction to Biological Science	4	--	--	4	2	65	--	--	--	65	20	15	35	--	35	100	4	4
	SEMVI																			
2	DJS22ADHN1C2	Algorithms for Computational Biology	4	--	--	4	2	65	--	--	--	65	20	15	35	--	35	100	4	5
3	DJS22ADHN1L1	Algorithms for Computational Biology Laboratory	--	2	--	1	--	--	--	--	--	--	--	--	--	25	25	25	1	
	SEM VII																			
4	DJS22ADHN1C3	Bigdata in Bioinformatics	4	--	--	4	2	65	--	--	--	65	20	15	35	--	35	100	4	5
5	DJS22ADHN1L2	Bigdata in Bioinformatics Laboratory	--	2	--	1	--	--	--	--	--	--	--	--	--	25	25	25	1	
	SEM VIII																			
6	DJS22ADHN1C4	Gneomic Data Science	4	--	--	4	2	65	--	--	--	65	20	15	35	--	35	100	4	4
		Total	16	4	0	18	8	260	0	0	0	260	80	60	140	50	190	450	18	18

Prepared by   Checked by   Head of the Department   Vice Principal   Principal

Program: Artificial Intelligence (AI) and Data Science   **Final Year B.Tech**   **Semester: VIII**

Course: Genomic Data Science (DJS22ADHN1C4)

Prerequisite: Basics of Computational Biology courses

Course Objectives: The Objectives of course are

1. To provide an understanding of the foundational concepts in genomics, including DNA, RNA, genes, and genomes, and the role of the central dogma of molecular biology in gene expression.
2. To introduce data science applications in genomics, covering common data formats, sequencing technologies, alignment, mapping, variant analysis, and the interpretation of genomic data for applications in medicine, research, and biotechnology.

Course Outcomes: Students will be able to

1. Explain Fundamental Concepts in Genomics and Data Science.
2. Analyze DNA Sequencing Technologies and Data Characteristics.
3. Apply Sequence Alignment and Mapping Techniques.
4. Perform Variant Calling and Genomic Data Annotation.
5. Investigate Population Genomics and Trait Associations.
6. Evaluate Ethical, Legal, and Social Implications of Genomics.

Genomic Data Science (DJS22ADHN1C4)		Duration
Unit	**Description**	
1	**Module 1: Overview of Genomics and Data Science** Introduction to Genomics- DNA, RNA, genes, genomes, Central Dogma of molecular biology, Applications of genomics in medicine, research, and biotechnology, Role of data science in genomics, Common data formats (FASTA, VCF, BAM)	09
2	**Module 2: DNA Sequencing Technologies** Types of Sequencing Technologies- Sanger sequencing, Next-Generation Sequencing (NGS), Single-molecule sequencing (e.g., PacBio, Oxford Nanopore), Data Generation and Characteristics- Sequencing depth, quality scores, Raw vs. processed genomic data	09
3	**Module 3: Sequence Alignment and Mapping** Basics of Sequence Alignment- Pairwise and multiple sequence alignment (Needleman-Wunsch, Smith-Waterman), Alignment algorithms (BLAST, BWA, Bowtie)	09
4	**Module 4: Variant Calling and Analysis** Detecting Genetic Variations- Single Nucleotide Polymorphisms (SNPs), Insertions/Deletions (INDELs) **Genomic Data Annotation**: Gene prediction methods, Annotating coding vs. non-coding regions	09
5	**Module 5: Population Genomics and GWAS** Allele frequency, population structure, Hardy-Weinberg equilibrium, GWAS- Methods for linking genetic variants to traits **Transcriptomics and RNA-Seq Data Analysis**: RNA sequencing and gene expression analysis, Differential expression analysis	10

6	**Module 6: Epigenomics and Regulatory Genomics** DNA methylation, histone modification, Regulatory elements in the genome (promoters, enhancers) **Introduction to Structural Genomics and Metagenomics:** Structural Variations in Genomics, Copy number variations (CNVs), translocations, inversions, Study of microbial communities from genomic data, Shotgun sequencing, 16S rRNA sequencing **Ethical, Legal, and Social Issues in Genomics:** Privacy, data sharing, and consent, Implications of genomic data in personalized medicine	10	
	Total	56	

Textbooks:

1. Durbin, "Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids", Cambridge University Press.
2. M. Zvelebil and J.O. Baum, "Understanding Bioinformatics", Published by Garland Science, 2008.
3. David Mount Bioinformatics: Sequence and Genome Analysis
4. Richard C. Deonier, Simon Tavare, Michael S. Waterman,"Computational Genome Analysis: An Introduction",Springer India

Reference Books:

1. Phillip Compeau and Pavel Pevzner, "Bioinformatics algorithms: an active learning approach", Published by Active Learning Pub.
2. Joseph Felsenstein, Inferring Phylogenies.
3. V. Makinen, D. Belazzougui, F. Cunial and A. Tomescu," Genome-Scale Algorithm Design", Cambridge University Press, 2015.

Web Links:

1. Big Data Analytics in Genomics
   https://www.google.co.in/books/edition/Big_Data_Analytics_in_Genomics/5_xRDQAAQBAJ?hl=en&gbpv=1&dq=Data+Science+for+Genomics&printsec=frontcover
2. Data Science for Genomics : https://www.sciencedirect.com/book/9780323983525/data-science-for-genomics

Online Resources:

1. Introduction to Genomic Data Science https://www.classcentral.com/course/bioinformatics-the-university-of-california- san-d-8962
2. Genomic Data Science Specialization: https://www.coursera.org/specializations/genomic-data-science
3. Computational Genomics: https://onlinecourses.nptel.ac.in/noc24_bt03/preview
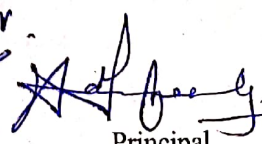
Prepared by   Checked by   Head of the Department   Vice Principal   Principal