Mathematical Methods				
Contact Hours /Week	: 3 Hrs.		Credits	: 4
Total Lecture Hours	:40		CIE Marks	: 50
	:0		SEE Marks	: 50
Total Tutorial Hours	:			
Sub. Code	:BRD01			

Module-I **07 Hrs**

Introduction: Scope of biostatistics, definition, data collection, presentation of data, graphs, charts (scale diagram, histogram, frequency polygon, frequency curve, logarithmic curves). Sampling & selection bias, probability sampling, random sampling, sampling designs. Descriptive statistics: Measure of central tendency (arithmetic mean, geometric mean, harmonic mean, median, quartiles, mode); Measure of dispersion (range, quartile deviation, mean deviation and standard deviation, coefficient of variation).

Module-II **08 Hrs**

Bi-Variate Distribution: Correlation and regression analysis (simple and linear) curve fitting (linear, non-linear and exponential).

Probability: Axioms, models, conditional probability, Bayes rule, Genetic Applications of Probability, Hardy - Weinberg law, Wahlund's Principle, Forensic probability determination, Likelihood of paternity, Estimation of probabilities for multilocus/ multi-allele finger print systems.

Module-III **08Hrs.**

Statistical Inference: Estimation theory and testing of hypothesis, point estimation, interval estimation, sample size determination, simultaneous confidence intervals, parametric and non-parametric distributions (T-test, F-test, Chi Squared distribution, goodness of fit test) analysis of variance (one-way and two-way classifications). Case studies of statistical designs of biological experiments (RCBD, RBD).

Module-IV **07 Hrs.**

Probability Distributions: Discrete probability distributions - Binomial, Poisson, geometric – derivations. Central limit theorem. Continuous probability distribution – normal, exponential, gamma distributions, beta and Weibull distributions, T & F distributions.

Design of Experiments: Sample surveys, comparisons groups and randomization, random

assignments, single and double blind experiments, blocking and extraneous variables, limitations of experiments.

Case Studies: Cigarette smoking, Lung cancer, endangered plants species, epidemics.

Module-V Problems for module I-IV

TEXT BOOKS

1	Wayne W Daniel	Biostatistics, Wiley 7th Edition
2	Ronadd N Forthofer and Eun Sul Lee	Introduction to Biostatistics, Academic Press.
3	Norman T J Bailey	Statistical methods in Biology, Cambridge Press.
4	J.N.Kapur.	Mathematical Models in Biology and Medicine
5	Ipsen, Feigl & Bancroff	Introduction to Biostatistics
6	Animesh K Datta	Basic Biostatistics & its Applications
7	Lehninger	Principles of Biochemistry

BASIC BIOINFORMATICS AND DATABASE MANAGEMENT

Contact Hours/ Week	: 04	Credits :	4
Total Lecture Hours	: 52	CIE Marks :	50
Sub. Code	: (BRD02)	SEE Marks :	50

Module-I

Introduction to Biological databases

Bioinformatics: (What is Bioinformatics, Goals, Scope, Application, Limitations and New Themes).Database: (What is a database, Types) Biological Database: Databases (Primary, Secondary and Specialized), Interconnection between the databases, Fit falls of Biological databases. Nucleotide and Protein sequence databases (NCBI, Genbank, EMBL, DDBJ, PDB and MMDB) Format of databases ☹GenBank flat file, FASTA Format, PIR Format) 10 Hours.

Module-II

Sequence Alignment and Database Similarity Searching

Sequence Alignment: Evolutionary Basis, Homology versus Similarity, Similarity versus Identity, Global alignment, Local alignment, Pairwise alignment,

Alignment algorithm:

Scoring Matrices: Amino acid scoring matrices; PAM, BLOSUM, Comparison between PAM and BLOSUM, Statistics of Sequence Alignment (e-Value, p-value)

Database Similarity Searching: BLAST. BLAST variants. Statistical significance. Low complexity Regions. BLAST output format. FASTA. Simple Alignment problems. Sequence alignment algorithms : DOT Matrix Dynamic programming algorithm: Comparative Genomics and proteomics ,Heuristic similarity searches: FASTA, BLAST, BLAT.

Module-III **10 Hours**.

Motif, Domain and Gene Prediction

Scoring Function , exhaustive algorithms, Heuristic algorithms, practical issues. Profiles and Hidden Markov Models: PSSM. Profiles. Markov Model and HMM. Zeroth, First and Higher order HMM. Protein Motif and Domain Prediction: Identification of Motif and Domains in MSA. PROSITE. Motif and Domain Databases using Statistical Models (PRINTS, BLOCKS, ProDom, Pfam, SMART)Protein Family databases (COG). Motif Discovery and Sequence Logos. Numerical conceptual on 0^{th} 1^{st} and Higher order HMM.

11 Hours

Module-IV

Molecular Phylogenetics and Protein structure visualization: Phylogenetics Basics, classification (Maximum likelihood and parsimony method). Terminologies, Gene versus species phylogeny, Forms of tree representation. Tree Construction: Choosing Molecular Markers. Alignment. Multiple Substitutions. DISTANCES, GROW TREE PAUP, PHYLIP. Choosing Substitution Models (Jukes Cantor Model).

Protein structural visualization: Rasmol, Swiss-PDB viewer, Chime, Cn3D. Protein structure comparison: Intermolecular, Intramolecular and combined methods. Multiple structure alignment. Protein structure classification: SCOP and CATH, Secondary structure Prediction: Ab Initio Methods, Secondary structure Prediction for transmembrane proteins: Helical membrane proteins, RNA structure prediction.

Module-V

Other Databases **11 Hours**

Introduction, Purpose of Database systems

Chemical compound databases (PubChem, NCI, ZINC, Drug Bank, Ligand)

Other Important Databases: KEGG, Pubmed, OMIM, Medline, TIGR. BioMed Central, Public Library of Sciences (PloS), CiteXplore.

TEXT BOOKS:

1.	Andreas D Baxevanis	Bioinformatics, Wiley Interscience, 1998.
2.	David W Mount	Bioinformatics, Cold spring arbour, 2001.
3.	Jin Xiong	Essentials Bioinformatics, Cambridge University Press

REFERENCE BOOKS:

1.	R F Doolittle	Computational methods for macromolecular sequence analysis, Academic Press, 1996.
2.	S C Rastogi, N Mendiratta & P Rastogi,	Bioinformatics, Methods and Applications- Genomics, Proteomics and Drug, Discovery –Phi, 2006.
3.	Arthur Lesk	Introduction to Bioinformatics, Oxford, 2006.
4.	Stuart M Brown	Bioinformatics, NYU Medical Center, NY USA. 2000.
5.	Middha Sk, T. Usha & Prashanth HP	Bioinformatics, CBH Publishers, 2012

Molecular Modeling and Drug design				
Contact Hours /Week	: 3 Hrs.		Credits	: 3
Total Lecture Hours	:40		CIE Marks	: 50
	:0		SEE Marks	: 50
Total Tutorial Hours	:0			
Sub. Code	:BRD 03			

Module I **07 Hours.**

Principles of protein structure; – Hierarchical organization of protein structure – Primary. Secondary, Super secondary, Tertiary and Quaternary structure; Hydrophobicity of amino acids, van der Waal and Solvent accessible surface, Internal coordinates of proteins; Derivation, significance and applications of Ramachandran Map, protein folding.

Protein structure prediction : basic principles of protein structure, first-principles methods for predicting protein structure, motifs of protein structure, alpha domain structures, alpha/beta structures, folding and flexibility, prediction, engineering and design of protein structures.

MODULE –II **8 Hours.**

Insilco Drug Design and Computer Assisted New Lead Design:

Introduction, historical perspective, drug compounds, reparation and organization for drug seeking, common stages in the drug seeking campaign, sources of hits, leads and candidate drugs, Natural products: higher plant and animal products, combinational libraries, Lead optimization. Introduction, Basic Concepts, Molecular Recognition by Receptor and Ligand Design, Active Conformation, Approaches to Discover New Functions, Approaches to the Cases with known and unknown receptor structure. Introduction to drug metabolism, toxicity and pharmacokinetics, toxicology considerations, problems and drawbacks on drug discovery and development.

MODULE –III **09 Hours**

Molecular Modelling and Simulation:

Introduction of molecular modelling – Coordinate system, Potential energy surfaces, Molecular modelling Literature, steps in molecular modelling: introduction to Quantum Mechanics, introduction to Molecular mechanics. Force filed, Types of force filed.

Example: Amber force filed, CHARMM force filed.

Introduction to Van der Waals interaction, Hydrogen bonding in molecular mechanics, Model for Molecular simulations, introduction about molecular dynamics.

MODULE –IV **07 Hours**

Structure Prediction and Drug Design

Drug design and discovery: an overview. Protein Structure Prediction - Introduction to Comparative Modelling. Sequence Alignment. Constructing and Evaluating a Comparative Model. Predicting Protein Structures by 'Threading', Comparative between Homology Modelling and Threading,

Basics of molecular modelling, methods, steps involved in MM, Selection of target and template, homology modelling, refinement and validation-SAVES Server, The Critical Assessment of protein Structure Prediction (CASP), Superposition of proteins using different tools, RMSD, Presentation of protein conformations, Hydrophobicity factor, Shape complementary

Molecular Docking studies, Structure based De Novo Ligand design, Drug Discovery - QSAR.

Module- V

Molecular Docking and case studies:

Different types of docking approaches (Structure based, Ligand based)

ADMET properties of drug, Lipinski Rules of 5

Docking and Virtual Screening: Using different docking algorithms, Optimization of docking algorithms based on different target (protein and DNA),Virtual screening-High throughput screening of Drugs, Post docking analysis, Mode of interaction studies, Pharmacophore prediction based on the docking analysis.

Computer-assisted drug design/Structural drug design:

Case studies (Any two or three)

- Aplaviroc: CCR5 receptor antagonists
- Capsazepine: TRPV1 antagonists
- Celecoxib:Cyclooxygenase 2 inhibitors
- Darunavir:HIV protease inhibitors

TEXT BOOKS

1	Povl Krogsgaard and Larsen	Molecular modelling, I edition, Multivista Global Ltd.2002.
2	Andrew R Leach	Molecular modelling: principles and applications, 2nd edition, Pearson education ltd., 2001.

Reference Books:

1. A.R.Leach, *Molecular Modelling Principles and Application*, Longman, 1996.

2. J.M.Haile, *Molecular Dynamics Simulation Elementary Methods*, John Wiley and Sons, 1997.

3. Satya Prakash Gupta, *QSAR and Molecular Modeling*, Springer - Anamaya Publishers, 2008

Computer Languages and Network Informatics				
Contact Hours /Week	: 4 Hrs.		Credits	: 4
Total Lecture Hours	:40		CIE Marks	: 50
			SEE Marks	: 50
Sub. Code	:BRD04			

Module I **07 Hrs**

Computer design of oligonucleotide probes: Introduction, Oligonucleotide probe design methodology, filtering of low-complexity sequence: RepeatMasker, prediction of cross hybridization to related genes, thermodynamics of nucleic acid duplexes and the prediction of melting temperature, probe secondary structure: computation of probe secondary structure, Mfold over the web.

Image processing: Introduction, Feature extraction: identifying the position of the feature, identifying the pixels that comprise the feature, identifying the background pixels.

Module II

Normalization: Introduction, Data cleaning and Transformation: removing flagged features, background subtraction, logarithms. Within array normalization: linear regression of Cy5 against Cy3, linear regression of log ratio against average intensity, non-linear regression of log ratio against average intensity, correcting for spatial effects. **Measuring and quantifying microarray variability:** Introduction, measuring and quantifying microarray variability: calibration experiments, pilot studies, quantifying the variability, log-normal distribution, methods for measuring the variability, variation between replicate features on the array, variability between hybridizations to different arrays, variability between the individuals.

Analysis of differentially expressed genes: Introduction, fundamental concepts: statistical inference, hypothesis tests, p-values, independence, classical parametric statistics: one sample t-test, two sample t-test, requirement of t-test, non-parametric statistics: bootstrap analysis

Module III

Computer concepts : Introduction to digital computer: Basic functional units of a digital computer – central processing unit, Arithmetic and Logic unit, control unit, Memory unit, Input and Output units, Stored Program concept, Types of computer.

Storage devices – Primary storage: Random access memory, Read only Memory, Secondary storage – Floppy disk, Hard disk, CD- ROM and its operation.

Computer Language – Machine language, Assembly language and assembler, High-level languages and compiler, Interpreter, Editor. System software, Application software.

Algorithm and Flow charts: The meaning of algorithms / flowcharts and their need, Example Algorithms and Flowcharts.

Module IV **07 Hrs**

History & Evolution of Java ,Overview of Java ,Data Types, Variables, and Arrays , Operators , Control statements ,Introducing Classes , A Closer Look at Methods and Classes , Inheritance Packages and Interfaces , Exception Handling , Multithreaded Programming , Enumerations, Autoboxing, and Annotations (Metadata) , I/O, Applets, and Other Topics

Module V **08 Hrs**

Biojava: sequence input, sequence output etc

Bioperl: Introduction to Perl, Unix basics, Sequences and Strings Variables, Arrays,Files ,Motifs and Loops ,Flow control ,String operators ,Writing files .

Introduction to Biopython, General overview of what Biopython provides , working with sequences , and usage example.

1	Dov Stekel	Microarray Bioinformatics –first edition, Cambridge university press, 2005.
2	Mark Schena	Microarray Analysis —new edition, J. Wiley & Sons, 2002.
3	C programming	Balguruswamy
4	James Tisdall	Beginning Perl for Bioinformatics
5	Herbert Schildt	Java The Complete reference

Computer Languages and Network Informatics				
Contact Hours /Week	: 4 Hrs.		Credits	: 4
Total Lecture Hours	:40		CIE Marks	: 50
	:0		SEE Marks	: 50
Sub. Code	:BRDL01			

LAB 1

Part A: BASICS OF BIOINFORMATICS AND SEQUENCE LAB

1. Sequence retrieval from nucleic acid and protein databases
2. Sequence (FASTA and BLAST) searches – Analysis of parameters affecting alignment.
3. Pair wise comparison of sequences – Analysis of parameters affecting alignment.
4. Multiple alignments of sequences – Analysis of parameters affecting alignment.
5. Evolutionary studies / Phylogenetic analysis – Analysis of parameters affecting trees.
6. Identification of functional sites in Genes / Genomes.
7. Secondary structure prediction of proteins.
8. Pattern elucidation in Proteins (PROSITE).
9. Primer Design.
10. Structure analysis using Rasmol and pymol
11. Comparative Modeling of homologous sequences.
12. Superposition of structures – Calculation of RMSD for main chain atoms.

Part B :

Lab name and objective(s) of the lab

1. Protein Structure Retrieval and Visualization
2. Active site Prediction Using Online server & Free tools
3. Active site Prediction Using Commercial software
4. Protein Minimization - CHARMm forcefield using
5. Ligand Retrieval and Visualization
6. To check Lipinski rule and ADMET Properties for small molecules
7. Ligand Minimization – CHARM and MMF forcefield
8. Molecular Docking using Free tool
9. Molecular Docking Using Commercial software
10. Visualization and Analysis of Docking Result
11. Homology Modeling
12. Creating and validating a 3D QSAR Pharmacophore

Computer Languages and Network Informatics lab				
Contact Hours /Week	: 4 Hrs.		Credits	: 4
Total Lecture Hours	:40		CIE Marks	: 50
	:0		SEE Marks	: 50
Sub. Code	:BRDL02			

Programming using PERL and BioPERL.

1. Storing and retrieving amino acid sequences using structure data types
2. Swapping two values using pointers
3. File Processing – Creation and simple processing
4. Computing amino acid composition of a given protein sequence.

Web Publishing

1. User interface design
2. Create a web page for your University / College using HTML - The opening page should provide hyperlinks to other
pages (add animation and sound effects appropriately).
3. Creating a web page to get protein sequence data and compute and display amino acid composition
4. Creating a web page to get nucleic acid sequence data and compute and display base composition.

BASIC OPERATIONS IN MATLAB AND R