# Syllabus for M.Sc. (Bioinformatics)

The M. Sc (Bioinformatics) Course is of four semesters

## Semester I

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIF 601</td>
<td>Basic Mathematics</td>
<td>3</td>
</tr>
<tr>
<td>BIF 611</td>
<td>Computer Fundamentals &amp; Biostatistics</td>
<td>3</td>
</tr>
<tr>
<td>BIF 621</td>
<td>Biomolecules</td>
<td>3</td>
</tr>
<tr>
<td>BIF 631</td>
<td>Cell Biology &amp; Genetics</td>
<td>3</td>
</tr>
<tr>
<td>BIF 641</td>
<td>Practicals</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Comprehensive Viva Voce</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td><strong>Total Credits</strong></td>
<td><strong>28</strong></td>
</tr>
</tbody>
</table>

## Semester II

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIF 602</td>
<td>Bioinformatics-I</td>
<td>3</td>
</tr>
<tr>
<td>BIF 612</td>
<td>Object Oriented Programming through 'C++'</td>
<td>3</td>
</tr>
<tr>
<td>BIF 622</td>
<td>Molecular Biology</td>
<td>3</td>
</tr>
<tr>
<td>BIF 632</td>
<td>Design &amp; Analysis of Algorithm</td>
<td>3</td>
</tr>
<tr>
<td>BIF 642</td>
<td>Internet &amp; Web Based Programming(CG1, Perl &amp; HTML)</td>
<td>3</td>
</tr>
<tr>
<td>BIF 652</td>
<td>Practicals</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Comprehensive Viva Voce</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td><strong>Total Credits</strong></td>
<td><strong>29</strong></td>
</tr>
</tbody>
</table>

## Semester III

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIF 603</td>
<td>Bioinformatics-II</td>
<td>3</td>
</tr>
<tr>
<td>BIF 613</td>
<td>Bioinformatics-III</td>
<td>3</td>
</tr>
<tr>
<td>BIF 623</td>
<td>Data Base Management System</td>
<td>3</td>
</tr>
<tr>
<td>BIF 633</td>
<td>Recombinant DNA Technology</td>
<td>3</td>
</tr>
<tr>
<td>BIF 643</td>
<td>Java Programming</td>
<td>3</td>
</tr>
<tr>
<td>BIF 653</td>
<td>Assignment</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Comprehensive Viva Voce</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td><strong>Total Credits</strong></td>
<td><strong>29</strong></td>
</tr>
</tbody>
</table>

## Semester IV

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIF 604</td>
<td>Project Work</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Comprehensive Viva-Voce</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td><strong>Total Credits</strong></td>
<td><strong>29</strong></td>
</tr>
</tbody>
</table>

**Total Credits for All Semesters in 2 Years**

115
M.Sc. Bioinformatics I Semester

BASIC MATHEMATICS

1. The set theory properties of subsets: Union, Intersection, De Morgan’s Law, Venn diagram, Complement, Set Difference, Linear and geometric functions

2. Limits of functions, derivatives of functions

3. The binomial theorem: Binomial Expansion, Finding middle term, general term. Probability calculations: Events and types of events, Probability of events, addition theorem on probability, conditional probability, Multiplication theorem, independent events

4. Differential calculus

4. Integration

COMPUTER FUNDAMENTALS AND BIOSTATISTICS

- Overview of computers: generations of computers, Classification of computer, Hardware, software, Operating system (Windows and Unix).
- Number systems, Computer Arithmetic, Logic Gates.
- Introduction to Programming: Development of Algorithms, flow charts and Decision table, low-level and high-level programming languages, 'C' language,
- Introduction to data structures and database concepts.
- Introduction to Internet and its applications.
- Introduction to MS-OFFICE: MS Word, MS Power point, MS Excel
- Computer-Oriented Statistical Techniques: SPSS, Mat lab.
- Sampling - Sampling procedure, homogenization of samples, samples size, Selection of random sample, Limitation of analytical methods, classification of errors, measurement of averages and variation, minimization of errors.
- Types of data, Frequency distribution, Frequency table of single discrete variable, Bubble sort.
- Graphical representation of Data, Histogram, Frequency polygon, Pie Chart.
- Measure of central values - Mean, median and mode, Measures of dispersion - range, mean deviation, standard deviation, coefficient of variation, moment, Skewness and kurtosis.
- Probability, Concept of Probability Theory, Events, Trials, Mutually exclusive events, favorable events, exhaustive events, Bayesian theorem of Probability, Addition theorem, Multiplication theorem
- Binomial distribution, Normal distribution, Poisson distribution & their applications.
- Design of experiments, ANOVA (one-way and two-way), F-test.
- Simple regression and correlation

BIOMOLECULES

3. Classes of organic compounds and functional groups - atomic and molecular dimensions, space filling and ball and stick models.
4. Amino acids and peptides - classification, chemical reactions and physical properties.
5. Sugars - classification and reactions.
6. Heterocyclic compounds and secondary metabolises in living systems - nucleotides, pigments, isoprenoids.
7. Separation techniques for different biomolecules.
8. Lipids - classification, structure and functions.
9. Proteins - classification and separation, purification and criteria of homogeneity, end group analysis, hierarchy in structure, Ramachandran map.
11. Analytical techniques in biochemistry and biophysics for small molecules and macro- molecules for quantization.

**CELL BIOLOGY AND GENETICS**

1. Diversity of cell size and shape.
2. Cell theory.
4. Cellular organelles- Plasma membrane, cell wall, their structural organization; Mitochondria, Chloroplast; Nucleus and other organelles and their organization.
5. Transport of nutrients, ions and macromolecules across membranes.
7. Cellular responses to environmental signals in plants and animals- mechanisms of signal transduction.
10. Cellular basis of differentiation and development - mitosis, gametogenesis and fertilization, development in Drosophila and Arabidopsis; Spatial and temporal regulation of Gene Expression.
11. Genes, Mutation and Mutagenesis - UV and chemical mutagens; Types of mutation; Ames test for mutagenesis; Methods of genetic analysis.
12. Bacterial Genetic System
   Transformation, Conjugation, Transduction, Recombination, Plasmids and Transposons. Bacterial genetics map with reference to E.Coli.
13. Virus and Their Genetic System
   Phage I and its life cycle; RNA phages; RNA Viruses; Retroviruses.
15. Extra-Chromosomal Inheritance.

**M.Sc. Bioinformatics II Semester**

**BIOINFORMATICS-I**

1. **Introduction and Bioinformatics Resources:** Knowledge of various databases and bioinformatics tools available at these resources, the major content of the databases, purpose and utility in life sciences, Literature databases:
   - Nucleic acid sequence databases: GenBank, EMBL, DDBJ
   - Protein sequence databases: SWISS-PROT, TrEMBL, PIR_PSD, PDB
   - Genome Databases at NCBI, EBI, TIGR, SANGER
   - Other Databases of Patterns/Motifs/System Biology(Gene and protein network database and resources)

2. **Sequence analysis:**
   - Various file formats for bio-molecular sequences: genbank, fasta, gcg, msf, nbrf-pir etc.
   - Basic concepts of sequence similarity, identity and homology, definitions of homologues, orthologues, paralogues.
• Scoring matrices: basic concept of a scoring matrix, Matrices for nucleic acid and proteins sequences, PAM and BLOSUM series and other related matrices.

• Sequence-based Database Searches: what are sequence-based database searches, BLAST and FASTA algorithms, Various versions of basic BLAST and FASTA,

3. Pairwise and Multiple sequence alignments: basic concepts of sequence alignment, Needleman & Wunch, Smith & Waterman algorithms for pairwise alignments, Progressive and hierarchial algorithms for MSA. use of pairwise alignments and Multiple sequence alignment for analysis of Nucleic acid and protein sequences and interpretation of results; Sequence patterns and profiles: PSI-Blast, PSSM.

4. Phylogenetics and Molecular Evolution
Basic concepts in systematics, taxonomy and phylogeny; molecular evolution; nature of data used in Taxonomy and Phylogeny, Definition and description of phylogenetic trees and various types of trees, Computational phylogenetics and Computational Molecular Evolution.

5. Current Advancements in Bioinformatics
Introduction to System Biology, Structural Biology, Structural bioinformatics, homology modeling, Chemioinformatics, Immunoinformatics etc.

Object Oriented Programming through 'c++'

1. Principles of Object-Oriented Programming : Software evolution, OOP Paradigm, Basic concepts of OOPs, Benefits of OOP, Object-Oriented languages, Application of OOP.


3. Tokens, Expressions and Control Structures : Tokens, Keywords, Identifiers and Constants, Basic data types, User defined data types, Declaration of variables, Reference variables, Operators, Scope Resolution Operator, Manipulators, Type cast operator, Operator overloading.


5. Classes and Objects : Defining member functions, A C++ program with class, Making an outside function Inline, Nesting of member functions, Private member functions, Arrays within a class, Memory allocation for objects, Static data members, Static member functions.

6. Constructors and Destructors : Constructors, Parameterized constructors, Multiple constructors, Constructors with default argument, Copy constructors, Dynamic constructors, Destructors.

7. Operators Overloading and Type Conversions : Defining Operators Overloading, Overloading Unary Operators, Overloading Binary Operators, type Conversions.


9. Pointers, Virtual Functions and Polymorphism : Introduction, Pointers to objects, this pointer, Pointers to derived classes, Virtual functions.
10. **Manipulating Strings** : Creating string objects, Manipulating string objects, Relational Operators, String characteristics, Comparing and Swapping.

---

**MOLECULAR BIOLOGY**

1. **Biosynthesis of Purines and pyrimidine nucleotides from ribose including regulation, salvage pathways.**

2. **DNA Replication**
   - Prokaryotic and eukaryotic DNA replication, Mechanics of DNA replication, Enzymes and accessory proteins involved in DNA replication. **DNA Repair.**

4. **Transcription**
   - Prokaryotic transcription, Eukaryotic transcription, RNA polymerase, General and specific transcription factors, Regulatory elements and mechanisms of transcription regulation, Transcriptional and post-transcriptional gene silencing. 5'-Cap formation, Transcription termination, 3'-end processing and polyadenylation, Splicing, Editing, Nuclear export of mRNA, mRNA stability.

5. **RNA splicing**
   - Nuclear Splicing, spliceosomes and small nuclear RNAs, group I and group II introns, Cis and trans-splicing reactions, tRNA splicing, alternate splicing.

6. **Translation**

7. **Protein Localization**
   - Synthesis of secretory and membrane proteins, Import into nucleus, mitochondria, chloroplast and peroxisomes, Receptor mediated endocytosis.

8. **Oncogenes and Tumor Suppressor Genes**
   - Viral and cellular oncogenes, tumor suppressor genes from humans, Structure, function and mechanism of action of pRB and p53 tumor suppressor proteins.

9. **Antisense and Ribozyme Technology**
   - Molecular mechanism of antisense molecules, inhibition of splicing, polyadenylation and translation, disruption of RNA structure and capping, Biochemistry of ribozyme; hammer- head, hairpin and other ribozymes, strategies for designing ribozymes, Applications of antisense and ribozyme technologies.

10. **Homologous Recombination**
    - Holliday junction, gene targeting, gene disruption, FLP/FRT and Cre/Lox recombination RecA and other recombinases.

11. **Molecular Mapping of Genome**
    - Genetic and physical maps, physical mapping and map-based cloning, choice of mapping population, Simple sequence repeat loci, Southern and fluorescence in situ hybridization for genome analysis, Chromosome microdissection and microcloning, Molecular markers in genome analysis: RFLP, RAPD and AFLP analysis, Molecular markers linked to disease resistance genes, Application of RFLP in forensic, disease prognosis, genetic counseling, Pedigree, varietal etc. Animal trafficking and poaching; Germplasm maintenance, taxonomy and Bio-diversity.

---

**DESIGN & ANALYSIS OF ALGORITHM**

1. **Basic Concepts of Algorithms**
2. Mathematical Aspects and Analysis of Algorithms
Mathematical Analysis of Non-recursive Algorithm – Mathematical Analysis of Recursive Algorithm –

3. Analysis of Sorting and Searching Algorithms
Brute Force – Selection Sort and Bubble Sort – Sequential Search and Brute-force string matching – Divide and
conquer – Merge sort – Quick Sort – Binary Search – Binary tree-Traversal and Related Properties – Decrease
and Conquer – Insertion Sort – Depth first Search and Breadth First Search.

4. Algorithmic Techniques
Transform and conquer – Presorting – Balanced Search trees – AVL Trees – Heaps and Heap sort – Dynamic
Programming – Warshall’s and Floyd’s Algorithm – Optimal Binary Search trees – Greedy Techniques – Prim’s
Algorithm – Kruskal’s Algorithm – Dijkstra’s Algorithm – Huffman trees.

5. Algorithm Design Methods
Backtracking – n-Queen’s Problem – Hamiltonian Circuit problem – Subset-Sum problem – Branch and bound

INTERNET & WEB BASED PROGRAMMING (CGI, PERL & HTML)

1. Internet Basics
The Basics of the Internet, Concepts of a Domain, Networking concepts, IP Addressing, Resolving Domain Names,

2. Hyper text markup language (HTML)
How a Web Browser communicates with a web server, what is HTML and various HTML tags, Commonly used
HTML commands, Lists, Adding Graphics to HTML documents, to create and use Tables, the concept of
Hyperlink, Types of Hyperlinks, Introduction to Frames, Using the <Frameset> and the <Frame> tag. Other tags
and versions of HTML such as DHTML and XML.

3. Common Gateway Interface (CGI)
The concept of CGI, Why CGI is used, How CGI works, The two methods of Data submissions, the differences
in the two methods of submissions, the importance of Environment variables in a CGI program, the basic steps
required to process from information in a CGI program, Why Perl is the language of choice for programming in
CGI.

4. Perl Language
The basics of the Perl Language, the concept of Perl Strings and their types, the values that can be stored in
scalar variables, Arrays, how to extract information from both types of Arrays, the importance of the special
Hash Array, Performing operations & Controlling program Flow, Perl Functions, File Handling.

5. Perl applications for biological data: BioPerl.

M.Sc. Bioinformatics Semester III

BIOINFORMATICS-II

1. Scoring Models and Matrices:
Scoring models for gap penalties, computational aspects and generation of PAM and BLOSUM matrices.
Applications of substitution matrices in protein sequence alignment and evolution.

2. Markov Chains & Hidden Markov Models:
Forward and backward algorithms. Parameters estimation for HMMs. HMMs for pairwise and multiple sequence alignments. Profile HMMs.

3. Machine Learning and Bioinformatics:
Introduction to various Machine Learning techniques and their applications in Bioinformatics. Genetic algorithms, Support Vector Machine, Neural Networks and their practical applications towards the development of new models, methods and tools for Bioinformatics.

4. Computational Models in Phylogenetics:

5. Computational RNA Structure analysis:

BIOINFORMATICS-III

1. Prediction of protein structure
   • Secondary structure: algorithms of Chou Fasman, GOR methods.
   • Tertiary Structure: basic principles and protocols, Methods to study 3D structure.
   • Protein structure comparison and classification: classes, folds; the concepts in 3D structure comparison, purpose of structure comparison, algorithms such as FSSP, VAST and DALI.

2. Molecular modeling and simulations
   • Visualization of structures using Rasmol or SPDBViewer or CHIME
   • Basic concepts in molecular modeling: different types of computer representations of molecules
   • Concepts of force fields: representations of atoms and atomic interactions, potential energy representation.

3. Homology modeling
   • Basic concepts in molecular modeling: different types of computer representations of molecules.
   • Visualization of structures using Rasmol or SPDBViewer or CHIME.
   • Principles of protein folding and methods to study protein folding.
   • Concepts of force fields: representations of atoms and atomic interactions, potential energy representation.

4. Computer aided drug design (CADD)
   Proteins as drug targets, docking, ligand, receptor and target. Analogy and structure based CADD. Computer aided molecular design.

5. Systems Biology

DATA BASE MANAGEMENT SYSTEMS

   • Introduction, data models- Entity Relationship Model, Relational Model
RECOMBINANT DNA TECHNOLOGY

1. Scope of Recombinant DNA Technology
2. Milestones In Genetic Engineering
3. Molecular Tools and Their Applications
   Restriction enzymes, modification enzymes, DNA, and RNA markers.
4. Nucleic Acid Purification, Yield Analysis
5. Nucleic Acid Amplification and Its Applications
6. Gene Cloning Vectors
   Plasmids, bacteriophages, phagemids, cosmids, Artificial chromosomes.
7. Restriction Mapping of DNA Fragments and Map Construction.
   Nucleic Acid Sequencing.
8. cDNA Synthesis and Cloning
   mRNA enrichment, reverse transcription, DNA primers, Linkers, adaptors and their chemical synthesis, Library construction and screening.
9. Alternative Strategies of Gene Cloning
   Cloning interacting genes- Two-and three hybrid systems, cloning differentially expressed genes, Nucleic acid microarray arrays.
10. Site-directed Mutagenesis and Protein Engineering
11. How to Study Gene Regulation?
    DNA transfection, Northern blot, Primer extension, S1 mapping, RNase protection assay, Reporter assays.
12. Expression Strategies for Heterologous Genes
    Vector engineering and codon optimization, host engineering, In vitro transcription and translation, expression in bacteria, expression in Yeast, expression in insects and insect cells, expression in mammalian cells, expression in plants.
13. Processing of Recombinant Proteins
    Purification and refolding, characterization of recombinant proteins, stabilization of proteins.
14. Phage Display
15. T-DNA and Transposon Tagging: Role of gene tagging in gene analysis, T-DNA and transposon tagging, Identification and isolation of genes through T-DNA or transposon.
Java Programming


2. **JAVA Evolution** : Java History, Java Features, How java Differs from C and C++, Java and Internet, Java and World Wide Web, Web Browsers, Hardware and Software Requirement, Java support system, Java environment.

3. **Overview Of Java language** : Introduction, Simple java program, More of java, Java program structure, Java tokens, Java statements, Implementing a java program, Java virtual machine, Command line arguments, Programming style.

4. **Constants, Variables and Data types** : Constants, Variables, Data types, Declaration of variables, Giving values to variables, Scope of variables, Type casting, Getting values of variables.

5. **Operators and Expressions** : Arithmetic operators, Relational operators, Logical operators, Assignment operators, Increment and Decrement operators, Conditional operators, Bitwise operators, Arithmetic expressions, Evaluation of expressions, Type conversions in expressions.


7. **Classes, Objects and Methods** : Defining a class, Adding variables, Adding methods, Creating objects, Accessing class members, Constructors, Methods overloading, Static members, Inheritance, Overriding methods, Final variables and methods, Final classes, Finalizer methods, Abstract methods and classes.

8. **Arrays, String and Vectors** : Arrays, one and two dimensional arrays, Strings, Vectors, Wrapper classes.

9. **Interfaces** : Multiple Inheritance : Defining Interfaces, Extending Interfaces, Implementing Interfaces, Accessing Interfaces variables.

10. **Packages** : Putting Classes Together : Java API packages, Using system packages, Creating packages, Accessing packages, Using packages.

11. **Multithreaded Programming** : Creating threads, Extending the thread class, Stopping and blocking a thread, Life cycle of a thread, Using thread methods.

12. **Managing Errors and Exceptions** : Types of errors, Exceptions, Syntax of exception handling code, Multiple catch statements, Using finally statement, Throwing our own exceptions.


14. **Managing input/Output Files in Java** : Concept of streams, Streams classes, Using streams, Using the file class, Creation of files.

---

**M.Sc. Bioinformatics Semester IV**

1. Research Project