M.Sc. BIOINFORMATICS

REGULATIONS AND SYLLABI
(Effective from 2011-2012)

Centre for Bioinformatics
SCHOOL OF LIFE SCIENCES
PONDICHERRY UNIVERSITY
PUDUCHERRY
Eligibility for M. Sc. Bioinformatics

Students from any of the below listed Bachelor degrees with minimum 55% of marks are eligible.

- Bachelor’s degree in any relevant area of Physics/Chemistry/Computers Science/Life Science/with a minimum of 55% of marks
# SYLLABUS FOR M. Sc. BIOINFORMATICS

(Pondicherry University  School of Life Sciences  Centre for Bioinformatics  Academic Year 2011-2012 onwards)

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
<th>H/S</th>
<th>Credits</th>
<th>Pg. No</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Semester I</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BINF 421</td>
<td>Cell and Molecular Biology</td>
<td>H</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>BINF 422</td>
<td>Bioinformatics Databases</td>
<td>H</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>BINF 423</td>
<td>C and Data Structures</td>
<td>H</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>BINF 424</td>
<td>Physics</td>
<td>S</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>BINF 425</td>
<td>Chemistry</td>
<td>S</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>BINF 426</td>
<td>Mathematics</td>
<td>S</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>BINF 427</td>
<td>Biology</td>
<td>S</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>BINF 428</td>
<td>Basics of Computer</td>
<td>S</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>BINF 419</td>
<td>Introduction to Bioinformatics*</td>
<td>S</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>BINF 471</td>
<td>Lab - Cell and Molecular Biology</td>
<td>H</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>BINF 472</td>
<td>Lab - Biological Databases</td>
<td>H</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>BINF 473</td>
<td>Lab - Basics of Computer &amp; Operating Systems</td>
<td>H</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>BINF 474</td>
<td>Lab - Programming in C/ C++</td>
<td>H</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>BINF 465</td>
<td>Lab - Bioinformatics databases and tools</td>
<td>S</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td><strong>Semester II</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BINF 451</td>
<td>Genomics and Proteomics</td>
<td>H</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>BINF 452</td>
<td>Bioinformatics: Sequence Analysis</td>
<td>H</td>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td>BINF 453</td>
<td>Probability and Statistics</td>
<td>S</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>BINF 454</td>
<td>Programming in Java</td>
<td>H</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>BINF 455</td>
<td>Database Management System</td>
<td>H</td>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>BINF 456</td>
<td>Fundamentals of Algorithms</td>
<td>S</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>BINF 457</td>
<td>Microscopic Techniques For Image Processing</td>
<td>S</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>BINF 475</td>
<td>Lab - Programming in Java</td>
<td>H</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>BINF 476</td>
<td>Lab – Programming in DBMS</td>
<td>H</td>
<td>1</td>
<td>26</td>
</tr>
<tr>
<td>BINF 477</td>
<td>Lab - Biosequence Analysis</td>
<td>H</td>
<td>1</td>
<td>27</td>
</tr>
<tr>
<td><strong>Semester III</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BINF 521</td>
<td>Structural Biology</td>
<td>H</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>BINF 522</td>
<td>Molecular Modeling and Drug Design</td>
<td>H</td>
<td>3</td>
<td>29</td>
</tr>
<tr>
<td>BINF 523</td>
<td>Programming in Perl</td>
<td>H</td>
<td>3</td>
<td>30</td>
</tr>
<tr>
<td>BINF 524</td>
<td>Systems Biology</td>
<td>H</td>
<td>3</td>
<td>31</td>
</tr>
<tr>
<td>BINF 525</td>
<td>Biological Spectroscopy</td>
<td>S</td>
<td>2</td>
<td>32</td>
</tr>
<tr>
<td>BINF 526</td>
<td>Data Communication and Networks</td>
<td>H</td>
<td>2</td>
<td>33</td>
</tr>
<tr>
<td>BINF 527</td>
<td>Scientific Presentation and Finishing School</td>
<td>H</td>
<td>3</td>
<td>34</td>
</tr>
<tr>
<td>BINF 571</td>
<td>Lab - Structural Biology</td>
<td>H</td>
<td>1</td>
<td>35</td>
</tr>
<tr>
<td>BINF 572</td>
<td>Lab - Molecular Modeling and Drug Design</td>
<td>H</td>
<td>1</td>
<td>36</td>
</tr>
<tr>
<td>BINF 573</td>
<td>Lab - Programming in Perl</td>
<td>H</td>
<td>1</td>
<td>37</td>
</tr>
<tr>
<td><strong>Semester IV</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BINF 551</td>
<td>Analytical Techniques</td>
<td>S</td>
<td>3</td>
<td>38</td>
</tr>
<tr>
<td>BINF 552</td>
<td>Bioethics, Biodiversity and Intellectual Property Rights</td>
<td>H</td>
<td>3</td>
<td>39</td>
</tr>
<tr>
<td>BINF 553</td>
<td>R language and Bioconductor</td>
<td>S</td>
<td>3</td>
<td>40</td>
</tr>
<tr>
<td>BINF 554</td>
<td>Project</td>
<td>H</td>
<td>8</td>
<td>41</td>
</tr>
</tbody>
</table>

* for other department students
BINF 421 - CELL AND MOLECULAR BIOLOGY

Total Credits: 3

Unit 1

Unit 2

Unit 3
Chloroplast structure and function – An overview of photosynthetic Metabolism – The absorption of light – Photosynthetic units and reaction centers – Photophosphorylation – Carbon dioxide fixation and the synthesis of carbohydrates. Chloroplast and its genome study.

Unit 4

Unit 5
DNA and Protein Synthesis - Structure of DNA - evidence for DNA as genetic material. Gene transfer in microorganisms – conjugation, transformation, transduction - protoplasmic fusion. The genomes of bacteria, viruses, plasmids. DNA Structural organization - DNA replication, Transcription – mRNA processing, Translation. Protein synthesis – Ribosomes, enzymes, Protein processing, Introduction to the methods of DNA sequencing

Text Book:

Reference Books:
2. Cell and Molecular Biology by De Robertes and De Robertis. Saunders College, Philadelphia, USA. 2002
BINF 422 - BIOINFORMATICS DATABASES

Total Credits: 3  
Total: 36 Hrs.

Unit 1  
6 lectures
Introduction to Bioinformatics data and databases:- Types of Biological data:-- Genomic DNA, Complementary DNA (cDNA), Recombinant DNA (rDNA), Expressed sequence tags (ESTs), Genomic survey sequences (GSSs). Primary Databases:- GenBank, EMBL, DDBJ, Composite Databases:-NRDB, UniProt, Literature Databases:- Open access and open sources, PubMed, PLoS, Biomed Central.

Unit 2  
8 lectures
Genome Databases:- Viral genome database (ICTVdb, VirGen), Bacterial Genomes database (Genomes OnLine Database –GOLD, Microbial Genome Database-MBGD), Organism specific Genome database (OMIM / OMIA, SGD, WormBase, PlasmoDB, FlyBase, TAIR), and Genome Browsers (Ensembl, VEGA genome browser, NCBI-NCBI map viewer, KEGG, MIPS, UCSC Genome Browser).

Unit 3:  
8 lectures
Sequence Databases :- Nucleotide sequence Databases (GenBank, EMBL, DDBJ). Protein sequences Databases (Swiss-ProtProt, TrEMBL, UniProt, UniProt Knowledgebase – UniProtKB, UniProt Archive –UniParc, UniProt Reference Clusters –UniRef, UniProt Metagenomic and Environmental Sequences –UniMES. Sequence motifs Databases:-Prosite, ProDom, Pfam, InterPro. Sequence file formats:- GenBank, FASTA, PIR, ALN/ClustalW2, GCG/MSF.

Unit 4:  
8 lectures

Unit 5  
6 lectures
Bioinformatics Database search engines:-Text-based search engines (Entrez, SRS, DBGET / LinkDB). Sequence similarity based search engines (BLASTBLAST and FASTA). Motif-based search engines (ScanPrositeScanProsite and eMOTIF). Structure similarity based search engines (VAST and DALI). Proteomics tools at the ExPASy server, GCG utilities and EMBOSS.

Text Books:
3. Introduction to bioinformatics by Teresa K. Attwood, David J. Parry-Smith. Pearson Education. 1999
BINF 423 – C AND DATA STRUCTURES

Total Credits: 3

Unit 1

Unit 2
Programming in C:

Unit 3
Procedural Concept:

Unit 4
String Handling & Sorting:
String declaration – String library functions - String Manipulation - Sorting: Bubble sort, Selection sort, Insertion sort – Searching: Linear search, Binary search

Unit 5
Object Oriented Programming: Programming in C++:
C++ programming – Object Oriented Concept: Encapsulation, Inheritance, Polymorphism – Different forms of Constructor – Destructor – Abstract class – Virtual function

Text Books:

Unit 1

Unit 2

Unit 3

Unit 4

Unit 5

Text Books:

Reference Books:
BINF 425 – CHEMISTRY

Total Credits: 2

Total: 24 Hrs.

Unit 1
5 Lectures

Atomic and Molecular Structure:

Unit 2
5 Lectures

Symmetry and Principles: Definitions and theorems of group theory, subgroups, Classes. Molecular symmetry and symmetry groups – symmetry elements and operations. Symmetry planes, reflections, inversion centre, proper/ improper axes of rotation, symmetry elements and optical isomerism, symmetry point groups, classes of symmetry operations, classification of molecular point groups.

Unit 3
5 Lectures

Introduction to Organic chemistry: Carbon and its compounds, Position of Carbon in periodic table, tetra covalency of carbon, catenation, functional groups, formal charge, oxidation number, aromaticity, electrophiles and nucleophiles, organic acids and bases, types of organic reactions.

Unit 4
5 Lectures

Stereochemistry: Concept of isomerism, types of isomerism, optical isomerism, elements of symmetry, molecular chirallity, enantiomers, stereogenic centres, optical activity, properties of enantiomers, chiral and achiral molecules with two stereogenic centres, distereoisomers, mesocompounds, resolution of enantiomers. Relative and absolute configurations, sequence rules, D & L , R & S systems of nomenclature.

Unit 5
4 Lectures

Heteroaromatics: Five membered and six membered hetero aromatics with one and two hetero atoms and their benannulated analogues, Nucleic acid bases, Structure, name and properties like acid base property, electron rich electron deficient heterocycles, hydrogen bonding etc. (Synthesis and reactions not necessary).

Text Books:

1. Organic Chemistry by Paula Yurkanis Bruice, Prentice Hall. 2010
Unit I
Matrices and Linear Algebra  
Matrices- Properties of Determinants, Minors and Cofactors, Multiplication of Determinants, Adjoint, Reciprocal, Symmetric Determinants, Cramer’s rule, Different types of matrices, Matrix Operations, Transpose of a matrix, Adjoint of a square matrix, Inverse of a matrix, Eigen values and eigen vector  
Linear Algebra - Definition of vector space, Subspaces, Linear independence and Bases.

Unit II

Unit III
Trigonometry and Analytical Geometry: Trigonometric ratios, De Moivre’s theorem, The general equation of a Straight line, slope of a line, intercepts of a line, Angle between two lines, Intersection of two lines, The general equation of a Circle.

Unit IV

Unit 5

Text Books:
1. Algebra by Serge A. Lang, Pearson Education. 2003

Reference Books:
1. Basic Mathematics by Serge A. Lang. Springer Publisher. 1988
BINF 427 – BIOLOGY

Total Credits: 2

Total: 24 Hrs.

Unit 1

Diversity in Living World: Diversity of living organisms-Classification of the living organisms (five kingdom classification, major groups and principles of classification within each kingdom). Systematics and binomial System of nomenclature - Salient features of animal (non-chordates up to phylum level and chordates up to class level) and plant (major groups; Angiosperms up to class) ) linnaean classification.

Unit 2

Structural Organisation in Animals and Plants: Morphology, anatomy and functions of different parts of flowering plants: Root, stem, leaf, inflorescence, flower, fruit and seed. Morphology, anatomy and functions of different systems of an annelid (earthworm), an insect (cockroach) and an amphibian (frog).

Unit 3


Unit 4


Unit 5


Text Book:
1. Molecular Biology of the cell by Bruce Alberts, Garland publishing Inc. 2002

Reference Books:
2. Cell and Molecular Biology by De Robertes and De Robertis. Saunders College, Philadelphia, USA. 2002
BINF 428 - BASICS OF COMPUTER

Total Credits: 2                 Total: 24 Hrs.

Unit 1  5 Lectures
Computer Organization

Unit 2  5 Lectures
Network Basics
Communication Technology – Networking Elements: Networking Hardware, Networking services: Types of Networks – LAN, WAN & MAN, Intranet–Wireless communication – Internet services, Uses of Internet

Unit 3  4 Lectures
Introduction to Database systems
Fundamentals of database - Database models (Hierarchical, Network, Relational and Object-Oriented Models) – RDBMS: Relational Database Management systems - Database System Applications and Security.

Unit 4  5 Lectures
Programming Language

Unit 5  5 Lectures
Internet Technologies
Web Services – WWW, URL, Servers: Client/ Server essentials - Domain Name Server, FTP server, E-mail server, WEB servers, Web publishing-Browsers-IP Addressing, IPV6

Text Books:

BINF 419 - INTRODUCTION TO BIOINFORMATICS

Total Credits: 3

Total: 36 Hrs.

Unit 1
Introduction: Aim and branches of Bioinformatics, Application of Bioinformatics, Role of internet and www in bioinformatics. Basic biomolecular concepts: Protein and amino acid, DNA & RNA, Sequence, structure and function. Forms of biological information, Types of Nucleotide Sequence: Genomic DNA, Complementary DNA (cDNA), Recombinant DNA (rDNA), Expressed sequence tags (ESTs), Genomic survey sequences (GSSs). DNA sequencing methods: Basic and Automated DNA sequencing, DNA sequencing by capillary array and electrophoresis, Gene expression data.

Unit 2
Bioinformatics Resources: NCBI, EBI, ExPASy, RCSB, DDBJ: The knowledge of databases and bioinformatics tools available at these resources, organization of databases: data contents, purpose and utility. Open access bibliographic resources and literature databases: PubMed, BioMed Central, Public Library of Sciences (PloS), CiteXplore.

Unit 3

Unit 4
Sequence Analysis: Basic concepts of sequence similarity, identity and homology, definitions of homologues, orthologues, paralogues and xenologues Scoring matrices: basic concept of a scoring matrix, Matrices for nucleic acid and proteins sequences, PAM and BLOSUM series, matrix derivation methods and principles.

Unit 5
Sequence alignment: Measurement of sequence similarity; Similarity and homology. Pairwise sequence alignment: Basic concepts of sequence alignment, Needleman and Wunsch, Smith and Waterman algorithms for pairwise alignments, gap penalties, use of pairwise alignments for analysis of Nucleic acid and protein sequences and interpretation of results.

Text Books:

Reference Book:
1. Introduction to bioinformatics by Teresa K. Attwood, David J. Parry-Smith, Pearson Education. 1999
2. Bioinformatics for Dummies by Jean-michel Claverie Cedric Notredame. Publisher: Dummies (Jan 2007)
BINF 471 - LAB - CELL AND MOLECULAR BIOLOGY

Total Credits: 1

**Exercises in Cell Biology**

Paper Chromatography of Chlorophyll pigments
Estimation of Chlorophyll
Ascorbic acid estimation in different tissues of plants and animals.
Growth curve of Bacteria.
Estimation of cell mass of bacteria.

**Exercises in Molecular Biology**

Isolation & Purification of genomic DNA from bacteria
Isolation & Purification of plasmid DNA
Agarose gel electrophoresis of chromosomal & plasmid DNA
Restriction Digestion of chromosomal & plasmid DNA
Isolation of DNA fragment from agarose gel
PCR for DNA amplification
Protein separation using HPLC (demo)
Protein separation using SDS-PAGE
Exercises:

1. Bioinformatics Resources: NCBI, EBI, DDBJ, RCSB, ExPASy

2. Open access bibliographic resources and literature databases
   a. PubMed
   b. BioMed Central
   c. Public Library of Sciences (PloS)
   d. CiteXplore.

3. Bioinformatics Resources at the species level
   a. ICTV Database
   b. AVIS
   c. VirGen
   d. Viral genomes at NCBI

4. Sequence databases:
   a. Nucleic acid sequence databases: GenBank, EMBL, DDBJ;
   b. Protein sequence databases: Uniprot-KB: SWISS-PROT, TrEMBL, UniParc;
   c. Repositories for high throughput genomic sequences: EST, STS, GSS.
   d. Genome Databases at NCBI, EBI, TIGR, SANGER


6. Derived Databases: InterPro, Prosite, Pfam, ProDom

7. Sequence file formats: GenBank, FASTA, GCG, MSF.

8. Protein and nucleic acid properties: Proteomics tools at the ExPASy server, GCG utilities and EMBOSS
BINF 473 - LAB – BASICS OF COMPUTER & OPERATING SYSTEMS

Total Credits: 1

Exercises:

1. Command Line Interface - Internal Commands- External commands

2. Graphical User Interface: Peer-to-Peer Operating System

3. Client- Server Operating System

4. Software Package:
   a. Create a manuscript using ms-word by applying relevant font styles, margins, bullets and tables.
   b. Prepare a call letter for the admission of MSC bioinformatics to all the selected students by using mail merge.
   c. Prepare a student’s fee table for four semesters in a excel sheet. Calculate the consolidated payment using links.
   d. Create all types of charts using excel for any clinical data.

5. Create a web page for an educational institution using HTML tags.

6. Create a web page to display your details by creating a model web site.
BINF 474 - LAB - PROGRAMMING IN C/ C++

Total Credits: 1

LINUX Operating System: Overview of Linux Architecture and Basic commands

C
1. Display a protein details using escape sequence
2. Calculate rotations per minute \[ \text{rpm} = 1000 \sqrt{\text{RCF}} / 11.17r \]
3. Create amino acid dictionary using switch construct
4. Identify the glucose level in a blood using if - else if construct
   \[ \text{The glucose level is identified by} \]
   \[ <70 – \text{hypoglycemia}, \text{70-180 hyperglycemia,} \text{> 180 diabetics} \]
5. Identify the type of two peptides using nested if
   \[ \text{peptide length is < 8 small, poly otherwise} \]
6. Count the number of base characters entered among n characters using loop
7. Implement stack operation
8. Count the number of positive, negative and zero energy molecules stored in an array
9. Find the transpose of the given matrix using two dimensional array
10. To calculate pH value for a given \[ [\text{OH}^-] \] concentration \[ \text{pH} = - \log_{10}(\text{OH}^-) \]
11. Draw a line in different pattern using user defined function
12. Write a user defined function to illustrate the storage class of the variables
13. Determine the percentage of a,t,g,c in the given sequence
14. Count the number of gaps in the given sequence using user defined function
15. Sort n names
16. Align two sequences
17. Count the number of motif in the given sequence
18. Swap two numbers using pointers
19. Process the organism details using structure
20. Convert the RNA sequence into DNA sequence using text file

C++
1. Create a class which shows the various forms of constructors
2. Inheritance implementation
3. Function overloading example
4. Operator overloading example
5. Dynamic polymorphism implementation
BINF 465 - LAB - BIOINFORMATICS DATABASES AND TOOLS

Total Credits: 1

Exercises:

1. Entrez and Literature Searches.
   a. PubMed
   b. PubMed central
   c. OMIM / OMIA
   d. Citation matcher

2. SRS of Biological Databases
   a. Nucleotide/ Genome Databases.
   b. Protein Sequence Database.
   c. Structure databases.
   d. Protein Pattern Databases

3. File format conversion
   a. FmtSeq
   b. ReadSeq
   c. Sequence manipulation Suite

4. Sequence Analysis
   a. Dot Plot
   b. Pairwise alignment
   c. Multiple Sequence Alignment

5. Softwares
   b. GeneDoc
   c. ClustalW / X, MEGA, MEME

6. Visualization Tool
   a. RasMol
   b. Cn3D
   c. MolMol
BINF 451 - GENOMICS AND PROTEOMICS

Total Credits: 3
Total: 36 Hrs.

Unit 1 8 Lectures
Genomics and Metagenomics: Large scale genome sequencing strategies. Genome assembly and annotation. Genome databases of Plants, animals and pathogens. Metagenomics: Gene networks: basic concepts, computational model such as Lambda receptor and lac operon. Prediction of genes, promoters, splice sites, regulatory regions: basic principles, application of methods to prokaryotic and eukaryotic genomes and interpretation of results. Basic concepts on identification of disease genes, role of bioinformatics-OMIM database, reference genome sequence, integrated genomic maps, gene expression profiling; identification of SNPs, SNP database (DbSNP). Role of SNP in Pharmacogenomics, SNP arrays. Basic concepts in identification of Drought stress response genes, insect resistant genes, nutrition enhancing genes

Unit 2 7 Lectures
Epigenetics: DNA microarray: database and basic tools, Gene Expression Omnibus (GEO), ArrayExpress, SAGE databases DNA microarray: understanding of microarray data, normalizing microarray data, detecting differential gene expression, correlation of gene expression data to biological process and computational analysis tools (especially clustering approaches)

Unit 3 7 Lectures
Comparative genomics: Basic concepts and applications, whole genome alignments: understanding the significance; Artemis, BLAST2, MegaBlast algorithms, PipMaker, AVID, Vista, MUMmer, applications of suffix tree in comparative genomics, synteny and gene order comparisons Comparative genomics databases: COG, VOG

Unit 4 7 Lectures
Functional genomics: Application of sequence based and structure-based approaches to assignment of gene functions – e.g. sequence comparison, structure analysis (especially active sites, binding sites) and comparison, pattern identification, etc. Use of various derived databases in function assignment, use of SNPs for identification of genetic traits. Gene/Protein function prediction using Machine learning tools viz. Neural network, SVM etc

Unit 5 7 Lectures
Proteomics: Protein arrays: basic principles. Computational methods for identification of polypeptides from mass spectrometry. Protein arrays: bioinformatics-based tools for analysis of proteomics data (Tools available at ExPASy Proteomics server); databases (such as InterPro) and analysis tools. Protein-protein interactions: databases such as DIP, PPI server and tools for analysis of protein-protein interactions

Text Books:

Reference books:
UNIT 1

**Sequence Analysis**: Basic concepts of sequence similarity, identity and homology, definitions of homologues, orthologues, paralogues and xenologues. **Scoring matrices**: basic concept of a scoring matrix, Matrices for nucleic acid and proteins sequences, PAM and BLOSUM series, matrix derivation methods and principles. **Database Searches**: Keyword-based Entrez and SRS; Sequence-based: BLAST & FASTA; Use of these methods for sequence analysis including the on-line use of the tools and interpretation of results from various sequence and structural as well as bibliographic databases.

UNIT 2

**Pairwise sequence alignment**: Basic concepts of sequence alignment, Needleman and Wunsch, Smith and Waterman algorithms for pairwise alignments, gap penalties, use of pairwise alignments for analysis of Nucleic acid and protein sequences and interpretation of results.

UNIT 3

**Multiple sequence alignments (MSA)**: The need for MSA, basic concepts of various approaches for MSA (e.g. progressive, hierarchical etc.). Algorithm of CLUSTALW and PileUp and their application for sequence analysis (including interpretation of results), concept of dendrogram and its interpretation, Use of HMM-based Algorithm for MSA (e.g. SAM method).

UNIT 4

**Sequence patterns and profiles**: Basic concept and definition of sequence patterns, motifs and profiles, various types of pattern representations viz. consensus, regular expression (Prosite-type) and sequence profiles; profile-based database searches using PSI-BLAST, analysis and interpretation of profile-based searches. **Algorithms for derivation and searching sequence patterns**: MeMe, PHI-BLAST, SCanProsite and PRATT. Algorithms for generation of sequence profiles: Profile Analysis method of Gribskov, HMMer, PSI-BLAST.

UNIT 5

**Taxonomy and phylogeny**: 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, Phylogenetic analysis algorithms such as maximum Parsimony, UPGMA, Transformed Distance, Neighbors-Relation, Neighbor-Joining, Probabilistic models and associated algorithms such as Probabilistic models of evolution and maximum likelihood algorithm, Bootstrapping methods, use of tools such as Phylip, Mega, PAUP.

**Text Books:**
BINF 453 - PROBABILITY AND STATISTICS

Total Credits: 2

Unit 1

Numerical descriptive techniques: Measures of central tendency: mean, median, mode, relation between mean, median and mode. Partition values: quartiles, deciles, percentiles; Measures of dispersion: Absolute and Relative Measures, Moments, skewness and kurtosis

Unit 2

Correlation and Regression: Principles of least squares, scatter diagram, correlation, covariance, correlation coefficient, properties of correlation coefficient, regression, properties of linear regression, rank correlation, multiple correlation

Unit 3

Probability Theory: Concept of probability: sample space and events, independent events, mutually exclusive events. axioms of probability, conditional probability, additional and multiplication theorem of probability, Baye's theorem, Introduction to Markov Chain Model.

Unit 4

Sampling Theory: Meaning and objective of sampling, Sampling Error, Types of Sampling, Sampling Distribution, Sampling Distribution of Sample Mean and Sample Proportion, Standard Error

Unit 5

Probability Distribution: Bernoulli trials, binomial distribution, normal distributions, Poisson distribution, Test of Hypothesis of Small and Large Samples- Standard Normal distribution, Chi-square distribution, Student’s t distribution, F distribution, Analysis of Variance

Text Books:


Reference Books:

BINF 454 - PROGRAMMING IN JAVA

Total Credits: 3

Unit 1
Java Basics - Importance and features of java, Modifiers, Access Controls, Data types, Expressions, Declarations, Statements & Control Structures, Program Structures, String handling, Packages, Interfaces, Working with java util Package, Garbage Collection

Unit 2
Exception Handling, I/O & JDBC – Exception Handling: built in exception, creating your own exceptions, Input Stream & Output Stream: Streams, Byte and Character stream, Predefined streams, Reading and Writing from Console and Files, BufferedReader & Writer, Serialization, Database: JDBC Basics

Unit 3

Unit 4
AWT & Event Handling – Creating User interface with AWT, Applets, Applet Life Cycle, Simple Graphics, Fonts and Colors, Events, Listeners, Components, Containers, Working with Layouts, Event Classes, Event Listener Interfaces, Adapter and Inner Classes

Unit 5
BioJava - Installing BioJava, Symbols, Basic Sequence Manipulation (DNA to RNA, Reverse Complement, motif as regular expression), Translation (DNA to Protein, Codon to amino acid, Six frame translation), Proteomics (Calculate the mass and pI of a peptide), Sequence I/O (File Formats conversions), Locations and Features (PointLocation, RangeLocation, Feature modifications), BLAST and FASTA (Blast and FastA Parser, extract information from parsed results), Counts and Distributions, Weight Matrices and Dynamic Programming, User Interfaces.

Text Books:

Reference Books:
2. Object Oriented Design and Applications (2nd Ed.) by Benjamin, Cummings and Booch, G., Addison Wesley Publishers. 1994
BINF 455 - DATABASE MANAGEMENT SYSTEM

Total Credits: 3                         Total: 36 Hrs.

Unit 1                                    7 Lectures
Introduction –, Database System Versus File Systems, Characteristics of Database, Database Concepts, Schemas & Instances, DBMS architecture and Data Independence, Data Models, Database Languages & Interfaces, View of Data, Database users and Administrators, Database System Structure, Database System Applications

Unit 2                                    7 Lectures
Data models – ER Model: Keys, Constraints, Design Issues, Extended ER features, Reductions of ER Schema to Tables. Relational Model: Structure, Relational Algebra; Hierarchical Model, Network Model, Object Oriented Model

Unit 3                                    6 Lectures
Structured Query Language – Basic Structure, Set Operations, Aggregate Functions, Null Values, Nested Sub queries, Views, Integrity: Domain constraints, Joined Relations, Data-Definition Language

Unit 4                                    8 Lectures
Relational Database and Storage – Pitfalls in Relational Design Database, Functional dependencies, Decomposition Normal Forms – 1NF, 2NF, 3NF & Boyce-Codd NF, Data Storage – Ordered indices, Hashing concepts - Security and Authorization.

Unit 5                                    8 Lectures

Text Books:

Reference books:
BINF 456 - FUNDAMENTALS OF ALGORITHMS

Total Credits: 2

Unit 1
Computing Algorithms
4 lectures

Unit 2
Sorting, Searching & Strings Matching
5 lectures

Unit 3
Graphs
5 lectures
Representation of Graphs, Breadth First Search, Depth First Search, Topological Sort, Connected Components, Minimum Spanning Tree, Single-Source Shortest Path: Dijkstra’s Algorithm, All-Pairs Shortest Paths, Coloring of Graphs

Unit 4
Trees
5 lectures
Forests, DAGs, Ancestors, and Descendants, Binary Search Trees, Querying a Binary search tree, Insertion and Deletion, Tree Traversals, AVL-Trees, Rotations, Insertion, Deletion, B-trees.

Unit 5
Algorithm Design and Analysis
5 lectures

Text Books:

Reference Books:
BINF 457 – Microscopic Techniques For Image Processing

Total Credits: 2
Total: 24 Hrs.

Unit 1
Transmission electron microscopy:
Wave nature of electrons – Electromagnetic lenses – Basic components of Transmission Electron Microscope – Alignment of TEM – Major operational modes of TEM.

Unit 2
Scanning electron microscopy:
Basic systems of the SEM – Contrast and three-dimensionality of the SEM image – Stereo imaging with the SEM

Unit 3
Specimen preparation for EM:

Unit 4
Image processing and image analysis by computer
Capturing the image – Conventional vs. digital – Image processing – Controlling contrast, brightness and gamma – Removing noise – Fast Fourier Transform – images for publication and presentation – Three dimensional imaging.

Unit 5
Atomic Force microscopy and Confocal Microscopy
Atomic force microscopy (AFM) including contact-mode, tapping-mode and lateral-force AFM
Confocal Microscopy: Basics of Confocal Microscopy, Sample Preparation, Confocal Optics, Resolution.

Text Book:

Reference Books:
BINF 475 - LAB - PROGRAMMING IN JAVA

Total Credits: 1

Exercise in JAVA

1. Working with Objects, Arrays, Conditionals and Loops.
2. Creating Classes and Applications in Java.
3. Java Exception handling
4. Streams and I/O, Using Native Methods and Libraries
5. Simple Animation and Threads, Advanced Animation, Images and Sound.
7. Local and global alignment of sequences
8. Creating User Interfaces with AWT, Modifiers.
9. Multithreading example
Total Credits: 1

Exercise in DBMS (MYSQL)

Data Definition Language (DDL) statements:
Creating database, Selecting database, Deleting database, Creating table, Modifying Table, Deleting table

Data Manipulation statements:
Inserting, updating and deleting records
Retrieving Records
Retrieving specific rows and columns
Use of MySQL operators – Arithmetic operators, Comparison
Operators, Logical operators
Math functions, Aggregate functions
String operations
Limiting, Sorting and grouping query results
Handling null values
Renaming or aliasing table and column names
Using subqueries
Using Joins – joining a table to itself, joining multiple tables
Use of Indexes
Security Management
Granting and Revoking rights on tables
Total Credits: 1

Exercises:

1. Sequence Analysis Packages: EMBOSS, NCBI ToolKit, SMS
2. Database search engines: Entrez, SRS, DBGET
3. Pairwise alignment:
   a. Search tools against Databases:
      i. BLAST
      ii. FASTA
4. Multiple sequence alignment:
   a. Clustal
   b. Dialign
   c. Multalign
5. Sequence patterns and profiles:
   a. generation of sequence profiles
      i. PSI-BLAST
   b. derivation of and searching sequence patterns:
      i. MeMe,
      ii. PHI-BLAST
      iii. SCanProsite
      iv. PRATT
6. Protein motif and domain analysis:
   a. MEME/MAST
   b. eMotif
   c. InterproScan
   d. ProSite
   e. ProDom
   f. Pfam
7. Phylogenetic analysis – MEGA, PAUP, PHYLIP
BINF 521 - STRUCTURAL BIOLOGY

Total Credits: 3

Unit 1

Unit 2

Unit 3
Phase Problem, What is phase problem, How to solve the phase problem, Patterson function, Direct methods, Isomorphism replacement method, heavy atom method. Nuclear Magnetic Resonance: Chemical Shift, Coupling constant, spin-spin relaxation, spin-lattice relaxation, COSY, NOESY and NOE.

Unit 4
Structure Prediction Strategies: Secondary structure prediction: Algorithms viz. Chou Fasman, GOR methods; analysis of results and measuring the accuracy of predictions using Q3, Segment overlap, Mathew’s correlation coefficient Identification/assignment of secondary structural elements from the knowledge of 3-D structure of macromolecule using DSSP and STRIDE methods

Unit 5
Classification and comparison of protein 3D structures: Purpose of 3-D structure comparison and concepts; Algorithms such as FSSP, CE, VAST and DALI, Fold Classes. Databases of structure-based classification: CATH and SCOP. Structures of oligomeric proteins and study of interaction interfaces

Text Books:
1. Molecular Modeling Principles and Applications (2nd Ed.) by Andrew R. Leach., Prentice Hall, USA. 2001
BINF 522 – MOLECULAR MODELING AND DRUG DESIGN

Total Credits: 3                  Total: 36 Hrs.

Unit 1                        8 Lectures
**Molecular Mechanics:** Introduction, The Morse Potential, The Harmonic Oscillator Model for Molecules, Comparison of Morse and Harmonic Potential, Two atoms connected by a bond, Polyatomic Molecules, Energy due to Stretch, Bend, Stretch-Bend, Torsional strain, van der Waals and Dipole-Dipole interactions. Types of Potentials: Lennard-Jones, Truncated Lennard-jones, Exponential-6, Ionic and Polar potentials. Types of Force Fields: AMBER, CHARMM, Merck Molecular Force Field, Consistent Force Field, MM2, MM3 and MM4 force fields.

Unit 2                        5 Lectures
**Potential Energy Surface:** Convergence Criteria, Characterizing Stationary Points, Search for Transition States. Optimization:- multivariable Optimization Algorithms, level Sets, Level Curves, Gradients, Optimization Criteria, Unidirectional Search, Finding Minimum Point, Gradient based Methods-Steepest Descent and Conjugate Gradient Methods

Unit 3                        8 Lectures
**Molecular Dynamics Simulation:** Introduction, Radial distribution functions, Pair Correlation function, Newtonian dynamics, Integrators- Leapfrog and Verlet algorithm, Potential truncation and shifted-force potentials, Implicit and explicit Solvation models, Periodic boundary conditions, Temperature and pressure control in molecular dynamics simulations

Unit 4                        8 Lectures
**Drug design:** Drug discovery process. Target identification and validation, lead optimization and validation. Methods and Tools in Computer-aided molecular Design, Analog Based drug design:- Pharmacophores (3D database searching, conformation searches, deriving and using 3D Pharmacophore, constrained systematic search, Genetic Algorithm, clique detection techniques, maximum likelihood method) and QSAR. Structure based drug design:- Docking, De Novo Drug Design (Fragment Placements, Connection Methods, Sequential Grow), Virtual screening.

Unit 5                        7 Lectures
**Structure Activity Relationship:** Introduction to QSAR, QSPR, Various Descriptors used in QSARs: Electronics; Topology; Quantum Chemical based Descriptors. Regression Analysis, The Significance and Validity of QSAR Regression Equations, Partial Least Squares (PLS) Analysis, Multi Linear Regression Analysis. Use of Genetic Algorithms, Neural Networks and Principle Components Analysis in the QSAR equations.

Text Books:

Reference:
BINF 523 - PROGRAMMING IN PERL

Total Credit: 3           Total hrs: 36

Unit 1  7 lectures

Unit 2  7 lectures
Modular Programming: Subroutines, Advantage of Subroutines, Scoping and Subroutines, Arguments, Passing Data to Subroutines, Modules and Libraries of Subroutines, Concept about File handle, Opening and Closing a File handle, Opening and Closing a Directory Handle, Reading a Directory Handle, File and Directory Manipulation.

Unit 3  7 lectures
Regular Expression and Pattern Matching: Concepts about Regular Expressions, Simple uses of Regular Expressions, Patterns, Matching Operator, Substitutions, Split and Join functions.

Unit 4  7 lectures
Common Gateway Interface (CGI) Programming: The CGI.pm Module, CGI program in Context, Simple CGI programs, Passing Parameters via CGI, Perl and the Web

Unit 5  8 lectures
Bioperl: Introduction to Bioperl, Installing procedures, Architectures, General Bioperl Classes, Sequences (Bio::Seq Class, Sequence Manipulation), Features and Location Classes (Extracting CDS), Alignments (AlignIO), Analysis (Blast, Genscan), Databases (Database Classes, Accessing a local database), Implementing REBASE

Text Books:

BINF 524 - SYSTEMS BIOLOGY

Total Credits: 3  
Total: 36 Hrs.


Unit 3: slow and auto–regulation The coherent FFL- temporal order, FIFO, DOR, Global, Development, memory and irreversibility- signaling networks and neuron circuits-robust adaptation–any model. Robustness and optimality in Biology: model and integral feedback-signaling/bifunctional enzymes. Perfect robustness- Role and its measurement. Linking models and measurement, concepts, calibration and identification, data Vs metadata

Unit 4: Introduction- databases KEGG, EMP, MetaCyc, AraCyc etc., Expression databases and various databases related to systems biology, introduction to iGEM. Optional design of gene circuits I- cost and benefit: gene circuits II- selection of regulation. Stochasticity in gene expression.


Text Books:
2. Synthetic Biology, A New Paradigm for Biological Discovery, a report by Beachhead Consulting. 2006

Reference Books:
3. Systems Biology and Synthetic Biology by Pengcheng Fu, Sven Panke., Wiley InterScience .2009
UNIT 1

UV-Visible Spectroscopy

UNIT 2

Infrared Spectroscopy

UNIT 3

Raman Spectroscopy
- Raleigh and Raman scattering - stoke's and anti stokes lines - instrumentation - block diagram - differences between IR and Raman spectroscopy - mutual exclusion principle - applications - structural diagnosis.

UNIT 4

Magnetic Resonance Spectroscopy: Nuclear Magnetic Resonance Spectroscopy
- Nuclear spin magnetic moment, Interaction of nuclear magnet with external magnetic field, NMR spectrometer, relaxation and dynamic processes, chemical shift, Heteronuclear NMR experiments.
- Electron Spin Resonance Spectroscopy: Electron spin and Magnetic moment, Resonance condition in ESR and significance of 'g' value, applications of ESR.

UNIT 5

X-ray Spectroscopy

Text books:
2. Introduction to molecular spectroscopy by G. M. Barrow., McGraw-Hill.1962

Reference Books:
1. Molecular spectroscopy by I. N. Levins, Wiley Interscience. 1975
BINF 526 – DATA COMMUNICATION AND NETWORKS

Total Credits: 2  
Total: 24 Hrs.

Unit 1  
**Nuts & Bolts in Networks**
Reference Model, Network Topologies and Protocols, Types of Networks: Local Area Network (LAN), Wide Area Network (WAN), Metropolitan Area Network (MAN), Network Security (Firewall, Packet Filtering, VPN), Uses of Computer Networks

Unit 2  
**Network Architecture**
OSI & Internet Architecture, IEEE 802 standards, Physical Layer - Transmission Media, Switching, Data Link Layer - Design Issues, Example Data Link Protocols, Data Link layer in the Internet, **Media Access Sub layer:** Static and Dynamic channel allocation – ALOHA – CSMA – CSMA / CD.

Unit 3  
**Network Layer**

Unit 4  
**Transport Layer**

Unit 5  
**Application Layer**
Design Issues, Conventional Encryption, Classical and Modern Techniques, Encryption and Decryption Algorithms (RSA), Confidentiality, DNS, SNMP, RMON, WWW, E-mail, Digital Signatures

**Text Book:**
1. Computer Networks (3rd Ed.) by Tananbaum A.S., PHI. 1999

**Reference Books:**
BINF 527 – SCIENTIFIC PRESENTATION AND FINISHING SCHOOL

Total Credits: 3

Unit 1
Scientific writing – Introduction- Types of scientific writings- Thesis or dissertation- Popular science writing- Types of publications- Open Access Resources- Choosing a journal- Instructions to authors - Scientific paper writing – Structure and Style- Authorships –figures tables with legends - References and citations - Acknowledgements- Conflict of interest - Peer review mechanism- Scientometric Analyses of a paper/ journal- Plagiarism issues

Unit 2

Unit 3
Oral presentation- Planning the oral presentations and visuals- In-class discussion (Students in small groups or individually will take up the assignments or select a research project/ topic and prepare oral presentations followed by a Q&A sessions)

Unit 4
Poster Presentation- Elements and Significance of poster presentations- Planning and designing a poster- Individual Poster presentation (Students select a research project/topic and prepare posters followed by a Q&A sessions)

Unit 5
Personality development, Ethics and values - elements of recruitment, selection, interview techniques- Personality Development - team work- inter personal and intra-personal interactions – human behavior at work- Time and human resources management- planning and scheduling, stress at work- work-life balance- Culture and cultural ethos- cultural diversity-diversity in organizations.

Text Books:

References:

Total: 36 Hrs.
Total Credits: 1

1. Advanced Visualization Software and 3D representations.

2. Small Molecule Structure determination
   a) Structure Solution: SHELXS
   b) Structure Refinement: SHELXL

3. Thermal Ellipsoid Plot:
   a) ORTEP

4. Structure analysis
   a) PARST
   b) Platon
   c) Mercury

5. Protein Structure Determination:
   a) Exploration of CCP4 website
   b) Protein Model building: COOT
   c) Structure Solution: AMoRe

6. Structure Validation
   Procheck, WHATIF, VERIFY 3D
**BINF 572 - LAB – MOLECULAR MODELING AND DRUG DESIGN**

**Total Credits:** 1

**Exercises**

1. Molecular Visualization Softwares: Pymol and Rasmol
2. Geometry Optimization
3. Tutorial on Molecular Dynamics: Gromacs
4. Binding Site Identification
5. Structure based Drug Design:- Molecular Docking
6. Ligand based Drug Design:- QSAR
BINF 573 - LAB - PROGRAMMING IN PERL

Total Credits: 1

1. Uses of Scalar and Array Variables to manipulate DNA/RNA/Protein sequence data
2. Concatenation DNA fragments, Transcribing DNA into RNA
3. Calculating the Reverse complement of a DNA strand
4. Uses of common Array Operators
5. Uses of Do-Until Loops
6. Uses of ‘substr’ function to look into the string
7. Reading a sequence data from a file and writing the results to a file
8. Opening and closing a Directory Handle, Reading a Directory and other directory manipulation functions.
9. Uses of Subroutines
10. Uses of Hashes for the genetic code: translating codons into amino acids
11. Uses of subroutine to read FASTA files
12. Translate a DNA sequence in all six reading frames
13. Uses of Regular Expressions
14. Extract annotation and sequence from GenBank file
15. Parsing GenBank annotation using arrays
16. Extract sequence chains from PDB file
17. Uses of CGI.pm Module and Passing Parameters via CGI, Debugging CGI programs
18. Installing Bioperl, Uses of Bioperl modules for sequence manipulation, accessing local database
BINF 551- ANALYTICAL TECHNIQUES

Total Credits: 3

Unit 1 7 lectures
Electrophoresis: Theory and types; moving boundary electrophoresis, zone electrophoresis, paper, cellulose acetate gel electrophoresis, Native PAGE, disc PAGE, Gradient PAGE, SDS PAGE, DNA agarose gel electrophoresis, Southern, Northern, Western blotting techniques, Isoelectric focusing, finger printing, DNA sequencing, Pulsed - field Electrophoresis, Capillary Electrophoresis.

Unit 2 7 lectures
Chromatography: Principles, methodology and applications of chromatography using paper, thin layer, column (gel filtration, ion exchange, affinity), gas and types of HPLC.

Unit 3 6 lectures
Centrifugation: Principles, types and applications. Ultracentrifugation- types, optical methods used and applications of preparative and analytical ultracentrifuges.

Unit 4 8 lectures
Enzyme kinetics: Membrane potential, Active site, Cofactors, apo-enzymes, Enzyme specificity, Factor affecting enzyme activity, Michaelis-Menten equation, LB Plot, Determination of Km, Types of inhibition, Allosteric enzymes.

Unit 5 8 lectures
Various Biophysical techniques to study interactions caused by the macromolecules: Isothermal Titration Calorimetry- instrumentation. Protocol and application in the study of Protein-ligand interactions. Optical and magnetic tweezers- principle, Instrumentation and modern developments. Role of Optical tweezers in studying the molecular motors and the properties of DNA. Fluorescence Resonance Energy Transfer: Principle, Instrumentation, Protocol and application in analyzing the macromolecular interactions. Dual Polarisation Interferometry [DPI]- Principle, Instrumentation and protocol. Application of DPI in one dimensional determination of protein structures and in studying the interactions and activity of biomolecules. ORD and CD, DLS

Text Books:

Reference Books:
1. Physical Biochemistry (2nd Ed) by D. Freifelder., Freeman. 1982
BINF 552 – BIOETHICS, BIODIVERSITY AND INTELLECTUAL PROPERTY RIGHTS

Total Credits: 3
Total: 36 Hrs.

Unit 1
8 Lectures
Regulatory Procedures: Good laboratory practice, Good manufacturing practice and FDA regulations - Regulations for recombinant DNA research and manufacturing process - Bio-safety and Bioethics - Regulations for clinical trials, Documentation and Compliance, in India and selected countries - Rules for import and export of biological materials.

Unit 2
8 Lectures
Biotechnology Processes and Products: Techniques used in Biotechnology, with special emphasis on molecular and recombinant DNA techniques - Cloning Strategies and Tissue culture procedures for plant cells, animal and stem cells - Transgenic plants, animals, genetically modified organisms (GMO) and GM food etc. - Large scale production of recombinant proteins, Processes for separation and purification - Medical Biotechnology: gene therapy, tissue engineering and xenotransplantations - Biotechnology Products: Health care products – Vaccines – Diagnostics - Recombinant therapeutic proteins - Agricultural : Hybrid and modified seeds - Bio-pesticides - Bio-fertilizers

Unit 3
6 Lectures
IPR - Definition - Forms of IPR Protection, WTO - Definition — Functions- International treaties for IPR Protection

Unit 4
7 Lectures
Patents - Definition - conditions for patentablity - test of novelty of patents – composition of a patent - Patenting of Biotechnological discoveries

Unit 5
7 Lectures
Other forms of IPR protection: Copyright - Trademark - Designs - Importance in Indian Scenario & laws in India for IPR protection.

Text Books:
2. Intellectual Property Rights by Deborah E. Bouchoux., Delmar Cenage Learning. 2005

Reference Books:
1. The Indian Environmental Protection Act (EPA), 1986
2. Rules for manufacture, use/import/export and storage of hazardous microorganisms or cells Act, 1989
3. Food Safety and Standards act (Government of India), 2006
4. Intellectual Property Rights on Biotechnology by Singh, KC, BCIL, New Delhi
BINF 553 – R Language and Bioconductor

Total Credits: 3                              Total: 36 Hrs.

Unit 1                                                                                                                   6 Lectures
Overview of the R language: Defining the R project, Obtaining R, Generating R codes, Scripts, Text editors for R, Graphical User Interfaces (GUIs) for R, Packages.

Unit 2                                                                                                                  7 Lectures
R Objects and data structures: Variable classes, Vectors and matrices, Data frames and lists, Data sets included in R packages, Summarizing and exploring data, Reading data from external files, Storing data to external files, Creating and storing R workspaces.

Unit 3                                                                                                                   8 Lectures
Manipulating objects in R: Mathematical operations (recycling rules, propagation of names, dimensional attributes, NA handling), Basic matrix computation (element-wise multiplication, matrix multiplication, outer product, transpose, eigenvalues, eigenvectors), Textual operations, Basic graphics (high-level plotting, low-level plotting, interacting with graphics).

Unit 4                                                                                                                   7 Lectures
Hypothesis testing and data handling: Parametric and nonparametric tests, Chi-square test, t-tests, ANOVA, Correlation and regression, Principal component Analysis

Unit 5                                                                                                                   8 Lectures
Bioconductor: Introduction, Bioconductor packages, ExpressionSet Class, Data annotation, biomaRt, Network analysis.

Text Books:

Reference books:
BINF 554 - PROJECT

Total Credits: 8

The course is designed to result in the satisfactory completion and defense of the Masters dissertation.

This process includes

a) the conceptualization of the independent research that will comprise the dissertation,
b) the preparation of and satisfactory defense of the dissertation proposal,
c) the collection, analysis, and interpretation of data,
d) presentation of findings in the dissertation format, and
e) oral defense of the dissertation.

Dissertation activity must be completed within prescribed time frame for the semester.