Central University of Punjab

M.Sc. Bioinformatics

Batch 2025

Department of Computational Sciences

School of Basic Sciences

Graduate Attributes

Graduates of the M.Sc. Bioinformatics program will possess a diverse set of competencies that prepare them to contribute meaningfully to science, society, and industry. The program nurtures a balanced integration of biological sciences, computational methods, and analytical thinking, enabling students to tackle real-world challenges in health, agriculture, and the environment.

Upon successful completion of the programme, graduates will be able to:

1. Possess interdisciplinary knowledge of biology, computational tools, and mathematics to solve complex biological problems.
2. Apply bioinformatics tools and techniques for data analysis, interpretation, and modelling of complex biological systems.
3. Think critically and analytically to address challenges in health, agriculture, and the environment.
4. Solve real-world problems using innovative computational approaches.
5. Demonstrate effective teamwork across multidisciplinary domains, communicate scientific ideas clearly while upholding professional ethics, engage in lifelong learning, and contribute responsibly to society through research and innovation.

SEMESTER I							
S.No.	**Course Code**	**Course Title**	**Course Type**	**L**	**T**	**P**	**Cr**
1	MBIM.401	Sequence Analysis	CC	3	0	0	3
2	MBIM.402	Python Programming	CC	2	0	2	3
3	MBIM.403	Structural Bioinformatics	CC	3	0	0	3
4	MBIM.404	Omics Cascade and Biological Data Analysis	CC	3	0	0	3
5	MBIM.405	Genome Data Analysis Lab (Practical)	SBC	0	0	4	2
6	MBIM.406	Biostatistics using R Lab (Practical)	SBC	0	0	4	2
7	MBIM.XXX	Discipline Elective	DE	3	0	0	3
8		Individualized Education Plan/ Tutorial		0	2	0	0
List of Discipline Electives (Choose anyone)							
	MBIM.407	Big Data Analytics in Healthcare	DE	3	0	0	3
	MBIM.408	Cheminformatics	DE	3	0	0	3
	MBIM.409	Linux and Perl	DE	3	0	0	3
Total				**14**	**2**	**10**	**19**

SEMESTER II							
S. No.	**Course Code**	**Course Title**	**Course Type**	**L**	**T**	**P**	**Cr**
1	MBIM.516	Biomolecular Simulations and Drug Design	CC	3	0	0	3
2	MBIM.517	Next-Generation Sequencing Data Analysis	CC	3	0	0	3
3	MBIM.518	Systems Biology	CC	3	0	0	3
4	MBIM.519	Data Mining and Machine learning	CF	2	0	0	2
5	MBIM.520	Data Mining and Machine learning Lab (Practical)	SBC	0	0	4	2
6	MBIM.521	Biological Databases and Management Systems Lab (Practical)	SBC	0	0	4	2
7	MBIM.522	Biomolecular Simulations and Drug design Lab (Practical)	SBC	0	0	4	2
8		Individualized Education Plan/ Tutorial		0	2	0	0
Interdisciplinary course offered for other departments							
9	MBIM.506	Introduction to Structural Biology	IDC	2	0	0	2
Value-Added Courses (Select One)							
10	MPCP.511	Research Methodology	VAC	2	0	0	2
11	MBIM.512	Bioinformatics for Next-Generation Sequencing Data	VAC	2	0	0	2
Total				**15**	**2**	**12**	**21**

SEMESTER III							
S.No.	**Course Code**	**Course Title**	**Course Type**	**L**	**T**	**P**	**Cr**
1	MBIM.599-1	Dissertation Part I	SBC	0	0	40	20
Total				**0**	**0**	**40**	**20**

SEMESTER IV							
S.No.	**Course Code**	**Course Title**	**Course Type**	**L**	**T**	**P**	**Cr**
1	MBIM.599-2	Dissertation Part-II	SBC	0	0	40	20
Total				**0**	**0**	**40**	**20**

Semester	**L**	**T**	**P**	**Credits**
I	14	2	10	19
II	15	2	12	21
III	0	0	40	20
IV	0	0	40	20
Total	**29**	**4**	**102**	**80**

L: Lectures; T: Tutorial; P: Practical

CC: Core Course, DE: Discipline Elective, CF: Compulsory Foundation, IDE: Inter-Disciplinary Elective, SBC: Skill-based Core, DEC: Discipline Enrichment Course, SBE: Skill- based Elective, VAC: Value Added Courses

Total Credits Distribution:

S.No	**Types of Courses**	**No. of courses (Total credits)**			
		Sem-I	Sem-II	Sem-III	Sem-IV
1	Discipline Specific Core	4 (12)	3 (9)		
2	Skill enhancement (Practical, Dissertation/Internship)	2 (4)	3 (6)	1 (20)	1 (20)
3	Discipline Elective	1 (3)			
4	Interdisciplinary/ Multidisciplinary Course		1 (2)		
5	Foundation Course (Ability Enhancement / Compulsory Foundation)		1 (2)		
6	Value based		1 (2)		

One non-credit hour (two contact hours) for tutorial will be added for remedial teaching to cater to the learning needs of all the learners. The objective of this class is to facilitate the students to understand the concepts better and absorb and assimilate the content more effectively during extra hours.

MOOC: MOOCs may be taken up to 40% of the total credits (excluding dissertation credits). MOOC may be taken in lieu of any course, but the content of that course should match a minimum of 70%. The department will do mapping, and students will be informed accordingly. Students are encouraged to take skill-based courses on SWAYAMPLUS platform.

Students will have an option to carry out dissertation work in industry, national institutes or Universities in the top 100 NIRF ranking.

Mode of Transaction: Lecture, Laboratory-based Practical, Seminar, Group discussion, Team teaching, Self-learning, Online tools.

Examination pattern and evaluation:

Formative Evaluation: Internal assessment shall be 25 marks using any two or more of the given methods: tests, open book examinations, assignments, term paper, etc. The Mid-semester test shall be a descriptive type of 25 marks, including short answer and essay type. The number of questions and distribution of marks shall be decided by the teachers.

Summative Evaluation: The End semester examination (50 marks) with descriptive type (up to 100%) and objective type (up to 30%) shall be conducted at the end of the semester. The objective type shall include one-word/sentence answers, fill-in-the-blanks, MCQs', and matching. The descriptive type shall include short answer and essay-type questions. The number of questions and distribution of marks shall be decided by the teachers. Questions for exams and tests shall be designed to assess course learning outcomes along with the focus on knowledge, understanding, application, analysis, synthesis, and evaluation.

The evaluation for IDC, VAC and entrepreneurship, innovation and skill development courses shall include MST (50 marks) and ESE (50 marks). The pattern of examination for both MST and ESE shall be the same as ESE described above for other courses.

Evaluation of dissertation proposal in the III semester shall include 50% weightage by supervisor and 50% by HoD and senior-most faculty of the department. The evaluation of the dissertation in the IV semester shall include 50% weightage for continuous evaluation by the supervisor/co-supervisor for regularity in work, mid-term evaluation, report of the dissertation, presentation, and final viva-voce; 50% weightage based on average assessment scores by an external expert. The distribution of marks is based on the report of dissertation (25), Presentation (10), Novelty/originality (5) and Final viva-voce (10). The external expert may attend the final viva-voce through offline or online mode.

Examination pattern from 2025-26 session onwards

Core, Discipline Elective, and Compulsory Foundation Courses			IDC, VAC, Entrepreneurship, Innovation and Skill Development Courses (<2 credits) or any other theory course of <2 credits	
	Marks	Evaluation	Marks	Evaluation
Internal Assessment	25	Various methods	-	-
Mid-semester test (MST)	25	Descriptive	50	Descriptive (up to 100%) Objective (up to 30%)
End-semester exam (ESE)	50	Descriptive (up to 100%) Objective (up to 30%)	50	Descriptive (up to 100%) Objective (up to 30%)

Dissertation Proposal (Third Semester)			**Dissertation (Fourth Semester)**		
	Marks	Evaluation		Marks	Evaluation
Supervisor	50	Dissertation proposal and presentation	Supervisor / Co-Supervisor	50	Continuous assessment (regularity in work, mid-term evaluation) dissertation report, presentation, final viva-voce

HoD and senior-most faculty of the department	50	Dissertation proposal and presentation	External expert	50	Report of dissertation (25), Presentation (10), Novelty /originality (5) and Final viva-voce (10).

Marks for internship shall be given by the supervisor/internal mentor and external mentor.

Some Guidelines for Internal Assessment:

1. The components/pattern of internal assessment/evaluation should be made clear to students during the semester.
2. The results of the internal assessment must be shown to the students.
3. The question papers and answers to the internal assessment should be discussed in class.
4. The internal assessment shall be transparent and student-friendly, and free from personal bias or influence.

Evaluation Criteria for Practical Courses:

Evaluation	Marks
Maintaining the lab records/notebooks	10
Continuous assessment	20
Attendance	10
Final practical examination	50
Viva-voce	10

***Multiple entry and exit*:** As per UGC guidelines, students who exit after the first year of M.Sc. Bioinformatics shall be awarded the PG Diploma in Bioinformatics, provided the candidate fulfils the following eligibility.

Eligibility: Successfully completing the first year (two semesters) courses of M.Sc. Bioinformatics degree programme and earning 4 credits from any of the following skill development / Experiential learning options.

- Completion of Skill based course(s) from MO0C (approved in CDDC/AAC)
- Mini Project in the proposed specialized area of the PG Diploma
- Industrial training or Internship in the relevant domain

SEMESTER-I

Course Title: Sequence Analysis
Course Code: MBIM.401
Course type: CC
Total Hours: 45

L	T	P	Cr
3	0	0	3

Course Learning Outcomes (CLO):
On completion of this course, students will be able to:

CLO1: Apply the algorithms for pairwise and multiple alignment.
CLO2: Choose a suitable scoring matrix for sequence alignment.
CLO3: Perform multiple sequence alignment.
CLO4: Perform searching in the sequence database and identify the motifs and profiles in the given biological sequences.

Units/ Hours	Contents	Mapping with CLO
I 13 Hours	Basic concepts of sequence similarity, identity and homology, homologues, orthologues, paralogues and xenologues Pairwise sequence alignments: basic concepts of sequence alignment, Needleman and Wunsch, Smith and Waterman algorithms for pairwise alignments, gap penalties. **Learning Activities**: Peer discussion, hands-on training in global and local alignments, Problem based learning.	CLO1
II 10 Hours	Scoring matrices: basic concept of a scoring matrix, PAM and BLOSUM series, Comparison of PAM and BLOSUM, Tools such as BLAST (various versions of it) and FASTA. **Learning Activities**: Hands-on training in PAM and BLOSUM matrix, Classroom presentation, Class quiz.	CLO2
III 12 Hours	Multiple sequence alignments (MSA): basic concepts of various approaches for MSA (e.g., progressive, hierarchical etc.). Algorithm of CLUSTALW (including interpretation of results), concept of dendrogram and its interpretation. **Learning Activities**: Research paper presentation, Peer discussion on various algorithms of MSA.	CLO3
IV 10 Hours	Sequence patterns and profiles: Basic concept and definition of sequence patterns, motifs and profiles, profile-based database searches using PSI- BLAST, analysis and interpretation of profile-based searches. **Learning Activities**: Case studies, research paper discussion, Student-generated test questions.	CLO4

Transactional Modes: Lecture, Laboratory-based Practical Seminar, Group discussion, Team teaching, Self-learning, Online tools.

Suggested Readings

1. David W. Mount (2001). Bioinformatics. Cold Spring Harbor Laboratory Press, ISBN 0-87969-608-7
2. Gayathri N. Silva (2021). Bench guide for Protein Expression and Purification. Bluerose Publishers Pvt. Ltd, ISBN-10: 935427725X.
3. Baxevanis, A. D., Bader, G. D., & Wishart, D. S. (2020). Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins 4th Edition. Wiley Publishers, ISBN-10: 1119335582
4. Pevzner P. A. (2004). Computational Molecular Biology. Prentice Hall of India Ltd, ISBN81-203-2550-8
5. Krane D.E. and Raymer M.L. (2003). Fundamental concepts of Bioinformatics. Pearson Education ISBN 81-297-0044-1
6. Gautham N. (2006). Bioinformatics. Narosa publications ISBN-13: 9781842653005
7. Labrou, N. E., and Labrou. (2021). Protein downstream processing. Springer US. ISBN: 978-1-62703-977-2.
8. Singh, V., and Dhar, P. K. (2020). Genome engineering via CRISPR-Cas9 system. Academic Press. ISBN: 9780128181409.
9. Yoon B.J., Qian,X (2021). Recent Advances in Biological Network Analysis. Springer International Publishing, ISBN 978-3-030-57172-6.

Course Title: Structural Bioinformatics
Course Code: MBIM.403
Course type: CC
Total Hours: 45

L	T	P	Cr
3	0	0	3

Course Learning Outcomes (CLO):

On completion of this course, students will be able to:

CLO1: Describe the macromolecular structures and the experimental methods to determine it
CLO2: Apply various structure analysis tools to assign protein fold, biological function and analyse the intramolecular interactions.
CLO3: Predict the protein tertiary structures from its amino acid sequence.
CLO4: Identify the membrane bound regions, stabilizing interactions and biological function using computational tools
CLO5: Analyse various forces, which stabilize macromolecular structure and biological assembles.
CLO6: Describe the role of structural genomics in explaining life process and drug discovery

Units/ Hours	Contents	Mapping with CLO
I 12 Hours	**Macromolecular Structures:** Introduction to biological macromolecules, Structural organization of proteins, Forces Stabilizing Proteins, Structure determination methods - X-ray crystallography, NMR, Cryo-electron microscopy, small angle X-ray scattering and Neutron scattering. Introduction to PDB Data-file formats, visualizing Structures, reading coordinate files. Structure validation- Ramachandran plot. **Learning Activities**: Peer discussion, visualization and analysing of protein structures, training on various structural databases and validation, Visit and demonstration of X-ray, NMR and EM facilities.	CLO1
II 10 Hours	**Structure comparison and alignment:** Protein structure comparison and alignment, sequence-structure relationship, multiple structure alignment, software/tools for structure comparison and alignment **Structure-function relationship:** Secondary structure assignment, Structural classification of proteins, Structural domains, Relationship between structure and function, assigning function from structure. **Learning Activities:** Problem based learning, Students Teams, hands-on training on structure comparison and derive structure-function relationships, Case studies, research paper discussion	CLO2

III 12 Hours	**Structure prediction:** Secondary structure prediction methods, Protein tertiary structure prediction- Comparative modeling, Threading, Ab initio modeling, introduction to artificial intelligence, Deep- learning algorithms in protein structure prediction, CASP experiments. Membrane Protein Structure Prediction, Transmembrane region prediction, difficulty in solving membrane protein structures; tools and databases for identification of membrane proteins and the prediction of their structures. **Learning Activities**: Hands-on training on structure prediction, membrane structure prediction, Peer group discussion on structural genomics findings, Student seminars on recent developments	CLO3 CLO4
IV 11 Hours	**Protein-Protein/DNA/RNA Interaction**: Evolutionary features related to structure and function, prediction of interacting regions and interaction partners. Structural analysis on protein– DNA interactions, protein–DNA binding specificity, Inter- and intramolecular interactions in protein–DNA recognition-prediction of DNA-binding sites, Introduction to RNA Structural Bioinformatics. AI/ML-based tools and algorithms for predicting interactions **Structural Genomics:** Structural annotation of genomes and structural genomics, structural genomics initiatives, impact of structural genomics on drug discovery, structural genome annotation resources. **Learning Activities**: Peer discussion, Visualizing Protein-protein/DNA interactions Exploring the outcomes of various Structural genomics initiatives, Case studies on the role of structural genomics in drug discovery.	CLO5 CLO6

Transactional Modes: Lecture, Laboratory based Practical, Seminar, Group discussion, Team teaching, Self-learning, Online tools.

Suggested Readings

1. Branden, C. and Tooze, J. (1999). Introduction to Protein Structure, Garland Publishing Inc., 2e
2. Pal, S. (2020). Fundamentals of Molecular Structural Biology. Academic Press. ISBN: 9780128148556.
3. Gu, J. and Bourne, P. E. (2009) Structural Bioinformatics, John Wiley & Sons, 2e,
4. Liljas, A and L. Liljas, J. Piskur, G. Lindblom, P. Nissen and M. Kjeldgaard. (2016) Textbook of Structural Biology. World Scientific Publishing Co.
5. Timir Tripathi, Vikash Dubey, (2022) Advances in Protein Molecular and Structural Biology Methods, Academic Press, 1e.
6. Rupp, B. (2009) Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology. Garland Science.
7. Gromiha, M.M., (2010) Protein Bioinformatics, From Sequence to Function, Academic Press, USA, 1e.
8. Zhou, Y., Kloczkowski, A., Faraggi, E., and Yang, Y. (2017). Prediction of protein secondary structure. Humana Press. ISBN: 9781493964048, 9781493964048.
9. Anderson, W. F. (2014). Structural genomics and drug discovery. Methods and Protocols, 1. Humana Press. ISBN: 9781493903535, 9781493903535.

Web resources:

NPTL http://nptel.ac.in/syllabus/syllabus.php?subjectId=104102016

MIT OpenCourseWare
http://ocw.mit.edu/courses/biological-engineering/20-442-molecular-structure-of-biological-materials-be-442-fall-2005/

Course Title: Python Programming
Course Code: MBIM.402
Course type: CC
Total Hours: 45

L	T	P	Cr
2	0	2	3

Course Learning Outcomes (CLO):
On completion of this course, students will be able to:

CLO1: Describe Python Programming to build applications in their core domain.
CLO2: Write Python scripts by using decision-making command and loops
CLO3: Apply functions, modules and external packages in Programming
CLO4: Write programs by calling inputs and specific output files

Units/ Hours	Contents	Mapping with CLO
I 8 Hours	**Introduction, Data Types and Operators:** Installation and working with Python, Variables and data types in python, perform computations and create logical statements using Python's operators: Arithmetic, Assignment, Comparison, Logical, Membership, Identity, Bitwise operators, list, tuple and string operations **Learning Activities**: Peer discussion, brainstorming and Problem based learning sessions, Class quiz	CLO1
II 7 Hours	**Python Decision making and Loops:** Write conditional statements using If statement, if ...else statement, elif statement and Boolean expressions, while loop, for loop, Nested Loop, Infinite loop, break statement, continue statement, Pass statement, Use for and while loops along with useful built-in functions to iterate over and manipulate lists, sets, and dictionaries. Plotting data, Programs using decision making and loops **Learning Activities**: Peer discussion on different code designing, Decoding quiz	CLO2
III 8 Hours	**Python Functions and Modules:** Defining custom functions, Organizing Python codes using functions, Create and reference variables using the appropriate scope, Basic skills for working with lists, tuples, work with dates and times, get started with dictionaries, importing own module as well as external modules, Programming using functions, modules and external packages **Learning Activities**: Brainstorming and Problem-Solving code designing.	CLO3
IV 7 Hours	**Python File Operations:** An introduction to file I/O, use of text files, use CSV files, use binary files, handling a single exception, handle multiple exceptions, and Illustrative programs. **Learning Activities**: Problem-Solving code practice based on CSV files.	CLO4

Lab Exercises:

1. Installation and basic demonstration of python package: PyCharm/IDLE
2. Different data types in Python: Variables, identifiers.
3. Perform basic arithmetic and logical operations in Python.
4. Apply Python's operators: Arithmetic, logical, membership, Identity, Bitwise operators.
5. Apply lists, tuples and string operations in python.
6. Apply conditional statements using If, if ...else and if-elif statements.
7. Apply looping statements in python: For loop, nested loops, continue and break statements
8. Employ and explore various built-in functions and modules in Python.
9. Writing own functions and modules in python.
10. Apply different file processing options in python.
11. Writing Python codes with AI chatbots

Transaction Mode: Lecture, tutorial, Laboratory- based Practical, problem solving

Suggested Readings

1. Gowri Shankar, S., Veena, A. (2018). Introduction to Python Programming, 1st Edition, CRC Press/Taylor & Francis.
2. VanderPlas, J. (2016). Python Data Science Handbook: Essential Tools for Working with Data, 1st Edition, O'Reilly Media.
3. Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd Edition, O'Reilly Media.
4. Chun, W. J. (2015). Core Python Applications Programming, 3rd Edition, Pearson Education India.
5. Tiago Antao (2022) Bioinformatics with Python Cookbook: Use Modern Python Libraries and Applications to Solve Real-World Computational Biology Problems, Packt Publishing
6. Ken Youens-Clark (2022) Mastering Python for Bioinformatics, O'Reilly Media

Course Title: Omics Cascade and Biological Data Analysis
Course Code: MBIM.404
Course type: CC
Total Hours: 45

L	T	P	Cr
3	0	0	3

Course Learning Outcomes (CLO):
On completion of this course, students will be able to:

CLO1: Describe the DNA-Protein Interactions and Transcription process
CLO2: Use the bioinformatics-OMIM database and identify SNPs
CLO3: Describe integrated genomic maps, gene expression profiling
CLO4: Apply probabilistic modeling techniques for building transcriptional regulatory networks.

Unit/ Hours	Contents	Mapping with CLO
I 12 Hours	**Proteomics and Metabolomics**: DNA- Protein interaction-Process of DNA-Protein Interactions during Transcription, Network identification and processing. Protein- Protein Interactions: Processing of the Proteome, Prediction of Post-translational modifications, Protein Degradation Introduction to metabolomics and its applications in health, and microbial systems. **Learning Activities**: Classroom presentation and discussion, hands-on training on protein-DNA interaction analysis, Problem based learning.	CLO1
II 10 Hours	**Genomics:** Role of bioinformatics-OMIM and ClinVar databases, integrated genomic maps, gene expression profiling, identification of SNPs, Vector Screening. **Learning Activities**: Research paper discussion, Hands-on training on biological databases, Problem based learning.	CLO2
III 12 Hours	**Transcriptomics:** Database and basic tools, Gene Expression Omnibus (GEO), ArrayExpress, understanding of RNA-seq data, normalizing RNA-seq data, Raw read quality assessment, Mapping data visualization, detecting differential gene expression. **Learning Activities**: Peer discussion, Hands-on training on Genomic databases and data visualization, Problem based learning.	CLO3
IV 11 Hours	**Metagenomics and Predictive Modelling:** Building predictive models of transcriptional regulatory networks using Machine Learning techniques. Viral Informatics, Metagenomics, repositories for comparative metagenomics analysis, host-microbe interactions. **Learning Activities**: Peer discussion on probabilistic approaches for transcriptional networks, Hands-on training on biological networks, quiz.	CLO4

Transactional Modes: Lecture, Laboratory based Practical, Seminar, Group discussion, Team teaching, Self-learning, Online tools.

Suggested Readings

1. Suhai, S (2002). Genomics and Proteomics. Springer US
2. Campbell (2007). Discovering Genomics, Proteomics and Bioinformatics. Pearson Education
3. Informatics, G., and Berger, S. (2021). Viral Meningitis: Global Status: 2021 edition. GIDEON Informatics Inc. ISBN: 1498833098, 9781498833097.
4. Lesk, A. (2019). Introduction to bioinformatics. Oxford university press. ISBN: 9780198794141.
5. Grant, P.R. (2004). Computational Genomics: Theory and Application. Horizon Bioscience
6. Gracia ,K.C, Husi H. (2019) Computational Approaches in Proteomics. Brisbane (AU): Codon Publications, PMID: 31815395.
7. Guzzi, P.H. (2016). Microarray Data Analysis: Methods and Applications. Humana Press. ISBN: 1493931725, 9781493931729.

Web Resources

1. http://virome.dbi.udel.edu

Course Title: Genome Data Analysis Lab (Practical)
Course Code: MBIM.405
Course type: SBC
Total Hours: 60

L	T	P	Cr
0	0	4	2

Course Learning Outcomes (CLO):
On completion of this course, students will be able to,

CLO1: Perform database search and retrieve genomic and proteomic data for analysis
CLO2: Utilize the tools available to analyse genome and protein sequences
CLO3: Identify different types of protein–DNA interactions and characterize different biological networks.
CLO4: Perform protein structure classification
CLO5: Perform protein structure prediction and function from the amino acid sequence

Course Content

The following experiments to be conducted:

1. Retrieve the gene sequences by exploring and querying the nucleic acid databases.
2. Find the chromosomal location of gene sequence and basic experiments in the NCBI map viewer.
3. Genome Annotation i.e. Mining Genomic Sequence Data, gene prediction methods
4. Gene expression data analysis: Differential gene expression analysis, Peak finding, Motif Discovery
5. Next Generation Sequencing File Formats
6. Next Generation Sequencing public data-repositories
7. Quality analysis of Reads – FastQC
8. Trimming and Alignment of short read with reference genome
9. NGS sequencing Demonstration

Transactional Modes: Laboratory- based Practical, Seminar, Group discussion, Team teaching, Self-learning, Online tools.

Suggested Readings

1. Mani, K. & Vijayaraj-Aparna, N. Bioinformatics: A Practical Approach- Publishers New Delhi
2. Informatics, G., and Berger, S. (2021). Viral Meningitis: Global Status: 2021 edition. GIDEON Informatics Inc.
3. Lesk, A. (2019). Introduction to bioinformatics. Oxford university press.
4. Deitel, P. J. (2002). How to program, Pearson

Course Title: Biostatistics using R Lab (Practical)
Course Code: MBIM.406
Course type: SBC
Total Hours: 60

L	T	P	Cr
0	0	4	2

Course Learning Outcomes (CLO):
On completion of this course, students will be able to,

CLO1: Use R programming to analyse the biological data.
CLO2: Perform descriptive statistical analysis on a given data set
CLO3: Perform point and interval estimation of statistical parameters
CLO4: Fit appropriate probability distributions for a provided data set

Course Content

1. How to generate statistics summary (location, spread, shape, and peakedness)
2. How to estimate data parameters: point and interval estimation- estimating the mean, proportion, variance, and variance ratio.
3. Performing One sample and two Sample t-tests on a given dataset
4. Performing F-test and ANOVA on a given dataset
5. Performing the chi–square test of independence and Fisher's test.
6. Practical understanding of Sampling distributions, Fitting of distributions (discrete and continuous)

Transactional Modes: Laboratory- based Practical, Seminar, Group discussion, Team teaching, Self-learning, Online tools.

Suggested Readings

1. Gupta S. C and Kapoor V. K.(2019). Fundamentals of Mathematical Statistics, 11/e. New Delhi: Sultan Chand and Sons publications
2. Daniel, W. W. and Cross C.L. (2018) Biostatistics: A Foundation for Analysis in the Health Sciences, 11/e. New York: John Wiley & Sons Inc.
3. Zar, J. H. (2010) Biostatistical Analysis, 5/e, Pearson Education.
4. Triola. M.M., M. F.Triola and J. Roy (2019). Biostatistics for the Biological and Health Sciences, 2/e, Pearson Education Ltd.
5. van Belle, G., L. D. Fisher, P. J. Heagerty, and T. Lumley (2004) Biostatistics: A Methodology for the Health Sciences, 2/e. John Wiley & Sons.

ONLINE MATERIAL

NPTEL
1. https://nptel.ac.in/courses/111/106/111106112/
2. https://nptel.ac.in/courses/111/102/111102111/

Course Title: Big Data Analytics in Healthcare
Course Code: MBIM.407
Course type: DE
Total Hours: 45

L	T	P	Cr
3	0	0	3

Course Learning Outcomes (CLO):
On completion of this course, students will be able to,

CLO1: Discuss the fundamental concepts of big data
CLO2: Explain the methods used for large-scale biological data handling
CLO3: Describe various approaches to generate protein big data
CLO4: Outline the role of big data analytics in genomic research
CLO5: Explain the importance of big data analytics in healthcare

Units/ Hours	**Contents**	**Mapping with CLO**
I 11 Hours	**Fundamentals in Big Data** Introduction to Big data and big data in bioinformatics. Techniques for handling big data- Data management, Data analytics. Map-Reduce Fundamentals, Technologies for handling big data- Vertical and Horizontal scaling techniques- Foundations of the Hadoop Ecosystem, Apache Spark **Learning Activities**: Group discussion, hands-on training in data handling, Peer discussion and student presentations	CLO1 & CLO2
II 12 Hours	**Big Data Analytics in Protein Bioinformatics** Protein Structure Alignment and Similarity Searching, functional assignment. Prediction of intrinsically disordered proteins using cloud computing. Mass spectrometry and NMR data analysis using inferential structure determination, low resolution crystallographic data analysis using probabilistic, Bayesian methods from SAXS and SANS data. **Learning Activities**: Hands-on training on biological data analysis, student presentations, research paper discussion	CLO3
III 11 Hours	**Big Data Analytics in Genomics** Challenges of Handling Genomic and Clinical Data, Big Data on the Cloud, Big data in NGS Read Alignment, Big Data Analytics in Calling Variants, Statistical Analysis of Genomic Data, Data mining using RNA seq data, Analysis of Genomic and Clinical Data. AI applications in clinical genomics. **Learning Activities**: Hands-on training on biological data analysis, student presentations, research paper discussion	CLO4

IV 11 Hours	**Big Data Analytics in Health Research:** Case studies on big data analytics for preventive and personalized medicine. Mining Massive Genomic Data for Therapeutic Biomarker Discovery in Cancer, Medical image processing and its role in healthcare data analysis. Big Data analytics-based models for various stages in healthcare. **Learning Activities**: Classroom presentation on real world application and peer discussion, Case studies, research paper discussion.	CLO5

Transactional Modes: Lecture, Seminar, Group discussion, Team teaching, Self-learning, Online tools.

Suggested Readings

1. Mrozek, D. (2018). Scalable Big Data Analytics for Protein Bioinformatics. Springer Press.
2. Hurwitz, J. S., Nugent, A., Halper, F., & Kaufman, M. (2013). Big data for dummies. John Wiley & Sons.
3. Elmasri, R. (2021). Fundamentals of database systems seventh edition.
4. Lytras, M. D., & Papadopoulou, P. (Eds.). (2017). Applying big data analytics in bioinformatics and medicine. IGI Global.
5. Mrozek, D. (2018). Scalable big data analytics for protein bioinformatics. Computational Biology.
6. Li, S. (Ed.). (2020). Computational methods and data analysis for metabolomics. Totowa, NJ, USA:: Humana Press.
7. Wong, K. C. (2016). Big data analytics in genomics. Springer. Springer Press, 1st ed. 2016
8. Wong, K. C. (Ed.). (2016). Big data analytics in genomics. Springer.
9. Dey, N., Das, H., Naik, B., & Behera, H. S. (2019). Big data analytics for intelligent healthcare management. Academic Press. 1st ed.

Web resources
1. https://doi.org/10.3390/ijms18020412
2. https://doi.org/10.1016/j.copbio.2019.03.004
3. https://doi.org/10.1007/s00521-019-04095-y

NPTEL
1. https://onlinecourses.nptel.ac.in/noc20_bt10/preview
2. https://onlinecourses.nptel.ac.in/noc20_cs92/preview
3. https://onlinecourses.nptel.ac.in/noc22_bt20/preview

Course Title: Cheminformatics
Course Code: MBIM.408
Course type: DE
Total Hours: 45

L	T	P	Cr
3	0	0	3

Course Learning Outcomes (CLO):
On completion of this course, students will be able to,

CLO1: Decode the molecular structure from various structural representations retrieved from the database
CLO2: Identify specific descriptors for small molecular compounds
CLO3: Develop a pharmacophore model from a set of drug molecule and quantify the structure activity relationship
CLO4: Virtual screening tools and efficiency assessments

Units/ Hours	**Contents**	**Mapping with CLO**
I 12 Hours	Cheminformatics as a theoretical chemistry discipline: definition, main concepts and areas of application. Representing chemical structures on computers. Molecular graphs. Connectivity tables. Adjacency and distance matrices. Linear representations SMILES and SMIRKS. Hashed fingerprints. Exchange formats for chemical structures (MOL, SDF) and reactions (RXN et RDF). Chemical Databases. Different types of searching structures in the databases: exact match, sub-structural, super- structural and by similarity. **Learning Activities**: Classroom presentation and discussion on the topic, hands-on training in the chemical structure file formats and databases	CLO1
II 11 Hours	Molecular descriptors. Definition and main requirements. Different types of descriptors: constitutional, topological indices, geometry-based, surface-based, substructural fragments, lipophilicity, etc. Development and validation of QSAR/QSPR models. Data preparation. Statistical parameters assessing model performance. Cross-validation. Models applicability domain. Ensemble modelling. **Learning Activities:** Problem-based learning, Improved discussion, research paper discussion	CLO2
III 11 Hours	Molecular Interaction Fields. 3D QSAR. Molecular fields' similarity Pharmacophore approach Pharmacophore features. Ligand- and structure-based pharmacophores. Merged and shared pharmacophores. Pharmacophore-based virtual screening **Learning Activities**: Peer discussion, Seminars on application of Pharmacophores, real-world application, Problem Solving.	CLO3

IV 11 Hours	Chemical Space concept. Graph-based chemical space: scaffolds, frameworks and R-groups. Scaffold tree approach. Descriptor-based chemical space: distance and similarity metrics. Data visualization: Generative Topographic Mapping. Network-like similarity graphs. Activity landscapes. Bioisosteres. Virtual screening workflow. Drug-likeness filters and structural alerts. Parameters of screening efficiency. **Learning Activities**: small group projects, Case studies, research paper and peer discussion	CLO4

Transactional Modes: Lecture, Laboratory-based Practical, Seminar, Group discussion, Team teaching, Self-learning, Online tools.

Suggested Readings

1. Leach, A. R., & Gillet, V. J. (2017). An Introduction to Cheminformatics, Springer,
2. Varnek, A. (2017). Tutorials in cheminformatics. John Wiley & Sons.
3. Engel, T., & Gasteiger, J. (2018). Chemoinformatics: basic concepts and methods. John Wiley & Sons.
4. Sharma, N., Ojha, H., Raghav, P., & Goyal, R. K. (2021). Chemoinformatics and Bioinformatics in the Pharmaceutical Sciences. Elsevier.
5. Stromgaard, K., Krogsgaard-Larsen, P., & Madsen, U. (2009). Textbook of drug design and discovery. CRC press. ISBN: 9780429111242.

Course Title: Linux and Perl
Course Code: MBIM.409
Course type: DE
Total Hours: 45

L	T	P	Cr
3	0	0	3

Course Learning Outcomes (CLO):
On completion of this course, students will be able to,
CLO1: Demonstrate Linux packages and use various compressed file formats
CLO2: Describe Perl programming to build application
CLO3: Demonstrate web-based PERL CGI programming concepts
CLO4: Develop scripts for Bioinformatics applications using PERL programming.

Units/ Hours	Contents	Mapping with CLO
I 15 Hours	**Basics of Linux shell, BASH, CSH, PERL Basics** Scalar variables, Syntax and semantics, processing scalar variables, Iteration with while construct, Variable containers, Loops, Conditional statements, Introducing Patterns, Reading and writing files, Case study: Making Motif Search tool. **Learning Activities:** Peer discussion, brainstorming and Problem based learning sessions, Class quiz	CLO1
II 11 Hours	**Advance data structure and programming in PERL** Arrays, Hashes, Sub-routines, Regular expression and Text mining: The Match Operator, Match Operator Modifiers, The Substitution Operator, Substitution Operator Modifiers, Translation, Translation Operator Modifiers, more complex regular expressions **Learning Activities:** Peer discussion on different code designing, Decoding quiz	CLO2
III 8 Hours	**PERL CGI Programming** Introduction to HTTP, HTTP Methods, generating web pages with Perl, Generating web pages with Perl, CGI. pm, HTML forms and input fields. **Learning Activities:** Brainstorming and Problem-Solving code designing.	CLO3
IV 10 Hours	**Bioperl** Introduction to Bioperl, creating a sequence and an object, manipulating sequence data with Seq methods, writing a sequence to a file, retrieving a sequence from a file, retrieving a sequence and multiple sequences from a database. Obtaining basic sequence statistics, identifying restriction enzyme sites, Identifying amino acid cleavage sites,Running applications: BLAST,Searching for genes in genomic DNA,Using EMBOSS applications with Bioperl, Aligning sequences with Smith-Waterman. **Learning Activities:** Peer discussion on different code designing, Decoding quiz	CLO4

Transactional Modes: Lecture, Laboratory-based Practical, Seminar, Group discussion, Team teaching, Self-learning.

Suggested Reading

1. Moorhouse M, Barry P (2005): Bioinformatics Biocomputing and Perl: An Introduction to Bioinformatics Computing
2. Skills and Practice, Book, John Wiley & Sons
3. Dwyer R. A. (2003): Genomic Perl: From Bioinformatics Basics to Working Code, Volume 1, Book, Cambridge University Press
4. Tisdall J (2003): Mastering Perl for Bioinformatics, Book, O'Reilly
5. Hietaniemi J, John Macdonald J, Orwant J (1999): Mastering Algorithms with Perl, Book, O'Reilly
6. William "Bo" Rothwell (2020) Advanced Perl Programming: From Advanced to Expert, Apress, 1st ed.
7. John Bach (2020) Programming Perl: The Ultimate Beginner's Guide to Learn Perl Programming Step by Step.
8. Rosenfeld L, Downey A.B, (2017) Think Perl 6: How to Think Like a Computer Scientist, O'Reilly Media

Web Resources

Bradnam K & Korf I (2012): Unix and Perl Primer for Biologists, Web tutorial at http://korflab.ucdavis.edu/Unix_and_Perl/current.html

Robert’s PERL tutorial http://www.physics.rutgers.edu/~kotliar/

SEMESTER II

Course Title: Biomolecular Simulations and Drug Design
Course Code: MBIM.516
Course type: CC
Total Hours: 45

L	T	P	Cr
3	0	0	3

Course Learning Outcomes (CLO):
On completion of this course, students will be able to,
CLO1: Describe the modelling of small to large molecular environments
CLO2: Describe the concept of molecular mechanics force fields and select an appropriate energy function/force field for a given problem
CLO3: Choose an appropriate energy minimization method for the required simulation study.
CLO4: Apply different methods for simulating large molecular systems
CLO5: Describe a pharmacophore model from a set of drug molecule and quantify the structure-activity relationship
CLO6: Perform and evaluate different virtual screening methods using large datasets

Units/ Hours	Contents	Mapping with CLO
I 11 Hours	**Molecular Modeling and Structure:** Introduction to Molecular modelling, Coordinate systems, potential energy surfaces for simple molecules. Tutorials with protein and nucleic acid structure as examples. **Force Fields and Molecular Representation:** The molecular mechanics force field and general features– Intramolecular Interactions, Non-bonded Interactions – London (van der Waals) Interactions, Electrostatic Interactions, Hydrogen Bonds, Constraints and Restraints, United Atom and Coarse-Grained Approaches, Non-pairwise Interactions, accuracy of the force fields, Water models. **Learning Activities**: Peer discussion, Demonstration using physical and computer model, Problem based learning, Analyzing Experimental data.	CLO1 & CLO2
II 12 Hours	**Energy Minimization:** Steepest Descent, Conjugate Gradient, Newton- Raphson, Comparison of Methods. Significance of Energy Minima. **Methods for Simulating Large Systems:** Non-bonded Cutoffs, Periodic Boundary Conditions, Long-range Interactions. **Learning Activities**: Hands-on training, Peer discussion, Research paper presentation	CLO3
III 11 Hours	**Simulation Methods:** Introduction to Molecular Dynamics**,** Molecular dynamics using simple models, finite difference and predictor-corrector integration methods, choosing the time step. Introduction to Metropolis Monte Carlo method. ML-assisted acceleration of MD simulations. **Learning Activities**: Hands-on training on simulation, Peer discussion, Case studies, research paper discussion.	CLO4

IV 11 Hours	**Drug design:** Introduction to drug designing, ADMET, drug metabolism, toxicity and pharmacokinetics. Lipinski rule of five, Identification and validation strategies. Drug Target classification, Concept of Pharmacophore, Structure and Ligand based drug design, QSAR, AI-accelerated molecular docking, scoring, drug discovery and development **Learning Activities**: Peer discussion, Case studies, research papers, Hands-on training on drug deigning	CLO5 & CLO6

Transactional Modes: Lecture, Laboratory based Practical, Seminar, Group discussion, Team teaching, Self-learning, Online tools.

Suggested Readings

1. Leach R. (2001). Molecular Modelling Principles and Applications 2nd Edition. Pearson.
2. Frenkel D. and Smit B. (2001). Understanding Molecular Simulation 2nd Edition, Academic Press.
3. Buschmann, H., & Holenz, J. (2018). Biomolecular Simulations in Structure- Based Drug Discovery. Germany: Wiley. ISBN:9783527342655
4. Alavi S. (2020). Molecular Simulations: Fundamentals and Practice 1st Edition, Wiley-VCH.
5. Bhatt, T. K., & Nimesh, S. (Eds.). (2021). The design and development of novel drugs and vaccines: Principles and protocols. Academic Press. ISBN: 9780128214718.
6. Renaud, J. P. (Ed.). (2020). Structural biology in drug discovery: Methods, Techniques, and Practices. John Wiley & Sons. ISBN: 9781118900406.
7. Allen, M. P., & Tildesley, D. J. (2017). Computer Simulation of Liquids 2nd Edition, Oxford University Press.
8. Stromgaard, K., Krogsgaard-Larsen, P., & Madsen, U. (2016). Textbook of Drug Design and Discovery. CRC Press, 5e, 2016.

Course Title: Next-Generation Sequencing Data Analysis
Course Code: MBIM.517
Course type: CC
Total Hours: 45

L	T	P	Cr
3	0	0	3

Course Learning Outcomes (CLO):
On completion of this course, students will be able to,

CLO1: Describe the experimental procedures involved in DNA sequencing
CLO2: Compare various NGS chemistries and NextGen Sequencing platforms
CLO3: Explain the principles and application of various sequencing methods.
CLO4: Describe the different workflows in the analysis of NGS data
CLO5: Describe complexity of NGS data, steps involved in NGS data analysis, and algorithms used for this purpose

Units/ Hours	Contents	Mapping with CLO
I 11 Hours	**Next-Generation Sequencing** NGS basic terminology: genome, transcriptome, exome, metabolome and metagenome; evolution of sequencing technologies: first, second and third generation; sequencing methods; Samples and preparation; different NGS platforms (Roche 454, AB SOLiD, Illumina, Ion Torrent, PacBio, Nanopore sequencing); challenges in the NGS field; sanger vs NGS sequencing; read length and error rates. **Learning Activities**: Peer discussion, Seminars on application of the techniques, Visit and demonstration of NGS platforms	CLO1 & CLO2
II 12 Hours	**NGS methods** Genomics: Whole genome sequencing, Target sequencing, Exome sequencing, Pooled Sequencing and De novo sequencing; Transcriptomics: Total RNA and mRNA seq, targeted RNA-seq and small noncoding RNA-seq, single-cell RNA-sequencing; Epigenomics: Methylation sequencing (Methylated DNA immunoprecipitation sequencing (MeDIP-Seq), miRNA-seq, ATAC-Seq, ChIP-Seq. **Learning Activities**: Discussion and seminars on application of the techniques, Case studies, research paper discussion	CLO3
III 11 Hours	**NGS data Analysis workflow and file formats** Primary, secondary and Tertiary analysis; Sequence quality evaluation - Phred Scores; Quality Control and Preprocessing; post alignment quality and bias control; Types of NGS data; File formats - FASTQ, SAM, BAM, VCF; Online NGS databases. **Learning Activities**: Hands-on training in sequence analysis, Seminars on application of the techniques, Case studies, research paper discussion	CLO4
IV 11 Hours	**NGS Data Analysis** RNA-Seq data analysis-expression estimation, Differential Expression Analysis Workflow; Aligning Reads to Reference; Transcriptome Assembly; Alignment programs; Spliced alignment Quantitation and Annotation-Based Quality	CLO5

	Control, data visualization, Downstream analysis such as pathway analysis, clustering and Gene ontology. **Learning Activities**: Peer discussion, hands-on training in sequence analysis, Seminars on application of the techniques, Case studies, research paper discussion	

Transactional Modes: Lecture, Laboratory-based Practical, Seminar, Group discussion, Team teaching, Self-learning, Online tools.

Suggested Readings:

1. Brown, S. M. (2013). Next-Generation DNA Sequencing Informatics. Cold Spring Harbor Laboratory Press, 1/e.
2. Xinkun Wang. (2016). Next-Generation Sequencing Data Analysis. Talyor and Francis Group.
3. Low, L., & Tammi, M. (2017). Introduction to next-generation sequencing technologies. Bioinformatics: A Practical Handbook Of Next Generation Sequencing And Its Applications.
4. Korpelainen et al (2015). RNA-seq Data Analysis – A Practical Approach, CRC Press, Taylor & Francis Group, 1e.
5. Jianping Xu. (2014). Next-generation sequencing: Current Technologies and Applications. Caister Academic Press, 1/e.

Course Title: Systems Biology
Course Code: MBIM.518
Course type: CC
Total Hours: 45

L	T	P	Cr
3	0	0	3

Course Learning Outcomes (CLO):
On completion of this course, students will be able to,

CLO1: Relate the various biological pathways with consensus network motifs and therefore decipher the functioning
CLO2: Correlate the various biological pathways with consensus network motifs and therefore decipher the functioning
CLO3: Identify the optimal structure for analyzing deep sequencing data and understand the main features of biological networks.
CLO4: Apply mathematical modelling to discuss relevant issues in Biology

Units/ Hours	Contents	Mapping with CLO
I 11 Hours	Introduction to system biology, networks, path, degree, cluster coefficient, scale-free networks and relevance, power law. Transcription networks, basic concepts, elements of Transcription networks, logic input functions. **Learning Activities**: Case study, research paper discussion, Quiz.	CLO1
II 12 Hours	Lac operon, Auto-regulation, a network motif: patterns, randomized networks and network motif, the feed forward loop (FFL) network motif, Case study in lactose system and arabinose system of E. coli, Dynamics of CI-FFL, IIFFL with logical gates, Case study of galactose utilization in E. coli, flagella system of *E.coli*, Convergent evolution of FFLs. **Learning Activities**: Problem Solving on various network study, case study, seminars on related topic	CLO2
III 12 Hours	Temporal programs and the global structure of transcription networks: SIM, BiFans, DORs, Interlocked FFL in B. subtilis sporulation network, Network motifs in developmental transcription networks: Network motif in PPI: Hybrid network motifs, hybrid FFL, Network motifs in neuronal networks: C elegans multi-input FFLs, Network motif in signal transduction networks. **Learning Activities**: Peer discussion, brainstorming discussion on related research papers.	CLO3
IV 10 Hours	Kinetic proofreading and conformational proofreading: genetic code, Precision of Translation, tRNA study, recognition of self and non-self by the immune system; Demand rule for gene regulation, optimal gene circuit design: cost of LacZ protein, fitness function and optimal expression level. **Learning Activities**: Problem Solving, Peer discussion and brainstorming quiz	CLO4

Transactional Modes: Lecture, Laboratory-based Practical, Seminar, Group discussion, Team teaching, Self-learning, Online tools.

Suggested Readings

1. Ringrose, L. (2017). Epigenetics and Systems Biology, Elsevier Science.
2. Markus W. Covert (2017). Fundamentals of Systems Biology: From Synthetic Circuits to Whole-cell Models, CRC Press.
3. Chen, B. S., & Li, C. W. (2016). Big Mechanisms in Systems Biology: Big Data Mining, Network Modeling, and Genome-Wide Data Identification, Elsevier Science.
4. Klipp, E., Liebermeister, W., Wierling, C., & Kowald, A. (2016). Systems Biology, A Textbook, Wiley.
5. Alon, U. (2020). An Introduction to Systems Biology: Design Principles of Biological Circuits. Chapman & Hall, Second Edition.
6. Hake, S. and Wilt, F. (2003). Principles of Developmental Biology. W.W. Norton and Company, New York, USA.

Web resources:

1. The lac operon: https://ocw.mit.edu/courses/biology/7-01sc-fundamentals-of-biology-fall-2011/molecular-biology/gene-regulation-and-the-lac-operon/#?w=535

NPTEL

https://nptel.ac.in/courses/102106035
https://onlinecourses.nptel.ac.in/noc20_bt08/preview

MIT Open Courseware

1. https://ocw.mit.edu/courses/biology/7-91j-foundations-of-computational-and-systems-biology-spring-2014/

Course Title: Data Mining and Machine Learning Course
Code: MBIM.519
Course type: CF
Total Hours: 30

L	T	P	Cr
2	0	0	2

Course Learning Outcomes (CLO):
On completion of this course, students will be able to,

CLO1: Perform data cleaning, cross-validation, and application of regression analysis
CLO2: Use various methods of clustering and dimensionality reduction for data analysis
CLO3: Explain classification as a tool to develop predictive platform
CLO4: Apply SVM and Neural network methods data analysis.

Units/ Hours	**Contents**	**Mapping with CLO**
I 7 Hours	**Introduction:** Overview of Machine Learning field, Artificial intelligence Vs Machine learning, Regression and classification problem, cross validation, Terminology in ML: true positive, false positive, Specificity, Sensitivity, Accuracy, Recall, Precision, AUC-ROC curves, Confusion matrix, Errors in ML: Bias and variance. **Learning Activities**: Demonstration using random dataset, Problem based learning,	CLO1
II 8 Hours	**Unsupervised Methods:** Clustering: Distance Metrics, K-Means, leader, Jarvis-Patrick, hierarchical clustering; Dimensionality Reduction: Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Partial Least Squares – Discriminant Analysis (PLS-DA). **Learning Activities:** Peer discussion, Problem based learning, Analysing Experimental data for unsupervised model training.	CLO2
III 8 Hours	**Supervised Methods:** Univariate and multivariate linear regression: Model representation and cost function; Classification: k-nearest neighbors' algorithm (K-NN), naïve Bayes, decision trees, boosting and bagging **Learning Activities:** Case study, Peer discussion, brainstorming and Problem Solving.	CLO3
IV 7 Hours	**Other ML algorithms:** Ensemble methods, random Forests; Support vector machines, Neural networks, Recommendation systems; Outlier detection. Bonferroni Correction. Reinforcement learning and introduction to deep learning. **Learning Activities:** Problem Solving, case study, Peer discussion on various classification methods, brainstorming discussion	CLO4

Transactional Modes: Lecture, Laboratory-based Practical, Seminar, Group discussion, Team teaching, Self-learning, Online tools.

Suggested Readings

1. Helder I., N. (2021). Bioinformatics. Exon Publications.
2. Alonso-Betanzos, A., & Bolón-Canedo, V. (2018). Big-Data Analysis, Cluster Analysis, and Machine-Learning Approaches. Advances in experimental medicine and biology, 1065, 607–626.
3. Applied Predictive Modeling by Max Kuhn and Kjell Johnson; 2013.
4. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2014). An introduction to statistical learning: With applications in R.
5. McKinney, W. (2013). Python for data analysis
6. Han, J., Kamber, M., & Pei, J. (2011). Data Mining: Concepts and Techniques, Third Edition.

Course Title: Data Mining and Machine Learning Lab (Practical)
Course Code: MBIM.520
Course type: SBC
Total Hours: 60

L	T	P	Cr
0	0	4	2

Course Learning Outcomes (CLO):

On completion of this course, students will be able to,

CLO1: Perform dimensionality reduction to solve biological problems
CLO2: Build and evaluate regression models
CLO3: Handle large data by applying PCA and machine learning methods
CLO4: Apply classification as a tool to develop a predictive platform
CLO5: Implement k-means and Hierarchical clustering procedure for the given biological data

Course Content

- i) Applying dimensionality reduction analysis on a given dataset: PCA, LDA using any programming language/tool
- ii) How to generate Heat-Maps for RNA-Seq dataset
- iii) Working with Hierarchical clustering on RNA-Seq data
- iv) Performing linear regression analysis on training and estimating accuracy on testing set using any programming language/tool
- v) Using k-Nearest Neighbors algorithm (K-NN) to classify the given dataset
- vi) Using Random Forest to classify the given dataset
- vii) Using Support Vector Machine (SVM) to make predictions for unknown datasets.
- viii) Applying the concepts of Decision trees learning algorithm to predict the unknown target variable

Transactional Modes: Laboratory based Practical, Seminar, Group discussion, Team teaching, Self-learning, Online tools.

Suggested Readings

1. Leskovec, J., Rajaraman, A. & Ullman, J. (2014). Mining of Massive Datasets.
2. Doupe, P., Faghmous, J., & Basu, S. (2019). Machine Learning for Health Services Researchers. Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research, 22(7), 808–815.
3. Lan, K., Wang, D. T., Fong, S., Liu, L. S., Wong, K., & Dey, N. (2018). A Survey of Data Mining and Deep Learning in Bioinformatics. Journal of medical systems, 42(8), 139.
4. Bishop, C. (2007). Pattern Recognition and Machine Learning.

Course Title: Biological Database and Management System Lab (Practical)
Course Code: MBIM.521
Course type: SBC
Total Hours: 60

L	T	P	Cr
0	0	4	2

Course Learning Outcomes (CLO):

On completion of this course, students will be able to,

CLO1: Apply DBMS principles to solve problems in biological sciences.
CLO2: Design an approach to create a Relational DBMS
CLO3: Define and enforce integrity constraints on a database using a state-of-the-art RDBMS
CLO4: Create and query a database DBMS concepts
CLO5: Create non-redundant databases

Course Content

1. Creating and working with databases using MySQL, command to know the available database, Creating and Deleting a Database, Setting the Default Database.
2. Working with tables using MySQL: Creating and Deleting a Table - CREATE TABLE and DROP TABLE, adding content into the created table.
3. Working with MySQL codes to expand and edit the table: Inserting Rows, using insert, select, update and delete commands.
4. Codes for Querying the Database, making use of Comparison Operators, operators for String Pattern Matching, Arithmetic Operators, Logical Operators, querying by using different clause
5. How to produce a summary report using MySQL keywords, GROUP BY clause, GROUP BY Aggregate Functions and having clause
6. Commands to modify and delete data using MySQL: UPDATE-SET command, DELETE FROM command
7. MySQL command to work with Nested Queries & Join Queries, Views, Functions [skip first/last] and Reports.
8. Command to Work with multiple tables, setting keys (primary, foreign etc.)

Transactional Modes: Laboratory-based Practical, Seminar, Group discussion, Team teaching, Self-learning, Online tools.

Suggested Readings

1. Negi, M.C. (2019). Fundamentals of Database Management System: Learn Essential Concepts of Database Systems. BPB PUBN publications
2. Chopra, R. (2016). Database Management System (DBMS): A Practical Approach, 5th Edition. CHAND & Company Limited publications.
3. Ayyavaraiah, M., & Gopi, A. (2017). Database Management System. Horizon Books (A Division of Ignited Minds Edutech P Ltd) publications
4. Silberschatz, A., Korth, H. F., & Sudarshan, S (2005) Database system concepts. 5 Ed McGraw Hill Publications.
5. Elmasri, R., & Navathe, S. B. (2007). Fundamentals of Database systems. Benjamin Cummings Publishing Company.
6. Rao, P. R. (2003). Database Management system, 3rd Ed McGraw Hill Publications. 9780071230575

Course Title: Biomolecular Simulations and Drug Design Lab (Practical)
Paper Code: MBIM.522
Course type: SBC
Total Hours: 60

L	T	P	Cr
0	0	4	2

Course Learning Outcomes (CLO):
On completion of this course, students will be able to,

CLO1: Use the Linux environment for scientific computing
CLO2: Use remote computing as a tool for high-performance computation
CLO3: Use different energy minimisation techniques
CLO4: Create macromolecular models from the amino acid sequence and high-definition images using various graphics tools
CLO5: Use of different *in-silco* techniques for biomolecular simulations CLO6: Perform molecular docking and high-throughput virtual screening.

Course Content

1. Linux basics and remote computing
2. Coordinate generations and inter-conversions of small molecules
3. Modelling macromolecular structure using homology modelling and *ab-initio* method.
4. Energy minimizations and optimization, *ab initio methods*
5. Advanced Visualization Software and 3D representations
6. Molecular Dynamics with GROMACS (a) Water structure and dynamics, (b) Binary Mixtures (c) HP36 in Water (d) Serotonin1A in Membrane Bilayers
7. Review of Molecular Dynamics Principles
8. Molecular Recognition (a) Prediction of Protein-ligand interaction sites and (b)Prediction of Protein-protein interaction sites
9. Protein Ligand Docking using AutoDock
10. High throughput virtual screening using High throughput virtual screening using AutoDock Tools (ADT).

Transactional Modes: Laboratory based Practical, Seminar, Group discussion, Team teaching, Self-learning, Online tools.

Suggested Readings

1. Allen, M. P., & Tildesley, D. J. (2017). Computer Simulation of Liquids 2nd Edition, Oxford University Press.
2. Frenkel, D., & Smit, B. (2001). Understanding Molecular Simulation 2nd Edition, Academic Press.
3. Leach, A. R., & Leach, A. R. (2001)., Molecular Modelling Principles and Applications 2nd Edition. Pearson.
4. Alavi, S. (2020). Molecular Simulations: Fundamentals and Practice 1st Edition, Wiley-VCH.

Course Title: Introduction to Structural Biology
Course Code: MBIM.506
Course type: IDC
Total Hours: 30

L	T	P	Cr
2	0	0	2

Course Learning Outcomes (CLO):

On completion of this course, students will be able to,

CLO1: Describe the physical and structural properties of proteins and the forces that stabilize protein structure
CLO2: Describe the principles and the practical aspects of macromolecular crystallography
CLO3: Describe the principles of Macromolecular NMR spectroscopy and Cryo EM and their application to study the structures of macro molecules
CLO4: Describe the applications bioinformatics tools in macromolecular structure determination.

Units/ Hours	Contents	Mapping with CLO
I 8 Hours	**Introduction to Protein structure:** Protein structure: Amino acids and their properties, peptide bond, primary, secondary and tertiary structure - structural properties required for complex formation, Ramachandran Plot, Stabilizing forces, evolutionary conservation of structure-function in proteins. **Learning Activities**: Demonstration using physical and computer models, Student seminars, Peer group discussion	CLO1
II 8 Hours	**MACROMOLECULAR CRYSTALLOGRAPHY** Protein crystallization, Introduction to X-ray sources and detectors, X-ray diffraction - Bragg's law, structure factor and electron density map, phase problem, Molecular Replacement, experimental phasing techniques, structure validation. **Learning Activities**: Peer discussion, Visit and demonstration of X-ray diffractometer, demonstration using protein models, Student seminars	CLO2
III 7 Hours	**NMR SPECTROSCOPY** Principles of NMR spectroscopy, Chemical shift, Relaxation parameters, spin-spin coupling; multidimensional NMR; NMR Structure determination. **CRYO ELECTRON MICROSCOPY** Basic scanning and transmission microscopy, Introduction to cryo EM, determining the structure using cryo EM. Learning Activities: Peer discussion, Visit and demonstration of NMR and EM facilities, Classroom presentation, Peer group discussion	CLO3

IV 7 Hours	**STRUCTURAL BIOINFORMATICS** Homology modelling, Fold recognition methods, *ab initio* methods, Prediction of membrane helices and accessibility. Introduction to structural genomics **Learning Activities**: Hands-on training on the bioinformatics tools, Peer discussion, research paper discussion	CLO4

Suggested Reading

1. Kessel, Amit and Nir Ben Tal, *Introduction to Proteins: Structure, Function and Motion,* CRC Press, 2e, 2018.
2. Pal, Subrata, *Fundamentals of Molecular Structural Biology*, Academic Press, 2020.
3. Anders Liljas, Lars Liljas , Jure Piskur, Göran Lindblom, Poul Nissen, Morten Kjeldgaard, *Textbook of Structural Biology*, World Scientific Publishing Co Pte Ltd, 2/e, 2017
4. Timir Tripathi and Vikash Kumar Dubey, *Advances in Protein Molecular and Structural Biology Methods*, Academic Press, 2022

Web Resources

1. https://ocw.mit.edu/courses/biology/7-91j-foundations-of-computational- and-systems-biology-spring-2014/video-lectures/leture-12-introduction-to-protein-structure-structure-comparison-and-classification/

2. https://nptel.ac.in/courses/102107086

Course Title: Bioinformatics for Next Generation Sequencing Data
Course Code: MBIM.512
Course type: VAC
Total Hours: 30

L	T	P	Cr
2	0	0	2

Course Learning Outcomes (CLO):
On completion of this course, students will be able to,

CLO1: Compare various NGS chemistries and NextGen Sequencing platforms CLO2: Outline the complexity of NGS data, steps involved in NGS data analysis, and algorithms
CLO3: Describe data analysis approaches used for DNA and RNA sequencing using NGS platforms
CLO4: Examine the real-time application of NGS in medicine and agriculture using case studies

Units/ Hours	Contents	Mapping with CLO
I 8 Hours	Introduction to NGS techniques: Illumina sequencing, Roche 454 sequencing, Ion torrent, SOLiD sequencing, Nanopore sequencing. **Learning Activities**: Visit and demonstration of NGS platforms, Seminars on application of the techniques,	CLO1
II 8 Hours	Sequence quality evaluation - Phred Scores, Sequence file formats, Alignment of Next-Gen sequences to reference sequences **Learning Activities**: Hands-on training in sequence analysis, discussion and seminars on application of the techniques, Case studies	CLO2
III 7 Hours	NGS methods: Whole genome sequencing, Target sequencing, Exome sequencing, RNA seq, analysis-Transcriptome mapping, Methylation sequencing, ChIP- Seq. **Learning Activities**: Peer discussion, hands-on training in sequence analysis, research paper discussion	CLO3
IV 7 Hours	RNA-Seq data analysis, Differential Expression Analysis Workflow, Aligning Reads to Reference, Transcriptome Assembly, Applications of NGS **Learning Activities**: Peer discussion, hands-on training in sequence analysis, research paper discussion	CLO4

Transactional Modes: Lecture; Tutorial; Problem solving; Self-learning.

Suggested Readings

1. Brown, S. M. (2013). Next-Generation DNA Sequencing Informatics. Cold Spring Harbor Laboratory Press, 1/e.
2. Head, Steven R., Ordoukhanian, Phillip, Salomon, Daniel R (2018) Next Generation Sequencing Methods and Protocols, Humana Press Inc.
3. Xinkun Wang. (2016). Next-Generation Sequencing Data Analysis. Talyor and Francis Group.
4. Low, L., & Tammi, M. (2017). Introduction to next generation sequencing technologies. Bioinformatics: A Practical Handbook Of Next Generation Sequencing And Its Applications.
5. Lloyd Low and Martti Tammi (2017) A Practical Handbook of Next Generation Sequencing and Its Applications, World Scientific Publishing Co Pte Ltd
6. Beginner's Handbook of Next Generation Sequencing, Genohub
7. Shawn E. Levy and Richard M. Myers, "Advancements in Next-Generation Sequencing". Annu. Rev. Genom. Hum. Genet. 2016. 17:16.1–16.21

Course Title: Research Methodology
Course Code: MPCP.511
Course type: VAC
Total Hours: 30

L	T	P	Cr
2	0	0	2

Course Learning Outcomes (CLO):

On completion of this course, students will be able to:

CLO1: Perform a Literature survey, critically analyse the scientific problem and develop a research plan

CLO2: Use reference management systems and perform literature reviews using online resources

CLO3: Describe the importance of IPR and develops interest in entrepreneurship

CLO4: Write a good technical report, manuscripts, and scientific proposals

CLO5: Appreciate the importance of Research and Academic Integrity and follow safety protocols

Units/ Hours	Contents	Mapping with CLO
I 8 Hours	Introduction: Meaning and importance of research, Different types and styles of research, Role of serendipity, Critical thinking, Creativity and innovation, Hypothesis formulation and development of research plan, Art of reading and understanding scientific papers, Literature survey, Interpretation of results and discussion. **Learning Activities**: Research paper presentation, Writing and Evaluation of research proposals, Peer discussion.	CLO1
II 7 Hours	Library: Classification systems, e-Library, Reference management, Web-based literature search engines, Intellectual property rights (IPRs). Entrepreneurship and Business Development: Importance of entrepreneurship and its relevance in career growth, Types of enterprises and ownership. **Learning Activities**: Concept built with real examples, case studies, Student presentation Perform literature survey, Research paper presentation and group discussion.	CLO2, CLO3
III 8 Hours	Scientific and Technical Writing: Role and importance of communication, Effective oral and written communication, Scientific writing, Research paper writing, technical report writing, Making R and D proposals, Thesis writing, Oral and poster presentation, Seminars, Group discussions, Use of modern aids. **Learning Activities**: Project report /research article preparation as a group activity, Research paper presentation	CLO4
IV 7 Hours	Research and Academic Integrity: Plagiarism, Copyright issues, Ethics in research, and case studies. Laboratory Safety Issues: Lab, Workshop, Electrical, Health and fire safety, Safe disposal of hazardous materials. **Learning Activities:** Case studies, Peer discussion, brainstorming, spontaneous quizzes	CLO5

Transaction Mode: Lecture, demonstration, PPT.

Suggested Readings

1. Kumar, R. (2012). Research Methodology, SAGE Publications India Pvt. Ltd., New Delhi, India.
2. Gupta, S. (2005). Research Methodology and Statistical techniques, Deep and Deep Publications (P) Ltd. New Delhi, India.
3. Kothari, C.R. (2008). Research Methodology, New Age International, New Delhi, India.
4. Standard /Reputed Journal authors' instructions.
5. Denisova-Schmidt, E. (2021). Book Review: A Roadmap to the Future of Academic Integrity Research. Academy of Management Learning & Education.
6. Sutherland-Smith, W. (2008). Plagiarism, the Internet, and student learning: Improving academic integrity. Routledge.
7. Bretag, T. (Ed.). (2020). A research agenda for academic integrity. Edward Elgar Publishing.
8. Gould, J. R. (2020). Directions in Technical Writing and Communication. Routledge.

Semester III

Course Title: Dissertation Part-I
Course Code: MBIM.599-1
Course type: SBC
Total Hours: 600

L	T	P	Cr
0	0	40	20

Course Learning Outcomes (CLO):
On completion of this course, students will be able to,

CLO1: Demonstrate an in-depth knowledge of scientific research about the area of study
CLO2: Demonstrate experimental/theoretical research capabilities based on rigorous hands-on training
CLO3: Critically analyse, interpret, and present the data considering existing scientific knowledge to arrive at specific conclusions
CLO4: Develop higher order thinking skills required for pursuing higher studies (Ph.D.)/research-oriented career options

Students will carry out their research work under the supervision of a faculty member. Students will interact with the supervisors through meetings and presentations on a regular basis. After completion of the research work, students will complete the dissertation under the guidance of the supervisor. The dissertation will include literature review, hypothesis, objectives, methodology, results, discussion, and bibliography.

Evaluation Criteria: The evaluation of dissertation in the fourth semester will be as follows: 50% weightage for continuous evaluation by the supervisor which includes regularity in work, mid-term evaluation, report of dissertation, presentation, and final viva-voce; 50% weightage based on average assessment scores by an external expert, HoD and senior-most faculty of the department. Distribution of marks will be based on report of dissertation (25), Presentation (10%), Novelty /originality (5) and Final viva-voce (10). The final viva-voce will be through offline or online mode.

Dissertation (Third Semester)		
	Marks	Evaluation
Supervisor	50	Continuous assessment (regularity in work, mid-term evaluation) dissertation report, presentation, final viva-voce
HoD and senior-most faculty of the department	50	Report of dissertation (25), Presentation (10%), Novelty /originality (5) and Final viva-voce (10).

Transactional Modes: Lecture, Laboratory based Practical, Seminar, Group discussion, Team teaching, Self-learning, Online tools

Semester IV

Course Title: Dissertation Part-II
Course Code: MBIM.599-2
Course type: SBC
Total Hours: 600

L	T	P	Cr
0	0	40	20

Course Learning Outcomes (CLO):
On completion of this course, students will be able to,

CLO1: Demonstrate an in-depth knowledge of scientific research about the area of study
CLO2: Demonstrate experimental/theoretical research capabilities based on rigorous hands-on training
CLO3: Critically analyse, interpret, and present the data considering existing scientific knowledge to arrive at specific conclusions
CLO4: Develop higher order thinking skills required for pursuing higher studies (Ph.D.)/research-oriented career options

Students will carry out their research work under the supervision of a faculty member. Students will interact with the supervisors through meetings and presentations on a regular basis. After completion of the research work, students will complete the dissertation under the guidance of the supervisor. The dissertation will include literature review, hypothesis, objectives, methodology, results, discussion, and bibliography.

Evaluation Criteria: The evaluation of dissertation in the fourth semester will be as follows: 50% weightage for continuous evaluation by the supervisor which includes regularity in work, mid-term evaluation, report of dissertation, presentation, and final viva-voce; 50% weightage based on average assessment scores by an external expert, HoD and senior-most faculty of the department. Distribution of marks will be based on report of dissertation (25), Presentation (10%), Novelty /originality (5) and Final viva-voce (10). The final viva-voce will be through offline or online mode.

Dissertation (Fourth Semester)		
	Marks	Evaluation
Supervisor	50	Continuous assessment (regularity in work, mid-term evaluation) dissertation report, presentation, final viva-voce
External expert	50	Report of dissertation (25), Presentation (10%), Novelty /originality (5) and Final viva-voce (10).

Transactional Modes: Lecture, Laboratory based Practical, Seminar, Group discussion, Team teaching, Self-learning, Online tools