**THIRUVALLUVAR UNIVERSITY**

**M.Sc. BIOINFORMATICS**

**SYLLABUS**

**UNDER CBCS**

**(with effect from 2020-2021)**

**The Course of Study and the Scheme of Examination**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Sl.***  ***No.*** | ***Study Components*** | | ***ins. hrs / week*** | | | ***Credit*** | | ***Title of the Paper*** | | ***Maximum Marks*** | | |
| ***Course Title*** | | ***CIA*** | ***Uni. Exam*** | ***Total*** |
| **SEMESTER I** | | | | | | | |  | |
|  | Core | **Paper-1** | 5 | | | 4 | | Basic Bioinformatics | | 25 | 75 | 100 |
|  | Core | **Paper-2** | 5 | | | 4 | | Cell and Molecular Biology | | 25 | 75 | 100 |
|  | Core | **Paper-3** | 5 | | | 4 | | Programming in C & C++ | | 25 | 75 | 100 |
|  | Practical | **Paper-1** | 3 | | | 0 | | Biochemistry, Biophysics and Molecular Biology (LAB) | | - | - | - |
|  | Practical | **Paper-2** | 3 | | | 0 | | Sequence Analysis (LAB) | | - | - | - |
|  | Practical | **Paper-3** | 3 | | | 0 | | Programming in C & C++ (LAB) | | - | - | - |
| Internal Elective for same major students | | | | | | | | | | | | |
|  | **Core Elective** | **Paper-1** | 3 | | | 3 | | **(to choose one out of 3)**  A. Mathematics & Statistical methods in Bioinformatics  B. Biodiversity Informatics  C. Genetics & Evolution | | 25 | 75 | 100 |
| External Elective for other major students (Inter/multi disciplinary papers) | | | | | | | | | | | | |
|  | **@ Open Elective** | **Paper-1** | 3 | | | 3 | | Biological Databases | | 25 | 75 | 100 |
|  |  |  | **30** | | | **18** | |  | | **125** | **375** | **500** |
|  | | | | | | | | | | | | |
| **SEMESTER II** | | | | | | | |  | | ***CIA*** | ***Uni. Exam*** | ***Total*** |
|  | Core | **Paper-4** | 5 | | | 4 | | Genomics & Proteomics | | 25 | 75 | 100 |
|  | Core | **Paper-5** | 4 | | | 4 | | Relational DBMS & MySQL | | 25 | 75 | 100 |
|  | Core | **Paper-6** | 4 | | | 4 | | Structural Biology | | 25 | 75 | 100 |
|  | Practical | **Paper-1** | 3 | | | 3 | | Biochemistry, Biophysics and Molecular Biology (LAB) | | 25 | 75 | 100 |
|  | Practical | **Paper-2** | 3 | | | 3 | | Sequence Analysis (LAB) | | 25 | 75 | 100 |
|  | Practical | **Paper-3** | 3 | | | 3 | | Programming in C & C++ (LAB) | | 25 | 75 | 100 |
| Internal Elective for same major students | | | | | | | | | | | | |
|  | Core Elective | Paper-2 | 3 | | | 3 | | **(to choose one out of 3)**  A. Biophysics & Biochemistry  B. Biological Algorithms in Computational Biology  C. Cheminformatics | | 25 | 75 | 100 |
| External Elective for other major students (Inter/multi disciplinary papers) | | | | | | | | | | | | |
|  | Open Elective | Paper-2 | 3 | | | 3 | | | Biological Sequence Analysis | 25 | 75 | 100 |
|  | \*Field Study |  | - | | | 2 | | |  | 100 | - | 100 |
|  | Compulsory Paper | | 2 | | | 2 | | | **Human Rights** | 25 | 75 | 100 |
|  |  |  | **30** | | | **22** | | |  | **325** | **675** | **1000** |
| **SEMESTER III** | | | | | | | | |  |  |  |  |
|  | Core | Paper-7 | | | 6 | | 4 | | Molecular Modeling & Drug Design | 25 | 75 | 100 |
|  | Core | Paper-8 | | | 6 | | 4 | | Systems Biology | 25 | 75 | 100 |
|  | Core | Paper-9 | | | 6 | | 4 | | Advanced Programming in Bioinformatics | 25 | 75 | 100 |
|  | Practical | Paper-4 | | | 0 | | 0 | | Molecular Modeling (LAB) |  |  |  |
|  | Practical | Paper-5 | | | 0 | | 0 | | Perl & Python (LAB) |  |  |  |
| Internal Elective for same major students | | | | | | | | | | | | |
|  | Core Elective | Paper-3 | | | 6 | | 3 | **(to choose one out of 3)**  A. Data warehouse & Data mining  B. Genetic Engineering  C. R Programming | | 25 | 75 | 100 |
| External Elective for other major students (Inter/multi disciplinary papers) | | | | | | | | | | | | |
|  | Open Elective | Paper-3 | | | 6 | | 3 | Introduction to Drug Design & Discovery | | 25 | 75 | 100 |
|  | \*\*MOOC Courses |  | | | - | | - |  | | - | - | 100 |
|  |  |  | | | **30** | | **18** |  | | **125** | **375** | **600** |
| **SEMESTER IV** | | | | | | | |  | | ***CIA*** | ***Uni. Exam*** | ***Total*** |
|  | Core | Paper-10 | | | 5 | 5 | | Research Methodology | | 25 | 75 | 100 |
|  | Practical | Paper-4 | | | 4 | 3 | | Molecular Modeling (LAB) | | 25 | 75 | 100 |
|  | Practical | Paper-5 | | | 4 | 3 | | Perl & Python (LAB) | | 25 | 75 | 100 |
|  | Core | Project | | 5 | | 5 | | Project with viva voce (Compulsory) | | 100  (75 Project +25 viva) | | 100 |
| Internal Elective for same major students | | | | | | | | | | | | |
|  | Core Elective | Paper-4 | | | 6 | 3 | | **(to choose one out of 3)**  A. Bioethic, Biosafety & IPR  B. Medical Biotechnology  C. Big Data Analytics & NGS | | 25 | 75 | 100 |
| External Elective for other major students (Inter/multi disciplinary papers) | | | | | | | | | | | | |
|  | Open Elective | Paper-4 | | | 6 | 3 | | Fundamental of Algorithms | | 25 | 75 | 100 |
|  |  |  | | | **30** | **21** | |  | | **125** | **375** | **600** |
|  |  |  | | | **120** | **90** | |  | |  |  | **2700** |

**\* Field Study**

There will be field study which is compulsory in the first semester of all PG courses with 2 credits. This field study should be related to the subject concerned with social impact. Field and Topic should be registred by the students in the first semester of their study along with the name of a mentor before the end of the month of August. The report with problem identification and proposed solution should be written in not less than 25 pages in a standard format and it should be submitted at the end of second semester. The period for undergoing the field study is 30 hours beyond the instructional hours of the respective programme. Students shall consult their mentors within campus and experts outside the campus for selecting the field and topic of the field study. The following members may be nominated for confirming the topic and evaluating the field study report.

(i). Head of the respective department

(ii). Mentor

(iii). One faculty from other department

\*\***Mooc Courses**

Inclusion of the Massive Open Online Courses (MOOCs) with zero credits available on SWAYAM, NPTEL and other such portals approved by the University Authorities.

**THIRUVALLUVAR UNIVERSITY**

**M.Sc. BIOINFORMATICS**

**SYLLABUS**

**UNDER CBCS**

**(with effect from 2020-2021)**

**SEMESTER III**

**PAPER - 7**

**MOLECULAR MODELING AND DRUG DESIGNING**

**Objective**

To familiarize the students in using computer techniques for molecular modeling and drug designing.

**UNIT-I**

Introduction to drug action. Physico-chemical properties and drug action. Pharmacological approaches of modern medicine. Historical approaches in drug discovery. Pro drug design. New approaches in drug discovery: combinatorial chemistry, high throughput screening, ultra high throughput screening and high content screening, technologies for high throughput screening, pharmacogenomics, proteomics and array technology.

**UNIT-II**

Quantitative analysis of Structure Activity Relationships (QSAR): Introduction, Parameters, quantitative models, design of test series in QSAR, Applications of Hansch Analysis, Applications of Free Wilson Analysis and related models, 3D QSAR approaches-COMFA (Comparative Molecular Field Analysis).

**UNIT-III**

Molecular Modeling in Drug Design: Introduction, Background and methods-molecular mechanics, quantum mechanics, Known receptors - Definition of site, site characterization, Design of ligands, affinity calculation, multiple binding modes, homology modeling, Unknown receptors - Pharmacophore versus binding site models, searching for similarity, molecular comparisons, field effects, volume mapping, Directionality, Locus maps. Vector maps, Conformational mimicry. Finding the common pattern - Constrained minimization, systematic search and active analog approach, alternative approach, receptor mapping, model receptor sites.

**UNIT-IV**

Natural products as Leads for New pharmaceuticals: Introduction. Drugs affecting the Central Nervous System - Morphine alkaloids, Cannabinoids, Asperlicine. Neuromuscular blocking drugs such as Curare, decamethonium, atracurium. Anticancer Drugs-*Catharanthus* (vinca alkaloids), Taxol and Taxotere, Podophyllotoxin, Etoposide and teniposide. Antibiotics - Beta lactams, Erythromycin macrolides, Echinocardins. Cardiovascular drugs - Lovastatin, Simvastatin, Pravastatin, Teproside, Captopril, Dicoumarol, Warfarin. Antiparasitic drugs - quinine, chloroquin, mefloquine, Artemisinin, artemether and arteether.

**UNIT-V**

Intellectual Property in Drug Discovery and Biotechnology: Patent protection and strategy - patent strategy, first to invent versus first to file, absolute novelty. Requirements for patents - patentable subject matter in the United States, patentable subject matter outside the United States. Patent specifications - written description, enablement, best made claims, procedure for obtaining patents in US, interference proceedings, correlation of patents. Worldwide patent protection - International agreements, PCT patent practice, other aspects of patent laws in other countries. Trademarks-trademarks as marketing tools, selection of trademarks, registration process, worldwide trade mark rights. Trade secrets-definition, requirements for protection of trade secrets, enforcement of relationship of trade secrets and patents, freedom of information act, trade secret protection outside the United States.

**Text Books**

1. Bohlin, L. and Bruhn, J.G. 1999. Bioassay methods in Natural Product research and Drug Development. Kluwer Academic Publishers, Netherlands.
2. Busse, W. D. and Ganellin, C. R. 1993. Views from Industry on the Medicinal Chemistry Curriculum: Answers to a Questionnaire. In Trends in Drug Research, (Ed.) V. Claassen, Pharmacochemistry Library, 20, Elsevier, Amsterdam.
3. Kulkarni V.M.1995. Drug Design. Nirali Prakashan, New Delhi.
4. Lawrence, D.R. and Bacharach, A.L. 1980. Evaluation of Drug activities: Pharmacometrics Vol. 1, 5th Edn. Academic Press, New York.
5. Vogel, G. H. 2002. Drug Discovery and Evaluation: Pharmacological Assays, 2nd Edn. Springler-Verlag, Berlin Heidelberg, Germany.
6. Wolff, M.E. 1995. Burgers’s Medical Chemistry and drug discovery, Vol. 1: principles and practice, 5th Edn. John Wiley & Sons, New York.

**PAPER - 8**

**SYSTEMS BIOLOGY**

**Objective**

The Objective of this subject is to understand the basics of systems biology and associated network approaches, databases and software tools.

**UNIT-I**

Introduction to Systems Biology: What is Systems Biology? Integrating Networks approaches, Dynamic Analysis, Organization of living cells, Components vs. Systems, Links and functional states, Links to Networks.

**UNIT-II**

Biochemical Reaction Kinetics – Rate equation approach, Biochemical Reaction Modeling, Basics principles and assumptions, elementary reactions, complex reaction, Michaelis-Menten equation for EK, Stochastic Modelling and Simulation, Modelling of Cell communication networks

**UNIT-III**

Reconstruction of Biochemical Networks : Basic features, Reconstruction methods, Organism specific source of information, Strategies relating to In silico Modeling of biological processes, Metabolic Networks, Regulation of metabolic networks, Signaling Networks.

**UNIT-IV**

Database and Software for Systems Biology: KEGG, EMP, MetaCyc, Gene expression and microarray databases related to systems biology, Cytoscape and Cell Designer.

**UNIT-V**

Software for Modeling and Simulation. E‐CELL, V‐CELL and GROMOS.

**REFERENCES**

1. Foundation of Systems Biology – Hi Roaki Kitano

2. Introduction to Systems Biology – Sangdun Choi

**PAPER - 9**

**ADVANCED PROGRAMMING IN BIOINFORMATICS**

**Objective**

To enable the students to acquire advanced training in Java, BioJava, Perl, Bioperl, Python

**UNIT-I**

JAVA - Introduction to object oriented programming - Basic Syntax - Control Structures - Arrays - Strings - Files and Streams - Applets; Introduction to JDBC (Java Database Conectivity) - JDBC Architecture, JDBC Drivers, Connecting to Database and accessing databases - Threads - Java beans.

**UNIT-II**

BIOJAVA: Introduction - Sequence Manipulation. Translation: DNA to Protein, Codon to amino acid, Six frame translation. Proteomics: mass and pI of a peptide - Sequence File Format conversions, Locations and Features. BLAST and FASTA parsing, Weight Matrices and Dynamic Programming

**UNIT-III**

PERL: Modules:- defining, storing and using modules; Data and control structures:- Operators and Control Flow, Data Types & structure, Function/Subroutines - File handling; Regular Expressions:- patterns - grouping and anchoring - string matching; Object oriented programming:- Classes, Objects, Methods.

BIOPERL: General Bioperl Classes, Sequences, Sequence Manipulation, Features and Location Classes. Alignments: AlignIO. Analysis: Blast, Genscan; Databases: Database Classes, Accessing a local database.

**UNIT-IV**

Common Gateway Interface:- HTML form elements, GET, POST & HEAD Method, CGI Environment Variables, Handling forms, Passing Parameters via CGI, Debugging CGI programs.

WEB DESIGNING & XML: HTML Specifications and Syntax, XML Basics, Style Sheets, XML Applications, Java Script.

**UNIT-V**

PYTHON: Overview, Data structures, Control Flow, Modules, Basic I/O, Regular Expressions, File Manipulation, Classes, Standard library.

**Text Books**

1. Herbertz Schildt, The complete Reference Java J2SE 5 Edition, Mc Graw Hill, Osborne, 2005.
2. E Balaguruswamy, Programming with Java, Tata Mc Graw Hill, New Delhi, 1999.
3. Larry Wall, Tom Christiansen & John Orwant, Programming Perl –3rd ed, O’Reilly, 2000.
4. James D. Tisdall, Beginning Perl for Bioinformatics, O'Reilly, 2001
5. Mark Lutz , Programming Python - 2nd Ed., O’ Reilly, 2003.
6. Eric Ladd, J.O’Donnell, Using HTML 4, XML and JAVA, Prentice Hall of India QUE, 1999.
7. Brown, The complete reference - Perl, Tata Mc Graw Hill, New Delhi, 2004.

**CORE PRACTICAL - 4**

**MOLECULAR MODELING**

1. Simple Genetic algorithm
2. Genetic algorithm and problem solving
3. Genetic algorithm in scientific models
4. Sequence alignment algorithm
5. Structure prediction - HM Model
6. Determination of Partition coefficient of established drugs by shake flask method to find out the distribution of drug.(2 to 3 experiments)
7. QSAR related experiments - calculation of molecular connectivity index values for Ibuprofen, Atropine, Propranalol, Epinephrine and methadone (Minimum of 3 experiments will be given for work out).
8. Molecular modeling experiments- Molecular Graphics (3D structure) - conformational analysis to correlate physicochemical parameters with biological activity using various molecular modeling software such as AMBER, CAMSEQ, FRODO and SYBL (minimum of 5 experiments will be given).
9. Study of Pharmacophore models by using software(2 experiments)
10. Designing a novel molecule and fitting with receptor for invitro activity evaluation - experiments related to this is given (minimum of 3 experiments)
11. Calculation of F and R substituent constants (2 experiments)

**CORE PRACTICAL- 5**

**PERL & PHYTON**

1. Program to store a DNA sequence using Perl
2. Program to concatenate DNA fragments using Perl
3. Program to convert DNA to RNA using Perl
4. Program to calculate reverse compliment of DNA sequence using Perl
5. Program to read protein sequence data from a file using Perl
6. Program to find motifs in a protein sequence using Perl
7. Program to count nucleotides in a sequence using Perl
8. Program to find the percentage of hydrophobic amino acids in a sequence using Perl
9. Program to find the percentage of G and C in a DNA sequence using Perl
10. Program to append ATGC to a DNA sequence using subroutines using Perl
11. [Counting Letters in DNA Strings](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#counting-letters-in-dna-strings) using Phyton
12. [Python Program to find sum of array](https://www.geeksforgeeks.org/python-program-to-find-sum-of-array/)
13. [Python Program to find largest element in an array](https://www.geeksforgeeks.org/python-program-to-find-largest-element-in-an-array/)
14. [Efficiency Assessment](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#efficiency-assessment) using Phyton
15. [Verifying the Implementations](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#verifying-the-implementations) using Phyton
16. [Computing Frequencies](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#computing-frequencies) using Phyton
17. [Analyzing the Frequency Matrix](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#analyzing-the-frequency-matrix) using Phyton
18. [Dot Plots from Pair of DNA Sequences](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#dot-plots-from-pair-of-dna-sequences) using Phyton
19. [Finding Base Frequencies](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#finding-base-frequencies) using Phyton
20. [Translating Genes into Proteins](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#translating-genes-into-proteins) using Phyton
21. [Random Mutations of Genes](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#random-mutations-of-genes) using Phyton

**CORE ELECTIVE**

**PAPER - 3**

**(to choose one out of 3)**

**A. DATA WAREHOUSING AND DATA MINING**

**Objective:** To learn the classification and clustering techniques that help to extract hidden knowledge in a domain.

**UNIT-I**

INTRODUCTION: Relation to Statistics, Databases - Data Mining Functionalities - Steps In Data Mining Process-Architecture Of A Typical Data Mining Systems- Classification Of Data Mining Systems - Overview Of Data Mining Techniques.

**UNIT-II**

DATA PREPROCESSING AND ASSOCIATION RULES: Data Preprocessing -Data Cleaning, Integration, Transformation, Reduction, Discretization Concept Hierarchies-Concept Description: Data Generalization And Summarization Based Characterization- Mining Association Rules In Large Databases.

**UNIT-III**

PREDICTIVE MODELING: Classification And Prediction: Issues Regarding Classification And Prediction - Classification By Decision Tree Induction - Bayesian Classification - Other Classification Methods – Prediction - Clusters Analysis: Types Of Data In Cluster Analysis - Categorization Of Major Clustering Methods: Partitioning Methods -Hierarchical Methods

**UNIT IV**

DATA WAREHOUSING: Data Warehousing Components - Multi Dimensional Data Model - Data Warehouse Architecture - Data Warehouse Implementation - Mapping The Data Warehouse To Multiprocessor Architecture - OLAP -Need - Categorization Of OLAP Tools.

**UNIT-V**

APPLICATIONS: Applications of Data Mining - Social Impacts Of Data Mining - Tools - An Introduction To DB Miner - Case Studies-Mining WWW - Mining Text Database - Mining Spatial Databases.

**Text Books**

1. Jiawei Han, Micheline Kamber, "Data Mining: Concepts and Techniques", Morgan Kaufmann Publishers, 2002.
2. Alex Berson,Stephen J. Smith, “Data Warehousing, Data Mining,& OLAP”, Tata Mcgraw- Hill, 2004.
3. Mohanty - Data warehousing, 2006, TMH, New Delhi.

**References**

1. Usama M.Fayyad, Gregory Piatetsky - Shapiro, Padhrai Smyth And Ramasamy Uthurusamy, "Advances In Knowledge Discovery And Data Mining", The M.I.T Press, 1996.
2. Ralph Kimball, "The Data Warehouse Life Cycle Toolkit", John Wiley & Sons Inc., 1998.
3. Sean Kelly, "Data Warehousing In Action", John Wiley & Sons Inc., 1997.

**CORE ELECTIVE**

**PAPER - 3**

**B. GENETIC ENGINEERING**

**Unit 1:** Role of genes within cells, genetic code, genetic elements that control gene expression, Method of creating recombinant DNA molecules, Types, biology and salient features of vectors in recombinant DNA technology–I: Plasmids, Phages, Cosmids, Fosmids, Phagemids, and Artificial chromosomes, Safety guidelines for recombinant DNA research, Control of spills and mechanism of implementation of biosafety guidelines

**Unit 2:** Enzymes in genetic engineering: Restriction nucleases: exo & endo nucleases, Enzymes in modification- Polynucleotide phosphorylase, DNase and their mechanism of action, Enzymes in modification- Methylases and phosphatases and their mechanism of action, Enzymes in modification- Polynucleotide kinase, Ligases, RNase and their mechanism of action.

**Unit 3:** Methods of nucleic acid detection, Polymerase chain reaction (PCR) and its applications, Variations in PCR and their applications, Methods of nucleic acid hybridization, Probe and target sequences, Nucleic acid mutagenesis in vivo and in vitro

**Unit 4:** Isolation and purification of nucleic acid (genomic/plasmid DNA and RNA), Quantification and storage of nucleic acids, Construction of cDNA library, Construction of Genomic library, Screening and preservation of DNA libraries, DNA Sequencing and cloning strategies.

**Unit 5:** Gene transfer techniques: biological methods, Gene transfer techniques: chemical methods, Gene transfer techniques: physical or mechanical methods,Agrobacterium- mediated gene transfer in plants, Chloroplast transformation Gene therapy: Introduction and Methods, Gene targeting and silencing, Gene therapy in the treatment of diseases, Challenges and future of gene therapy

**Texts & References:**

1) Fundamental Molecular Biology ; Allison LA; 2007

2) Recombinant DNA, Watson et al ; 5th Ed; 2006

3) Techniques for Engineering Genes ; Curell BR et al;2004

4) Techniques for Molecular Biology ; Tagu D & Moussard C; INRA; 2006

5) Gene Cloning and DNA Analysis ; 5th Ed ; Brown TA ; 2006

6) Analysis of Genes and Genomes ; Reece RJ ; Wiley; 2004

7) Recombinant DNA and Biotechnology ; 2nd Ed ; Kreuzer H and Massey A ;ASM;2006

8) Human Genetics and Genomics ; Korf BR ; 3rd Ed ; Blackwell; 2007

9) Molecular Cloning; 3rd Ed; Sambrook & Russel : Cold Spring Harbour Laboratory press, NY ; 2001

10) ICRF Handbook of Genome Analysis ; Spurr NK , Young BD , Bryant SP;1998 MSGEN-302 : Genomics, Proteomics & Bioinformatics Module I : GENOMICS (14 Periods)

**CORE ELECTIVE**

**PAPER - 3**

**C. R PROGRAMMING**

**COURSE OBJECTIVE:** The main goal of this course is to introduce the student with the R environment for biological big data analysis using various statistical methods.

**Unit 1 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 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 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 Hypothesis testing and data handling:** Hypothesis testing, Parametric and nonparametric tests, Chi-square test, t-tests, ANOVA, Correlation and regression, Principal component Analysis

**Unit 5 Big Data Analytics in Bioinformatics using R:** Introduction to Big data: Characteristics, data structures and data repositories; exploratory analysis of big data in R environment, Bioconductor, Microarray and next-generation sequencing (NGS) data analysis in R environment.

**Text Books:**

1. Paul Gerrard and Radia M. Johnson. Mastering Scientific Computing with R. Packt Publishing, UK, 2015.

2. P.P. Sinha. Bioinformatics with R Cookbook. Packt Publishing, UK, 2014.

**Reference Books:**

1. Florian Hahne, Wolfgang Huber, Robert Gentleman, Seth Falcon. Bioconductor case studies. Springer, 2008.

2. Paul D. Lewis, R for Medicine and Biology, Jones and Bartlett Series, 2010.

**COURSE OUTCOME:** The student will have an understanding of various statistical methods employed in biological data analysis. He/she will be able to perform statistical modelling and analysis of microarray and next-generation data in the R environment.

**OPEN ELECTIVE**

**PAPER - 3**

**INTRODUCTION TO DRUG DESIGN AND DISCOVERY**

**Course Objectives:**

* To utilize the available sequences to model the target and use computational tools and software to design a drug.
* To acquire knowledge on the computational software to visualize and compare the protein structure and sequences.
* To analyze the conformational properties of protein using Ramachandran plot.

**SYLLABUS**

**Unit-I:** Introduction to drug discovery History of drug design, Drug properties, likeness; Principles of Protein structure - Helix, Sheet, Strand, Loop and Coil, Torsion angles, Active site, Domains, Fold, Motif, PSSM; Structural databases- PDB, CATH, SCOP; Chemical Databases – ZINC, Pubchem, Chembl.

**Unit-II:** Macromolecular modeling Ab initio modeling; Homology Modeling; Threading; Fold Recognition. Model refinement and validation – Ramachandran Plot, PROCHECK. Prediction of Binding site; ADME prediction; Rasmol viewer.

**Unit-III:** Quantitative Structure Activity Relationship (QSAR) SAR, QSAR, Types of physicochemical parameters, experimental and theoretical approaches for the determination of physicochemical parameters. 3D-QSAR software COMFA.

**Unit-IV:** Molecular docking and Virtual screening Structure-based drug design and Ligand based drug design; Virtual Screening, Pharmacophore design and identification. Molecular docking- AutoDock, Drug-receptor interaction.

**Unit-V:** Molecular mechanics and dynamics General features of molecular mechanics; Energy Minimization - local and global energy minima, applications. Molecular dynamics simulation.

**REFERENCE BOOKS:**

* Molecular Modeling: Basic Principles and Applications, 3rd Edition,Hans-Dieter Höltje, Wolfgang Sippl, Didier Rognan, GerdFolkers
* Andrew R. Leach Molecular Modeling: Principles and Applications.

**SEMESTER IV**

**PAPER - 10**

**RESEARCH METHODOLOGY**

**UNIT 1**

Foundations of Research: Meaning, Objectives, Motivation, Utility. Concept of theory, empiricism, deductive and inductive theory. Characteristics of scientific method – Understanding the language of research – Concept, Construct, Definition, Variable. Research Process

**UNIT 2**

Problem Identification & Formulation – Research Question – Investigation Question – Measurement Issues – Hypothesis – Qualities of a good Hypothesis –Null Hypothesis & Alternative Hypothesis. Hypothesis Testing – Logic & Importance.

**UNIT 3**

Research Design: Concept and Importance in Research – Features of a good research design – Exploratory Research Design – concept, types and uses, Descriptive Research Designs – concept, types and uses. Experimental Design: Concept of Independent & Dependent variables.

**UNIT 4**

Qualitative and Quantitative Research: Qualitative research – Quantitative research – Concept of measurement, causality, generalization, replication. Merging the two approaches.

**UNIT 5**

Interpretation of Data and Paper Writing – Layout of a Research Paper, Bioinformatics and Computational Biology Journals, Impact factor of Journals, When and where to publish ? Ethical issues related to publishing, Plagiarism and Self-Plagiarism. Use of tools / techniques for Research: methods to search required information effectively, Reference Management Software like Zotero/Mendeley, Software for paper formatting like LaTeX/MS Office, Software for detection of Plagiarism.

**Reference Books**

1. Business Research Methods – Donald Cooper & Pamela Schindler, TMGH, 9th edition

2. Business Research Methods – Alan Bryman & Emma Bell, Oxford University Press.

3. Research Methodology – C.R.Kothari

**CORE PRACTICAL-4**

**MOLECULAR MODELING**

1. Simple Genetic algorithm
2. Genetic algorithm and problem solving
3. Genetic algorithm in scientific models
4. Sequence alignment algorithm
5. Structure prediction - HM Model
6. Determination of Partition coefficient of established drugs by shake flask method to find out the distribution of drug.(2 to 3 experiments)
7. QSAR related experiments - calculation of molecular connectivity index values for Ibuprofen, Atropine, Propranalol, Epinephrine and methadone (Minimum of 3 experiments will be given for work out).
8. Molecular modeling experiments- Molecular Graphics (3D structure) - conformational analysis to correlate physicochemical parameters with biological activity using various molecular modeling software such as AMBER, CAMSEQ, FRODO and SYBL (minimum of 5 experiments will be given).
9. Study of Pharmacophore models by using software(2 experiments)
10. Designing a novel molecule and fitting with receptor for invitro activity evaluation - experiments related to this is given (minimum of 3 experiments)
11. Calculation of F and R substituent constants (2 experiments)

**CORE PRACTICAL- 5**

**PERL & PHYTON**

1. Conversion of DNA to RNA sequence using Perl program.
2. Conversion of a DNA to RNA sequence using subroutine in Perl
3. Reverse complement of DNA sequence using Perl program.
4. Reading DNA sequence data from files using Perl program.
5. Finding motif in a protein sequence using Perl program.
6. Translation of given DNA sequence to protein sequence using Perl program.
7. Translation of DNA (taken from files) to protein sequence using Perl program.
8. Reading DNA sequence in all six reading frames using Perl program.
9. [Counting Letters in DNA Strings](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#counting-letters-in-dna-strings) using Phyton
10. [Python Program to find sum of array](https://www.geeksforgeeks.org/python-program-to-find-sum-of-array/)
11. [Python Program to find largest element in an array](https://www.geeksforgeeks.org/python-program-to-find-largest-element-in-an-array/)
12. [Efficiency Assessment](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#efficiency-assessment) using Phyton
13. [Verifying the Implementations](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#verifying-the-implementations) using Phyton
14. [Computing Frequencies](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#computing-frequencies) using Phyton
15. [Analyzing the Frequency Matrix](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#analyzing-the-frequency-matrix) using Phyton
16. [Dot Plots from Pair of DNA Sequences](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#dot-plots-from-pair-of-dna-sequences) using Phyton
17. [Finding Base Frequencies](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#finding-base-frequencies) using Phyton
18. [Translating Genes into Proteins](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#translating-genes-into-proteins) using Phyton
19. [Random Mutations of Genes](http://hplgit.github.io/bioinf-py/doc/pub/html/main_bioinf.html#random-mutations-of-genes) using Phyton

**CORE ELECTIVE**

**PAPER - 4**

**(to choose one out of 3)**

1. **BIOETHICS BIOSAFTEY AND IPR**

**Unit I Introduction to Intellectual Property:** Types of IP: Patents, Trademarks, Copyright & Related Rights, Industrial Design, Traditional Knowledge, Geographical Indications, Protection of GMOs IP as a factor in R&D; IPs of relevance to Biotechnology and few Case Studies

**Unit II Agreements and Treaties:** History of GATT & TRIPS Agreement; Madrid Agreement; Hague Agreement; WIPO Treaties; Budapest Treaty; PCT; Indian Patent Act 1970 & recent amendments

**Unit III Basics of Patents and Concept of Prior Art:** Introduction to Patents; Types of patent applications: Ordinary, PCT, Conventional, Divisional and Patent of Addition; Specifications: Provisional and complete; Forms and fees Invention in context of “prior art”; Patent databases; Searching International Databases; Country-wise patent searches (USPTO, esp@cenet(EPO), PATENTScope(WIPO), IPO, etc.)

**Unit IV Patent filing procedures:** National & PCT filing procedure; Time frame and cost; Status of the patent applications filed; Precautions while patenting–disclosure/non-disclosure; Financial assistance for patenting-introduction to existing schemes, Patent licensing and agreement Patent infringement- meaning, scope, litigation, case studies

**Unit V Biosafety:** Introduction; Historical Backround; Introduction to Biological Safety Cabinets; Primary Containment for Biohazards; Biosafety Levels; Biosafety Levels of Specific Microorganisms; Recommended Biosafety Levels for Infectious Agents and Infected Animals.

**Unit VI Biosafety guidelines:** Government of India; Definition of GMOs & LMOs; Roles of Institutional Biosafety Committee, RCGM, GEAC etc. for GMO applications in food and agriculture; Environmental release of GMOs; Risk Analysis; Risk Assessment; Risk management and communication; Overview of National Regulations and relevant International Agreements including; Cartegana Protocol. Module VII Bioethics: Ethical implications of biotechnological products and techniques. Social and ethical implications of biological weapons.

**REFERENCES:**

1. Beier F.K, Crespi R.S and Straus T. Biotechnology and Patent protection, Oxford and IBH Publishing Co. New Delhi.

2. Jeffrey M. Gimble, Academia to Biotechnology, Elsevier Academic Press.

3. Rajmohan Joshi (Ed.). 2006. Biosafety and Bioethics. Isha Books, Delhi.

4. Sasson A, Biotechnologies and Development, UNESCO Publications.

5. Senthil Kumar Sadasivam and Mohammed Jaabir M. S. (2008). IPR, Biosafety and Biotechnology Management, Jasen Publications, India.

6. BAREACT, Indian Patent Act 1970 Acts & Rules, Universal Law Publishing Co. Pvt. Ltd., 2007.

7. Kankanala C., Genetic Patent Law & Strategy, 1st Edition, Manupatra Information Solution Pvt. Ltd., 2007.

**CORE ELECTIVE**

**PAPER - 4**

1. **MEDICAL BIOTECHNOLOGY**

**Unit I Introduction to brain and neurobiology:** Sight and perception, hearing and balance, smell, taste, touch, pain, analgesics. Skin, hair.Muscles, movement, rheumatoid disorders. nervous system, skin, glands. Heart and blood circulation, blood clotting, microvasculature. Lungs, surfactants.Body fluids, fluid balance, parenteral solutions, renal physiology.

**Unit 2 Hormones and homeostasis:** Digestive system, reproductive system, nervous system. Genital system, reproductive biology and contraception. Diseases of the digestive system, breathing, circulation, Mechanisms of drug action.

**Unit 3 Immune System:** Structure, function and Cells of the immune system: The classification of human immune response: Humoral and cellular immunity, Innate and Adaptive immune response, Cellular components of the adaptive immune system, Phases of adaptive immune responses, Clonal expression, Toll like receptors, ABO blood Group. Lymphoid cells, clinical focus on the stem cells. Clinical uses and potential.B-lymphocytes and T-lymphocytes. Primary and Secondary lymphoid organs.

**Unit 4 Molecular Basis of cancer:** Types of Cancer – Stages of cancer development – Properties of cancer cells - Cell transformation – Tumor viruses – Genetic basis of cancer – Oncogenes – Tumor suppressor genes – Care taker genes.

**Unit 5 Designing vaccines for active immunization:** Live, attenuated vaccines. Subunit vaccines. Conjugate vaccines. DNA vaccines.Recombinant vector vaccines. Allergic responses in host defense Gell and Coombs Classification. IgE- Mediated (Type-I) hypersensitivity. Antibody-Mediated Cytotoxic (Type-II) hypersensitivity. Immune complex-Mediated (Type-III) Hypersensitivity.Type-IV or Delayed –type Hypersensitivity (DTH).Anaphylaxis. Pathogenesis and management of allergic asthma and rhinitis Clinical manifestations.

**Texts / References:**

1. Kuby, RA Goldsby, Thomas J. Kindt, Barbara, A. Osborne Immunology, 6th Edition, Freeman, 2002.

2. Brostoff J, Seaddin JK, Male D, Roitt IM., Clinical Immunology, 6th Edition, Gower Medical Publishing, 2002.

3. Janeway et al.,Medical Immunobiology, 4th Edition, Current Biology, publications., 1999.

4. Paul, Fundamental of Immunology, 4th edition, Lippencott Raven, 1999.

5. Goding, Monoclonal antibodies, Academic Press. 1985.

**CORE ELECTIVE**

**PAPER - 4**

1. **BIG DATA ANALYTICS AND NGS**

**COURSE OBJECTIVES:** To understand basics Big data and Next generation seqe3uncing and tools to analyze the huge biological data set.

**Unit 1**

Overview of Big Data, Evolution of Big data - Best Practices for Big data Analytics - Big data characteristics - Validating - The Promotion of the Value of Big Data - Big Data Use Cases.

**Unit 2**

Characteristics of Big Data - Four V's, Basic operations of in big data, Datasets, Data Analytics, different data types of big data, Awareness of Architecture- Big Data with Hadoop, Introduction to Hadoop,

**Unit 3**

Emergence of Next generation sequencing, Illumina Genome Analyzer, Applied Biosystems Sequencing, Ion Torrent Sequencing, Polonator Technology, Nanopore Sequencing, Single Molecule Real Time DNA sequencing, Comparison of Next generation sequencing techniques, NGS File formats, & applications.

**Unit 4**

De novo Genome sequence assembly, Reference sequence assembly, Challenges of Genome assembly, Use of paired – end reads in the assembly, Data Preprocessing methods and sequencing read correction methods, Assembly Errors, Evaluation of assembly methods.

**Unit 5**

Transcriptome (RNA) sequencing, Exome sequencing, Genome Annotation, Using NGS to detect sequence variants, ChIP-sequence, Biological theories on ChIPsequence analysis, Understanding the non – coding genome, Disease gene identification, DNA fragment evaluation, Peak identification, Two condition comparison, Saturation analysis, Motif finding and related theories.

**Text Books**

1. Big Data Fundamentals, Concepts, Drivers & Techniques Concepts, Drivers & Techniques by Thomas Erl, Wajid Khattak, and Paul Buhler, PRENTICE HALL, 2012.

2. Analytics in a Big Data World: The Essential Guide to Data Science and its Applications (WILEY Big Data Series) – Bart Baysen, 2014

3. Ali Masoudi-Nejad, Zahra Narimani, Nazanin Hosseinkhan; “Next Generation Sequencing and Sequence Assembly”, Methodologies and Algorithms, Springer; 2013.

4. Stuart M. Brown, “Next-Generation DNA Sequencing Informatics”, Cold Spring Harbor Laboratory Press, 2013.

**Reference Books**

1. Big Data Now, O’Reilly Radar, O’Reilly Media, 2012

2. Big Data for Dummies, Wiley, Judith Hurwitz, Alan Nugent, Fern Halper, Marcia Kaufman, 2012 COURSE OUTCOME: Able to analyze the data using R language. Able to work in NO SQL databases. Able to find hidden patterns in Big data

3. Mark I. Rees, “Challenges and Opportunities of Next-generation Sequencing for Biomedical Research”, Academic Press, 2012.

**OPEN ELECTIVE**

**PAPER - 4**

**FUNDAMENTALS OF ALGORITHMS**

**UNIT 1**

Notion of an Algorithm – Fundamentals of Algorithmic Problem Solving – Important Problem Types – Fundamentals of the Analysis of Algorithmic Efficiency –Asymptotic Notations and their properties. Analysis Framework – Empirical analysis – Mathematical analysis for Recursive and Non-recursive algorithms – Visualization

**UNIT 2**

Brute Force – Computing an– String Matching – Closest-Pair and Convex-Hull Problems -Exhaustive Search – Travelling Salesman Problem – Knapsack Problem – Assignment problem. Divide and Conquer Methodology – Binary Search – Merge sort – Quick sort – Heap Sort -Multiplication of Large Integers – Closest-Pair and Convex – Hull Problems.

**UNIT 3**

Dynamic programming – Principle of optimality – Coin changing problem, Computing a Binomial Coefficient – Floyd‘s algorithm – Multi stage graph – Optimal Binary Search Trees – Knapsack Problem and Memory functions. Greedy Technique – Container loading problem – Prim‘s algorithm and Kruskal’s Algorithm – 0/1 Knapsack problem, Optimal Merge pattern – Huffman Trees.

**UNIT 4**

The Simplex Method – The Maximum-Flow Problem – Maximum Matching in Bipartite Graphs, Stable marriage Problem.

**UNIT 5**

Lower – Bound Arguments – P, NP NP- Complete and NP Hard Problems. Backtracking – n-Queen problem – Hamiltonian Circuit Problem – Subset Sum Problem. Branch and Bound – LIFO Search and FIFO search – Assignment problem – Knapsack Problem – Travelling Salesman Problem – Approximation Algorithms for NP-Hard Problems – Travelling Salesman problem – Knapsack problem.

**REFERENCE BOOKS**

1. Introduction to Algorithms, 3rd Edition by Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein

2. A Guide to Experimental Algorithmics, by McGeoch.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*