Faculty of Science, Technology and Architecture | School of Engineering
Department of Biotechnology & Chemical Engineering

M.Tech in Computational Biology syllabus (Batch 2026 and onwards)

FIRST SEMESTER

BIO6101 Cell and Molecular Biology [3-1-0-4]

Cell Structure: Comparison between plant and animal cells; Plasma membrane; Electrical properties of membrane, Modification of plasma membrane and intracellular junctions; Organization of plant cell wall. Cell Signaling: Communication between cells and their environment. Introduction to Cytoplasmic Organelles and Cytoskeleton: Protoplasm; Mitochondria; Chloroplast; ER; Golgi complex; Lysosome, Endosome, Ribosome; Centriole; Nucleus. Chromosomes, Chromatin and Nucleosome: Chromosome structure in bacteria and eukaryotes, Centromere, Telomere, Hetero- and euchromatin, Nucleosome model and radial-loop scaffold model. Overview of Cell Cycle: Stages of cell cycle, Cell cycle control, Mitotic and meiotic cell division; Distinction between mitosis in plant and animal. Cell Death: Apoptosis and Necrosis. Nucleic Acids Structure: DNA as genetic material, Watson-Crick model, A, B and Z forms of DNA; RNA types, Distinctions between RNA and DNA. The Central Dogma: Overview of synthesis of DNA, RNA and protein. The Genetic Code: Genetic code and its properties. Gene Cloning: Restriction endonuclease and cloning vector, Screening of cloned DNA.

References

1. B. Alberts, A. Jahusan, J. Levis, M. Raff, K. Roberts and P. Walter. *Molecular Biology of Cell*, Garland Science, USA, 2002.
2. H. Lodish and D. Baltimore. *Molecular Cell Biology,* WH Freeman and Company, USA, 2012.
3. B. Lewin. *Genes XII*, Jones and Bartlett Publishers, USA, 2014.
4. G. Karp. *Cell and Molecular Biology – Concepts and Experiments*, John Wiley & Sons, Inc. USA, 2009.
5. T.A. Brown. *Genomes*, Garland Science, New York, 2006.

CSB6102 Bioinformatics [3-1-0-4]

Introduction to bioinformatics: Scope and applications of bioinformatics in genomics, transcriptomics, proteomics, drug discovery, and precision medicine. Biological databases: Primary, secondary, and composite databases; nucleotide databases, protein databases, pathway and disease databases. Sequence data formats and standards: FASTA, FASTQ, GFF, VCF. Sequence alignment: Pairwise sequence alignment concepts, scoring matrices, global and local alignment algorithms. Database searching: BLAST algorithms and variants, interpretation of BLAST results and E-values. Multiple sequence alignment: Progressive alignment methods, evaluation of alignments. Phylogenetic analysis: Phylogenetic tree construction methods (distance-based, maximum parsimony, maximum likelihood), tree visualization and interpretation. Genome annotation: Structural and functional annotation, gene prediction methods, annotation pipelines. Introduction to transcriptomics, proteomics and metabolomics informatics. Ethical issues and data sharing in bioinformatics.

CSB6102	Bioinformatics

CO.1	Apply bioinformatics concepts to classify biological databases, data formats, and applications.
CO.2	Apply sequence alignment and database search techniques to analyze nucleotide and protein sequences.
CO.3	Apply phylogenetic, motif, and genome annotation methods to study evolutionary and functional relationships.
CO.4	Apply integrative bioinformatics tools for functional annotation and biological data interpretation.

References

1. D. W. Mount, *Bioinformatics: Sequence and Genome Analysis*, 3rd ed. Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press, 2024.
2. A. D. Baxevanis and B. F. F. Ouellette, *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*, 4th ed. Hoboken, NJ, USA: Wiley-Blackwell, 2020.
3. P. Pevsner, *Bioinformatics and Functional Genomics*, 4th ed. Hoboken, NJ, USA: Wiley-Blackwell, 2022.
4. R. Durbin, S. R. Eddy, A. Krogh, and G. Mitchison, *Biological Sequence Analysis*. Cambridge, U.K.: Cambridge University Press, 2018.

CSB6103 AI/ML for Biological systems [3-1-0-4]

Introduction to machine learning (ML) and artificial intelligence (AI) in biological sciences: Role of ML/AI in genomics, transcriptomics, proteomics, imaging, and drug discovery. Types of learning: Supervised, unsupervised, semi-supervised, and reinforcement learning. Data preprocessing: Data cleaning, normalization, feature selection, dimensionality reduction, handling class imbalance. Classical machine learning algorithms: Linear and logistic regression, k-nearest neighbors, naïve Bayes, decision trees, random forests, support vector machines (SVM). Model evaluation and validation: Cross-validation, performance metrics, overfitting and underfitting. Deep learning fundamentals. Applications in biology and medicine: Gene expression analysis, biomarker discovery, disease classification, survival analysis, protein structure prediction, molecular property prediction, drug–target interaction prediction. Explainable AI (XAI) in biological data analysis. Ethical considerations and challenges: Data bias, reproducibility, model generalization, transparency, and responsible AI in biomedical research.

CSB6103	**AI/ML for Biological systems**
CO.1	Apply machine learning and artificial intelligence techniques to preprocess and analyze biological datasets.
CO.2	Analyze biological and biomedical data using classical machine learning algorithms and performance metrics.
CO.3	Develop and evaluate deep learning models for complex biological data such as gene expression and imaging datasets.
CO.4	Interpret ML and deep learning models using explainable AI methods for biological and medical applications.

References

1. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 3rd ed. New York, NY, USA: Springer, 2023.
2. K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2022.
3. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2024.
4. A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 3rd ed. Sebastopol, CA, USA: O'Reilly Media, 2023.

CSB6104 Genomics and Transcriptomics [3-1-0-4]

Introduction to genomics and transcriptomics: Scope, applications, and role in biomedical research and precision medicine. Genome organization and sequencing strategies: Structural and functional genomics, whole-genome sequencing, whole-exome sequencing, comparative genomics, and genome assembly concepts. Genome annotation. Epigenomics. Principles of transcriptomics. Transcriptomics technologies: Microarray-based gene expression profiling, experimental design, normalization, and differential gene expression analysis. RNA sequencing. Introduction to single-cell transcriptomics. Functional interpretation of genomics and transcriptomics data: Gene ontology analysis, pathway enrichment, gene set enrichment analysis, co-expression analysis, and basic integration of genomics and transcriptomics data. Applications in disease biology: Cancer genomics, transcriptomic signatures, biomarker discovery, and case studies from human diseases.

CSB6104	Genomics and Transcriptomics
CO.1	Apply genomics and transcriptomics principles to analyze genome structure, gene organization, and transcriptome complexity.
CO.2	Analyze sequencing- and microarray-based transcriptomics data to identify differentially expressed genes and transcriptional patterns.
CO.3	Evaluate genomic and epigenomic variations associated with biological processes and disease states.
CO.4	Interpret functional and pathway-level insights from genomics and transcriptomics data for biomedical applications.

References

1. A. D. Baxevanis and B. F. F. Ouellette, *Bioinformatics and Functional Genomics*, 4th ed. Hoboken, NJ, USA: Wiley, 2023.
2. I. Shmulevich and E. R. Dougherty, *Genomic Signal Processing*. Princeton, NJ, USA: Princeton University Press, 2022.
3. R. K. Dash, A. M. Dutta, and S. K. Sahoo, *Systems Biology: Concepts and Applications*. London, U.K.: Academic Press, 2023.
4. S. Wang, X. Li, Y. Zhang, and H. Chen, *Multi-Omics Data Analysis*. Boca Raton, FL, USA: CRC Press, 2024.

CSB6130 Computational biology lab- I [0-0-4-2]

List of Experiments:

1. Retrieval and exploration of nucleotide and protein sequences from biological databases (NCBI, UniProt, PDB, KEGG).
2. Study and conversion of biological data formats: FASTA, FASTQ, GFF, and VCF.
3. Pairwise sequence alignment using global (Needleman–Wunsch) and local (Smith–Waterman) algorithms.
4. BLAST analysis (BLASTn, BLASTp) and interpretation of E-value, identity, and bit score.
5. Multiple Sequence Alignment (MSA) using Clustal Omega or MUSCLE and identification of conserved regions.
6. Construction and interpretation of phylogenetic trees using distance-based and maximum likelihood methods.
7. Motif and domain identification using Pfam/InterPro and functional annotation of proteins.
8. Gene prediction and basic genome annotation using available annotation tools.
9. Differential gene expression analysis using a public dataset (GEO) and identification of significant genes.
10. Pathway and protein–protein interaction network analysis using KEGG/STRING.
11. Data preprocessing for biological datasets: normalization, feature scaling, handling missing values.
12. Dimensionality reduction using Principal Component Analysis (PCA) on gene expression data.
13. Supervised machine learning for disease classification using Logistic Regression or SVM.
14. Unsupervised learning using K-means or hierarchical clustering for biological data analysis.
15. Random Forest model development and feature importance analysis for biomarker discovery.

CSB6130	**Computational biology lab- I**
CO.1	Apply bioinformatics databases and data formats to retrieve, organize, and manage nucleotide, protein, and genomic datasets.
CO.2	Analyze sequence alignment, BLAST, phylogenetic, and annotation results to interpret functional and evolutionary relationships.
CO.3	Evaluate transcriptomics and systems biology datasets to identify differentially expressed genes, pathways, and interaction networks.
CO.4	Develop machine learning models for biological data classification and biomarker discovery using appropriate preprocessing and validation techniques.

References

1. D. W. Mount, *Bioinformatics: Sequence and Genome Analysis*, 3rd ed. Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press, 2024.
2. A. D. Baxevanis and B. F. F. Ouellette, *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*, 4th ed. Hoboken, NJ, USA: Wiley-Blackwell, 2020.
3. P. Pevsner, *Bioinformatics and Functional Genomics*, 4th ed. Hoboken, NJ, USA: Wiley-Blackwell, 2022.
4. E. Alpaydin, *Introduction to Machine Learning*, 4th ed. Cambridge, MA, USA: MIT Press, 2020.

CSB6131 Cell and Molecular biology lab [0-0-4-2]

List of Experiments

1. Laboratory safety training, GLP, and biosafety practices
2. Isolation and quantification of genomic DNA
3. Isolation and quality assessment of total RNA
4. Agarose gel electrophoresis of nucleic acids
5. Polymerase Chain Reaction (PCR) and optimization
6. Protein extraction and quantification (Bradford/Lowry)
7. SDS-PAGE for protein separation
8. Enzyme-linked immunosorbent assay (ELISA)
9. UV–Visible and fluorescence spectroscopy experiments
10. Enzyme kinetics and biochemical assays

CSB6131	Cell and Molecular biology lab
CO.1	Apply laboratory safety guidelines, Good Laboratory Practices (GLP), and biosafety procedures while handling biological samples and reagents.
CO.2	Perform isolation, quantification, and quality assessment of nucleic acids and proteins using standard molecular biology techniques.
CO.3	Analyze nucleic acids and proteins using electrophoresis, PCR, spectroscopy, and immunoassay techniques to interpret experimental results.
CO.4	Evaluate enzyme kinetics and biochemical assay data to determine reaction parameters and biological significance.

References

1. J. Sambrook and D. W. Russell, *Molecular Cloning: A Laboratory Manual*, 4th ed. Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press, 2012.
2. D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, 8th ed. New York, NY, USA: W. H. Freeman, 2021.
3. K. Wilson and J. Walker, *Principles and Techniques of Biochemistry and Molecular Biology*, 8th ed. Cambridge, U.K.: Cambridge University Press, 2018.
4. M. R. Green and J. Sambrook, *PCR Protocols: A Guide to Methods and Applications*. Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press, 2019.
5. B. R. Glick, J. J. Pasternak, and C. L. Patten, *Molecular Biotechnology: Principles and Applications of Recombinant DNA*, 5th ed. Washington, DC, USA: ASM Press, 2018.

SECOND SEMESTER

CSB6201 Biostatistics with R [3-1-0-4]

Introduction to biostatistics: Role of statistics in biological and medical research, types of data, scales of measurement, descriptive statistics. Probability and distributions,Sampling and estimation, Hypothesis testing: Null and alternative hypotheses, p-values, Type I and Type II errors, parametric tests, non-parametric tests. Analysis of variance. Correlation and regression: Pearson and Spearman correlation, simple and multiple linear regression, interpretation of regression coefficients. Introduction to statistical modeling in biology: Logistic regression and basic survival analysis concepts.

R programming for data analysis: Introduction to R and RStudio environment, data types and structures, data import/export, data manipulation using tidyverse (dplyr, tidyr), exploratory data analysis and visualization using ggplot2, boxplots, histograms, scatter plots. Statistical analysis using R: Implementation of hypothesis tests, ANOVA, correlation and regression in R, multiple testing correction (Bonferroni, FDR). Case studies using biological datasets: Gene expression data, clinical datasets, reproducible analysis using R Markdown.

CSB6201	Biostatistics with R
CO.1	Apply biostatistical concepts, probability models, and sampling methods to analyze biological and medical data.
CO.2	Analyze biological datasets using hypothesis testing, ANOVA, correlation, regression, and basic survival analysis techniques.
CO.3	Perform statistical analysis, data manipulation, and visualization using R and tidyverse packages.
CO.4	Interpret and communicate statistical results using reproducible workflows and real biological datasets.

References

1. R. R. Sokal and F. J. Rohlf, *Biometry: The Principles and Practice of Statistics in Biological Research*, 5th ed. New York, NY, USA: W. H. Freeman and Company, 2023.
2. P. Dalgaard, *Introductory Statistics with R*, 3rd ed. New York, NY, USA: Springer, 2022.

3. H. Wickham and G. Grolemund, *R for Data Science*, 2nd ed. Sebastopol, CA, USA: O'Reilly Media, 2023.
4. J. H. McDonald, *Handbook of Biological Statistics*, 4th ed. Baltimore, MD, USA: Sparky House Publishing, 2022.

CSB6202 Advanced Multi-Omics & Systems Biology [3-1-0-4]

Introduction to advanced omics and systems biology. Overview of omics technologies: Genomics, transcriptomics, proteomics, metabolomics, epigenomics, lipidomics, and microbiomics. Experimental design considerations and data preprocessing. Multi-omics data integration strategies: Horizontal and vertical integration approaches, statistical and probabilistic methods, correlation-based analysis, network-based integration, and machine learning–based multi-omics frameworks. Dimensionality reduction and feature selection techniques for high-dimensional omics data. Single-cell and spatial multi-omics. Systems biology foundations: Biological networks and analysis. Dynamic and predictive systems biology modeling: Mathematical modeling approaches including Boolean models, ordinary differential equation (ODE) models, stochastic models, and flux balance analysis. Applications of advanced omics and systems biology: Disease mechanism elucidation, cancer systems biology, biomarker discovery, drug target identification, drug response prediction, and personalized medicine.

CSB6202	Advanced Multi-Omics & Systems Biology
CO.1	Apply advanced omics technologies and preprocessing methods for multi-layer biological data analysis.
CO.2	Analyze and integrate multi-omics datasets using statistical, network-based, and machine learning approaches.
CO.3	Evaluate high-dimensional, single-cell, and spatial omics data using dimensionality reduction and feature selection techniques.
CO.4	Develop systems biology models (Boolean, ODE, stochastic, FBA) to predict biological network behavior and disease mechanisms.

References

1. A. D. Baxevanis, D. B. Davison, and G. D. Page, *Bioinformatics and Functional Genomics,* 4th ed. Hoboken, NJ, USA: Wiley, 2023.
2. R. K. Dash and K. A. Kahn, *Systems Biology: Concepts and Applications*. London, U.K.: Academic Press, 2023.
3. S. Wang, X. Li, Y. Zhang, and H. Chen, *Multi-Omics Data Analysis*. Boca Raton, FL, USA: CRC Press, 2024.
4. E. Klipp, W. Liebermeister, C. Wierling, A. Kowald, H. Lehrach, and R. Herwig, *Systems Biology: A Textbook*, 2nd ed. Weinheim, Germany: Wiley-VCH, 2022.

CSB6203 Structural Biology [3-1-0-4]

Introduction to structural bioinformatics: Scope and applications in protein function prediction, drug discovery, and molecular biology. Protein structure fundamentals: Levels of protein structure, folding principles, secondary structure elements, motifs and domains. Structural databases: Protein Data Bank, SCOP, CATH; structure classification and annotation. Experimental methods for structure determination. Protein structure visualization and analysis: Molecular visualization tools, structure quality assessment, Ramachandran plots, structure alignment and comparison. Protein structure prediction. Protein-protein and protein-ligand interactions: Interaction interfaces, binding site identification, docking overview. Molecular simulations: Basics of molecular dynamics simulations, force fields, energy minimization, simulation workflows, trajectory analysis. Structural variation and dynamics: Conformational flexibility, intrinsically disordered proteins. Case studies: Structural analysis in enzyme function, disease mutations, and structure-based drug design.

CSB6203	Structural Biology
CO.1	Apply structural bioinformatics concepts and structural databases to analyze protein structure and function.
CO.2	Analyze and evaluate protein structures using visualization tools, quality assessment methods, and structure alignment techniques.
CO.3	Develop and validate protein structure prediction and interaction models using homology modeling, docking, and related approaches.
CO.4	Apply molecular dynamics simulations to investigate protein conformational dynamics, structural variations, and structure-based drug design applications.

References

1. A. M. Lesk, *Introduction to Protein Structure*, 3rd ed. Oxford, U.K.: Oxford University Press, 2016.
2. T. E. Creighton, *Proteins: Structures and Molecular Properties*, 2nd ed. New York, NY, USA: W. H. Freeman and Company, 2019.
3. Z. Gáspári, Ed., *Structural Bioinformatics: Methods and Protocols*, Methods in Molecular Biology. New York, NY, USA: Springer, 2020.
4. N. Haspel, F. Jagodzinski, and K. Molloy, Eds., *Algorithms and Methods in Structural Bioinformatics*. Cham, Switzerland: Springer, 2022.

CSB6204 Clinical Data Management and Analytics [3-1-0-4]

Overview of clinical and biomedical data types, sources, and challenges in healthcare analytics. Clinical data standards and terminologies. Fundamentals of Database Management Systems including relational data models, ER diagrams, SQL queries, indexing, transactions, and data integrity. Clinical data warehousing, ETL pipelines, and data integration from heterogeneous sources. Concepts of data privacy, security, and regulatory frameworks such as HIPAA and GDPR. Introduction to High-Performance Computing architectures, parallel computing concepts, and their role in large-scale clinical and omics data processing. Cloud computing for clinical data management: cloud service models, deployment models, and overview of platforms such as AWS, Google Cloud, and Azure for biomedical applications. Clinical data analytics techniques including data preprocessing, quality control, normalization, exploratory data analysis, and visualization. Reproducible data analysis workflows.

CSB6204	Clinical Data Management and Analytics
CO.1	Apply clinical data standards, database design principles and ETL processes for managing healthcare datasets.
CO.2	Analyze heterogeneous clinical data using preprocessing, quality control, normalization, and exploratory data visualization techniques.
CO.3	Evaluate data privacy, security regulations, and regulatory considerations in clinical data management.
CO.4	Apply HPC and cloud computing frameworks to process large-scale clinical and omics datasets and develop reproducible analytics workflows and decision support systems.

References

1. R. Elmasri and S. B. Navathe, *Fundamentals of Database Systems*, 8th ed. New Delhi, India: Pearson Education, 2021.
2. A. Silberschatz, H. F. Korth, and S. Sudarshan, *Database System Concepts*, 7th ed. New York, NY, USA: McGraw-Hill, 2019.
3. M. G. Kahn, J. S. Brown, and C. J. Chun, *Secondary Use of Electronic Health Records*, 2nd ed. Cambridge, MA, USA: Academic Press (Elsevier), 2022.
4. T. Erl, R. Puttini, and Z. Mahmood, *Cloud Computing: Concepts, Technology & Architecture*. Upper Saddle River, NJ, USA: Prentice Hall, 2013.
5. R. Buyya, J. Broberg, and A. M. Goscinski, Eds., *Cloud Computing: Principles and Paradigms*. Hoboken, NJ, USA: Wiley, 2011.

CSB6230 Computational biology lab- II [0-0-4-2]

List of Experiments

1. Descriptive statistics and data visualization in R using biological datasets (mean, SD, boxplots, histograms, ggplot2).
2. Probability distributions and sampling techniques in R (normal, binomial, Poisson distributions).
3. Hypothesis testing and ANOVA in R (t-test, chi-square test, one-way ANOVA).
4. Correlation and regression analysis in R (Pearson, Spearman, simple and multiple linear regression).
5. Logistic regression and introduction to survival analysis using clinical datasets in R.
6. Quality control, normalization, and batch effect correction of transcriptomics datasets.
7. Differential gene expression analysis and visualization (volcano plots, heatmaps).
8. Multi-omics data integration using correlation-based and network-based approaches.

9. Construction and analysis of biological networks (PPI or gene co-expression networks) and hub gene identification.
10. Dimensionality reduction and feature selection techniques (PCA/t-SNE) for high-dimensional omics data.
11. Retrieval and visualization of protein structures from PDB using molecular visualization tools (PyMOL/Chimera).
12. Structure quality assessment and Ramachandran plot analysis.
13. Homology modeling and structure validation of a target protein.
14. Protein–ligand docking and analysis of binding interactions.
15. Design of ER diagrams and implementation of clinical database schema using SQL.
16. Execution of SQL queries, indexing, and transaction management for clinical datasets.
17. Development of ETL pipeline for integrating heterogeneous clinical datasets and data preprocessing.
18. Exploratory clinical data analytics, visualization, and reproducible reporting.

CSB6230	**Computational biology lab- II**
CO.1	Apply statistical methods and R programming techniques to analyze, visualize, and interpret biological and clinical datasets.
CO.2	Analyze high-throughput omics datasets to identify key molecular features.
CO.3	Perform structural bioinformatics analyses to investigate biomolecular interactions.
CO.4	Develop and manage clinical databases using SQL and ETL workflows, ensuring data integrity, privacy, and reproducibility.

References

1. G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning with Applications in R*, 2nd ed. New York, NY, USA: Springer, 2024.
2. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*, 3rd ed. New York, NY, USA: Springer, 2023.
3. R. Gentleman, V. Carey, W. Huber, and F. Hahne, *Bioconductor Case Studies*. New York, NY, USA: Springer, 2008.
4. P. Dalgaard, *Introductory Statistics with R*, 3rd ed. New York, NY, USA: Springer, 2022.

CSB6270 Research Practice [0-0-8-4]

Introduction to research methodology, identification of research problems in computational biology, formulation of objectives and hypotheses, literature survey using scientific databases, reference management, and research ethics. Application of statistical methods, bioinformatics approaches, and machine learning techniques for biological data analysis. Feature selection, model development, and interpretation of analytical results in biological context. Model validation using cross-validation methods, performance metrics such as accuracy and ROC analysis, comparison with published studies, biological interpretation, and reproducibility assessment. Preparation of structured research report, data visualization, plagiarism awareness, publication ethics, and oral presentation of research findings demonstrating readiness for dissertation work.

CSB6270	Research Practice
CO.1	Design and implement computational workflows for biological data analysis using appropriate algorithms, tools, and programming languages.
CO.2	Apply statistical, bioinformatics, and machine learning techniques to analyze and interpret biological datasets.
CO.3	Evaluate and validate results using suitable performance metrics, biological interpretation, and comparison with existing studies.
CO.4	Prepare a structured research report and presentation demonstrating readiness for independent dissertation work.

THIRD SEMESTER

CSB7170 Dissertation-I [0-0-0-15]

Introduction to advanced research methodology and critical analysis of recent literature in computational biology for identification of a relevant and novel research problem. Systematic literature review using scientific databases such as PubMed, Scopus, Web of Science, and Google Scholar, synthesis of research gaps, and use of reference management tools. Formulation of clear research objectives, hypotheses, and research questions based on identified gaps. Design of appropriate experimental and/or computational methodology including selection of datasets, tools, algorithms, statistical techniques, and validation strategies. Planning of workflow, feasibility analysis, timeline preparation, and consideration of ethical guidelines and reproducibility standards. Preparation of a structured research proposal including problem statement, literature review, methodology, expected outcomes, and work plan. Oral presentation and defense of the proposal demonstrating clarity, scientific rigor, and readiness for Dissertation II.

CSB7170	Dissertation-I
CO.1	Identify a relevant research problem in computational biology through critical analysis of recent literature.
CO.2	Conduct a systematic literature review and synthesize research gaps using appropriate scientific databases and tools.
CO.3	Formulate research objectives, hypotheses, and design a suitable experimental and/or computational methodology.
CO.4	Prepare and present a structured research proposal demonstrating clarity, feasibility, and scientific rigor.

FOURTH SEMESTER

CSB7270 Dissertation-II [0-0-0-15]

Implementation of the approved research methodology for systematic investigation of the identified research problem in computational biology. Acquisition and preprocessing of relevant datasets, development and execution of computational workflows, and application of advanced statistical, bioinformatics, and machine learning tools to generate meaningful results. Analysis and interpretation of results using appropriate statistical measures and biological validation techniques. Comparison with existing literature, performance evaluation, reproducibility assessment, and refinement of models or methods where necessary. Preparation of a comprehensive dissertation report including introduction, methodology, results, discussion, and conclusions following standard scientific format. Presentation and defense of research findings through viva voce, demonstrating technical competence, analytical ability, and contribution to the field.

CSB7270	Dissertation-II
CO.1	Implement the approved research methodology to analyze computational biology data effectively.
CO.2	Apply advanced computational, statistical, and bioinformatics tools to address the research problem.
CO.3	Interpret and validate results with appropriate statistical and biological justification.
CO.4	Compile a comprehensive dissertation and defend the research outcomes through viva voce.

Program Electives

CSB6141 Deep Learning for Biological Data [3-1-0-4]

Introduction to deep learning: Neural networks, backpropagation, optimization techniques, regularization, and model evaluation. Deep learning architectures: Convolutional Neural Networks, Recurrent Neural Networks, LSTM/GRU, Autoencoders, Graph Neural Networks, and Transformers. Applications in biology: Genomics sequence modeling, variant prediction, gene expression modeling, protein structure/function prediction, medical image analysis, and multi-omics data integration. Model interpretability and performance evaluation in biomedical contexts. Implementation using Python frameworks and case studies in biomedical AI.

CSB6141	Deep Learning for Biological Data
CO.1	Apply deep learning architectures to analyze biological sequence, imaging, and omics datasets.

CO.2	Develop and optimize neural network models for biomedical prediction tasks.
CO.3	Evaluate model performance and interpretability in biological applications.
CO.4	Assess real-world case studies of AI-driven biomedical research.

References

1. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
2. A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow*, 3rd ed. Sebastopol, CA, USA: O'Reilly Media, 2022.
3. K. P. Murphy, *Probabilistic Machine Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2022.
4. C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

CSB6142 Single-Cell & Spatial Transcriptomics [3-1-0-4]

Introduction to single-cell omics technologies including scRNA-seq principles, experimental workflows, library preparation methods, and major platforms. Data preprocessing steps such as quality control, filtering, normalization, detection of highly variable genes, and batch effect correction using commonly used tools. Dimensionality reduction and clustering techniques including PCA, t-SNE, UMAP, and graph-based clustering. Identification of marker genes, differential expression analysis, and strategies for cell type annotation using reference datasets. Trajectory and pseudotime analysis for studying cellular differentiation and dynamic transitions. Spatial transcriptomics technologies and integration of spatial data with scRNA-seq datasets. Cell–cell communication analysis using ligand–receptor interaction frameworks and identification of spatial expression patterns. Applications in developmental biology, tumor microenvironment analysis, immune profiling, disease heterogeneity, and precision medicine.

CSB6142	**Single-Cell & Spatial Transcriptomics**
CO.1	Apply preprocessing and clustering methods to single-cell transcriptomic datasets.
CO.2	Analyze cellular heterogeneity and lineage trajectories using dimensionality reduction techniques.
CO.3	Integrate spatial and single-cell omics data for biological interpretation.
CO.4	Evaluate applications of single-cell and spatial transcriptomics in disease research.

References

1. A. Kolodziejczyk and J. C. Marioni, Eds., *Single-Cell RNA Sequencing: Methods and Protocols*, Methods in Molecular Biology. New York, NY, USA: Springer, 2018.

2. S. L. Salzberg and M. Pop, *Next-Generation Genome Sequencing: Towards Personalized Medicine*. Hoboken, NJ, USA: Wiley, 2010.
3. T. Stuart and R. Satija, *Single-Cell Genomics: Methods and Protocols*. New York, NY, USA: Springer, 2022.
4. A. M. Lesk, *Introduction to Bioinformatics*, 5th ed. Oxford, U.K.: Oxford University Press, 2019.
5. L. Haghverdi, A. T. L. Lun, and J. C. Marioni, *Orchestrating Single-Cell Analysis with Bioconductor*. New York, NY, USA: Springer, 2023.

CSB6143 Explainable AI for Biomedical Applications [3-1-0-4]

Introduction to Explainable Artificial Intelligence (XAI): importance of interpretability, transparency, and trust in healthcare and biomedical decision-making systems. Distinction between black-box and interpretable models. Global vs. local interpretability concepts. Model-agnostic and model-specific explanation techniques. Evaluation of explanation quality, robustness, and reliability in biomedical datasets. Fairness, bias detection, and mitigation strategies in AI models. Regulatory and governance frameworks for AI in healthcare, including FDA guidance and GDPR considerations. Applications of interpretable AI in genomics, radiology and histopathology image analysis, clinical risk prediction, and electronic health record analytics. Case studies and development of reproducible XAI workflows using Python-based tools.

CSB6143	Explainable AI for Biomedical Applications
CO.1	Apply explainability techniques to interpret biomedical AI models.
CO.2	Analyze fairness, bias, and ethical considerations in healthcare AI systems.
CO.3	Evaluate model transparency and regulatory compliance requirements.
CO.4	Develop interpretable AI workflows for clinical and omics datasets.

References

1. C. Molnar, *Interpretable Machine Learning*, 2nd ed. New York, NY, USA: Lulu.com, 2022.
2. Z. C. Lipton, *The Mythos of Model Interpretability*. New York, NY, USA: Queue Books, 2018.
3. D. Gunning and D. W. Aha, Eds., *Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models*. Cham, Switzerland: Springer, 2019.
4. W. Samek, G. Montavon, A. Vedaldi, L. K. Hansen, and K.-R. Müller, Eds., *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*. Cham, Switzerland: Springer, 2019.

CSB6241 Computational Drug Design & Discovery [3-1-0-4]

Overview of the drug discovery and development pipeline. Biological databases and computational resources for target discovery. Structure-based and ligand-based drug design approaches. Virtual screening techniques. Evaluation of docking results and binding interaction analysis. ADMET prediction and drug-likeness assessment, and in silico pharmacokinetic modeling. Introduction to molecular dynamics simulations. Role of artificial intelligence and machine learning in drug discovery. Case studies in oncology, infectious diseases, and personalized therapeutics.

CSB6241	Computational Drug Design & Discovery
CO.1	Apply computational approaches for target identification and validation.
CO.2	Develop docking and QSAR models for virtual screening.
CO.3	Evaluate ADMET properties and drug-likeness parameters.
CO.4	Assess AI-based strategies in modern drug discovery.

References

1. A. R. Leach and V. J. Gillet, *An Introduction to Chemoinformatics*. Dordrecht, The Netherlands: Springer, 2007.
2. T. L. Blundell, S. Jhoti, and C. Abell, *Protein Structure and Drug Discovery*. London, U.K.: Royal Society of Chemistry, 2021.
3. D. C. Young, *Computational Drug Design: A Guide for Computational and Medicinal Chemists*. Hoboken, NJ, USA: Wiley-Interscience, 2009.
4. H. J. Böhm and G. Schneider, Eds., *Protein-Ligand Interactions: From Molecular Recognition to Drug Design*. Weinheim, Germany: Wiley-VCH, 2003.

CSB6242 Vaccine & Antibody Engineering [3-1-0-4]

Immunological foundations: Overview of innate and adaptive immune responses, antigen presentation, B-cell and T-cell activation, cytokine signaling, and immunological memory. Antibody structure and engineering: Structure of immunoglobulins, antigen-antibody interactions, affinity maturation, humanization, chimeric and fully human antibodies, bispecific antibodies, antibody-drug conjugates, and Fc engineering strategies. Monoclonal and therapeutic antibodies: Hybridoma technology, phage display, recombinant antibody production, mechanisms of action, and clinical applications in cancer, autoimmune, and infectious diseases. Epitope prediction and reverse vaccinology: B-cell and T-cell epitope mapping, MHC binding prediction, antigenicity and immunogenicity assessment, population coverage analysis, and in silico screening of vaccine candidates. Computational vaccine design and immunoinformatics: Multi-epitope vaccine construction, adjuvant selection, molecular docking of vaccine-receptor interactions, immune simulation, and structural validation. mRNA vaccines and novel delivery systems. Case studies: Computational and structural approaches in vaccines and immunotherapeutic strategies.

CSB6242	Vaccine & Antibody Engineering

CO.1	Apply immunoinformatics tools for epitope and vaccine design.
CO.2	Analyze antibody structure and engineering strategies.
CO.3	Evaluate computational and experimental vaccine development approaches.
CO.4	Assess emerging technologies in therapeutic antibody and vaccine engineering.

References

1. J. Murphy and C. Weaver, *Janeway's Immunobiology,* 9th ed. New York, NY, USA: Garland Science, 2017.
2. P. J. Delves, S. J. Martin, D. R. Burton, and I. M. Roitt, *Roitt's Essential Immunology*, 13th ed. Hoboken, NJ, USA: Wiley-Blackwell, 2017.
3. R. Rappuoli and F. Bagnoli, Eds., *Reverse Vaccinology 2.0: Human Immunology in Vaccine Design*. London, U.K.: Academic Press, 2021.
4. L. Brusic and D. Petrovsky, Eds., *Immunoinformatics: Predicting Immunogenicity In Silico*. New York, NY, USA: Springer, 2003.

CSB6243 Molecular Simulations [3-1-0-4]

Fundamentals of molecular mechanics: Classical description of biomolecular systems; potential energy functions; bonded and non-bonded interactions; overview of major force fields. Energy minimization and molecular dynamics: Minimization methods; integration algorithms; periodic boundary conditions; treatment of long-range interactions; temperature and pressure control; equilibration and production runs. Simulation workflows: System preparation; topology generation and ligand parameterization; use of MD software packages such as GROMACS, AMBER, and NAMD.Trajectory analysis and free energy calculations: Structural stability metrics; hydrogen bonding and secondary structure analysis; principal component analysis; binding free energy estimation methods. Monte Carlo simulations: Random sampling approaches and comparison with MD for conformational exploration. Applications: Protein folding, protein–ligand interactions, membrane and nucleic acid simulations, and structure-based drug design. HPC considerations: Parallel computing, GPU acceleration, scalability, and data management for large-scale simulations.

CSB6243	**Molecular Simulations**
CO.1	Apply molecular mechanics principles and force fields in biomolecular simulations.
CO.2	Develop molecular dynamics simulation workflows and perform trajectory analysis.
CO.3	Evaluate simulation results for protein folding and ligand binding studies.
CO.4	Assess applications of molecular simulations in drug discovery and biomolecular research.

References

1. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, 2nd ed. Oxford, U.K.: Oxford University Press, 2017.
2. D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed. San Diego, CA, USA: Academic Press, 2002.
3. A. R. Leach, *Molecular Modelling: Principles and Applications*, 2nd ed. Harlow, U.K.: Pearson Education, 2001.
4. T. Schlick, *Molecular Modeling and Simulation: An Interdisciplinary Guide*, 2nd ed. New York, NY, USA: Springer, 2010.