PATNA UNIVERSITY

Memorandum for Syllabus of Post Graduate Diploma (Vocational) Course in Bio-informatics

Information Technology and Biology are shaking hands with each other to form interdisciplinary field of study which is known as Bio-informatics. It is growing at very fast pace and is generating demands and necessities for new events to come up. Molecular Biology along with Computer Science, Statistics and Mathematics forged together to develop new studies like visualization pattern, graph theory, complex systems, development of databases, algorithms, computational and statistical techniques/theory, analysis of biological data and their management in order to provide solutions to many societal problems.

Keeping this inter-disciplinary nature of Bio-informatics in mind, this Post Graduate Diploma (Vocational) Course in Bio-informatics has an essential component from basic science having two groups that is Life Sciences and Physical Sciences. This course will include necessary features of Information Technology, Computer Science, Statistics, Mathematics, Chemistry, Physics, Botany, Zoology, Bio-chemistry, Molecular Biology, Computational Biology etc.

Goal of this Course

The goal is to impart fruitful standard education according to the needs of the society to the graduates from Science, Medicine, Engineering, Agriculture, Pharmacy, Biotechnology or Biochemistry who want to contribute and pursue career in the field of Bio-informatics.

Course Details

Eligibility

One who has acquired graduation degree in Science, Medicine, Engineering, Agriculture, Pharmacy, Biotechnology or Biochemistry from a recognized university with at least 45% marks.

Ordinance and Regulations:		As approved by the Hon'ble Chancellor of Universities of Bihar and communicated vide letter No. 2743 dated 09.12.2011
Admission Procedure :		Aptitude (Entrance) Test
Total no. of Seats	:	50
Duration	:	Three Semesters

Course Structure – PGDBIN				
Paper	Paper Title	Internal	Semester	Total
Code		Assessment	Examination	
Semester - I				
BIN101	Bioinformatics - I	20	80	100
BIN102/	Elective – I *	20	80	100
BIN107				
BIN103/	Elective – II**	20	80	100
BIN108				
BIN104	Fundamentals of Computing	20	80	100
BIN105	Computer Programming in C (Theory)	20	80	100
BIN106	Computing Lab – I (Practical): C and	20	80	100
	Software packages.			
Semester - I	I			
BIN201	Bioinformatics -II	20	80	100
BIN 202	Computational Methods for Sequencing	20	80	100
	Analysis			
BIN 203	Molecular Imaging and Drug Design	20	80	100
BIN 204	Database Management System	20	80	100
BIN 205	PERL Programming and Applications to	20	80	100
	Bioinformatics			
BIN 206	Computing Lab – II (Practical): Sequence	20	80	100
	alignment and database search techniques.			
Semester – I	Ш			
BIN301	Project	20 (Viva-Voce)	80	100
	Grand Total			1300

* Note: - Keeping in view the background of the candidates, the Department shall do the allocation of elective course to students from Physical Sciences and Biological Sciences as follows:

Elective	Students from Physical Sciences	Students from Biological Sciences		
	to offer	to offer		
Elective - I	BIN102 Modern Biology	BIN107 Basic Mathematics		
Elective - II	BIN103 Structural Biology	BIN108 Statistics		

<u>Note:-</u> The regulations not covered above and common to other courses will be applicable as per the rule of the University.

Syllabi placed before the Academic Council for consideration and approval

SEMESTER – I

PAPER CODE - BIN101 PATER TITLE - BIOINFORMATICS – I

Introduction to Bioinformatics : Definition and Scope of Bioinformatics, Applications of Bioinformatics in various areas, Overview of available Bioinformatics Resources on the Web, Protein and Genome; Information Resources and Analysis Tools; Established Techniques and Methods; Sequence File Formats FASTA, GenBank and Structured File Formats.

Biological Databases: Protein Sequence and Structural Databases, Nucleotide Sequence Databases; NCBI, PubMed, Protein Data Bank(PDB), PIR, Swiss Prot, EMBL. GenBank, DDBJ, UniGene, SGD and EMI Genomes.

Specialized Databases: Pfam, SCOP, GO, GenBank, Genome Net, EST, SNP, Metabolic Pathways Databases, EMBL, Similar Sequence Search BLAST, Gene/Protein Sequences and its Implications. Pair-wise Alignments, Scoring Matrix, PAM, BLOSUM and Gap Penalty.

Secondary Structure Analysis Tools: Sequence Motif Databases, Pfam, PROSITE, Protein Structure Classification; SCOP, CATH, Other Relevant Databases, KEGG, PQS, PMDB, MPDB. Protein Structure Alignments; Structure Superposition, RMSD, Different Structure Alignment Algorithms, DALI, and TM-align.

Methods of Sequence Analysis: Heuristic Methods; FASTA, Statistics of Sequence Alignment Score; E-Value, P-Value, Multiple Sequence Alignments; ClustalW, Profile, Profile-Sequence Alignment, Profile-Profile Alignment, PSI-BLAST, Hidden Markov Models, Viterbi Algorithm and HMM Based Multiple-Sequence Alignment.

Phylogenetic Analysis: Distance and Character Based Methods and Software, Computing Tools for Phylogenetic Analysis, Distances, GROWTREE, PAUP, PHYLIP and MEGA; Construction and Visualization of Phylogenetic Tree; and Application of Phylogenetic Analysis.

Practicals :

- 1. Exploring the integrated database system at NCBI server and querying the PUBMED and GenBank databases using the ENTREZ search engine:
- 2. Exploring the integrated database system at EBI server and searching the EMBL. Nucleotide database using the SRS search engine.
- 3. Exploring & querying SWISSPROT & UniProtKB. Sequence Formats & Format converters. Exploring tools on ExPASY.
- 4. Homology search using BLAST.
- 5. Introduction to "R".

Text / Reference Books:

- 1. Anna Tramontano, 2007; "Introduction to Bioinformatics", Chapman and Hall Series.
- 2. Jason T.L .Wang, Mohammed J. Zaki, Hannu T.T. Toivonene and Dennis Shasha, 2005; "Data Mining in Bioinformatics", Springer International Edition.
- 3. Yi-Ping Phoebe Chen, 2007; "Bioinformatics Technologies", Springer International Edition.
- 4. Mount, D.W., 2001; "Bioinformatics: Sequence and Genome analysis", Cold Spring Harbour Laboratory Press.
- 5. Teresa Attwood, Parry-Smith David J., 2001; "Introduction to Bioinformatics", Pearson Education (Singapore) Pvt. Ltd.

ELECTIVE – I* PAPER CODE - BIN102 PAPER TITLE - MODERN BIOLOGY

Introduction to Modern Biology

Prokaryotic and eukaryotic cells : Structure and function.

Cell cycle and its regulation.

Nucleic acids : Structure and types of DNA and RNA.

Packaging of genome.

Genome organization: C-value paradox, Genetic Code, Genes, Gene Clusters and families, Transposons, Organelle genomes.

DNA replication – Prokaryotes and Eukaryotes.

DNA recombination.

Gene Expression – Prokaryotes and Eukaryotes, Motifs

Transcription – Prokaryotes and Eukaryotes, Reverse Transcription and cDNA, Gene Cloning.

Translation – Prokaryotes and Eukaryotes.

Post-translational modifications of proteins.

Targetting gene therapy,

Text/ Reference Books:

- 1. Alberts Bruce, Johnson Alexander, Lewis Julian, Raff Martin, Keith Roberts, Walter Peter, 2008; "Molecular Biology of the Cell", 5th Edition. Garland Science.
- Watson James D., Baker Tania A., Bell Stephen P., Alexander Gann, Levine, Michael Losick Richard, 2008; "Molecular Biology of the Gene", 6th Edition. New York, Cold Spring Harbor Laboratory Press.
- 3. Alberts Bruce, Bray Dennis, Hopkin Karen, Johnson Alexander D, Julian, Lewis Raff Martin, Keith Roberts, Walter Peter, 2010; "Essential Cell Biology", 3rd Edition. Garland Science.

PAPER CODE - BIN107 PAPER TITLE - BASIC MATHEMATICS

Function : Definition of Function, Relations and functions, Domain, Range and Image of function; Types of Function: Increasing Function, Decreasing Function, Exponential Function, Logarithmic Function, Polynomial Function, Operation on function : Addition, Sum, Product, Division; Inverse of Function. Graphs of important functions.

Set Theory: Definition of sets and its notation, Finite set, Infinite set, Null or Empty set, Singleton set, Subset, Equivalent sets, Equal sets, Power set, Universal set, Venn Diagram, Operations of sets: Union, Intersection, Difference, Symmetric Difference of sets; Disjoint sets, Compliment of a set, Algebra of set.

Limits and Continuity: Intervals: closed and open; Left Hand Limit, Right Hand Limit, Infinite Limits, Limits at Infinity, Evaluation of Limits, Algebraic Limits, Trigonometric Limits, Logarithmic Limits, Exponential Limits, Other Useful Limits, Continuity at a point, Left Hand Continuity, Right Hand Continuity, Continuity at End Points, Continuity over an interval, Conditions for Continuity.

Differential Calculus: Differentiation, Derivative from First Principle, One Side Derivative, Differentiability at End Points, Differentiability over an Interval, Chain Rule of Differentiation, Logarithmic Differentiation, Derivatives of Trigonometric Functions, First Order Derivatives, Second Order Derivatives and Derivatives of Higher Order, Maxima and Minima.

Integral Calculus: Integration of Trigonometric Functions, Indefinite Integral, Integration by Substitution, Integration by Parts, Definite Integral, Properties of Definite Integral, Summation of series using Definite Integral, Application of Integrals.

Differential equation: Definition, Order of Differential Equation, Degree of Differential Equation, Linear and Non-linear Differential Equation, Solution of a Differential Equation.

Matrices and Determinant : Definition and types of a Matrix, Operation on Matrices: addition, subtraction and multiplication, Transpose of a Matrix, Cofactor of a Square Matrix, Determinant of a Square Matrix, Rank of a Matrix, Inverse of a matrix.

Text/ Reference Books:

- 1. Frank Ayres, Jr., Matrices, Schaum's Outline Series.
- 2. Grewal, B.S., "Engineering Mathematics",
- 3. Das and Mukherjee, "Differential Calculus", U.N. Dhar & Co., Kolkata
- 4. Das and Mukherjee, "Integral Calculus", U.N. Dhar & Co., Kolkata
- 5. Sancheti, and Kapoor, K. K., "Business Mathematics", Sultandard and Company, New Delhi.

ELECTIVE – II*

PAPER CODE - BIN103 PAPER TITLE - STRUCTURAL BIOLOGY

Biological Macromolecules: Composition, Structure and Function of Biomolecules (Carbohydrates, Lipids, Proteins, Nucleic Acids and Vitamins). Principles of Biophysical Chemistry (pH, Buffer, Reaction Kinetics, Thermodynamics, Colligative Properties).

Structure of Carbohydrates: Structure of Monosaccharide, Disaccharide and Polysaccharide. Conformations of Monosaccharide; Stereo Isomerism, Mutarotation, Reducing and Non-Reducing Sugars, Nomenclature etc. Major Monosaccharide and Disaccharide and its Potential Uses.

Structure of Nucleic Acid: The Nature of Chemical Bonds; Nucleotides and Nucleic Acids, Structure of DNA (B, A and Z forms), DNA-Binding Proteins, Protein-DNA Recognition in Euraryotic and Prokaryotic Organisms, Protein-DNA Interactions, Supercoiled DNA, Nucleosome Structure, Histones.

Structure of Enzymes: Enzyme nomenclature and classification; units of enzyme activity; specificity; coenzymes and metal cofactors; temperature and pH effects; Michaelis-Menten kinetics; Inhibitors and activators; active site and catalytic mechanisms; covalent and non-covalent regulations; isoenzymes.

Structure of Proteins: Properties of Amino Acids, Peptides and Proteins, Conformation of Peptide Bonds. Characterization of Peptides. Protein Architecture, Sequence, Secondary Structure, Surersecondary Structures, Structural Domains, Tertiary structure, Quartenary structure, Protein-protein interactions, Protein–Drug interactions

Text/ Reference Books:

- 1. Donald Voet, Judith G. Voet, 2004; "Biochemistry", 3rd Ed, Wiley.
- 2. Lubert Stryer, 2007; "Biochemistry", 4th ed, W. H. Freeman and Company
- 3. C. Stan Tsai, 2003; "An introduction to computational biochemistry", 4th ed, Wiley.

PAPER CODE - BIN108 PAPER TITLE - STATISTICS Introduction : Definition and

Probability and Distributions: Sample space, events, equally likely events. Definitions of Probability, Independent Events. Addition and multiplication rules, Conditional probability, Random Variables, Mathematical Expectation, Continuous and discrete distribution: Binomial, Poisson, Uniform, Exponential and Normal distribution.

Central Tendency and Dispersion: Measures of central tendency and dispersion.

Correlation Analysis: Methods of studying simple correlation: Scatter Diagram, Karl Pearson's Coefficient of Correlation, Spearman's Rank Correlation. Multiple and Partial Correlation, intra-class correlation (without proof).

Regression Analysis: Introduction: Meaning and Purpose of Regression. Simple and Multiple Linear Regressions using Least Square Principle. Coefficient of Determination.

Analysis of Variance: Assumptions for Use of the ANOVA, One-Way ANOVA, F-test, Two-Way ANOVA for the Randomized Block Design, Linear Model Representation of the ANOVA.

Testing Hypothesis: Sampling & Test of Significance, Testing Hypotheses about the Variance, Testing Hypotheses about the Difference of Two Means, Testing Hypotheses about the Difference of Two Proportions, Tests of Hypotheses and Sample Size. Parametric and nonparametric test. The Goodness-of-Fit Test; Chi Square Test.

Text/Reference Book:

- 1. W.Evans & G.Grant, 2001; "Statistical Methods in Bioinformatics": An Introduction, Springer.
- 2. Levin and Rubin (). "Statistics for Management" Prentice Hall of India, New Delhi.
- 3. S. Prasad, 2007; "Elements of Biostatistics", Rastogi publication.
- 4. Probability and Statistics, Schaum's Outline Series.
- 5. Gupta S.C. : Statistics, Himalaya Publication, New Delhi.

PAPER CODE- BIN104 PAPER TITLE - FUNDAMENTALS OF COMPUTING

Computers and Data Representation Systems: Computing Systems; Layers of a Computing Systems; Abstraction; Computers as a Tool and a Discipline; Ethical Issues and Concept of Digital Divide. Binary Values and Computers. Data and Computers; Analog and Digital Data; Binary Representation; Representation of Numeric Data – Negative, Real, Text; The ASCII and Unicode Character, Set; Audio Formats – MP3 ; Representing Images, Graphics and Video; and Video Codes.

Programming Languages: Computer Operations; Levels of Abstraction; Machine Language; Virtual Computers; Assembly Language; Pseudo-Operations; Low Level Programming Example; Compilers; Interpreters; Programming Language Paradigms; Functionality of Imperative Languages; Boolean Expressions; Strong Typing; Input-Output Structures; Control Structures; Composite Data Types and Object Oriented Concepts.

Information Systems: Data vs Information; Information Systems and Technologies; Spreadsheets; Spreadsheet Formulas; Circular References; Spreadsheet Analysis,; Database Management System; The Relational Model; Relationships; Structured Query Language; Database Design; Information Security; Confidentiality, Integrity and Availability, File Systems; File Types; File Operations and Directories.

Operating System: Introductory concepts, Operating system structures, System components, OS Services, System calls, System programs, System structure, Virtual machines, Processes, Process Concepts, Process scheduling, Operations on processes, Process states, Cooperating processes, Interprocess communication, Common OS problems. DOS and Windows, Introductory concepts, Architecture and special Features, Useful DOS Commands, DOS - Editor EDIT & Writing Batch Files, Different Versions and their Distinguishing Features, Managing Operations with Windows. Linux : Introductory concepts, Architecture and special features, Useful commands, Related operating system and their Introduction.

Feasibility Analysis: Project Fundamentals; Project Initiation; Problems with the Organization; Determining Feasibility – Defining Objectives, Determining Resources and Judging Feasibility; Activity Planning and Control; Computer Based Project Scheduling- Time boxing; Managing Analysis and Design Activities; Goal Tending – Setting.

Text / Reference Books:

- 1. Dale & Lewis, 2007; "Computer Science Illuminated", 3rd ed, Narosa Publishing House.
- 2. Kedall & Kendall, 2005; "Systems Analysis and Design", 5th ed, Prentice Hall India.

- 3. Rajaraman, 2007; "Fundamentals of Computers", 4th ed, Prentice Hall India
- 4. ITL Esl, 2009; "Introduction to Computer Science", Pearson Education.

PAPER CODE- BIN105 PAPER TITLE- COMPUTER PROGRAMMING IN C (THEORY)

Problem Solving Approach: Introduction to Programs and Algorithms; Problem Solving Aspect (Algorithm Devising); Algorithm Design Aspect (Top-down Design); Algorithm Implementation; Program Verification. Fundamental Algorithms, Flowchart.

Basic Concepts of C Programming Language: Character Set; Keywords; Identifier, Constants, and Variables; Constant Types – Numeric and Character Constants; Data Types and Range of Values – Character, Integer and Floating Point; Signed, Unsigned, Short, and Long Integers; Data Declaration and Definition, Operator & Expression – Arithmetic, Relational, Logical, Increment, Decrement, Assignment, Conditional, and Bitwise Operators; Precedence & Associability of Operators; Managing Console I/O – Reading and Writing Characters, Integers, Floating Point Numbers and Strings; Formatted I/O.

Control Structures: Decision Making (Branching) Structures – If Statement, If-Else Statement, Nested If-Else Statement, Else-If Ladder, Switch Statement, Goto Statement; Looping Structures – While Statement, Do-While Statement, For Statement, Continue and Break Statements.

Functions: Library Functions; User-Defined Functions; Function Declaration (Prototype) and Function Definition; Function Arguments – Dummy, Actual and Formal Arguments; Local and Global Variables; Function Calls – Call by Value and Call by Reference; Returning Multiple Values from a Function, Recursion and Recursive Functions, Storage Class & Scope of Variables – Automatic Storage, Extern Storage, Static Storage, Register Storage.

Arrays, Strings & Pointers: Single Dimensional Arrays; Accessing Array Elements; Initializing an Array; Multidimensional Arrays; Initializing Multidimensional Arrays; Memory Representation; Accessing Multidimensional Array Elements; Array of Characters; String Manipulation Functions; Introduction to Pointers; Pointer Variable Declarations and Initializations; Null Pointer; Constant Pointers; Void Pointer; Pointer Operators; Pointer Arithmetic; Application of Pointers; Dynamic Memory Allocations – malloc, calloc, realloc and free functions; Implementation of One Dimensional Array Using Pointers; Implementation of Two Dimensional Array Using Pointers – Array of Pointers and Pointers to Arrays Representations; Pointers and Strings; De-Referencing Pointers; Pointer to Functions.

Structure, Union, Enumeration and Files: Structure Declaration and Initialization; Accessing Structure Members, Structure Assignments; Array of Structures and Arrays within Structures, Nested Structures; Structure as Function Arguments; Structure Pointer; Unions; Difference between Structure and Union; Bit-Fields; Introduction to File; Text and Binary Files; Defining, Opening and Closing Files; I/O Operations on Files, Error Handling During I/O Operations, Random Access to Files, Command Line Arguments.

Text / Reference Books:

- 1. E. Balagruswamy, 2008; "Programming in ANSI C", 3rd Ed., Tata McGraw Hill.
- 2. R. G. Dromey, 2007; "How to Solve by Computer", 5th Ed., Pearson Education.
- 3. Deitel & Deitel, 2007; "C How to Program", 3rd Ed., Pearson Education, 2007

PAPER CODE	-	BIN106
PAPER TITLE	-	COMPUTING LAB – I (PRACTICAL): C AND SOFTWARE
		PACKAGES

SEMESTER – II

PAPER CODE - BIN201 PAPER TITLE - BIOINFORMATICS – II

Basic Concept : DNA, RNA, Genetic Code, Genes, ORFs, Motifs, Proteins, Translation, Transcription, Junk DNA, Tandem Repeat, The Central Dogma,

Biological databases: Primary and Secondary Database, File Formats, Accession Number, Gi Number, Data Submission Tools for DNA, BankIt, Sequin, Webin, Sakura, Data Submission Tool for Proteins, SPIN, AutoDep, Sequence Retrieval Tool, Entrez, Nucleotide Database, GenBank, EMBL, DDBJ, Protein Databases, SwissProt, Structure Database, PDB, MMDB, CATH, DALI, SCOP, FSSP, Other Databases, KEGG, MEROPS, BRENDA, ORF Finding Tool- ORF Finder.

Sequencing Maps: Maps and Sequences, RFLPs, SNPs, Contigs, Restriction Maps, Graphs, Interval Graphs, Genetic Mapping, Genetic Mapping Problem, Physical Maps and Clone Library, Single Digest, Multiple Digest, Double Digest Problem, Travelling Salesman Problem, Partial Digest Problem, Mapping with Unique Probes, C1P Problem.

Genome Sequencing and Genome Maps: Shotgun Sequencing, Sequencing by Hybridization, DNA Arrays, Shortest Superstring Problem, Overlap, Layout, Consensus.

Sequence Comparison: Similarity and Distance, Sequence Similarity, Homology, Sequence Alignment, Pairwise Alignment, Dynamic Programming, Local Sequence Alignment, Global Alignment, Longest Common Subsequence Problem, Heuristic Alignment Algorithms.

Multiple Sequence Alignment: Definition, Multiple Sequence Alignment using Sum of Pair Methods, Progressive Alignment Method, Multiple Alignment via Dot-Matrix Multiplication, Graph Consistency Problem. Multiple Alignment by Profile HMM Training.

Text / Reference Books:

- 1. Setubal, Meidanis, "Introduction to Computational Biology", Thomsom Brooks, Cole.
- 2. Computational Molecular Biology: An Algorithmic Approach, Pavel L. Pevzner.
- 3. Biological Sequence Analysis : Probablilistic Models of Proteins and Nucleic Acid, R. Durbin et al., Publisher: Cambridge University Press
- 4. Bioinformatics: Databases, Tools, and Algorithms, Orpita Bosu, Simminder Kaur Thukral, Publisher: Oxford University Press

PAPER CODE- BIN202 PAPER TITLE- COMPUTATIONAL METHODS FOR SEQUENCE ANALYSIS

Pair-wise Alignment: The Dotplot and Sequence Alignment, Dynamic programming: Local Alignment (Smith Waterman algorithm) and Global Alignment (Needlemann-Wunch Algorithm), Semi global alignment. Gaps and Indels .Alignment Score and Gap Penalties. Searching Sequence Database by Sequence Similarity Criteria. Database Searches: FASTA, BLAST and its Variants.

Statistical Analysis of Alignment Score and Statistical Significance of Search Result. E-Value, Pvalue and Z-score., Multiple Sequence Analysis:Scoring a Multiple Alignment. Minimum Entropy, SP Scores. HMM based Multiple Sequence Alignment. Applications of MSA.

Scoring Matrices: Derivation of PAM and BLOSUM Matrices, Comparison and their Advantages. Relationship Between BLOSUM and PAM .Selection of Scoring Matrices for Different Types of Alignment. Searching using Motifs and Patterns.

3 Protein secondary structure predictions: Prediction of Secondary Structures of Proteins using at Least Different Methods with Analysis and Interpretation of theRresults. Chou and Fasman Method. Comparison of the Performance of the Different Methods for Various Classes of Proteins

Phylogenetic Analysis: The Structure and interpretation of phylogenetic Trees. Phenotypic and Molecular phylogeny. Types of Trees: Rooted, Unrooted, Additive, Ultrametric, Condensed, Consensus tree. Phylogenetic tree reconstruction. Cladogram, Similaraties and Distances. Distance and Character Based tree Construction Methods. Orthologs, Paralogs and Xenologs. Tree evaluation. Computer tools for phylogenetic Analysis, MEGA, GROWTREE, PAUP, PHYLIP. Construction and visualization of phylogenetic tree. Application of Phyolgenetic Analysis.

Markov Chains and Hidden Markov Model: Markov Chains, Hidden Markov Model., Viterbi Algorithm. The Forward and Backward Algorithm. Pairwise alignment using HMMS, Profiles, HMMER.

Structure Analysis: Covariance models: SCFG-based RNA profiles. Application of RNA Structure Prediction. Protein folding problem, Affinsen hypothesis. Protein structure alignment. Protein structure analysis, Accessible surface area, Solvent content, contact map and Temperature factor analysis.

Text / Reference Books:

- 1. Andreas D. Baxevanis, B.F. Francis Ouellette, 2006; "Bioinformatics, A Practical Guide to the Analysis of Genes and Proteins", Wiley
- 2. S.C. Rastogi, N. Mendiratla and P. Rastogi, 2004; "Bioinformatics methods and applications-Genomics, Proteomics and Drug Discovery", Prentice Hall (India)
- 3. Richard Durbin, Sean Reddy, Andes Krogh, Graeme Mitchison, 2008; "Biological Sequence Analysis Probability Models of Proteins and Nucleic Acids", Cambridge University Press.
- 4. David. W Mount, 2005; "Bioinformatics: Sequence and Genome Analysis", CBS.

PAPER CODE-BIN203 PAPER TITLE- MOLECULAR IMAGING AND DRUG DESIGN

Biomolecular Simulations: Basic concepts, Geometry optimization algorithms: Steepest descent, Conjugate Gradient. Various Simulation Techniques: MD, Monte Carlo, Docking Strategies etc. Molecular Mechanics, Conformational Searches. Units and Derivatives, Force field and Energy Landscape, Introduction, Minimization and Algorithms, Solvation; Introduction, Periodic Boundary Condition, Ewald summation, Implicit solvent Model and continuum Electrostatics, Introduction, Replica-exchange simulations, Restraint potentials, Free Energy Calculations, Membrane Simulations.

Molecular Modeling: Elements of Molecular Modelling. Basic Principles of Modeling. Modelling Softwares; Sali Modeler, Swiss-Modeler, GenThreader, PROSPECTOR, FFAS03.Abinitio modeling, Combined Modeling Approaches. Prediction of Tertiary Structures of Proteins using Homology Modeling approach: SWISSMODEL, SWISS-PDB Viewer. Minimization of a Peptide Energy using Appropriate Force Field. Ramachandran Plot, Torsional Space Minimization, Energy Minimization in Cartesian Space, Molecular Mechanics-Basic Principle, Molecular Dynamics Basic Principles.

Modeling Genetic Networks: Modeling, Types of Modeling, Modeling the Activity of a Single Gene, Gene Regulatory Network(GRN) Understanding Gene Regulation, Understanding the Biology, Biochemical Processes; Transcription, Exons & Introns, Splicing, Translation, Post Ttranslation Modification. Overview of Models; Boolean, Differential Equation, Stochastic Models, Kinetic Logic Model etc.

Drug design: Drug discovery process, Role of Bioinformatics in Drug Design, Target

Identification and Validation, Lead Optimization and Validation. Structure Activity Relationship QSARs and QSPRs, QSAR Methodology, Various Descriptors used in QSARs: Electronics; Topology; Quantum Chemical based Descriptors. Use of Genetic Algorithms, Neural Networks and Principle Components Analysis in the QSAR equations.

Structure-Based Drug Design and Ligand Based Drug Design: Introduction to Docking, Principle of Rigid and Flexible Docking, Lock and key Hypothesis, Ligand Docking into Active sites of Proteins. High Throughput Screening, Virtual Screening, Protein-Ligand Docking.

ADMET in Drug Design, Lipinski" rule of 5", Principle and Methods of Docking, Ligand basedDrug Designing, Target identification, Drug Metabolism, Denovo design of Drug. Protein-Protein Docking, Energy Minimization. Various Available Docking Software.

Text / Reference Books:

- 1. D. M, 2000; "Protein structure prediction: methods and protocols", Webster, Totowa Humana Press.
- 2. Andrew R. Leach, 2009; "Molecular Modelling: Principal and Applications", Longman Singapore Publisher.
- 3. Richard S. Larson, 2006; "Bioinformatics and drug discovery", Humaana Press.
- 4. Daniel I. Chasman, 2007; "Protein structure: determination, analysis and applications for drug discovery", CRCpress.
- 5. P.E. Bourne and B.H. Wessing, 1999; "Structural Bioinformatics", Wiley VCH, 1999
- 6. S.C. Rastogi, N. Mendirtla and P.Rastogi, 2004; "Bioinformatics Methods and Applications Genomics, Proteomics and Drug Discovery", Prentice Hall (India).4

PAPER CODE- BIN204 PAPER TITLE- DATABASE MANAGEMENT SYSTEM

Basic Concepts: Data, Database and Database Management System (DBMS); Database vs. Traditional File System Approach; Three Tier Architecture of DBMS and Data Independence; Categories of Database Management Systems: Hierarchical, Network and Relational Database Systems.

Entity Relationship Diagram: Data and Database Models, Categories of Database Models, Entity relationship (ER) Model – Basic Concepts and their Representations – Entity, Entity Type and Entity Set, Attributes and Keys, Relationships, Relationship Types, and Structural Constraints, Weak Entity, Naming Conventions & Design Issues in ER Model. ER Diagrams.

Relational Database Model: Structure of Relational Model; Domains, Attributes, Tuples, and Relations; Characteristics of Relations; Relational Constraints – Domain Constraints, Key Constraints, Entity Integrity, and Referential Integrity Constraints; Relational Database Schema;

Functional Dependencies and Normalization: Informal Design Guidelines for Relation Schemas; Functional Dependencies; Inference Rules for Functional Dependencies; Normalization using Functional Dependencies – First Normal Form (INF), Second Normal Form (2NF), Third Normal Form (3NF), and Boyce-Codd Normal Form (BCNF);

Relational Algebra and SQL: Relational Algebra Operations – Select, Project, Rename, Union, Intersection, Set Difference, Join, and Division Operations; Aggregate Functions and Groupings. SQL Concepts, Schema and Table Deletion; Table Modification; Insert, Delete, and Update Statements; SELECT-FROM-WHERE Structure; Renaming Attributes; Nested Queries and Set Comparisons; EXISTS and UNIQUE Functions; Aggregate Functions; Creating and Updating Views.

Data Mining: Introduction to Data Mining, Knowledge Discovery from Databases (KDD), KDD Steps, DM Functionalities, Data Preprocessing - Data Cleaning, Data Integration and Transformation, Data Reduction, Discretization and Concept Hierarchy Generation. Classification Rule Mining Techniques – Decision Tree and Naïve Bayes Classification. Association Rule Mining – Itemsets, Support, Confidence, Apriori Algorithm. Clustering Techniques.

Text / Reference Books:

- 1. R. Elmasri and S. B. Navathe, "Fundamentals of Database Systems", 5th Ed., Pearson Education.
- 2. J. Han and M. Kamber, "Data Mining Concepts and Techniques", Morgan Kaufman Publisher
- 3. K. Loney and G. Koch. "ORACLE 9i The Complete Reference", Tata McGraw-Hill.

PAPER CODE-BIN205 PAPER TITLE-PERL PROGRAMMING AND APPLICATIONS TO BIOINFORMATICS

Introduction: History of Perl, Availability, Support, Versions, Installation. Basic Concepts, Significance of Perl in Bioinformatics.

Basics Constructs: Scalar Data, Numbers, Strings, Scalar Operators, Scalar Variables, Scalar Operators and Functions. Arrays and List Data: What is a List or Array? Literal Representation, Variables, Array Operators and Functions, Scalar and List Context. Hashes: What is a Hash? Hash Variables, Literal Representation of a Hash, Hash Functions, Hash Slices. Control Structures: Statement blocks, Loops and conditions. Basic Input / Output.

Advance Constructs and Features: Regular Expressions: Concepts About Regular Expressions, Simple Uses of Regular Expressions, Patterns, Matching Operator, Substitutions, The split and join functions, Subroutines: System and User Functions, The local Operator, Variable-length, Parameter Lists, Lexical Variables, File handles and File Tests: File Handle, Opening and Closing a File handle, Using Pathnames and Filenames, die, Using File handles.

Object-Oriented Perl: Introduction to Modules, Creating Objects and References. CGI Programming: The CGI.pm Module, CGI Program in Context, Simple CGI Programs, Passing Parameters via CGI, Perl and the Web.

Application of Perl in Bioinformatics: Representing String and Sequenced Data in Perl,

Program to Store a DNA Sequence, Concatenating DNA Fragments, DNA to RNA Transcription, Reading Proteins from File, Finding Motifs, Counting Nucleotides, Explodind Strings into Arrays, Operating on Strings, Generating random DNA, Analyzing DNA, Genbank, Sequence and Annotation, Parsing PDB Files, Parsing BLAST Output, Reading DNA from Files in FASTA format.

BioPerl: BioPerl Overview and Installation Procedures; Fundamental Constructs and Special Features; BioPerl Modules, Creating BioPerl Objects. Applications of BioPerl, Utility and its Applications.

Text / Reference Books:

- 1. James D. Tisdall, 2001; "Beginning Perl for Bioinformatics", O'Rilley and Association.
- 2. Cynthi Gibas & Per Jamesbeck, 2000; "Developing Bioinformatics Computer Skills", O'Rilley & Association.
- 3. Harshawardhan P Bal, 2003; "Perl Programming for Bioinformatics", Tata McGraw Hill.
- 4. Randal L. Schwartz and Tom Phoenix, 2005; "Learning Perl", 3rd Edition, O'Rilley.

PAPER CODE- BIN206 PAPER TITLE -COMPUTING LAB – II (PRACTICAL): SEQUANCE ALIGNMENT AND DATABASE SEARCH TECHNIQUES.

SEMESTER -III PAPER CODE - BIN301 PAPER TITLE - PROJECT

Topics and supplements are to be prescribed by concerned teacher.