Semester-1

S.No.	Course No.	Subject	Periods			Evaluation Scheme		
						Sessionals		
			L	T	P	Th	IA	Total
1.	BIN - 131	Cell biology	4	0	4	80	20	100
2.	BIN - 132	Biomolecule and metabolism	4	0	4	80	20	100
3.	BIN - 133	Microbiology	4	0	4	80	20	100
4.	BIN - 134	Biostatistic	4	0	4	80	20	100
5.	BIN-135	Molecular Biology	4	0	4	80	20	100
6	BIN-136	Communicative skills						25
7	BIN-137	Lab Course-1			4x3			50
8	BIN-138	Lab Course-2						50

Total= 625

Semester-2

S.No.	Course No.	Subject	Periods			Evaluation Scheme		
						Sessionals		
			L	T	P	Th	IA	Total
1.	BIN - 231	Immunology	4	0	4	80	20	100
2.	BIN - 232	Bioinformatics	4	0	4	80	20	100
3.	BIN - 233	Enzymology	4	0	4	80	20	100
4.	BIN - 234	Human Physiology and Developmental Genetics	4	0	4	80	20	100
5.	BIN-235	Nano Biotechnology	4	0	4	80	20	100
6	BIN-236	Seminar						25
7	BIN-237	Lab Course-1			4x3			50
8	BIN-238	Lab Course-2						50

Total= 625

3rd Semester

S.No.	Course No.	Subject	Periods			Evaluation Scheme		
							Sessionals	
			L	T	P	Th	IA	Total
1.	BIN - 331	Taxonomic & Phylogenic analysis	4	0	4	80	20	100
2.	BIN - 332	Object Oriented and Relational Databases	4	0	4	80	20	100
3.	BIN - 333	Tools & techniques for biological data mining	4	0	4	80	20	100
4.	BIN - 334	Drug design and chemoinformatics	4	0	4	80	20	100
5.	BIN-335	Biological Databases and Data Analysis	4	0	4	80	20	100
6	BIN-336	Seminar			4x3			25
7	BIN-337	Lab Course-1			4x3			50
8	BIN-338	Lab Course-2						50

Total= 625

4TH Semester

S.No.	Course No.	Subject	Periods			Evaluation Scheme		
							Sessionals	
			L	T	P	Th	IA	Total
1.	BIN - 431	Advanced Techniques for Sequence and Structure Analysis	4	0	4	80	20	100
2.	BIN - 432	Comparative Genomics and Proteomics	4	0	4	80	20	100
3.	BIN - 433	Structural Biology & Molecular Modeling	4	0	4	80	20	100
4.	BIN - 434	Programming in object oriented languages	4	0	4	80	20	100
5.	BIN-435	Project work				-	-	100
6	BIN-436	Seminar						25
7	BIN-437	Lab Course-1			4x3			50
8	BIN-438	Lab Course-2			4x3			50

Total= 625

BIM. 301 (T+P) Taxonomic & Phylogenic analysis

Theory

Syllabus:

Unit -1

• Basic concepts in Systematics, Taxonomy and Phylogeny: Species concept; kingdom to species; the five kingdoms; classical, phenetic and cladistic approaches; taxonomic information on viruses, microbes, plants & animals .

Unit- 2

• Concepts in Molecular Evolution

• Nature of data used in Taxonomy and Phylogeny: Morphological and molecular character data

• Phylogenetic trees and their comparison: Definition and description, various types of trees; Consensus (strict, semi-strict, Adams, majority rule, Nelson). Data partitioning and combination. Tree to tree distances, similarity

Unit 3

• Probabilistic models and associated algorithms

      o Probabilistic models of evolution

      o Maximum likelihood algorithm

• Phylogenetic analysis algorithms

      o Maximum Parsimony

      o Distance-based: UPGMA, Transformed Distance, Neighbors-Relation, Neighbor-Joining

Unit-4

• Approaches for tree reconstruction

      o Character optimization; delayed and accelerated transformation.

      o Reliability of trees. Bootstrap, jackknife, decay, randomization tests.

Unit-5

• Applications of phylogeny analyses

      o Comparison of Phylogenetic Trees obtained using DNA seq. Vs. protein seq. Vs. Full genomes. Need for addition of other properties towards total phylogenetic analysis

      o Comparative methods for detection of species / organism relationships

      o Gene duplication, Horizontal transfer, Domain evolution

      o Study of co-evolution: Plant-insect interactions. Host-parasite interactions.

      o Viral evolution.

Practicals

Syllabus:

• Explore & study the taxonomic resources available for

       o Animals, Plants, Microbes, Viruses etc

• Numerical taxonomy & Bacterial identification using matrices

• Viral identification: ICTV key & probabilistic

• Survey of software programs available for phylogenetic analysis

       o Installation of at least 2 public domain packages for both Windows & Unix environment: Phylip, PAUP, MEGA

• MSA using ClustalW: writing files for phylogenetic analysis

• Exploring Phylip package, its utilities and process flow

• Preparing input files using sample datasets

       o Numerical, binary & molecular data

       o Formats & format conversions

• Reconstruction of phylogenetic tress using molecular data (at least 2 datasets)

       o Distance-based methods: UPGMA, Neighbor-joining, Neighbor-relations & Transformed distance

       o Maximum Parsimony

       o Maximum likelihood

• Using bootstrapping tool to generate multiple datasets from the original input data & generation of consensus tree

• Plotting, visualizing & printing phylogenetic trees: TreeView and other tools

       o Various rendering

       o Formatting & labeling

       o Interpretation of trees

• Reconstruction of phylogenetic trees using whole genome data of viruses

References:

• Fundamentals of Molecular Evolution by D. Graur and W-H Li, 2nd Edition, Sinauer Associates.

• Molecular Evolution a Phylogenetic Approach by R. D. M. Page and E.C. Holmes, Blackwell Scientific, 1998.

• Protein Evolution by L. Patthy, Blackwell Scientific, 1999.

• Practical taxonomic computing by Pankhurst, R.J., 1991

BIM 302 (T) Object Oriented and Relational Databases
Theory
Syllabus:

Unit 1

Basic Concepts

      o Database System Concepts and Architecture
      o Entity-Relationship Model
      o EER and Object Modeling

  Relational Databases
      o Relational Model, Algebra, Calculus
      o ER- and EER-to-Relational Mapping
      o Relational languages SQL and QBE
      o RDBMS Systems: SQL server and MS Access

Unit-2

• Object-Oriented Database Systems
      o Object-oriented concepts
      o Object Modeling
      o Object-Oriented Databases
      o Object Database Languages
      o Object Database Design
      o Object-Relational and Extended Relational Database Systems

• Database Design
      o Functional Dependencies
      o Normalization
      o Design Algorithms and Further Dependencies

Unit-3

• Part 5: System Implementation Techniques
      o Query Processing and Optimization
      o Transaction Processing
      o Concurrency Control
      o Recovery
      o Security and Authorization

• Distributed Databases and Client-Server Architecture
      o Distributed Database Architecture
      o Data Fragmentation, Replication, and Allocation
      o Distributed Query Processing
      o Distributed Concurrency Control
      o Client-Server Architecture

Unit-4

• Active Databases
      o Event Languages, Event Detection and Delivery
      o Rule Processing

  o Applications of Active Databases
  o Data Management in Mobile Computing
  o Caching and Prefetching
  o Data Replication
  o Speculative Data Dissemination and Broadcast Disk
  o Mobile and Distributed Query Processing
  o Mobile Transactions

Unit-5

• Multimedia Databases
  o Multimedia Networking
  o Multimedia Storage and Indexing
  o Content-based Multimedia Information Retrieval
• Data Warehousing and Data Mining
  o Data Warehousing
  o Data Mining
  o OLAP
• Database Systems and the World-Wide-Web)
  o Connecting Database to the Web
  o Web Search
  o XML and the New Generation Web

References:

• Database Management and Design by G.W. Hansen and J.V. Hansen; Prentice-Hall of India
• Database System Concepts by A. Silberschatz, H.F. Korth and S. Sudarshan; McGraw-Hill
• Database Systems: The Complete Book by Garcia-Molina, J. D. Ullman, and J. Widom; Prentice Hall.
• Fundamentals of Database Systems by Ramez Elmasri and Shamkant B. Navathe, Addison-Wesley.
• Database Management Systems by R. Ramakrishnan and J. Gehrke; McGraw-Hill
• Database Systems by T. Connolly and C. Begg.; Addison-Wesley.
• Database Management Systems by A.K. Majumdar and P. Bhattacharyya; Tata McGraw-Hill
• An Introduction to Database Systems by C.J. Date, Addison-Wesley.
• Fundamentals of Database Systems by R. Elmasri and S.B. Navathe; Addision-Wesley.
• Modern Database Management by R.F. McFadden and J.A. Hoffer, Benjamin-Cummins (Narosa).
• Data structures, algorithms, and object oriented programming by Heileman, G.L New Delhi, Tata McGraw-Hill Publishing Company Limited, 2002.

BIM 303(T+P) Tools & Techniques for biological data mining

Theory
Syllabus:

Unit-1

• Quality of Biological Data & Data Accuracy

• General issues regarding Biological Databases; Representation of errors due to (machines, 3D structural and sequence data of proteins and nucleic acid, Proteomics and Microarray data)

Unit-2

• Optimization Techniques

- Steepest Descent, Conjugate Gradient, Newton-Raphson
- Simulated annealing in Biomolecular Structure Optimization

- Genetic Algorithms
- *Ab initio* methods for structure prediction
o Lattice, SOM, etc.
o Information theory, entropy and relative entropy
o Stochastic Grammars & Linguistics

Unit-3

- Clustering & Classification Algorithms
o Hierarchical and non-hierarchical Clustering
o K-Means clustering
o Grid based clustering
o Analysis of MD trajectories
o Microarray and Protein Array data Analysis

Unit-4

- Dynamic Programming & application in
o Sequence Alignments
o Structure Alignments
- Foundations for Machine learning Techniques:
o Hidden Markov Model
o Neural Network
o Bayesian modeling
o The Cox-Jaynes Axiomes

Unit-5

- Support Vector machine & Ant colony optimization applied to
o Multiple Sequence Alignments
o Biomolecular Structure Prediction
- Fuzzy logic system & application in
o Clustering and classifications
o Microarray and Protein Array data Analysis

Practicals
Syllabus:

- Neural networks:
o Use of neural network tools like BrainBox, MATLAB etc.
o Estimator of transition probabilities for markov models based on various sample sizes.
o Hidden Markov model implementation in C/ Java.
o Ant colony algorithm for the Travelling salesman problem (TSP) & Implementation of SVM
- Optimization Algorithms:
o Programming in 'C' for implementation of Golden section algorithm, Steepest descent, Newton Raphson, Conjugate gradient etc for energy minimization applications.
o Implementation of random walk and Monte Carlo algorithm.
o Simulated Annealing algorithm for energy minimization.

References:

• Data Mining: Concepts and Techniques by Han and Kamber, Morgan Kaufmann.
• Machine Learning by Tom Mitchell, McGraw Hill.
• Data Mining: Practical Machine Learning Tools and Techniques by Witten and Frank, Elsevier.
• Biological Sequence Analysis: probabilistic models of proteins and nucleic acids by Durbin, R., Eddy, S., Krogh, A. & Mitchison, G.Cambridge Univ. Press, 1998.
• Optimization Theory and Application by Rao, S.S., 1984.
• Discrete optimization by Parker, R. G. & Rardin, R. L., 1988.
• Stochastic simulation by Repley, Brian D, Wiley series, 1987.
• Methods of microarray data analysis III by Johnson, K.F. & Lin, S.M. Boston. Kluwer academic publishers, 2003.
• Exploration and analysis of DNA microarray and protein array data by Amaratunga, D. & Cabrera, J.New Jersey. John Wiley & Sons Inc., 2004.
• Ant colony optimization by Dorigo, Marco & Stutzle, Thomas New Delhi, Prentice-Hall of India Pvt Ltd, 2004.
• Data mining: introductory and advanced topics by Dunham, M.H.:New Delhi, Pearson Education, 2003.
• An introduction to bioinformatics algorithms by Jones, Neil.C. & Pevzner, Pavel A. New Delhi, Anne Books, 2005.

BIM 304 (T + P) : Chemoinformatics and Drug design
Theory
Syllabus:
Unit-1

• Role of Chemoinformatics in pharmaceutical/chemical research
    •Integrated databases
    •HTS analysis
    •Ligand based design of compounds
    •Structure based design of compounds
     • Structure representation systems, 2D and 3D structures
     • General introduction to chemical structure-hybridization, tetrahedron geometry etc
     • The degeneracy of isomeric SMILES and introduction to unique SMILES. Reaction transformations notation like SMIRKS.

Unit-2

•Introduction to graph theory, vertex partitioning algorithms- Morgan's and CANGEN algorithms and canonical labeling of the symmetrical vertex

• Introduction to conformation generating methods. Various ring conformation (sugar) and ring closure problem. Method to identify SSR (smallest subset of ring)

• Internal co-ordinates and introduction to calculation of Z matrix of simple small organic molecules.

Unit-3

• Chemical Databases – Design, Storage and Retrieval methods

• Introduction to database filters, property based & (drug-like)-Lipinski Rule of Five

• Search techniques, similarity searches and clustering

   Introduction to molecular pattern finding language- SMARTS

   o Introduction to distance measurement methods from the bit-strings of fingerprints-Tanimoto index and Tversky Index

   o General introduction to clustering- K means and Hierarchical clustering of chemical database

   o Diversity analysis- BCUT descriptors

Unit-4

• Modeling of small molecules and methods for interaction mapping

• Chemical properties 2D and 3D

   o Introduction to adjacency, distance matrix and use of these matrices for calculating Weiner Index, Hosoya Index, Balban Index, Shultz Index, Randic Index .
   o Introduction to shape indices- Kappa Shape index and calculation of molecular shape.

• Characterization of chemicals by Class & by Pharmacophore.

   o Introduction to pharmocophore

   o Identification of pharmacophore features

   o Building pharmacophore hypothesis

   o Searching databases using pharmocophores

Unit-5

• Design & Analysis of combinatorial libraries

   o Reagent and product base combinatorial library generation

   o Focus library and HTS library

• Chemoinformatics tools for drug discovery

   o Integration of active drugs

   o Optimization techniques

   o Filtering chemicals

o In silico ADMET; QSAR approach, Knowledge-based approach

Practicals

Syllabus:

• Importance of storing chemical in the form of graph, linear notation (SMILES,WLN, ROSDAL-
with special emphasis on SMILES and stereochemistry- both optical and geometrical
isomerism), connection tables-sd and mol files. (1)

• Graph data structure and its implementation in the context of chemistry.

• Importance of 3D structure and methods available for 3D structure generation- CORINA and
CONCORD

• A brief introduction to database (ISIS Base) with special emphasis on the storage of chemical in
the database format.

References:

• Chemoinformatics by Johann Gasteiger and Thomas Engel, 2004.

• An introduction to Chemoinformatics by Andrew R. Leach and Valerie J. Gillet, Kluwer
Academic Publisher, 2003.

• Handbook of Chemoinformatics. From Data to Knowledge by Johann Gasteiger.

• Chemometrics and Chemoinformatics by Barry K. Lavine, ACS Symposium series 894.

• Molecular modelling and prediction of bioactivity by Gundertofte, K. & Jorgensen, F.S. New
York. Kluwer academic publishers, 2000.

BIM 305 (T+P): Biological Databases and Data Analysis

Theory
Syllabus:

Unit-1

• Nature of biological data

• Overview of available Bioinformatics resources on the web

      o NCBI/EBI/EXPASY etc

• Biological Databases: Nucleic acid sequence databases

      o GenBank/EMBL/DDBJ

• Biological Databases: Protein sequence databases

      o PIR-PSD

      o SwissProt, UniProtKB

Unit-2

• Database search engines

      o Entrez

      o SRS

• Overview/concepts in sequence analysis

• Pairwise sequence alignment algorithms
      o Needleman & Wunsch

      o Smith & waterman

Unit-3

• Scoring matrices for Nucleic acids and proteins

      o MDM

      o BLOSUM

      o CSW

• Database Similarity Searches
      o BLAST

      o FASTA

Unit-4

• Multiple sequence alignment

     o PRAS

     o CLUSTALW

   • Biological databases: Genome & genetic disorders

     o Genome databases: Human, model organisms, microbes & viral

     o OMIM

Unit-5

   • Biological databases: structural databases
     o PDB

     o NDB

     o CCSD

• Derived databases

     o Prosite

     o BLOCKS

     o Pfam/Prodom

Practicals
Syllabus:

• Exploring the integrated database system at NCBI server and querying the PUBMED and GenBank databases using the ENTREZ search engine

• Exploring the integrated database system at EBI server and searching the EMBL Nucleotide database using the SRS search engine

• Exploring & querying SWISSPROT & UniProtKB

• Exploring and querying the PIR database

• Pair-wise global alignments of protein and DNA sequences using Needleman-Wunsch algorithm & interpretation of results to deduce homology between the sequences, use of scoring matrices

• Pair-wise local alignments of protein and DNA sequences using Smith-Waterman algorithm and interpretation of results

• Database (homology) searches using different versions of BLAST and interpretation of the results to derive the biologically significant relationships of the query sequences (proteins/DNA) with the database sequences

• Database (homology) searches using different versions of FASTA & interpretation of the results to derive the biologically significant relationships of the query sequences (proteins/DNA) with the database sequences.

• Multiple sequence alignments of sets of sequences using web-based and stand-alone version of CLUSTAL. Interpretation of results to identify conserved and variable regions and correlate them with physico-chemical & structural properties .

• Studying the format & content of structural databases & visualization of structures using Rasmol, Cn3D and other utilities.

References:

• Bioinformatics: A Practical Guide to the analysis of Genes and Proteins (2nd Ed.) by Baxevanis, A.D. & Ouellettee, B., F. F., New York, John Wiley & Sons, Inc. Publications, 2002.

• Introduction to Bioinformatics by Attwood, T.K. & Parry-Smith, D.J., Delhi, Pearson Education (Singapore) Pte.Ltd., 2001.

• Bioinformatics: Sequence and Genome Analysis by Mount, David, New York, Cold Spring Harbor Laboratory Press, 2004.

BIM 302 (T) Object Oriented and Relational Databases
Theory
Syllabus:

Unit 1

Basic Concepts

- Database System Concepts and Architecture
- Entity-Relationship Model
- EER and Object Modeling

Relational Databases
- Relational Model, Algebra, Calculus
- ER- and EER-to-Relational Mapping
- Relational languages SQL and QBE
- RDBMS Systems: SQL server and MS Access

Unit-2

• Object-Oriented Database Systems
- Object-oriented concepts
- Object Modeling
- Object-Oriented Databases
- Object Database Languages
- Object Database Design
- Object-Relational and Extended Relational Database Systems

• Database Design
- Functional Dependencies
- Normalization
- Design Algorithms and Further Dependencies

Unit-3

• Part 5: System Implementation Techniques
- Query Processing and Optimization
- Transaction Processing
- Concurrency Control
- Recovery
- Security and Authorization

• Distributed Databases and Client-Server Architecture
- Distributed Database Architecture
- Data Fragmentation, Replication, and Allocation
- Distributed Query Processing
- Distributed Concurrency Control
- Client-Server Architecture

Unit-4

• Active Databases
- Event Languages, Event Detection and Delivery
- Rule Processing

- o Applications of Active Databases
- o Data Management in Mobile Computing
- o Caching and Prefetching
- o Data Replication
- o Speculative Data Dissemination and Broadcast Disk
- o Mobile and Distributed Query Processing
- o Mobile Transactions

Unit-5
- Multimedia Databases
  - o Multimedia Networking
  - o Multimedia Storage and Indexing
  - o Content-based Multimedia Information Retrieval
- Data Warehousing and Data Mining
  - o Data Warehousing
  - o Data Mining
  - o OLAP
- Database Systems and the World-Wide-Web)
  - o Connecting Database to the Web
  - o Web Search
  - o XML and the New Generation Web

References:
- Database Management and Design by G.W. Hansen and J.V. Hansen; Prentice-Hall of India
- Database System Concepts by A. Silberschatz, H.F. Korth and S. Sudarshan; McGraw-Hill
- Database Systems: The Complete Book by Garcia-Molina, J. D. Ullman, and J. Widom; Prentice Hall.
- Fundamentals of Database Systems by Ramez Elmasri and Shamkant B. Navathe, Addison-Wesley.
- Database Management Systems by R. Ramakrishnan and J. Gehrke; McGraw-Hill
- Database Systems by T. Connolly and C. Begg.; Addison-Wesley.
- Database Management Systems by A.K. Majumdar and P. Bhattacharyya; Tata McGraw-Hill
- An Introduction to Database Systems by C.J. Date, Addison-Wesley.
- Fundamentals of Database Systems by R. Elmasri and S.B. Navathe; Addision-Wesley.
- Modern Database Management by R.F. McFadden and J.A. Hoffer, Benjamin-Cummins (Narosa).
- Data structures, algorithms, and object oriented programming by Heileman, G.L New Delhi, Tata McGraw-Hill Publishing Company Limited, 2002.

BIN 432 (T+P) Comparative Genomics and Proteomics

Theory
Syllabus:

Unit-1

• Objective and Overview of Genome Comparisons

• Genome Alignments

    o BLAST2

    o MUMmer

    o PipMaker

    o VISTA

Unit-2

• Comparison of Gene Order

    o Gene Order

• Comparative Genomics

    o Viruses

    o Microbes

    o Pathogens

    o Eukaryotes

• Comparative Genomics Databases

    o VirGen

    o CORG

    o HOBACGEN

    o Homophila

    o XREFdb

    o Gramene

Unit-3

• Single Nucleotide Polymorphism, db SNP and other SNP-related databases

• Overview of Proteomics

     o Experimental Techniques

     o Bioinformatics Approaches

Unit-4

• Protein-Protein Interaction Networks, databases and software

     o DIP (Database of Interacting Proteins
     o PPI Server

     o BIND - Biomolecular Interaction Network Database

     o PIM -Hybrigenics

Unit-5

     o Path Calling Yeast Interaction Database

     o MINT - a Molecular Interactions Database

     o GRID - The General Repository for Interaction Data sets

     o Inter Pre TS - protein interaction prediction through tertiary structure

Practicals
Syllabus:

• Explore comparative genomics resources and NCBI and EBI

• Comparison of full / partial genomic sequences using following methods to identify conserved genes and map/compare the annotations of the two sequences

     o BLAST2

     o MegaBLAST and Discontiguous MegaBLAST

     o MUMmer

     o PipMaker

   o VISTA

• Compare gene order of given genomic sequences using the GeneOrder tool

• Explore and query the comparative genomics databases: COG, VirGen, CORG, HOBACGEN, Homophila, XREFdb, Gramene etc.

• Explore and query SNP and SNP-related databases

• Explore and query the protein-protein interaction databases: DIP, PPI Server, BIND, PIM, PathCalling, MINT, GRID, InterPreTS

References:

• Proteomics of microorganisms by Hecker, M. & Mullner, S., Berlin, Springer-Verlag, 2003.

• Introduction to proteomics: tools for the new biology by Liebler, D.C. & Yates, J.R. III, New York. Humana Press, 2002.

• Proteomics: from protein sequence to function by Pennington, S. R. & Dunn, M. J.: New Delhi, Viva Books Private Ltd, 2002.

• Bioinformatics: sequence and genome analysis by David Mount, cold springer harbour press, 2004.

• Comparative genomics: empirical and analytical approaches to gene order dynamics, map alignment and the evolution of gene families by Sankoff, D. & Nadeau, J.H., Netherlands, Kluwer Academic Publishers, 2000.

BIM 303(T+P) Tools & Techniques for biological data mining

Theory
Syllabus:

Unit-1

• Quality of Biological Data & Data Accuracy

• General issues regarding Biological Databases; Representation of errors due to (machines, 3D structural and sequence data of proteins and nucleic acid, Proteomics and Microarray data)

Unit-2

• Optimization Techniques

  o Steepest Descent, Conjugate Gradient, Newton-Raphson

  o Simulated annealing in Biomolecular Structure Optimization

• Genetic Algorithms

• *Ab initio* methods for structure prediction

  o Lattice, SOM, etc.

  o Information theory, entropy and relative entropy

  o Stochastic Grammars & Linguistics

Unit-3

• Clustering & Classification Algorithms

  o Hierarchical and non-hierarchical Clustering

  o K-Means clustering

  o Grid based clustering

  o Analysis of MD trajectories

  o Microarray and Protein Array data Analysis

Unit-4

• Dynamic Programming & application in

- o Sequence Alignments

- o Structure Alignments

• Foundations for Machine learning Techniques:

- o Hidden Markov Model

- o Neural Network

- o Bayesian modeling

- o The Cox-Jaynes Axiomes

Unit-5

• Support Vector machine & Ant colony optimization applied to

- o Multiple Sequence Alignments

- o Biomolecular Structure Prediction

• Fuzzy logic system & application in

- o Clustering and classifications

- o Microarray and Protein Array data Analysis

Practicals
Syllabus:

• Neural networks:

- o Use of neural network tools like BrainBox, MATLAB etc.

- o Estimator of transition probabilities for markov models based on various sample sizes.

- o Hidden Markov model implementation in C/ Java.

- o Ant colony algorithm for the Travelling salesman problem (TSP) & Implementation of SVM

• Optimization Algorithms:

- Programming in 'C' for implementation of Golden section algorithm, Steepest descent, Newton Raphson, Conjugate gradient etc for energy minimization applications.

- Implementation of random walk and Monte Carlo algorithm.

- Simulated Annealing algorithm for energy minimization.

References:

• Data Mining: Concepts and Techniques by Han and Kamber, Morgan Kaufmann.

• Machine Learning by Tom Mitchell, McGraw Hill.

• Data Mining: Practical Machine Learning Tools and Techniques by Witten and Frank, Elsevier.

• Biological Sequence Analysis: probabilistic models of proteins and nucleic acids by Durbin, R., Eddy, S., Krogh, A. & Mitchison, G.Cambridge Univ. Press, 1998.

• Optimization Theory and Application by Rao, S.S., 1984.

• Discrete optimization by Parker, R. G. & Rardin, R. L., 1988.

• Stochastic simulation by Repley, Brian D, Wiley series, 1987.

• Methods of microarray data analysis III by Johnson, K.F. & Lin, S.M. Boston. Kluwer academic publishers, 2003.

• Exploration and analysis of DNA microarray and protein array data by Amaratunga, D. & Cabrera, J.New Jersey. John Wiley & Sons Inc., 2004.

• Ant colony optimization by Dorigo, Marco & Stutzle, Thomas New Delhi, Prentice-Hall of India Pvt Ltd, 2004.

• Data mining: introductory and advanced topics by Dunham, M.H.:New Delhi, Pearson Education, 2003.

• An introduction to bioinformatics algorithms by Jones, Neil.C. & Pevzner, Pavel A. New Delhi, Anne Books, 2005.

BIM 304 (T + P) : Chemoinformatics and Drug design
Theory
Syllabus:

Unit-1

• Role of Chemoinformatics in pharmaceutical/chemical research

   •Integrated databases

   •HTS analysis

   •Ligand based design of compounds

   •Structure based design of compounds

   • Structure representation systems, 2D and 3D structures

   • General introduction to chemical structure-hybridization, tetrahedron geometry etc

   •The degeneracy of isomeric SMILES and introduction to unique SMILES. Reaction transformations notation like SMIRKS.

Unit-2

   •Introduction to graph theory, vertex partitioning algorithms- Morgan's and CANGEN algorithms and canonical labeling of the symmetrical vertex

   • Introduction to conformation generating methods. Various ring conformation (sugar) and ring closure problem. Method to identify SSR (smallest subset of ring)

   • Internal co-ordinates and introduction to calculation of Z matrix of simple small organic molecules.

Unit-3

• Chemical Databases – Design, Storage and Retrieval methods

• Introduction to database filters, property based & (drug-like)-Lipinski Rule of Five

• Search techniques, similarity searches and clustering

   Introduction to molecular pattern finding language- SMARTS

   o Introduction to distance measurement methods from the bit-strings of fingerprints- Tanimoto index and Tversky Index

- General introduction to clustering- K means and Hierarchical clustering of chemical database

- Diversity analysis- BCUT descriptors

Unit-4

• Modeling of small molecules and methods for interaction mapping

• Chemical properties 2D and 3D

- Introduction to adjacency, distance matrix and use of these matrices for calculating Weiner Index, Hosoya Index, Balban Index, Shultz Index, Randic Index .
- Introduction to shape indices- Kappa Shape index and calculation of molecular shape.

• Characterization of chemicals by Class & by Pharmacophore.

- Introduction to pharmocophore

- Identification of pharmacophore features

- Building pharmacophore hypothesis

- Searching databases using pharmocophores

Unit-5

• Design & Analysis of combinatorial libraries

- Reagent and product base combinatorial library generation

- Focus library and HTS library

• Chemoinformatics tools for drug discovery

- Integration of active drugs

- Optimization techniques

- Filtering chemicals

- In silico ADMET; QSAR approach, Knowledge-based approach

Practicals

Syllabus:

• Importance of storing chemical in the form of graph, linear notation (SMILES,WLN, ROSDAL-with special emphasis on SMILES and stereochemistry- both optical and geometrical isomerism), connection tables-sd and mol files. (1)

• Graph data structure and its implementation in the context of chemistry.

• Importance of 3D structure and methods available for 3D structure generation- CORINA and CONCORD

• A brief introduction to database (ISIS Base) with special emphasis on the storage of chemical in the database format.

References:

• Chemoinformatics by Johann Gasteiger and Thomas Engel, 2004.

• An introduction to Chemoinformatics by Andrew R. Leach and Valerie J. Gillet, Kluwer Academic Publisher, 2003.

• Handbook of Chemoinformatics. From Data to Knowledge by Johann Gasteiger.

• Chemometrics and Chemoinformatics by Barry K. Lavine, ACS Symposium series 894.

• Molecular modelling and prediction of bioactivity by Gundertofte, K. & Jorgensen, F.S. New York. Kluwer academic publishers, 2000.

BIM 305 (T+P): Biological Databases and Data Analysis

Theory
Syllabus:
Unit-1
• Nature of biological data

• Overview of available Bioinformatics resources on the web

     o NCBI/EBI/EXPASY etc

• Biological Databases: Nucleic acid sequence databases

     o GenBank/EMBL/DDBJ

• Biological Databases: Protein sequence databases

     o PIR-PSD

     o SwissProt, UniProtKB

Unit-2
• Database search engines

     o Entrez

     o SRS

• Overview/concepts in sequence analysis

• Pairwise sequence alignment algorithms
     o Needleman & Wunsch

     o Smith & waterman

Unit-3

• Scoring matrices for Nucleic acids and proteins

     o MDM

     o BLOSUM

     o CSW

• Database Similarity Searches
     o BLAST

o FASTA

Unit-4

• Multiple sequence alignment

    o PRAS

    o CLUSTALW

   • Biological databases: Genome & genetic disorders

    o Genome databases: Human, model organisms, microbes & viral

    o OMIM

Unit-5

• Biological databases: structural databases
    o PDB

    o NDB

    o CCSD

• Derived databases

    o Prosite

    o BLOCKS

    o Pfam/Prodom

Practicals
Syllabus:

• Exploring the integrated database system at NCBI server and querying the PUBMED and GenBank databases using the ENTREZ search engine

• Exploring the integrated database system at EBI server and searching the EMBL Nucleotide database using the SRS search engine

• Exploring & querying SWISSPROT & UniProtKB

• Exploring and querying the PIR database

• Pair-wise global alignments of protein and DNA sequences using Needleman-Wunsch algorithm & interpretation of results to deduce homology between the sequences, use of scoring matrices

• Pair-wise local alignments of protein and DNA sequences using Smith-Waterman algorithm and interpretation of results

• Database (homology) searches using different versions of BLAST and interpretation of the results to derive the biologically significant relationships of the query sequences (proteins/DNA) with the database sequences

• Database (homology) searches using different versions of FASTA & interpretation of the results to derive the biologically significant relationships of the query sequences (proteins/DNA) with the database sequences.

• Multiple sequence alignments of sets of sequences using web-based and stand-alone version of CLUSTAL. Interpretation of results to identify conserved and variable regions and correlate them with physico-chemical & structural properties .

• Studying the format & content of structural databases & visualization of structures using Rasmol, Cn3D and other utilities.

References:

• Bioinformatics: A Practical Guide to the analysis of Genes and Proteins (2nd Ed.) by Baxevanis, A.D. & Ouellettee, B., F. F., New York, John Wiley & Sons, Inc. Publications, 2002.

• Introduction to Bioinformatics by Attwood, T.K. & Parry-Smith, D.J., Delhi, Pearson Education (Singapore) Pte.Ltd., 2001.

• Bioinformatics: Sequence and Genome Analysis by Mount, David, New York, Cold Spring Harbor Laboratory Press, 2004.

BIN 431 (T+P) Advanced Techniques for Sequence and Structure Analysis

Theory
Syllabus:

Unit-1

• Advanced Techniques for Sequence Analysis

    o Sequence Profiles: Derivation, Databases, Application

    o Gapped BLAST, PSI-BLAST, PHI-BLAST

Unit-2

• Advanced Techniques for Structure Analysis

    o Molecular replacement method, Direct method & Fiber diffraction

    o Methods for Comparison of 3D structures

• Application of the optimization techniques in

    o Sequence Alignments

    o Prediction of Protein Structure

    o Docking Simulations

Unit-3

• Advance techniques in Prediction of 3D Structure

    o Hidden Markov Model

    o Neural networks

    o Rossetta Stone

Unit-4

• Molecular Dynamics Simulations & Monte Carlo Methods

    o Electrostatics of biomolecules

    o Simulations of Bio-macromolecular Structures in Water & membrane

    o Free energy perturbation method

Unit-5

• Simulated Annealing

  o Multiple Sequence Alignments

  o Simulations of Bio-macromolecular Structures

• Designing of molecules like drug, inhibitors using

  o Structure based & ligand based docking methods

  o Different Scoring schemes

Practicals
Syllabus:

• Data mining:

  o Determine the entropy in a DNA sequence, relative entropy in 2 sequences etc.

  o Implement the clustering algorithm for biological data like Phylogenetic Analysis

  o MD trajectories & Microarray and Protein Array data Analysis

  o Program the UPGMA and EM algorithm for Clustering.

  o Dynamic programming application to structural alignment.

• Molecular simulation in water and membrane

• Characterization of active site, Substrate based design , docking & scoring of chemical data base .

References:

• Hidden markov models for Bioinformatics by Koski, T, Dordrecht Kluwer, Academic Publishers, 2001.

• Neural Networks: A Comprehensive Foundation (2nd Edition) by Simon Haykin.

• Genetic algorithms in search, optimization an machine learning by Goldberg, D.E. 2002 Delhi, Pearson Education Pte. Ltd.

• Molecular dynamics of lipid bilayers by Nagumo, M. & Bruce, P.G.,1996.

• Genetic algorithms in molecular modeling by Devillers, J., Ed, 1996.

• Molecular dynamics: applications in molecular biology by Goodfellow, Julia M., 1991.

• Biological Sequence Analysis: probabilistic models of proteins and nucleic acids by Durbin, R., Eddy, S., Krogh, A. & Mitchison, G.: Cambridge Univ. Press, Cambridge, 1998.

• Methods in Enzymology, (V. 183), Molecular evolution: computer analysis of protein & nucleic acid seqs. 1990.

• Sequence analysis in molecular biology: treasure trove or trivial pursuit by Heijne, Gunnar von, 1987.

• Simulation Modeling and Analysis by Averill M. Law d and W. David Kelton, Tata MacGraw – Hill, 2000.

• Protein Bioinformatics: An Algorithmic Approach to Sequence and Structure Analysis by Ingvar Eidhammer, Inge Jonassen, William R. Taylor John Wiley, 2003.

BIN 433(T) Structural biology & Molecular Modelling

Theory
Syllabus:

Unit-1

• Macromolecular Structure

Protein - Primary, Secondary, Supersecondary, Tertiary and Quaternary structure

Nucleic acid – DNA and RNA

Carbohydrates

3D Viral structures

Unit-2

• Overview of macromolecular x-ray crystallography

Methods to study 3D structure

Principles of crystallography

Co-ordinate systems

Fitting and refinement,Validation

Analysis of 3D structures

• Principles of protein folding and methods to study protein folding

• Structure of Ribosome

• Macromolecular interactions

Protein – Protein

Protein – Nucleic acids

Protein - carbohydrates

Unit-3

• Mass spectrometry and computational approaches in structural biology

• Overview of molecular modelling - Introduction and challenges

• Molecular modelling methods

Conformational searching, Potential energy maps,

Ramachandran maps, Ab-initio methods, Semi-empirical methods

Empirical methods , Conformational analysis

Introduction and Methods:

• Molecular fitting

• Energy Minimisation .

Unit-4

•Dynamics of Bio-macromolecules.

• Molecular visualization tools .Visualization of tertiary structures, quaternary structures, architectures and topologies of proteins and DNA using molecular visualization softwares such as RasMol, Cn3D, SPDBV, Chime, Mol4D, etc.

Unit-5

 Structure prediction tools and homology modeling.

Prediction of secondary structures of proteins using at least 5 different methods with analysis and interpretation of the results. Comparison of the performance of the different methods for various classes of proteins.

Internet based modeling tools.

Prediction of tertiary structures of proteins using Homology Modeling approach:

SWISSMODEL, SWISS-PDB Viewer

Prediction of tertiary structures of proteins using at least 3 methods for fold recognition along with analysis and interpretation of results

References:

• Structural Bioinformatics - Methods of biochemical Analysis V. 44 by Philip E. Bourne (Editor), Helge Weissig (Editor) New Jersey. Wiley-Liss, 2003.

• Principles of protein X-ray Crystallography by Jan Drenth, Springer-Verlag, 1994.

• Introduction to Protein Structure by Branden, Carl & Tooze, John, Garland Publishing, 1991.

• Molecular Modeling: Principles and Applications by Andrew Leach, Prentice Hall, 2001.

• Computational methods for protein folding : advances in chemical physics vol. 120 by Friesner, R.A. Ed., Prigogine, L. Ed. & Rice, S.A.New York. John wiley & sons, Inc. publication, 2002.

- Dynamics of Proteins and Nucleic Acids by J.A. McCammon and S.C. Harvey Cambridge University Press, 1087.
- Protein Structure: A Practical approach by Creighton T. E., 1989.
- Protein Folding by Creighton T., 1992.
- Protein Structure Prediction: A practical approach by Sternberg M.J.E., 1996.
- Molecular Modeling: Basic Principles and application by Hans Dieter and Didier Rognan. Wiley VeH Gmbh and Co. KGA, 2003.

BIN 434 (T+P) Programming in Object Oriented Languages

Theory

Syllabus:

Unit-1

• An introduction to JAVA programming

• Object-oriented programming and Java

• Java Basics

• Working with objects

• Arrays, Conditionals and Loops

• Creating Classes and Applications in Java

• Graphics, Fonts and Color

Unit-2

• Simple Animation and Threads

• Advanced Animation, Images and Sound

• Managing Simple Events and Interactivity

• Creating User Interfaces with AWT

Unit-3

• Windows, Networking and other Tidbits

• Modifiers, Access Control and Class Design

• Packages and Interfaces

• Exception

• Multithreading

• Streams and I/O

Unit-4

• Using Native Methods and Libraries

• Java Programming Tools

• Working with Data Structures and Java

• Image Filters

Unit-5

Oracle Architecture :

- Oracle objects - Tables, Views, Indexes, Sequences; Synonyms, Snapshots, Clusters
- Database - Table space, Data files, Blocks, Extents, Segments; Oracle Background Processes, control files; Oracle Memory Management; Rollback Segments; Redo logs/Archival; Security, Grants, Roles, Privileges
- Oracle Utilities & SQL *DBA - Oracle Server Manager; Export-Import/SQL Monitor Backup & Recovery (Archiving); Physical Storage & Logical Storage
- Oracle * Reports - Reports Features; Full Integration with Forms and Graphics; Data Model and layout editors

Practicals

Syllabus: ***Programs related to biological data to be implemented using:***

- Java Basics
- Working with objects
- Arrays, Conditionals and Loops
- More about methods
- Java Applets Basics • Graphics, Fonts and Color
- Simple Animation and Threads
- Advanced Animation, Images and Sound
- Managing Simple Events and Interactivity
- Windows, Networking and other Tidbits
- Modifiers, Access Control and Class Design
- Packages and Interfaces
- Using Native Methods and Libraries
- Java Programming Tools
- Working with Data Structures and Java
- Image Filters

References:

- Java In a Nutshell by David Flanagan, Oreilly Publications.
- Java Examples in a Nutshell by David Flanagan, Oreilly Publications.
- Java 2: The Complete Reference by Patrick Naughton and Herbert Schildt, McGraw Hill.