

## **Syllabus of B.Sc. (Bioinformatics)**

Subject- Bioinformatics (as one subject)

B.Sc. I Year

### Semester I

Paper I: Basic of Bioinformatics 85 marks  
Practicals 50 marks

### Semester II

Paper II: Mathematics I 85 marks  
Practicals 50 marks

B.Sc. II Year

### Semester III

Paper III: Structural Bioinformatics 85 marks  
Practicals 50 marks

### Semester IV

Paper IV: Mathematics II 85 marks  
Practicals 50 marks

B.Sc. III Year

### Semester V

Paper I: Computer graphics, Machine learning and Bioperl 85 marks  
Practicals 50 marks

### Semester VI

Paper I: Informatics in omics and application 85 marks  
Practicals 50 marks

15 marks CCE as per higher education norms in each semester.  
Practicals based on theory are to be designed by the concerned teacher.

# **B.Sc. BIOINFORMATICS**

## **B.Sc. Part - I Semester - I**

### **Session 2011-12**

#### **Paper I –Basic of Bioinformatics**

**Max. marks-85**

#### **Unit I Introduction to bioinformatics and data generation**

What is bioinformatics and its relation with molecular biology. Examples of related tools(FASTA, BLAST, BLAT, RASMOL), databases(GENBANK, Pubmed, PDB ) and software(RASMOL,Ligand Explorer).

Data generation; Generation of large scale molecular biology data. (Through Genome sequencing, Protein sequencing, Gel electrophoresis, NMR Spectroscopy, X-Ray Diffraction, and microarray).

Applications of Bioinformatics.

#### **Unit II Biological Database and its Types**

Introduction to data types and Source. Population and sample, Classification and Presentation of Data. Quality of data, private and public data sources.

General Introduction of Biological Databases;

Nucleic acid databases (NCBI, DDBJ, and EMBL).

Protein databases (Primary, Composite, and Secondary).

Specialized Genome databases: (SGD, TIGR, and ACeDB).

Structure databases (CATH, SCOP, and PDBsum)

#### **Unit III Data storage and retrieval and Interoperability**

Flat files, relational, object oriented databases and controlled vocabularies. File Format (Genbank, DDBJ, FASTA, PDB, SwissProt).

Introduction to Metadata and search; Indices, Boolean, Fuzzy, Neighboring search.

The challenges of data exchange and integration. Ontologies, interchange languages and standardization efforts.

General Introduction to XML, UMLS, CORBA, PYTHON and OMG/LIFESCIENCE.

#### **Unit IV Sequence Alignments and Visualization**

Introduction to Sequences, alignments and Dynamic Programming;

Local alignment and Global alignment (algorithm and example),

Pairwise alignment (BLAST and FASTA Algorithm) and multiple sequence alignment (Clustal W algorithm).

Methods for presenting large quantities of biological data: sequence viewers (Artemis, SeqVISTA), 3D structure viewers (Rasmol, SPDBv, Chime, Cn3D, PyMol), Anatomical visualization.

## **Unit V Gene Expression and Representation of patterns and relationship**

General introduction to Gene expression in prokaryotes and eukaryotes, transcription factors binding sites. SNP, EST, STS.

Introduction to Regular Expression, Hierarchies, and Graphical models (including Markov chain and Bayes notes).

Genetic variability and connections to clinical data.

# **B.Sc. BIOINFORMATICS**

## **B.Sc. Part - I Semester - II**

### **Session 2011-12**

#### **Paper II –Mathematics I**

**Max. marks-85**

#### **Unit I**

Sets, Types of Sets, Subsets, Complement of Sets, union and Intersection of Sets, Difference of Sets, Demorgan's Law, Cartesian product of Sets.

Basics of Probability, Permutation and Combination.

#### **Unit II**

Measure of central tendency and dispersion: Mean, median, mode, range, standard deviation, variance

#### **Unit III**

Correlation and Regression: Types, Karl-Pearson's correlation, Spearman's Rank correlation, Regression equation and fitting

#### **Unit IV**

Probability Distribution: Basics of Binomial, Poisson and Normal distributions and their application in biology.

Random Variable; Discrete and Continuous Probability Distribution, Probability mass function, probability Density function, Mathematical Expectation.

#### **Unit V**

Matrices, Types of Matrices, Addition of matrices, Subtraction of matrices and Product of matrices.

Properties of Matrix Multiplication. Transpose of Matrix, Symmetric and Skew-symmetric Matrices, Inverse of Matrix

**B.Sc. Part - II**  
**Semester - III**  
**Paper III –Structural Bioinformatics**

**Session 2012-13**

**Max. marks-85**

**Unit I**

Fundamentals of X-ray diffraction, NMR spectroscopy of macromolecules

Protein Structure: Primary, Secondary, Super Secondary, Domains, Tertiary, Quaternary, Ramachandran plot.

**Unit II**

Protein secondary structure classification databases: HSSP, FSSP, CATH, SCOP.

Protein secondary structure prediction methods: GOR, Chou-Fasman, PHD, PSI- PRED, J-Pred.

**Unit III**

Protein Tertiary structure prediction methods: Homology Modeling, Fold Recognition, Ab-initio Method.

Protein folding, Molecular Dynamics of Protein, Molecular Docking of Protein, Small molecule and Nucleotide, Concepts of Force Field

**Unit IV**

Motif and Domain: Motif databases and analysis tools.

Domain databases (CDD, SMART, ProDom) and Analysis tools.

HMM (Hidden Markov Model): Introduction to HMM, its application in Sequence alignment and Structure prediction, HMM based Softwares (HMMER and HMMSTR)

**Unit V**

Structural features of RNA: Primary, Secondary, Tertiary.

Introduction to RNA Secondary structure prediction, Methods for RNA Secondary structure prediction, Limitation of RNA Secondary structure prediction

**B.Sc. Part - II**  
**Semester - IV**

**Session 2012-13**

**Paper IV- Mathematics II**

**Max. marks-85**

**Unit I**

Law of Indices, trigonometric ratios, Inverse trigonometric functions.  
Binomial theorem.

**Unit II**

Sequences and series AP, GP, HP, Logarithmic and Exponential Series.

**Unit III**

Basics of Functions and Limits, Elementary Differentiation and Integration  
Fourier transform, Laplace transform and other standard transform.

**Unit IV**

Coordinate geometry: Distance between two points, section formula, Locus of points.  
Equation of lines, Circle, Ellipse, Parabola, Hyperbola

**Unit V**

Vector: Addition, subtraction, dot product, cross product, scalar triple product,  
Vector differentiation and vector integration, gradient, divergence, curl of a vector, equation  
of normal

**B.Sc. Part - III**  
**Semester - V**

**Session 2013-14**

**Paper V- Computer graphics, Machine learning and Bioperl**

**Max. marks-85**

**Unit I**

Color models: CMY, HSV, RGB, Visualization techniques.

Graphics display devices, Raster and Random scan devices, color CRT monitors, LCD and LED.

**Unit II**

Artificial Neural Networks, Genetic algorithm, Bayesian modeling, Monte Carlo Simulation Method, Markov Models and their application

**Unit III**

Perl basic: Variables, Perl operations, A Program to store DNA sequence, Concatenating DNA fragment, Transcription: DNA to RNA, Subroutines, scoping and subroutines, command-line arguments and arrays passing data to subroutines, modules and libraries of subroutines, fixing bugs.

**Unit IV**

Random number generators a program using randomization, a program to simulate DNA Mutation, generating random DNA, analyzing DNA, Hashes, data structures and algorithms for biology, the genetic code,

**Unit V**

Translating DNA into proteins reading DNA from files in FASTA format, reading frames, Regular expressions, restriction maps and restriction enzymes, Genbank files, Genbank libraries, separating sequence and annotation parsing, Annotations indexing, parsing PDB files, parsing BLAST files.

**B.Sc. Part - III**  
**Semester - VI**

**Session 2013-14**

**Paper VI- Informatics in omics and application**

**Max. marks-85**

**Unit I:**

Genomics: Genome Annotation, Genome Assembly, Structural and Functional Genomics.

Comparative Genomics

Microarray: technique, Design, Analysis, Drug target identification.

**Unit II:**

System biology: Introduction, Associated disciplines, Interactomics (PPI), Fluxomics, Biomics.

Metagenomics: Introduction, metagenome, shotgun metagenomics (pyrosequencing).

Tool's in metagnomics, MEGAN, MG- RAST, and SEED.

Application: Gene survey, Enviornmental genomes, Microbial diversity.

**Unit III:**

Metabolic pathway database (KEGG pathway database), Concept of metabolome and metabolomics.

Drug Discovery and design : Target identification , Target Validation , Lead Identification , lead optimization , preclinical Pharmacology & Taxology.

**Unit IV**

Chemoinformatics: Cheminformatics tools for drug discovery.

Chemical Structure Representation (SMILE & SMART).

Chemical databases: CSD, ACD, WDI, ChemBank, hazardous chemical database, PUBCHEM.

**Unit V:**

Quantitative Structure Activity Relationship (2D & 3D). Combinatorial libraries & their design. High throughput screening, virtual screening, Lipinski's rule of five.