Department: Bioinformatics Centre, Savitribai Phule Pune University

Course: M. Sc. Bioinformatics

Duration: 2 years

Total Number of Credits: 88

Semester I		
Mandatory Courses		
Subject Code	**Subject Title**	**Number of Credits**
BIN 101 (T+P)	Biological Data & Databases	2T + 2P
BIN 102 (T+P)	Biological Sequence Analysis	2T + 2P
BIN 103 (T+P)	Biomolecular Structure & Organization	2T + 2P
BIN 104 (P)	Python Programming	2P
Total Mandatory Credits		6T + 8P =14 credits
BIN 105 (T)	Research Methodology	4T
Elective Courses ***Courses worth of 4 credits from among the following:***		
Subject Code	**Subject Title**	**Number of Credits**
BIE 106 (T)*	Mathematics & Statistics	2T
BIE 107 (T)*	Basic Biology	2T
BIE 108 (T)	Molecular Biology	2T
BIE 109 (T)	Biochemistry and Cell Biology	2T
Total Elective credits		4 credits
Semester Total		**22 credits**

* Depending upon the UG training background, the students coming from Biology, Chemistry and all streams other than Mathematics and Statistics may opt for the course BIE 106 Mathematics & Statistics and the students coming from Mathematics and Statistics streams may opt for BIE 107 Basic Biology.

Semester II

Mandatory Courses

Subject Code	Subject Title	Number of Credits
BIN 201 (T+P)	Structural Bioinformatics	2T + 2P
BIN 202 (T+P)	Chemoinformatics	2T + 2P
BIN 203 (T+P)	Genomics	2T + 2P
BIN 204 (P)	Introduction to Database Systems	2P
	Total Mandatory Credits	6T + 8P = 14 credits
BIN 205 (P)	On Job Training/ Field Project	4P

Elective Courses

Courses worth of 4 credits from among the following:

Subject Code	Subject Title	Number of Credits
BIE 206 (P)	Biological Data Curation and Analysis	2P
BIE 207 (P)	Advanced Techniques in Experimental Biology	4P
BIE 208 (T+P)	Molecular Phylogenetics	1T + 1P
BIE 209 (T+P)	Programming in Java	2T + 2P
BIE 210 (P)	Linux Operating System	2P
BIE 211 (T+P)	Software Testing	1T + 1P
	Total Elective credits	4 credits
	Semester Total	**22 credits**

Semester III		
Core Courses		
Subject Code	**Subject Title**	**Number of Credits**
BIN 301 (T+P)	Immunology and Immunoinformatics	1T + 1P
BIN 302 (T+P)	Comparative Genomics	2T + 2P
BIN 303 (T+P)	Molecular Modeling & Simulations	2T + 2P
BIN 304 (T+P)	Scientific Data Mining and Visualization	2T + 2P
	Total Core credits	7T + 7P = 14 credits
BIN 305 (P)	Project	4P
Elective Courses ***One course worth 4 credits from among the following::***		
Subject Code	**Subject Title**	**Number of Credits**
Group A		
BIE 306 (T)	Biology of Diseases	2T
BIE 307 (T+P)	Pathways & Networks	1T + 1P
BIE 308 (P)	Laboratory Course in Experimental Immunology	2P
BIE 309 (P)	Web Design	2P
	Total Elective credits	4 credits
	Semester Total	**22 credits**

Semester IV

Core Courses

Subject Code	Subject Title	Number of Credits
BIN 401 (P)	Project	6P
BIN 402 (P)	Seminar	2P
BIN 403 (T+P)	Transcriptomics and Proteomics	2T + 2P
BIN 404 (T+P)	Metabolic Engineering & Systems Biology	2T + 2P
BIN 405 (T+P)	Modern Drug Design	1T + 1P
	Total Core credits	5T + 13P = 18 credits

Elective Courses

Courses worth of 4 credits from among the following::

Subject Code	Subject Title	Number of Credits
BIE 406 (T+P)	Translational Bioinformatics	1T + 1P
BIE 407 (T+P)	Advanced Genomics	1T + 1P
BIE 408 (T+P)	Advanced Molecular Phylogenetics	1T + 1P
BIE 409 (T+P)	Advanced Algorithms in Machine Learning	1T + 1P
	Total Elective credits	4 credits
	Semester Total	**22 credits**

UGC recommended courses(Additional 10 credits)

Subject Code	Subject Title	Number of Credits
	Cyber security/Information security	4C
	Skill based credits	4C
	Human rights education	2C

Semester I

BIN 101 (T+P): Biological Data & Databases (2T + 2P)

Objectives:

This course will enable the students to:
- understand the nature of biological data and need for biological databases
- to understand and explore a few major biological databases (organization and contents); search and retrieve data from the databases using their respective search engines

Theory

Syllabus:

- Introduction to Bioinformatics (2)
- Overview of Bioinformatics resources on the web (2)
  NCBI/EBI/EXPASY etc
- Nature of biological data and formats (2)
- Biological literature databases (2)
  PubMed

- Nucleic acid sequence databases (6)
  - GenBank, EMBL, DDBJ
  - RefSeq
- Protein sequence databases (6)
  - UniProtKB
  - UniRef, UniParc, Proteomes, NextProt

- Derived databases (5)
  - InterPro and constituent databases
  - Recent derived databases.
- RNA sequence databases (3)
  - miRBase, lncRNAdb, MIT/ICBP siRNA database, RNACentral.
- Species and Biodiversity databases / resources (2)
  - NCBI Taxonomy database, GBIF, Sahyadri

Practicals

Objectives:

This course will enable the students to:
- To understand and explore major biological sequence databases (organization and contents); search and retrieve data from the databases using their respective search engines.

Syllabus:
- Exploring the integrated database system at NCBI server and querying the PUBMED, GenBank, NCBI Taxonomy database and other databases using the ENTREZ search engine (6)
  - Use of operators (AND, OR & NOT)
  - Use of limits

    - Exporting GI list
    - Batch retrieval
- Exploring the integrated database system at EMBL-EBI server (4)
- Exploring & querying UniProtKB and other protein sequence databases. (4)
- Sequence Formats & Format converters (2)
- Exploring tools on ExPASY (2)
- Exploring and using InterPro, its constituent databases and recent derived databases. (6)
- Exploring and querying RNA sequence databases: miRBase, lncRNAdb, MIT/ICBP siRNA database, RNACentral. (4)
- Exploring and querying Biodiversity databases / resources: GBIF and Sahyadri (2)

References:

- Nucleic Acids Research - Database issue (Most recent issues relevant to appropriate databases)
- Baxevanis A.D., Davison D.B., Page R. D. M. & Petsko G.A. Current Protocols in Bioinformatics. New York, John Wiley & Sons Inc., Latest Edition
- Korf Ian, Yandell Mark, Bedell Joseph. BLAST: an essential guide to the basic local alignment search tool. Shroff Publishers and Distributors Pvt. Ltd., Latest Edition
- Baxevanis Andreas D. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, Latest Edition. Publisher: New York, John Wiley & Sons, Inc.
- Teresa Attwood, Parry-Smith David J. Introduction to Bioinformatics. Publisher: Pearson Education (Singapore) Pte.Ltd., Latest Edition.
- Mount David W.. Bioinformatics: Sequence and Genome Analysis. Publisher: Cold Spring Harbor Laboratory Press; Latest Edition
- Gibas Cynthia, Jambeck Per. Developing Bioinformatics Computer Skills. Publisher: Shroff Publishers and distributors O'Reilly Media, Inc., Latest Edition

BIN 102 (T+P): Biological Sequence Analysis (2T + 2P)

Objectives:

This course will enable the students to:

- understand the concepts and methods for biomolecular sequence analysis
- study various algorithms for sequence analysis and understand underlying statistics
- understand when and why to use these algorithms to answer some of the biological questions

Syllabus: Theory

- Biomolecular sequence analysis (2)
    - Overview
    - Concepts
- Analysis of single sequence (2)
    - Nucleotide
    - Protein
- Pairwise sequence alignment algorithms (3)
    - Needleman & Wunsch
    - Smith & Waterman
- Scoring matrices for Protein and Nucleotide sequences (3)
    - MDM/PAM series
    - BLOSUM series
    - CSW
- Database Similarity Searches (6)
    - FASTA
    - BLAST
- Multiple sequence alignment algorithms (6)
    - CLUSTALW
    - MUSCLE
    - DALIGN
    - T-Coffee
- Sequence logos, consensus & patterns (3)
- Basic concept of sequence profiles, Derivation of profiles; applications (5)
    - Gribskov's Profile Analysis method
    - PSI-BLAST

Practicals/Tutorials:

Objectives:

This course will enable the students to:

- understand how to use various algorithms for biomolecular sequence analysis
- understand use of various parameter in respective algorithms and their effects on the outcome
- learn to write codes and practice programming skills for some of the simple tasks and to understand how bioinformatics evolved

Syllabus:

- Computing & plotting mono-, di-, tri-mer frequencies (2)
- Analysis & comparison of sequence with itself (1)
- Compute properties of sequences using tools on ExPASY (2)
- Exploring utilities in EMBOSS packages (2)
- Pairwise global alignment using Needleman-Wunsch algorithm (2)

    - DNA
    - Protein
- Pairwise local alignment using Smith-Waterman algorithm (2)
    - DNA
    - Protein
- Pairwise alignments (2)
    - Use of scoring matrices
    - Interpretation of results: dos & don’ts
- Database (homology) searches using different versions of FASTA (2)
    - DNA/Protein ganizsequence as a query
    - Interpretation of results
- Database (homology) searches using different versions of BLAST (4)
    - DNA/Protein sequence as a query
    - Interpretation of results
- Multiple sequence alignment programs (5)
    - Compilation of dataset
    - Learning to use programs
    - Input/output formats & parameters
    - Analysis & interpretation of results
    - Alignment visualization and editing tools (JalView)
- Build sequence logos (1)
- Deriving consensus & Prosite patterns (2)
- Using PSI-BLAST for detection of distant homologs (3)

Reference books:

- Sung Wing-Kin, Algorithms in Bioinformatics: Pratical Introduction. Publisher: CRC Press, 2010. ISBN: 978 1420070330
- Gibas, Cyntbia, Developing bioinformatics computer skills. Publisher: Oreilly, 2001. ISBN: 8173662428
- Pevsner, Jonathan, Bionformatics and functional genomics 2nd ed.Publisher: John Wiley & Sons,2009. ISBN: 978 0470085851
- Posada, David, Bioinformatics for DNA sequences analysis. Publisher: Humana Press, 2009. ISBN 978 1588299109
- Haubold, Bernhard, Introduction to computational Biology: An Evalutionary Approach. Publisher: Birkhäuser, 2006, ISBN: 9783764367008
- Xiong, Jin, Essential Bioinformatics. Publisher: Cambridge University Press, 2006. ISBN: 978 0521706100
- Mount,David. Bioinformatics: sequence and genome analysis. Publisher: Cold Spring Harbor Laboratory Press, 2004. ISBN: 087967121
- Ian, Korf. Blast. Publisher: Oreilly. 2003. ISBN: 9788173665127
- Baxevanis, A.D. Bioinformatics: a practical guide to the analysis of genes and proteins. 2nd Ed.. Publisher: John Wiley & Sons, 2002, ISBN: 9814126756
- Attwood, T.K. Introduction to Bioinformatics. Publisher: Pearson Education Pte. Ltd, 2001. ISBN: 8178085070

BIN 103 (T+P): Biomolecular Structure & Organization (2T + 2P)

Objectives:

This course will introduce the students to:

- Basic physico-chemical principles involved in structural organization of macromolecules
- Organization of the structures of proteins, DNA, RNA, carbohydrates and lipids
- experimental methods of structure determination

Theory

Syllabus:

- **Physicochemical Principles of biomolecular structure organization** (1)
    - Basic concepts of atom structure, hybridization of atomic orbitals, valence, covalent bonds, atomic interactions and forces, formation of polymeric molecules
- Co-ordinate systems: Rectangular, Cylindrical and spherical coordinate systems (2)
- **Experimental Methods for determination of biomolecular structures** (4)
    - X-ray Diffraction
    - NMR Spectroscopy
- **Protein Structure** (10)
    - Internal Coordinates – Bond lengths, bond angles, torsional angles; peptide unit, Ramachandran Map; Calculation of dihedral angles; Fourth atom fixing

    - Hierarchical Organization of Protein structure - Primary, Secondary, Supersecondary, Tertiary and Quaternary structure; Membrane Protein Structures

    - Principles of protein folding and Energetics; Mechanism of Protein folding; methods to study protein folding

- **DNA and RNA Structure** (8)
    - Base pairing in DNA & RNA
    - Double Helix – Organization, types and structural features
    - Structural & Geometric parameters associated with DNA
    - Secondary structures of DNA (triple helices, quadruplex, cruciform)
    - Sequence-structure relationships in DNA
    - Secondary Structures in RNA; Representations of RNA structures
    - Energetics of RNA structure
    - tRNA structure

- **Carbohydrates** (3)
    - Conformations of mono and oligosaccharides and Correlation with properties
    - Glycoproteins, Proteoglycans and Glycolipids: Structural aspects

- **Lipids & Membranes** (3)
    - Membrane microdomains and organization

Practicals:

Objectives:

In this course, the students will learn

- Contents, organization of structure databases, searching and retrieving the structural data
- Use of computational tools for visualization, understanding the organization and various properties of macromolecular structures
- Computations of structural parameters

- Visualization and interpretation of molecular interactions

Syllabus:

- Protein Data Bank (2)
- Nucleic Acid Databank (2)
- Visualization of structures (SWISS-PDB Viewer, Discovery Studio) (5)
- Calculation of structural parameters of Proteins & Ramachandran Plot (3)
- Calculation of structural parameters of DNA & RNA (3)
- Calculation of structural parameters of Carbohydrates and Lipids (3)
- Understanding Macromolecular interactions through visualisation & structure analysis (6)

  Protein – Protein
  Protein – Nucleic acids
  Protein - carbohydrates
- Understanding assemblies of biomolecules through visualization: Ribosome, Nucleosome, & Viral particles (6)

Reference books:

- Carl Ivar & Branden, John Tooze, Introduction to Protein Structure; 1998 (2nd Edition) Garland Science, ISBN 9780815323051
- Arthur Lesk Introduction to Protein Architecture: The Structural Biology of Proteins; Oxford Press ISBN:978-0198504740
- Thomas E. Creighton, Proteins: Structures and Molecular Properties; W. H. Freeman, 1993, ISBN: 071677030X, 9780716770305
- G.N. Ramachandran, V. Sasisekharan, Conformation of Polypeptides and Proteins; Advances in Protein Chemistry, Volume 23, 1968, Pages 283-437
- Jane S. Richardson, The Anatomy and Taxonomy of Protein Structure; Advances in Protein Chemistry, Volume 34, 1981, Pages 167-339
- Thomas E. Creighton, Protein Folding; W. H. Freeman, 1992, ISBN: 071677027X, 9780716770275
- Victor A. Bloomfield, Donald M. Crothers, Ignacia Tinoco, Nucleic Acids: Structures, Properties and Functions; University Science Books, ISBN 978-0-935702-49-1
- Gale Rhodes, Crystallography Made Crystal Clear A Guide for Users of Macromolecular Models, Elsevier Science, 2010, ISBN: 9780080455549, 0080455549

BIN 104 (P): Python Programming (2P)

Objectives:

This course will enable the students to:

- Learn python programming concepts
- Analysis of biological data using Python codes

Syllabus:

Introduction, Features of Python, IDLE environment	(1)
Data types	(1)
Variables	(2)
Data operators	(2)
Input and Output processing	(2)
Control statements, loops, iterations	(3)
Exception handling	(1)
Collections: list, tuple, set, dictionary	(5)
File handling	(4)
Pattern matching	(2)
Object oriented concepts – classes, attributes, methods	(1)
Functions	(2)
Python libraries: using standard modules and creating a new module	(2)
Working with Biological data	(2)

References:

- Mitchell L Model, Bioinformatics programming using Python. Publisher: O’Reilly Media. ISBN: 9788184048988
- Paul Berry, Head First Python. Publisher: Shroff/ O'Reilly ISBN: 9350231883
- Wesley J Chun, Core Python Applications Programming. Publisher: Pearson Education ISBN: 8131791343
- Sebastian Bassi, Python for Bioinformatics. Publisher: CRC Press. ISBN: 9781584889298
- Jason Kinser, Python for Bioinformatics. Publisher: Jones and Bartlet. ISBN: 0763751863

BIN 105 (P): Research Methodology (4T)

Objectives:
This course will enable the students to gain insights into the detailed methodologies and various other aspects of research

Module 1: Foundation of Research Methodology (15)

1. History and Philosophy of Research: What is Research? Ancient methodologies and research in various fields of science. Ancient Indian research methodologies. Scientific Method in research. Philosophy of Research including Rationalism, Empiricism, Skepticism, Reductionism, Pseudoscience etc. Qualities of a Researcher.

2. **Problem Identification & Formulation**: definition and formulating the research problem, Necessity of defining the problem, Importance of literature review in defining a problem, Literature survey: primary and secondary; web sources; critical literature review. Design of Research Question..Hypothesis Generation

3. **Approaches to Research Design:** Induction and deduction in research. Variables in research. Features of a good research design - Exploratory Research Design - Concept, Types and uses, Descriptive Research Design - concept, types and uses. Experimental Design - Concept of Independent & Dependent variables. Biased and unbiased research design.

Module 2: Data Collection, Processing and Analysis (15)

1. **Types of Biological data and nature of research:** Qualitative data – Quantitative data Concept of measurement, causality, generalization, replication. Basic concepts in qualitative and quantitative data analysis.

2. **Data Collection:** Execution of the research - Observation and Collection of data - Methods of data collection Databases (ENA, UniProtKB, PDB, GEO etc) searching databases, data retrieval, preparing data sets, data curation - Positive and negative data sets. Basic concepts in organization of data into databases - hypothesis-testing - Generalization and Interpretation. Observation, Surveys, Monitoring research. records to be maintained by researcher.

3. **Sampling & Measurement**: Concept of Statistical population, Sample, Sampling Frame, Sampling Error, Sample size, Non-Response. Characteristics of a good sample. Probability Sample - Simple Random Sample, Systematic Sample, Stratified Random Sample & Multi-stage sampling etc. Determining size of the sample - Practical considerations in sampling and sample size. Concept of measurement - what is measured? Problem in measurement in research - Validity and Reliability. Levels of measurement - Nominal, Ordinal, Interval, Ratio.

5. **Data Analysis and Interpretation:** Graphical representation of data, Univariate analysis (frequency tables, bar charts, pie charts, percentages), Bivariate analysis – Cross tabulations and Chi-square test including testing hypothesis of association. Correlation and regression analysis. Probability distributions: Discrete distribution (bionomial & Poisson) – Continuous distribution (Normal & Exponential) Bayesian Modeling, Estimation of accuracy (sensitivity specificity, Mathew’s Correlation coefficient)

Module 3: Scientific Paper Writing and Communication (20)

1 **Literature Review and Writing a Review**: Planning and Writing a Research Proposal to seek funding

2. **Use of tools / techniques for referencing**: methods to search required information effectively, PubMed, effective literature search using Entrez, Google Scholar

3. **Scientific writing and communication**: Different types of publications, their structure and components, Choice of journals for publication based on various parameters such as impact factor, scope of journal etc. Types of report - Technical reports and thesis, Different steps in the preparation - Layout, Structure and Language of typical reports, Paraphrasing, Illustrations and tables, Bibliography, referencing. Oral presentation, Making presentation, Use of visual aids, Importance of effective communication. Tools/Software for paper formatting like MSOffice, software for detection of Plagiarism

Module 4: Research Ethics & Safety (10)

1. Ethics in Research: Ethical issues, ethical committees, Reproduction of published material - Plagiarism - citation and acknowledgement, Self-Plagiarism, Reproducibility and accountability

2. IPR: Commercialization, Copy right, Royalty, Intellectual property rights and patent law, Trade related aspects of intellectual property Rights

3. Biosafety: Good lab practices, Guidelines for Biosafety, Institutional Biosafety committee, Institutional Animal ethics committee.

References:

- ‘History of the Scientific Methods’ by Martin Shuttleworth, https://explorable.com/history-of-the- scientific-method.
- ‘The Statistical Analysis of Experimental Data’ by, John Mandel, ISBN: 0486646661, ISBN13: 9780486646664
- ‘Research Methodology and Scientific Writing’ by C. George Thomas, ISBN 978-3-030-64864-0
- ‘Research Methods for Science’ by Michael Marder, Online ISBN:9781139035118
- Research Methodology: An Introduction - Stuart Melville and Wayne
- Practical Research Methods - Catherine Dawson
- Research Methodology - C. R. Kothari Essential Bioinformatics – Jin Xiong (Cambridge University Press)

Elective Courses

Courses worth of 4 credits from among the following:

BIE 106 (T): Mathematics & Statistics (2T)

Objectives:

This course will enable the students to achieve skills in statistics and mathematics that are essential for application in bioinformatics.

Mathematics **(1 credit)**

- Introduction to Set Theory (1)
- Trigonometry: (2)
  Trigonometric Functions, Series Expansion, Inverse, General Values, Graphs, Taylor series
- Vector & Matrices (2)
  Vector Algebra, Vector Calculus, Basic Computations, Matrices
- Calculus: (4)
  Limits, Continuity, Analysis, Differentiation (1D & Partial), Reimann Integration, Definite Integrals.
- Ordinary & Partial Differential Equation (3)
  1st Order & 2nd Order Ordinary Differential Equations. Self-Adjoint Equations, Method of Separation of Variables
- Integral transform (3)
  Fourier Series, Fourier Transform, Laplace Transform

Statistics **(1 credit)**

- Overview of applications of statistics in Bioinformatics (1)
- Introduction to principles of statistical sampling from a population (1)
- Frequency Distributions and Statistical Measures: mean, mode, median, variance, standard deviation, coefficient of variation, measures of skewness

  and kurtosis. Computation of these measurements for the given data. (2)

- Introduction to theory of Probability, Conditional Probability, Bayesian Rules, Random variable, Distributions of random variables, Binomial, Poisson, Geometric, Normal and extreme value distribution, (5)
- Regression, correlation, fitting regression line. (2)
- Hypothesis testing: (4)
  - Test of significance viz. Z test, t test, paired t test, chi^2 test of goodness of fit

References:

- Isaev Alexander. Introduction to Mathematical Methods in Bioinformatics. Publisher: Berlin; New York: Springer, 2004. ISBN: 3540219730
- Raman K. V. & Pal Sourav. Mathematics in Chemistry. New Delhi, Vikas Publishing House Pvt Ltd., 2004. ISBN: 8125912886
- Jones D.S., Sleeman B.D.. Differential Equations and Mathematical Biology Publisher : Chapman & Hall. 2003. ISBN:1584882964
- Bracewell Ronald. The Fourier transform and it's applications 3rd edition. Publisher: New Delhi : McGraw Hill, 2000. ISBN: 0073039381.
- Stephenson G., Radmore P. M.. Advanced Mathematical Methods for Engineering and Science Students Cambridge: Cambridge University Press, 1990.
- Arfken George. Mathematical methods for physicists. Publisher: Orlando : Academic Press 1985. ISBN: 0120598205.

BIE 107 (T): Basic Biology (2T)

Objectives:

This course will enable the students to:

- Understand the basic principles of biology.
- Appreciate the basic nature and diversity of microbial, plant and animal life.
- Understand the classification of organisms and taxonomy.

Syllabus:

- **Origin and evolution of Life**: Living and nonliving things; theories and evidences of origin and evolution of life; chemical evolution- atom to procell and biological evolution- procell to human; cellular tree of life. (4)

- **Unit of life**: Basic unit of life; prokaryotic cell, eukaryotic cell, plant cell and animal cell; Structure and functions of cell organelles- membrane and non-membrane organelles. (4)

- **Diversity of life**: Classification of organisms- five-kingdom classification (Monera, Protista, Fungi, Plantae and Animalia); classification and characteristics of different kingdoms; Biodiversity- hotspots of biodiversity, ecological and economic role of biodiversity; threats to biodiversity. (16)

- **Ecology and nutrition**: Biotic and abiotic components; food chain; trophic levels; food webs; ecological pyramids; ecosystems- structure, types of ecosystem; Nutrition- macronutrients and micronutrients; modes of nutrition- parasitic, saprophytic, symbiotic and insectivorous. (3)

- **Viruses**: Biology of viruses; bacteriophages, plant and animal viruses. (3)

References:

- Wallace Robert A., Sanders Gerald P., Ferl Robert J. The science of life. Publisher: New York, NY : Harper Collins, 1991. ISBN: 0673380440.
- Solomon Eldra P., Berg Linda R., Martin Diana W. Biology 6th edition. Publisher: Pacific Grove, CA, Brooks/Cole Thomson Learning, 2002. ISBN: 0030335035.
- Recknagel F. 2002 Ecological Informatics: Understanding Ecology by Biologically-Inspired Computation, Springer, New York.
- Odum E.P. 1983 Basic Ecology. Saunders International Edition, Japan.
- The Cell: A Molecular Approach (5th edition) by Cooper, G.M. and Hoffman, R.E., A. S. M. Press, 2009.
- The world of the Cell (7th edition) by Becker, W.M., Kleinsmith, L.J., Hardin, J and Bertoni, G.P., Pearson / Benjamin Cummings, 2009
- Molecular Cell Biology (5th edition) by H. Lodish, A. Berk, P. Matsudaira, C.A. Kaiser, M. Krieger, M. P. Scott, S. L. Zipursky & J. Darnell, W. H. Freeman & Company, 2004.

BIE 108 (T): Molecular Biology (2T)

Objectives:

This course will enable the students to:

- understand the concepts in gene structure and expression which leads to specific functions.
- understand the concepts in vital processes such as replication, transcription, post-transcriptional modifications, translation etc.

Syllabus:

- Nucleic acid: Composition, Primary and Secondary structures, Circular DNA (2)
- Genome organization: Prokaryotic and eukaryotic genomes – C value paradox, repetitive and non-repetitive DNA., transposons and retroposons; Exons and introns – organization of interrupted genes, one gene-many proteins concept; Gene numbers – essential genes and total gene number, gene clusters, pseudogenes; Gene families – globin and rDNA gene families; Organelle genome – mitochondrial and chloroplast. (4)
- Packaging of genome – Bacterial genome as nucleoid; Eukaryotic genome – nucleosomes, chromatin, solenoids, loops, domains, scaffolds, chromosomes (2)
- DNA Replication- Details of prokaryotic and eukaryotic DNA replication: DNA polymerases, initiation, elongation and termination of replication; multiple origins of replication, Regulation of replication (4)
- DNA damage, repair and recombination- DNA damaging agents physical and chemical, types of DNA damages; DNA repair systems in prokaryotes and eukaryotes -Single step repair, Base excision repair, Nucleotide excision repair, Mismatch repair, Recombination repair; Recombination - homologous and non-homologous recombination. (2)
- Gene Expression - Transcription: Details of prokaryotic and eukaryotic transcription: RNA polymerase, promoters, initiation, elongation and termination of transcription; regulation of transcription, operons. (4)
- Processing of transcripts - 5’ capping, 3’ polyadenylation, splicing and editing, self-splicing (1)
- Translation: Details of prokaryotic and eukaryotic translation: Protein synthesis machinery, initiation, elongation and termination, Genetic code, accuracy of translation, regulation, Post-translational modifications of proteins (6)
- Recombinant DNA technology: Enzymes, cloning strategies, vectors, screening for recombinants (3)
- Small non-coding RNAs – micro-RNA, small nuclear RNA, small nucleolar RNA, si-RNA, piwi-RNA and Genome editing. (2)

References:

- Watson James D., Baker Tania A., Bell Stephen P., Alexander Gann, Levine, Michael Losick Richard. Molecular Biology of the Gene 6th Edition. Publisher: New York, Cold Spring Harbor Laboratory Press. 2008. ISBN: 9780321507815
- Benjamin Lewin, Jocelyn E. Krebs, Elliott S. Goldstein, Stephen T. Genes XIth Edition. Publisher: Kilpatrick Jones & Bartlett Publishers, 2014.

BIE 109 (T): Biochemistry and Cell Biology (2T)

Objectives:

This course will enable the students to:

- Understand the physicochemical properties of bio-macro-molecules with their building blocks and their interactions in an aqueous environment
- Understand the structure-function relationships of bio-macro-molecules, the principles of enzyme catalysis, regulation and inhibition.
- Study the principles of structure and function of cells, membranes and organelles.
- Study the cell division cycles with their regulation and mechanism of signal transduction.

Biochemistry (1 credit)

- **Chemistry of life**: Atoms, elements, ions, molecules and macromolecules of life. (2)
- **Water:** Water as the universal biological solvent; physicochemical properties of water; Ionization of water; concepts of pH, pI, pKa, pKb, buffer, acidosis, alkalosis and osmolarity; Henderson-Hasselbalch equation. (2)
- **Biomolecules:** Classification and physico-chemical properties of biomolecules; (2)
- **Enzymes**: Historical perspective, general characteristic of enzymes; enzyme cofactors, prosthetic groups, coenzymes, apoenzyme, holoenzyme, metalloenzymes, proenzymes, isoenzymes, enzyme activity; specificity activity; enzyme models; enzyme nomenclature and classification; enzyme catalysis; enzyme regulation- reversible and non-reversible regulations; Enzyme kinetics- basic enzyme reactions, energy levels, Factors affecting enzyme activity (temperature, pH, enzyme concentration, substrate concentration, inhibitors and activators); Michaelis-Menten kinetics- K_m, V_{max}, K_{cat} and $K_{cat/}K_m$; enzyme inhibition. (4)
- **Vitamins**: General characteristics of vitamins; classification- water-soluble and fat-soluble vitamins, sources, structures, dietary requirements and deficiency conditions; coenzyme forms; Role of vitamins in metabolism; therapeutic uses of vitamins; antivitamins and hypervitaminosis. (2)
- **Hormones**: General characteristics of hormones; classification, structures, functions, mechanism of action and abnormalities of the hormones (thyroid, pancreas, hypothalamus, pituitary and gonads); Plant hormones (auxin, gibberllic, abscisic acids and cytokinins); Pheromones. (2)
- **Minerals**: Structure, classification, daily requirement, sources, uptake, transport, excretion and functions of minerals (calcium, phosphorus, magnesium, iron, iodine, zinc and copper). (1)

Cell Biology (1 credit)

- **Cellular organization of Prokaryotic and eukaryotic cells**: Cell types; Cell theory; Structure and function of organelles- Double membrane bounds organelles (Nucleus, mitochondria, chloroplasts); Single membrane bounds organelles (endoplasmic reticulum, Golgi apparatus, lysosomes, peroxisomes, vacuoles) and Non-membrane bounds organelles (Ribosomes, nucleoles and centrioles). (4)
- **Membrane structure and function**: Structure, transports, channels, carriers, receptors, membrane potentials. (3)
- **Cell motility and shape**: Structure, polymerization, organization and functions cytoskeletal elements; Structure and functions of motor proteins. (2)
- **Signal transduction**: Types of signaling; signaling molecules; receptors and major signaling pathways. (3)

- **Cell cycle and its regulation**: events during mitosis and meiosis (3)

References:

Biochemistry-

- Voet, Donald, Voe Judith, Pratt, Charlotte W. Fundamentals of Biochemistry: Life at the molecular Level 2nd Edition. Publisher: Asia, John Wiley & Sons. 2006. ISBN: 0471753416.
- Nelson David L., Cox Michale. Lehninger Principles of Biochemistry 5th Edition. Publisher: New York. W. H. Freeman. 2008. ISBN 978 0716771081.
- Berg, Jeremy M, Tymoczko, John L. Stryer, Lubert. Biochemistry 6th Edition. Publisher: New York : W.H. Freeman. 2007.ISBN: 071676766X.
- Hames David, Hooper Nigel. Instant Notes in Biochemistry 3rd Edition. Publisher . Nodia, Taylor & Francis. 2007. ISBN: 185996 2491.
- Zubay, Geoffrey.Biochemistry 4th Edition, Publisher: Boston,Wm C. Brown, 1998, ISBN 0697219003.
- Horton, Robert, Moran, Laurence A, Scrimgeour, Perry Gray Marc, Rawn.David. Principles of biochemistry. Publisher: New Jersey, Pearson Prentice Hall, 2006. ISBN: 0131453068.
- Mathews, Christopher K. van Holde, K. E., Ahern Kevin G. Biochemistry. Publisher: San Francisco, Pearson Education 2000.ISBN: 81 29702150.
- Garrett, Reginald H., Grisham, Charles M. Principles of biochemistry: with a human focus. Publisher: Australia Brooks/Cole, Thomson Learning, 1997. ISBN:0030973694.

Cell Biology

- Essential Cell Biology (3rd edition), by B. Alberts, D. Bray, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts & P. Walter, Garland Science, 2010.
- Molecular Biology of the Cell (5th edition) by B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts & P. Walter Garland Science, 2008.
- The Cell: A Molecular Approach (5th edition) by Cooper, G.M. and Hoffman, R.E., A. S. M. Press, 2009.
- The world of the Cell (7th edition) by Becker, W.M., Kleinsmith, L.J., Hardin, J and Bertoni, G.P., Pearson / Benjamin Cummings, 2009
- Molecular Cell Biology (5th edition) by H. Lodish, A. Berk, P. Matsudaira, C.A. Kaiser, M. Krieger, M. P. Scott, S. L. Zipursky & J. Darnell, W. H. Freeman & Company, 2004.
- Molecular Biology of the Cell: A Problems Approach, (4th edition) by Wilson, J and Hunt, T., Garland Science publishers, 2002.

Semester II

BIN 201 (T+P): Structural Bioinformatics (2T+2P)

Objectives:

- Know how three dimensional structures can be analyzed for gaining insights into functions and other biological aspects
- know the computational approaches for structure analysis
- acquire knowledge of various algorithms & methods for structure prediction
- understand the principles of macromolecular interactions

Syllabus:

Theory:

- Overview of Structural Bioinformatics (2)
- Prediction of protein structure
    - secondary structure prediction methods (4)
        - First, second and third generation methods
    - Tertiary structure prediction
        - Homology modeling (4)
        - Fold Recognition: (5)
            - 1D-3D Profile-based methods,
            - Threading methods
        - *ab initio* methods (2)
        - AI-based methods of Structure prediction (2)
- Structural alignments of proteins (4)
    - Superimposition of structures & calculation of RMSD
    - Vector-based, distance matrix-based and combined algorithms for structural alignments
- Structure-based classification of proteins: SCOP & CATH (2)
- Prediction of binding pockets on protein structures (3)
- Structure-based function Prediction (2)
- Prediction of RNA structures (2)

Practicals:

- Write codes for Calculation of inter-atomic distances (4)
    - to find chain breaks in polypeptides
    - to detect presence of non-bonded interactions
- Write codes for computation of dihedral angles (3)
- Write code for 4th atom fixing (3)
- Prediction of Secondary structures of proteins using various methods and computation of prediction accuracies (2)
- Prediction of tertiary structures of proteins using Homology Modeling Protocol/s and validation of templates as well as predicted structures (4)
- Prediction of tertiary structures of proteins using Fold Recognition methods (2)
- Prediction of tertiary structures of proteins using *ab initio* methods (2)
- Write code for construction of distance matrix using atomic coordinates (e.g. C^{α} - C^{α} distance matrix) and analysis for detection of tertiary interactions (2)
- Detection of structural similarities among proteins and finding remote homologues using protein structural alignment methods (2)
- Searching and browsing structural classification databases (2)

  - Prediction of surface pockets of proteins using various methods (2)
  - Prediction of RNA structures using various methods (2)

References:

- Forbes Burkowski. Structural bioinformatics: An algorithmic approach. Publisher: CRC Press, 2009. ISBN: 9781584886839.
- Bourne Philip E., Weissig Helge. Structural Bioinformatics (Methods of Biochemical Analysis, V. 44), 2003. Publisher: Wiley-Liss. ISBN: 0471202002.
- Höltje Hans-Dieter, Sippl Wolfgang, Rognan Didier, Folkers Gerd. Molecular Modeling: Basic Principles and Applications. Publisher: New York, Wiley-VCH. 2003. ISBN: 3527305890.
- Friesner Richard A. Computational Methods for Protein Folding:advances in Chemical Physics Volume 120 Kindle Edition. Publisher: New York, John Wiley & Sons. 2002. ISBN: 0471209554.
- Heilmeyer L., Friedrich P. Protein Modules in Cellular Signalling. Publisher: Amsterdam, IOS Press. 2001. ISBN: 1586031805.
- Branden ,Tooze John. Introduction to Protein Structure. Publisher: New York, Garland Publishing Inc. 1999. ISBN: 0815323050.
- Hill H.A.O. Sadler P.J., A.J. Ed. Metal Sites in Proteins and Models Redox Centres Publisher: New York, Springer 1999. ISBN: 3540655564.
- Sternberg Michael J. E. Protein Structure Prediction: A Practical Approach. Publisher: USA, Oxford University Press. 1997. ISBN: 0199634953.
- Fasman G.D. Prediction of Protein Structure and the Principles of Protein Conformation. Publisher: New York, Plenum Press. 1989 ISBN: 0306431319.
- Creighton T. E. Editor. Protein Structure: A Practical Approach. Publisher: IRL Press at Oxford University Press. 1989. ISBN: 0199630011.

BIN 202 (T+P): Chemoinformatics **(2T+2P)**

Objectives:

This course will enable the students to:

- understand the representation and use of chemical information on computer
- appreciate the complementary aspects of cheminformatics and bioinformatics

Theory

Syllabus:

- Representation of chemical compounds: 1D, 2D, 3D (7)
  - SMILES, InChI, Fingerprints (Daylight), Matrices, Connection Tables, mol & sdf files, Markush structures, Rotatable bonds and conformers, Molecular surfaces
- Representation of reactions: (3)
  - SMARTS, SMIRKS, Matrices
- Searching chemical structures: (6)
  - Exact (Canonicalizing SMILES: Morgan, CANGEN), Substructure, Superstructure and Similarity search (Tanimoto, Euclidean and Tversky index)
  - Clustering small molecules : Hierarchical and non-hierarchical methods
  - 3D pharmacophore based searching: Common Pharmacophore features, building hypotheses and searching databases
- Chemical Databases: CSD, Pubchem, other relevant databases (5)
  - Information content and applications
  - Data mining
- Quantitative Structure Activity Relationship (QSAR) (4)
  - Structure descriptors: topological and shape indices
  - Training and test data set
  - Applicability domain
  - 2D QSAR
- Combinatorial chemistry and Library Design (3)
  - Historical methods: Mix and Split library design and Iterative deconvolution
  - Computer based methods
  - Diversity analysis
  - Lipinski Rules
- Tools for cheminformatics algorithm development (2)
- Future perspectives (1)

Practicals

Objectives:

This course will enable the students to

- understand the representation and use of chemical information on computer
- using chemical databases
- introduction to analysis and predictive modeling

Syllabus:

- Practical sessions on structure storage and representation in various formats: SMILES, InChI, mol & sdf files (2)
- Practical sessions on understanding reaction transforms and applying them to various small molecules (2)
- Building a database of small molecules (1)

- Structure searching: Substructure, Superstructure and Similarity searches	(1)
- Exploring and data mining of : CSD and Pubchem	(3)
- Pharmacophore hypothesis and searching	(2)
- Descriptors and 2D QSAR studies	(2)
- Tools for cheminformatics algorithm development	(2)

References:

Books

- Gasteiger Johann, Engel Thomas. Chemoinformatics: A Textbook. Publisher: Wiley-VCH; 1st edition. 2003. ISBN: 3527306811.
- Leach Andrew R., Valerie J. Gillet. An introduction to chemoinformatics. Publisher: Kluwer academic , 2003. ISBN: 1402013477.
- Gasteiger Johann, Handbook of Chemoinformatics: From Data to Knowledge (4 Volumes). Publisher: Wiley-VCH. 2003. ISBN:3527306803.
- Muthukumarasamy Karthikeyan, Renu Vyas. Practical Chemoinformatics. Publisher: Springer. 2014. ISBN: 9788132217794
- Bajorath Jürgen. Chemoinformatics and computational chemical biology. Publisher: Humana Press. 2011. ISBN: 9781607618386

Other learning resources:

- Indiana Cheminformatics Education Portal (http://icep.wikispaces.com/)
- Henry Stewart Talks Introducing Cheminformatics

BIN 203 (T+P): Genomics (2T + 2P)

Objectives:

This course will enable the students to:

- Understand concepts in genome organization, genome packaging.
- Understand concepts in recombinant DNA technology and DNA markers.
- understand genome sequencing, assembly, alignment and annotation etc. and related technologies.

Syllabus

Theory

- Genome projects – importance, objectives and strategies (1)
- Genome markers and mapping: STS, EST, RFLP, SNP (4)
- Genome sequencing – First, Second and Next generation sequencing platforms (5)
- Genome sequence assembly (5)
    - Output file formats
    - Quality assessment of sequence reads
    - Removal of adapter contamination
    - Reference-guided and *de novo* Assembly: Approaches and Tools
    - Visualization of alignments in Reference-guided Assembly
    - Removal of redundancy
    - Archival of assembled sequences in databases
- Genome Annotation (9)
    - Basic Aspects of Genome Annotation
    - Prediction of ORFs
    - Algorithms for gene prediction & gene modeling
    - Prediction of Promoters, splice sites, UTRs etc
- Structural Genomics (2)
- Functional Genomics (2)
- Gene Ontology (2)
    - Basic concepts
    - GO terms and relationships (DAGs)
    - Tools for GO data: (AmiGO, BioConductor)
    - GO Browsers and editors
    - GO visualization tools
    - Microarray-related tools
    - Protein interaction related tools
    - Statistical analysis
    - Term enrichment tools

References:

- Brown, T.A. Genomes 2 Publisher: New York, BIOS Scientific Publishers Ltd. 2002, ISBN: 1859960294
- Old R.W. and Primrose S.B. Principles of Gene Manipulation: An Introduction to Genetic Engineering Publisher: University of California Press, 1980
- Benjamin Lewin, Jocelyn E. Krebs, Elliott S. Goldstein, Stephen T. Genes XI^th^ Edition. Publisher: Kilpatrick Jones & Bartlett Publishers, 2014.

Practicals:

Objectives:

- This course intends to introduce the students with genome browsers and databases, assembly softwares and annotation procedures.
- Hands-on exercises will help the students to understand steps involved in comphrensive genomic study of an organism

Syllabus:

- Browsing & viewing genome data (3)
  - Ensembl@EBI
  - MapViewer@NCBI
  - UCSC Genome Browser

  Visualization, Browsing and searching of
  - Protein coding and non-coding Genes, ESTs, STSs, Retrotransposons, RFLPs, SNPs
  - regions exhibiting Synteny
- Exploration of Genome Databases (4)
- Genome assembly (8)
- Genome Annotation
  - Using integrated genome annotation servers such as the server developed at IMTech, Chandigarh (http://imtech.res.in/raghava/gp.html) (2)
  - Gene Prediction and Gene Modeling
  - Prediction of genes and gene structures (gene modeling) using online (web) servers of different methods tailored for prokaryotic and eukaryotic organisms such as GLIMMER, GeneMark, Grail, GENSCAN etc. Interpretation of results and comparison with known gene models (where available). Evaluation of accuracy of the methods. (4)
  - Prediction of promoters using methods such as Neural Network Promoter Prediction (NNPP) at Berkeley Drosophila Genome Project server, Genome inspector for combined analysis of multiple signals in genomes etc. Using Promoter databases. (2)
  - Prediction of alternate splice sites using methods such as Splice Site Prediction by Neural Network (at Berkeley Drosophila Genome Project server), GenScan, NetGene2, GeneSplicer etc. Prediction of PCR primers using Primer 3, ePCR etc. (2)
- Structural Genomics: PSI Knowledgebase, VIPs (2)
- Functional Genomics (3)
  - Using sequence -based and structure-based Function Annotation Servers such as
    - ProKnow (http://www.doe-mbi.ucla.edu/Services/ProKnow/)
    - Joined Assembly of Function Annotations (JAFA) at http://jafa.burnham.org/learnMore.html etc. which are integrated services for function annotation
    - ProFunc (http://www.ebi.ac.uk/thornton-srv/databases/ProFunc)

References:

- Brown, T.A. Genomes 2 Publisher: New York, BIOS Scientific Publishers Ltd. 2002, ISBN: 1859960294
- Brown, T A Genomes 4th Ed. Garland Science 2018
- Old R.W. and Primrose S.B. Principles of Gene Manipulation: An Introduction to Genetic Engineering Publisher: University of California Press, 1980
- Benjamin Lewin, Jocelyn E. Krebs, Elliott S. Goldstein, Stephen T. Genes XIth Edition. Publisher: Kilpatrick Jones & Bartlett Publishers, 2014.
- Sankoff, David and Nadean, Joseph Comparative Genomics. Khuwer Academic Publishers 2000
- Sharma, Prabhakar and Narain, Rabindra Handbook of Genomics 2015

- Primrose S. B. & Twyman, R. M. Principles of Genome Analysis and Genomics 3rd Ed. By Blackwell Publishing 2003
- Pagel, Mark and Pomiankowski, Andrew (Ed) Evolutionary Genomics and Proteomics 2008
- Peacock, Christopher Ed. Parasite Genomics Protocols 2nd Ed. Humana Press 2015
- Samuelsson, Tore Genomics and Bioinformatics: Introduction to Programming Tools for Life Scientists Cambridge University Press 2012
- Bergman, Nicholas Comparative Genomics Volume 2 Humana Press 2007
- Hall, Brian and Hallgrimsson, Benedikt. Strickberger’s Evolution 5th Ed. Jones & Bartlett Learning 2014
- Dale, Jeremy and Schantz, Malcolm, From Gens to Genomes: Concepts and Applications of DNA Technology. Wiley 2008
- Dutheil, Julien Statistical Population Genomics Humana Press, 2021.

BIN 204 (P): Introduction to Database Systems (2P)

Objectives:

This course will enable the students to:

- understand the concepts of data, data models and relationships be aware of various data representation techniques and various types of databases
- appreciate and implement relational database design create the database systems with user-friendly front-ends for fast and efficient data retrieval and storage
- acquire the skills of using MySQL, SQL and basic skills in creating front end applications

Syllabus:

- Database Management System basic concepts (1)
- Data abstraction, Data Models, Instance & Schema (2)
- Data Normalization (1)
- Overview of MySQL and its features (1)
- Installing MySQL Server and MySQL Workbench (1)
- Data Definition Language (DDL) (10)
    - Creating and altering database tables
    - Defining primary keys, foreign keys, and constraints
    - Understanding data types in MySQL
- Data Manipulation Language (DML) (10)
    - Inserting data into tables
    - Updating and deleting data
    - Retrieving data using SELECT statements
    - Arithmetic and logical operators.
    - Grouping and aggregate functions
    - Filtering grouped data with HAVING clause
    - Pattern matches
    - Performing subqueries for complex queries
- Joins (1)
- Importing and Exporting Data (2)
    - Importing data from CSV and other formats
    - Exporting query results to files
    - Backup and restore databases
- Advanced options: Defining Indices, Management using Grant and Revoke (1)

References:

- Silberschatz, A, Korth, H F & Sudarshan, S. Database system concepts; McGraw-Hill higher education, 2002. ISBN: 0071148108.
- Date C. J. An Introduction to Database Systems. 1999. Publisher: Addison Wesley. ISBN: 0201327546.
- Bayross Ivan. SQL, PL/SQL The Programming Language of Oracle. 2nd revised edition. Publisher: BPB ISBN: 8176560723.
- Ramez Elmasri, Shamkant Navathe, Fundamentals of Database Systems. Publisher: Addison Wesley. ISBN: 0136086209
- Vikram Vaswani, MySQL: The Complete Reference. Publisher: Tata Mcgraw Hill. ISBN: 0070586845

BIN 205 (P): On Job Training / Field Project (4P)

Students will undertake Internship/ Apprenticeship with a company related to Bioinformatics / Computational Biology. Alternately the student will engage in a Field Project that addresses the use of computational biology to solve applied problems.

Elective Courses

Courses worth of 4 credits from among the following:

BIE 206 (P): Biological Data Curation and Analysis (2P)

Objectives:

This course will enable the students to:

- Understand the Biological data as Big Data.
- Understand the significance of data curation.
- Understand and use the web services provided by the biological databases.
- Programmatic access to biological resources.
- Write advanced codes for biological data curation and analysis.
- Design workflows for data processing.

Syllabus:

- Biological Data (1)
    - Types, complexity & big data
- Biological data curation (1)
    - Need, issues & challenges
- Using APIs (5)
    - Introduction to APIs and their importance
    - Making API requests with Python
    - Parsing data from API responses
    - Web scraping and its applications
    - Use of requests, BeautifulSoup modules
- Web services of commonly used Bioinformatics databases: (10)
    - NCBI: E-utilities
    - EMBL-ENA: Programmatic access using ENA browser REST URLs
    - Uniprot: Uniprot resources programmatically access
    - PDB: The RCSB PDB RESTful Web Service interface
    - KEGG: KEGG API
- Data Analysis with Pandas: using data frames, data cleaning, filtering (2)
- Biopython: Introduction, installation, important components like seq, seqIO, alignIO, BLAST, ClustalW, PDB, SwissProt etc., parsing output (6)
- Workflows in Bioinformatics (5)
    - Known workflow environments
    - Case studies: Existing workflows (ex: Immunoinformatics)
    - Workflows design & implementation
        - Development of one server

References:

- Web services documentation of various biological databases & resources
- Review and research articles on Big data and data curation & web services
- Agostino, Michael, Practical Bioinformatics. Publisher: Garland Science, 2012. ISBN: 9780815344568
- Sung Wing-Kin, Algorithms in Bioinformatics: pratical introduction. Publisher: CRC Press, 2010. ISBN: 978 1420070330

BIE 207 (P): Advanced Techniques in Experimental Biology (4P)

Objectives:
This course will enable the students to:

- gain hands-on experience with some of the laboratory techniques necessary to understand genomics and proteomics.

Practical:
A. Microbiological techniques: (15)
  - Media preparation, Sterilization, Inoculum preparation.
  - Staining: Gram staining, endospore staining.
  - Streak, pour and spread plate methods.
  - Growth curve and kinetics
  - Screening of antibiotic producer, amylase producers, PHA producers
  - Fermentation

B. Biochemical techniques: (30)
  - Extraction and fractionation methods- Cell lysis, protein precipitations, dialysis, density gradient centrifugation/differential centrifugation, ultracentrifugation
  - Electrophoresis- Native-PAGE, SDS-PAGE, 2DE, 2D-DIGE, agarose gel
  - Chromatography- AC, IEC, TLC, HPLC, GC, gel filtration and paper chromatography
  - Isolation and Purification of enzymes (Invertase, Amylase, Alkaline Phosphatase, Acid Phosphatase etc.)
  - Enzymes kinetics- Enzyme activity, factors affecting enzyme activity (pH, Temp., Substrate, Inhibitors, Activators), Determination of pKa, pI, K_{cat}, K_m, K_I,$V_{max.}$
  - Other techniques- Western blot, In-gel digestion, metabolism assays, gel image analysis, CD/NMR/MS spectra

C. Molecular biology techniques: (15)
  - Bacterial Transformation
  - Plasmid isolation
  - Restriction digestion
  - *Polymerase Chain Reaction*
  - Gel Elution of DNA
  - TA/TOPO-TA cloning of PCR product

References:

- Microbiological Methods (Eighth edition) by C.H. Collins and P.M. Lyne, Arnold, a member of the Hodder Headline Group,338 Euston Road, London NW1 3BH
- Medical Microbiology: The practice of medical microbiology. Vol. 2 (Twelth edition) by Robert Cruickshank, Churchill Livingstone, 1975.
- Practical Handbook of Microbiology (Second Edition) by Emanuel Goldman, Lorrence H Green, CRC Press, 2008.
- Biochemical Techniques: Theory and Practice by John F. Robyt and Bernard J. White, Waveland Press, 1987.
- Introduction to Practical Biochemistry (Third edition) by Plummer Mu, David T. Plummer, Tata McGraw-Hill Education, 1988 - 332 oldal.
- Principles and Techniques of Practical Biochemistry (7th Edition) by Keith Wilson and John Walker, Cambridge University Press.

- Introductory Practical Biochemistry by S.K. Sawhney and R. Singh, Narosa Publishing House, 2000.
- Biochemical Methods (Second edition) by S Sadasivam, A. Manickam, New Age International (P) Ltd. Publisher, 1996.

BIE 208 (T+P): Molecular Phylogenetics (1T + 1P)

Objectives:
This course will enable the students to:
- understand concepts of molecular evolution and the nature of data for deriving molecular phylogeny
- learn and apply the statistical approaches and models for phylogenetic analysis and tree reconstruction

Theory
Syllabus:
- Molecular phylogenetics: Overview (2)
    - Concepts in molecular evolution
    - Nature of data
- Molecular Phylogeny (5)
    - Concept & overview
    - Distance-based methods: UPGMA & NJ
    - Character-based methods: Maximum Parsimony
- Probabilistic models and associated algorithms (3)
    - Probabilistic models of evolution
    - Maximum likelihood algorithm
- Phylogenetic analysis algorithms (5)
    - Distance-based: UPGMA, Transformed Distance, Neighbors-Relation, Neighbor-Joining
    - Character optimization; delayed and accelerated transformation
    - Maximum Parsimony
    - Reliability of trees: Bootstrap, jackknife, decay, randomization tests.
- Phylogenetic trees and their comparison: (2)
    - Definition and description, various types of trees;
    - Consensus (strict, semi-strict, Adams, majority rule, Nelson).
    - Data partitioning and combination.
    - Tree to tree distances, similarity
- Case studies (2)
- Numerical taxonomy & bacterial identification (1)

Practicals
Objectives:
This course will enable the students to:
- use computational approaches for phylogenetic analysis
- explore and use packages available for molecular phylogeny

Syllabus:
- Compilation & curation of dataset, format conversion (1)

- Survey of software programs available for phylogenetic analysis (1)
    - Installation of at least 2 public domain packages for both Windows & Unix environment: Phylip, PAUP, MEGA
- MSA using MUSCLE (1)
    - Informative and variable sites; Singleton sites etc.

- Reconstruction of phylogenetic trees using distance-based methods (1 datasets) (3)
    - Converting sequence data into distance data
    - UPGMA, Neighbor-joining, Neighbor-relations & Transformed distance
    - Difference between dendogram & phylogenetic tree

- Reconstruction of phylogenetic tress using character-based methods
    - Maximum Parsimony (1)
    - Maximum likelihood (1)

- Using bootstrapping tool to generate multiple datasets from the original input data & generation of consensus tree (1)
- Plotting, visualizing & printing phylogenetic trees: TreeView and other tools (1)
    - Various rendering
    - Formatting & labeling
    - Interpretation of trees
- Comparison of trees drawn using (2)
    - RNA, Nucleotide & protein data
    - Gene tress & species tree
- Un-rooted & rooted tree (1)
    - Rooting un-rooted tree using an out group
- Reconstruction of phylogenetic trees using whole genome data of viruses (1)

- Numerical taxonomy & Bacterial identification using matrices (1)

References:

- Bromham Lindell. Reading the Story in DNA: A Beginner's Guide to Molecular Evolution. Publisher: USA, Oxford University Press. 2008. ISBN: 9780199290918.
- Bernardi Giorgio. Structural and Evolutionary Genomics, Volume 37: Natural Selection in Genome Evolution (New Comprehensive Biochemistry). Publisher: Netherlands, Elsevier Science.2005. ISBN: 9780444521361.
- Marco Salemi, Anne-Mieke Vandamme. The phylogenetic handbook: a practical approach to DNA and protein phylogeny. Publisher: Cambridge University Press, 2003. ISBN: 052180390X.
- Patthy Laszlo. Protein Evolution. Publisher: London, Blackwell Science Ltd. 1999. ISBN: 0632047747.
- Takahata Naoyuki, Clark Andrew G.(Editor). Mechanisms of Molecular Evolution: Introduction to Molecular Paleopopulation Biology. Publisher: Japan, Japan Scientific Socities Press and Sinauer Associates, Inc. 1993. ISBN: 476226718X.
- Graur Dan, Wen-Hsiung Li. Fundamentals of molecular evolution Publisher: Sinauer Associates, 1991. ISBN: 0878932666.

BIE 209 (T+P): Programming in Java (2T + 2P)

***Objectives*:**

The course will enable the students to:

- Acquire skills in programming using object oriented language: Java
- Develop and implement programs to analyze biological data

***Syllabus*:**

- An introduction to Java programming
- Features of Object-Oriented programming
- Java Basics
- Working with objects
- Arrays & Control structures
- Modifiers, Access control and class design
- Working with data structures
- Packages & Interfaces
- Using Native methods & Libraries
- Exception handling in Java
- Multithreading
- File I/O
- Brief Introduction to Applets & AWT

***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
- Horstmann Cay S.Cornell Gary. Core Java Volume II Advanced Features. Publisher: New Delhi, Education
- Schildt Herbert. Java 2: The Complete Reference Publisher: New Delhi, Tata McGraw-Hill Publishing Company.
- Bruce Eckel. Thinking in Java. Publisher: Pearson Education Flanagan David. Java in a Nutshell: A Desktop Quick Reference. Publisher: Mumbai, Shroff Publishers & Dist
- Balagurusamy E. Programming With Java A Primer. Publisher: New Delhi, Tata McGraw Hill

BIE 210 (P): Linux Operating System (2P)

Objectives:

This course will enable the students to:

- Writing basic and advanced Linux/Unix commands
- Understand and write shell scripts

Syllabus:

Introduction to Unix/Linux, Linux distributions	(1)
Using graphical and command line environments	(1)
Linux file system and file/directory manipulation commands	(3)
File permissions and commands to modify permissions	(1)
File manipulation commands, redirection, pipes	(4)
Searching and filtering files, grep command	(4)
Monitoring system resources	(1)
Text editors: basic editing and navigation	(2)
Shell Basics, writing and executing scripts	(1)
Conditional statements: If-else-elif, test, logical operators-AND, OR, NOT	(3)
Loops: while, for, until, break & continue	(3)
Command line arguments	(1)
Advanced commands: process management, package management, sed and awk commands, archiving and compressing files	(5)

References:

- Richard Blum and Christine Bresnahan, Linux Command Line and Shell Scripting Bible, 3ed, Wiley publishers. ISBN: 8126554983
- Richard Petersen, Linux: The Complete Reference, Sixth Edition, McGraw Hill Education. ISBN: 0070222940
- Sumitabha Das, Unix Concepts And Applications, McGraw Hill Education. ISBN: 0070635463

BIE 211 (T+P): Software Testing (1T + 1P)

***Objectives**:*

- Enable students to become aware of errors and inaccuracies in software programs and software routines.
- To learn the process of making error free new or existing computing software programming systems/packages.
- To optimize the performance of software programming systems/packages.

Aim:

To enhance the quality of source code, software program, web-enabled databases released in open source or charged environment.

Syllabus:

- Principles of testing- test-case, test scenarios, different methods in testing. (1)
- Principle of Automation Testing (1)
- JUnit testing (Brief introduction) (1)
- Types of testing: black box testing, white box testing. (1)
- Defect life cycle, STLC (1)
- Different methodologies for: V model, water fall, Agile, continuous delivery, test driven development, extreme programming. (3)
- Databases Testing (1)
- Fundamentals of Server-side Testing: What is server, server types, types of testing servers (2)
- Web-services testing. (1)
- Algorithmic Testing (1)
- Some existing tools used in Software Testing industry (2)

Practicals:

- Perform testing of existing well known General Purpose applications, Life Sciences/ Bioinformatics applications. Store and analyze the results.
- Testing of Application of languages/interpreters (perl, Java, etc) and other standard services / desktop programs**.** (15)

Semester III

BIM 301 (T+P): Immunology and Immunoinformatics (1T+1P)

Theory

Objectives:

This course will enable the students to:

- understand the immune system, its components and their functions
- study immunological databases
- study informatics-based approaches for prediction of epitopes, design of vaccines and immuno-diagnostic tools

Syllabus:

- Immune systems: (3)
    - Innate and adaptive immunity in vertebrates
    - Antigen, processing and presentation
    - Autoimmune diseases
    - Categories of vaccines
- Antibodies (1)
    - Immunoglobulin classes, subclasses and diversity
    - Hybridoma technology: applications

- Membrane receptors for antigen (2)
    - The B-cell receptor (BCR)
    - The T-cell receptor (TCR)
    - The major histocompatibility complex (MHC)
    - Cluster of differentiation (CD)

- Immunoinformatics overview (1)
    - Reverse vaccinology
    - Rational Vaccine design
    - Evolution and escape due to variations

- Immunoinformatics resources: (3)
    - IMGT, IEDB, STCRDab

- Molecular interactions (1)
    - Ag-Ab: Sequential and conformational epitopes
    - MHC - peptide – T-cell receptor

- Epitope prediction algorithms: (3)
    - B-cell epitope prediction
    - T-cell epitope prediction

- Vaccine design: (1)
    - Pipeline & workflows
    - Prediction of immunogenicity

Practicals

Objectives:

This course will enable the students to:

- understand and explore immunoinformatics resources
- characterize and understand principles of antigen-antibody interactions
- understand algorithms and methods for prediction of epitopes
- explore and use approaches for vaccine design

Syllabus:

- Browsing and searching Immunoinformatics resources to retrieve data (6)
- Study of Ag-Ab interactions (3)
    - Searching for Ag-Ab complexes
    - Charaterisation of binding sites
        - Ab-protein complex
        - Ab-peptide complex
        - Geometric and molecular interaction
- Prediction of epitopes and estimation of accuracy of predicted epitopes using experimental data (6)
    - sequential B-cell epitopes
    - conformational B-cell epitopes
    - cytotoxic T-cell epitopes
    - helper T-cell epitopes

References:

- Flower Darren R. Bioinformatics for Vaccinology Publisher: UK, John Wiley & Sons Inc. Latest Edition.
- Flower Darren R. Immunoinformatics: Predicting Immunogenicity In Silico Publisher: New Jersey,Humana Press. Latest Edition
- Kindt, Thomas J.,Osborne Barbara A.Goldsby Richard A. Kuby Immunology 6th Edition. Publisher:New York, W. H. Freeman. Latest Edition.
- Lund Ole, Nielsen Morten, Lundegaard Claus, Kesmir Can, Brunak SÃ¸ren. Immunological Bioinformatics. Publishesr: London, MIT PRess Latest Edition
- Foundation Novartis.Immunoinformatics: Bioinformatic Strategies for Better Understanding of Immune Function. Publisher: Chichester,John Wiley & Sons Inc. Latest Edition.
- Roitt Ivan, Delves Peter. Roitt's Essential Immunology 10th Edition.Publisher: Canada,Blackwell. Latest Edition.
- Ellis Ronald W. Vaccines: New Approaches to Immunological Problems. Publisher: Stoneham, Butterworth-Heinemann Latest Edition
- Dattatraya V. Desai and Urmila Kulkarni-Kale. T-cell epitope prediction methods: an overview. Publisher: Springer. Latest Edition

BIN 302 (T+P): Comparative Genomics (2T + 2P)

Theory

Objectives:

This course will enable the students to:

- understand and explore the field of comparative genomics
- explore computational tools for genome alignments & analysis
- understand applications of comparative genomics

Syllabus:

- Objectives and Overview of Genome Comparisons (2)
- Methods and Tools for Comparative Genomics (6)
  - Genomic and Pairwise BLAST
  - PipMaker
  - LAST
  - Mauve
- Comparison of Gene Order & Synteny analysis (2)
- Recombination methods and analysis (2)
  - RDP
- Analysis of Population structure (2)
  - Prokaryotic
  - Eukaryotic
- Population diversification: dN/dS ratio and selection pressure analysis (2)
- Single Nucleotide Polymorphism, dbSNP and other SNP-related databases (4)
- GWAS resources and analysis of polymorphism in populations (5)
- Comparative Genomics (5)
  - Viruses
  - Microbes
  - Pathogens
  - Eukaryotes

Practicals

Objectives:

This course will enable the students to:

- understand use of comparative genomics for genome annotations
- explore computational methods tools for genome alignments & analysis
- explore methods to study viral population diversity, recombination analysis etc.

Syllabus:

- Comparative genomics resources and NCBI and EBI (2)
- Comparison of full / partial genomic sequences using following methods to identify conserved genes and map/compare the annotations of the two sequences
  - Ensembl Genome Browser and Comparisons (3)
  - BLAST2 (1)
  - MegaBLAST and Discontiguous MegaBLAST (2)
  - PipMaker (2)
  - LAST (1)
  - Mauve (1)
- RDP (1)
- STRUCTURE using case studies of - (3)
  - virus

    - microbes
    - model organisms
- Calculating dN/dS ratio and selection pressure analysis (2)
- Explore and query SNP and SNP-related databases (2)
- GWAS resources and analysis of polymorphism (3)
- Workout any one comparative genomics case study based on publication & reproduce the output. (5)

References:

- Research papers & review papers covering various algorithms & methods
- Manuals & documentations of various software programs.
- Xia, Xuhua, Comparative Genomics, 2013, Springer-Verlag Berlin Heidelberg, 978-3-642-37146-2
- Mount, David. Bioinformatics: sequence and genome analysis. Publisher: Cold Spring Harbor Laboratory Press, 2004. ISBN: 087967121
- Arcady R. Mushegian, Foundations of Comparative Genomics, 2007Publisher: Academic Press, ISBN-10: 0120887940

BIN 303 (T+P): Molecular Modeling & Simulations (2T + 2P)

Theory Syllabus:

- Overview of molecular modeling (1)
- Potential energy surfaces (1)
- Molecular mechanics (6)
  - Force fields
- Energy Optimization algorithms (4)
  - Golden Section
  - Derivative based (SD, CG)
- Molecular dynamics simulations (8)
  - Introductory Stat. Mech., ensembles
  - Newtons laws of motion and integration schemes
  - Analysis
  - Free energy calculations
  - Normal Mode /Principal Component Analysis
  - Case Studies
- Electrostatics (2)
  - Poisson Boltzmann
- Docking simulations (6)
  - Thermodynamics of protein-ligand biding
  - Overview and review of popular algorithms
  - Virtual Screening
- Multiscale modeling (2)

Practicals

Objectives:

This course will enable the students to

- get hands on experience on tools used in molecular modeling

Syllabus:

- Modeling basic building blocks to macromolecules (4)
- Plotting PES of simple molecules (2)
- Calculation of energy using FF terms simple molecules (2)
- Coding for a non-gradient based and a gradient based optimization algorithm (6)
- Running MD simulations and analysis (8)
- Calculating electrostatics for biomolecules (2)
- Docking simulations (6)

References:

- Schlick, T. Molecular modelling and simulation: an interdisciplinary guide. Publisher: Springer. 2002. ISBN: 0-387-95404-X
- Leach, Andrew. Molecular Modelling: Principles and Applications. Publisher: Prentice Hall. 2001. ISBN: 0582239338.
- Jensen, Jan H. Molecular Modeling Basics. Publisher: CRC Press. 2010. ISBN: 978-1420075267
- Hinchliffe Alan. Molecular modelling for beginners. Publisher: John Wiley and Sons Ltd. 2008. ISBN: 978 0470513149

BIN 304 (T+P): Scientific Data Mining and Visualization (2T + 2P)

Theory

Objectives:

The course will enable the students to:

- Handle large amounts of data generated in experiments
- Organize, perform, and write-up data analyses
- Mine the data to make sense out of it
- Get familiar with concepts of visualizing data
- Visualize the biological data using various tools/techniques & derive knowledge from it

Syllabus:

- Concepts of Shannon Entropy and Information Theory (2)
- Optimization techniques:
  - Random walk (1)
  - Monte Carlo (2)
  - Simulated Annealing (1)
  - Genetic Algorithm and Ant colony optimization (2)
- Introduction to Statistical & Machine Learning techniques
  - Supervised and unsupervised Classification (1)
  - Hierarchical and k-means clustering (1)
  - K Nearest Neighbour classifier (1)
  - Support vector Machines & Attribute Selection methods (2)
  - Decision Tree (1)
  - Random Forest (2)
  - Linear Regression (1)
  - Performance measures for classification and regression (1)
  - PCA (1)
  - ANOVA (one-way and two way) (2)
  - Markov models and Hidden Markov Models (3)
  - Artificial Neural Networks (3)
- Big Data: Concepts, Sources, Technological advances (2)
- Introduction to data visualization: (1)
  - What is data visualization?
  - Need for scientific data visualization
  - Advantages and Applications

Practicals Syllabus:

- Programs for Random walk, Monte Carlo and Simulated Annealing (3)
- Tools and Techniques for Classification of Data (3)
- Applications and case studies using kNN, Decision Tree and Random Forest (3)
- Applications of Classification Regression (3)
- Use of 'R' for biological data processing and data visualization: (6)
  - Introduction to R environment
  - Data types and their properties
    - Vectors, Factors, Arrays & Matrices, Lists & Data Frames
  - File IO
  - Data grouping and Control statements
  - Functions
- Different statistical plots with R/ R packages: (4)
  - Histogram, line, pie, box-whisker etc.
  - 3-D interactive plots
  - Heat maps, contour plots

        - Contour lines, Streamlines, Streaklines,
            - Utilities provided by “grid” package
            - Introduction to “lattice” & “ggplot2” packages
- Statistical data analysis with R/ R packages: (8)
    - PCA, ANNOVA, Clustering (K-means and hierarchical)
    - Regression, correlation, Fitting a regression line, Multiple regression
    - Packages:
        - Standard Packages in R
        - Creating custom packages
        - Case Studies using BioConductor Project (CRAN)

References:

- Curry Edward. Introduction to Bioinformatics with R: a practical guide for biologists Publisher: CRC Press. ISBN-13: 978-1138495715 ISBN-10: 1138495719
- Richard Cotton. Learning R: A Step-By-Step Function Guide to Data Analysis. Publisher(s): O'Reilly Media, Inc. ISBN: 1449357199, 9781449357191
- Flach, Peter. Machine Learning: the Art and Science of Algorithms that Make Sense of Data Cambridge University Press ,2012
- Theodoridis, Sergios. Machine Learning: A Bayesian and Optimization Perspective 2nd Ed. Academic Press, 2020.
- Zhang Yan-Qing , Rajapakse, Jagath C. Machine Learning in Bioinformatics John Wiley & Sons, 2009
- Goldberg, David. Genetic Algorithms: in Search, Optimization & Machine Learning Pearson Education Asia , 2002
- Foulkes, Andrea. Applied Statistical Genetics with R: for Population–based Association Studies. Springer, 2009
- Ortutay, Csaba and Ortutay, Zsuzsanna. Molecular Data Analysis Using R. Wiley Blackwell 2017
- Shahbaba B. Biostatistics with R. Springer, 2011
- Cormen, Thomas, Leiserson , Charles, Rivest, Ronald & Stein, Clifford. Introduction to Algorithms 3rd Ed. PHI Learning Private Ltd., 2014
- Horowitz, Ellis, Sahni, Sartaj, Rajasekaran Sanguthevar. Fundamentals of Computer Algorithms 2nd Ed. University Press, 2008
- Affenzeller, Michael, Winkler, Stephan, Wagner, Stefan & Beham, Andreas Genetic Algorithms and Genetic Programming: Modern Concepts and Practical Applications. CRC Press, 2009.
- Mitchell, Melanie. Introduction to Genetic Algorithms Prentice-Hall of India, 1998
- Prabhu, C. S. Chivukula, Aneesh, Mogadala, Aditya, Ghosh, Rohit, Livingston, Jenila, Big Data Analysis: Systems Algorithms Applications, Springer, 2019

BIN 305 (P): Project (4P)

Objectives:

Students will be able to

- explore field of interest and literature to gain state of art knowledge
- identify and define research problem for the project
- search and review literature

Contents:

- Identification of Research Problem
- Literature Review and Review Writing

Elective Courses

BIE 306 (T): Biology of Diseases (2T)

Theory

Objectives:

This course will enable the students to:

- understand the etiology, available therapies & future prospects for the enlisted diseases
- understand molecular level details of host-pathogen and vector-pathogen relationship where relevant.
- the role of bioinformatics in identification of disease markers, drug and vaccine development.

Syllabus:

- Infectious diseases and their causative agents
    - Bacterial- *Mycobacterium tuberculosis, M. leprae*, Antibiotic resistant medically relevant bacteria (5)
    - Viral- HIV, Influenza, Dengue, Other relevant viruses (5)
    - Parasitic- *Plasmodium, Leishmania*, Trypanosomes, filarial worms, other relevant organisms (5)
    - Prions (2)
- Non-infectious diseases:
    - Cancer (5)
    - Diabetes mellitus (2)
    - Cardiovascular diseases (2)
- Human microbiome & its relation to health and disease (4)

References:

- Graham F. Hatfull, William R. Jacobs, Jr. Molecular genetics of Mycobacteria. Publisher: ASM Press. 2014. ISBN: 978-1555818838
- Hays, J. D. Ed.;Wolf,C.R. Ed. Molecular genetics of drug resistance. Publisher: Harwood Academic 1997 ISBN: 9057021676
- Morse, Stephen S. Emerging Viruses. Publisher: Oxford University Press. 1996. ISBN: 978-0195104844

Other learning resources:

- Relevant Henry Stewart Talks
- Research papers relevant for the topics

BIE 307 (T+P): Pathways & Networks (1T + 1P)

Theory

Objectives:

This course will enable the students to:

- Understand the basic biochemical concepts of metabolic pathways.
- Understand metabolic pathway databases and its uses.
- Understand metabolic pathway networks to understand metabolic systems.

Syllabus:

- **Gene and Regulatory Networks:** (4)
    - Types of molecular interactions – overview / revision
    - Types of biological networks- Gene regulatory network, gene co-expression network, metabolic networks.
    - Properties of metabolic Networks
- Representation of biological networks: Graphs (3)
- Types of networks : Linear, Boolean, Bayesian (3)
- **Organization of metabolic systems**: Components of pathways; Modes of enzymes- enzyme chains, multi-enzyme complexes, multifunctional enzymes: Modes of pathways- linear, branched and cyclic pathways; Types of pathways-anabolic, catabolic, amphibolic, anaplerotic, cataplerotic, shuttle pathways, secretory pathways and signal transduction; Regulation of pathways- reversible and irreversible regulation, feedback and feedforward regulation; regulatory enzymes, pacemaker enzymes and linker enzymes. (5)

Practicals

Syllabus:

- Browse & search following metabolic pathway databases (3)
    - KEGG, BioCyc, MetaCyc, EcoCyc, Reactome, BiGG, Boehringer Mannheim - Biochemical Pathway database.
- Browse & search following metabolic disease databases (2)
    - HMDB, MPMP, RAMEDIS
- Browse & search following Enzymes, Compounds and Reactions databases
    - BRENDA, ExplorEnz, IntEnz, ExPASy Enzyme database, LIGAND, ICEED (2)
- Visualization and reconstruction of gene regulatory networks using Cytoscape. (4)
- Building of co-expression networks. (4)

References:

- Nelson David L., Cox Michale. Lehninger Principles of Biochemistry 5th Edition. Publisher: New York. W. H. Freeman. 2008. ISBN 978 0716771081.
- Berg, Jeremy M, Tymoczko, John L. Stryer, Lubert. Biochemistry 6th Edition. Publisher: New York: W.H. Freeman. 2007. ISBN: 071676766X.
- Hames David, Hooper Nigel. Instant Notes in Biochemistry 3rd Edition. Publisher. Nodia, Taylor & Francis. 2007. ISBN: 185996 2491.
- Voet, Donald, Voe Judith, Pratt, Charlotte W. Fundamentals of Biochemistry: Life at the molecular Level 2nd Edition. Publisher: Asia, John Wiley & Sons. 2006. ISBN: 0471753416.
- Horton, Robert, Moran, Laurence A, Scrimgeour, Perry Gray Marc, Rawn. David. Principles of biochemistry. Publisher: New Jersey, Pearson Prentice Hall, 2006. ISBN: 0131453068.

- Mathews, Christopher K. van Holde, K. E., Ahern Kevin G. Biochemistry. Publisher: San Francisco, Pearson Education 2000.ISBN: 81 29702150.
- Zubay, Geoffrey. Biochemistry 4th Edition, Publisher: Boston,Wm C. Brown, 1998, ISBN 0697219003.
- Garrett, Reginald H., Grisham, Charles M. Principles of biochemistry: with a human focus. Publisher: Australia Brooks/Cole, Thomson Learning, 1997. ISBN:0030973694.
- Nelson David L., Cox Michale. Lehninger Principles of Biochemistry 5th Edition. Publisher: New York. W. H. Freeman. 2008. ISBN: 9780716771081.
- Voet, Donald, Voe Judith, Pratt, Charlotte W. Fundamentals of Biochemistry: Life at the molecular Level 2nd Edition. Publisher: Asia, John Wiley & Sons. 2006. ISBN: 0471753416.
- David Greenberg Metabolic Pathways, Third Edition, ACADEMIC PRESS, 1970.
- Purich Daniel L., Allison R. Donald. The Enzyme Reference: A Comprehensive Guidebook to Enzyme Nomenclature, Reactions, and Methods. Publisher: California, Academic Press. 2002. ISBN: 0125680414.
- Collado-Vides Julio, Hofestadt Ralf. Gene Regulation and Metabolism: Post-Genomic Computational Approaches. Publisher: England, The MIT Press. 2002. ISBN: 026203297X.
- Computing for Biologists- A. Fielding, Addison Wesley Pub., UK (1985).
- Microcomputers in Biochemical Education- E. J. Wood (Ed), Taylor and Francis Ltd., UK (1984).
- Computer Games and Simulation for Biochemical Engineering- H. R. Bungay, John Wiley and Sons Ltd., New York (1985).
- Microcomputers in Biology- A practical approach- C. R. Ireland and S.P. Lang, IRL Press Ltd., (1985)

BIE 308 (P): Laboratory Course in Experimental Immunology (2P)

Practicals

Objectives:

- The course will help the students to understand tools and techniques used to generate immunological data

Syllabus:

A: Classical Techniques

1. Ouchterlony Double Diffusion (4)
2. Immunoelectrophoresis (4)
3. Rocket Immunoelectrophoresis (4)
4. Haemagglutination (4)

B: Advanced Techniques

1. Western blotting/Co-immunoprecipitation/Enzyme Linked Immunosorbent assay/Chromatin immunoprecipitation (14)

BIE 309 (P): Web Design (2P)

Objectives:

This course will enable the students to:

- Design web pages and forms
- Perform backend data processing using files and databases

Syllabus:

HTML tags for front end design	(3)
Web page styling using CSS	(2)
PHP scripting	(3)
Form design	(2)
Backend processing of form data in PHP	(2)
Javascript functions for data validation	(3)
MySQL database connectivity	(3)
Server side scripting in Python	(3)
Implementing online database and webserver	(2)
Mini project	(7)

References:

1. Learning Web Design: A Beginner's Guide to HTML, CSS, JavaScript, and Web Graphics by Robbins Jennifer, Shroff/O'Reilly; Fifth edition ISBN: 978-9352137381
2. Responsive Web Design with HTML5 and CSS: Build future-proof responsive websites using the latest HTML5 and CSS techniques by Ben Frain, Packt Publishing Limited; 4th edition ISBN: 978-1803242712
3. Secure Web Application Development: A Hands-On Guide with Python and Django by Matthew Baker, Apress 1st edition ISBN: 978-1484285954
4. PHP & MySQL: Server-side Web Development by Jon Duckket, Wiley 1st edition ISBN: 978-1119149217

Semester IV

Core Courses

BIN 401 (P): Project (6P)

Objectives:

This course will enable students to

- learn how to formulate, design and plan scientific research problems
- review the literature
- write project thesis in a structured manner
- present research work

BIN 402 (P): Seminars (2P)

Objectives:

This course will enable students to

- explore various topics of research in Bioinformatics and allied areas
- search, retrieve, read and understand research papers in the selected topic of interest
- prepare and deliver structured presentation on the selected research topic based on the relevant research papers
- raise and answer questions on the research work being presented

BIN 403 (T+P) : Transcriptomics and Proteomics (2T+2P)

Theory

Objectives:

This course will enable the students to:

- Understand the basic concepts of proteomics and its applications.
- Understand the role of Bioinformatics in the study of proteome analyses.

Syllabus:

Transcriptomics

- Regulation of transcription in prokaryotes and eukaryotes, dynamics in gene expression (2)
- Principle and methodology of tools to study comparative gene expression – RT-PCR, qRTPCR, differential display RT-PCR, cDNA micro-array, Northern blotting (4)
- Transcriptome analysis- RNA-seq, miRNA-seq (6)
- Computational methods for Transciptome analysis and case studies (3)

Proteomics

- **Organization of proteome**: Basic concept of proteome; history, types of proteomics and applications of proteomics. (3)
- **Experimental techniques in Proteomics:** (6)
  - Conventional techniques: Cell disruption methods, chromatography (GF, IEC, AC, HPLC, RP-HPLC, Mud-LC), dialysis, ELISA and Western blotting.
  - Advance techniques: Gel based techniques (1DE, 2DE, 2-D DIGE), In-gel digestion, Protein sequencing methods, Mass spectrometry (LC-MS, ESI-Q-MS, MALDI-TOF-MS, SELDI-TOF-MS, ESI-MS/MS, etc.), yeast two hybrids, Protein microarrays.
  - Quantitative techniques: ICAT, SILAC, iTRAQ.
  - High throughput techniques: X-ray crystallography, NMR, AFM, FRET, etc

- **Computational methods in Proteomics:** (6)
  - Proteomics databases- PRIDE, SWISS-2DPAGE, Human Protein Atlas.
  - MS data analysis tools- ProFound, MASCOT, SEQUEST, prospector, FindPept, InsPect, MALDIPepQuant, PepFrag, etc.
  - Protein identification tools- PeptideMass, AACompldent, Taglent, HAMAP.
  - Protein characterization tools- PeptideCutter, PredictProtein, PROPSEARCH, PSORT, TargetP, SMART, ELM,ProtParam, GPMAW lite,etc.
  - PTM analysis tools- FindMod, GPS, NetAcet, NetPhos,NetNGlyc, SUMOsp CSS-Palm, etc
  - Protein-protein interaction databases-STRING, DIP, BIND, BioGRID, MINT, MIPS-MPact, MIPS-MPPI; Meta-databases (APID, MPID8, PINA); PPIs prediction databases (PIPs, I2D, UniHI).

Practicals

Objectives:

- Proteomics concepts and technologies currently used to unravel protein functions and their role health and diseases.
- To use bioinformatics approaches for identification and characterization of proteins and peptides.

Syllabus:

Transcriptomics

- Browse and search in regulatory databases (both for prokaryotes and eukaryotes) (2)
- Identification of regulatory motifs using packages like MEME. (2)
- Browse and download RNA-seq data from NCBI GEO and Quality check. (4)
- Hands-on analysis of RNA-seq data (6)

Proteomics

- Browse and search the proteomics databases. (4)
- Identification of proteins with its molecular weight and pI from sequence using bioinformatics tools. (2)
- MS data analysis using ProFound/MASCOT/prospector tools. (3)
- Prediction of protein motifs and domain using online tools. (2)
- Prediction of protein-protein interaction and protein-ligand interaction. (3)
- Visit to virtual proteomic laboratory and ExPASy Bioinformatics Resource Portal. (2)

References:

- Introduction to Proteomics: Principles and Applications by Nawin Mishra, John Wiley & Sons.
- Principles of Proteomics Second edition by Richard Twyman, Garland Science, 2013.
- Introduction to Proteomics: Tools for the New Biology by Daniel C. Liebler, Springer Science & Business Media, 2002.
- Integrative Proteomics by Hon-Chiu Eastwood Leung, JanezaTrdine 9, 51000 Rijeka, Croatia.
- Proteomics: Biomedical and Pharmaceutical Applications by Hubert Hondermarck, Kluwer Academic Publishers, 2004.
- Bioinformatics for Comparative Proteomics by Cathy H. Wu, Humana Press, 2011.
- Eng JK, McCormack AL, Yates JR III. 1994. An Approach to Correlate Tandem Mass
- Spectral Data of Peptides with Amino Acid Sequences in a Protein Database. J. Am. Soc. Mass Spectrom. 5:976. PubMed
- Ficarro SB, McCleland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, Hunt
- DF, White FM. 2002. Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. Nature Biotechnology. 20(3): 301.

Web Resources of Mass Spectrometry:

- http://i-mass.com/guide/protocol.html - General proteomics site containing information about protocols and resources
- www.cellmigration.org –Research Resources - Consortium Activities – Protein Discovery - Proteomics
- http://us.expasy.org/ - Proteomics tools and resources
- http://prospector.ucsf.edu/ - Mass spec data analysis
- http://depts.washington.edu/~yeastrc/ms_home.htm- General tandem mass specinformation
- http://www.ionsource.com/Card/ester/esters.htm - Making methyl esters
- http://www.ionsource.com/Card/cmc/why.htm - Reduction of proteins
- http://www.ionsource.com/Card/cmc/intro.htm - Reduction/alkylation of proteins
- http://prowl.rockefeller.edu/recipes/sproteas.htm - Proteolysis

BIN 404 (T+P): Metabolic Engineering & Systems Biology (2T + 2P)

Theory

Objectives:

This course will enable the students to:

- learn how metabolic pathways can be modeled, analyzed and engineered using mathematical tools
- apply advanced methodologies to model pathways

Syllabus:

- Overview / Introduction to Metabolic Engineering (1)
- Why mathematical models (1)
- Mathematical modeling of biochemical pathways (3)
  - Pathway types - Metabolic, Signaling, Gene regulation
  - Introduction to types of mathematical modeling - Boolean, ODE, PDE, discrete stochastic, Constraint based modeling (FBA)
- Mathematical model database (1)
  - Biomodels database, PantherDB
- Boolean modeling details (4)
- Ordinary differential equation based deterministic modeling (4)
- Constraint based modeling - Flux balance analysis (4)
- Metabolic Control Analysis
- Properties of network (4)
  - Topological properties of network - load-choke point analysis, centrality analysis
- Metabolic Network Reconstruction
- Applications of systems biology
  - Metabolic engineering - Rational metabolic network design basics (eg. knock-out, knock-in), Biobricks database, open wet ware (4)
  - Drug discovery - drug target identification (eg. knock-out, knock-down prediction, vulnerability prediction) (4)

Practicals

Objectives:

This course will enable the students to:

- Use visualization tools for understanding properties of biochemical networks
- use advanced mathematical tools to model metabolic pathways

Syllabus:

- Cell Net Analyser - For boolean modeling (5)
- CellDesigner, Copasi, SBML format - For ODE based modeling (7)
- GSMN-TB - For FBA (3)
- Cytoscape - Topological properties of biochemical networks (4)
- MATLAB introduction, SBML toolbox, SBtoolbox, COBRA (6)
- Mini project (5)

References:

- Systems Biology: A Textbook, 2nd Edition by Edda Klipp, Wolfram Liebermeister, Christoph Wierling, Axel Kowald. ISBN: 978-3-527-33636-4. 2016. Wiley-Blackwell
- Modeling Biological Systems: Principles and Applications by James W. Haefner. ISBN: 978-0-387-25011-3. 2005. Springer New York, NY
- Elements of Computational Systems Biology by Huma M. Lodhi, Stephen H. Muggleton. ISBN: 978-0-470-18093-8. 2010. Wiley-Blackwell

- Fundamentals of Systems Biology From Synthetic Circuits to Whole-cell Models, 1st Edition. by Markus W. Covert. ISBN 9781138582927. 2015. CRC Press
- Systems Biology Philosophical Foundations by Frank J. Bruggeman, Fred Boogerd, H.V. Westerhoff, Jan-Hendrik S. Hofmeyr. ISBN: 9780080475271. 2007. Elsevier Science
- Big Mechanisms in Systems Biology: Big Data Mining, Network Modeling, and Genome-Wide Data Identification by Bor-Sen Chen, Cheng-Wei Li. ISBN: 9780128097076. 2016. Elsevier Science
- Systems Biology: Definitions and Perspectives by Hans V. Westerhoff, Lilia Alberghina. ISBN: 9783540742692. 2007. Springer
- Systems Biology by Ivan V. Maly. ISBN: 9781603278768. 2008. Humana Press
- Contextualizing Systems Biology: Presuppositions and Implications of a New Approach in Biology by Martin Döring, Imme Petersen, Anne Brüninghaus, Regine Kollek. ISBN: 9783319171067. 2015. Springer International Publishing
- Computational Systems Biology by Jason McDermott, Reneé Ireton. ISBN: 9781597452434. 2009. Humana Press
- Bioinformatics for Systems Biology by Stephen Krawetz. ISBN: 9781607612070. 2009. Humana Press
- Kinetic Modelling in Systems Biology by Oleg Demin, Igor Goryanin. ISBN: 9781420011661. 2008. CRC Press
- Systems Biology: Simulation of Dynamic Network States by Bernhard Ø. Palsson. ISBN: 9781139495424. 2011. Cambridge University Press
- An Introduction to Systems Biology: Design Principles of Biological Circuits by Uri Alon. ISBN: 9781584886426. 2007. Taylor & Francis

BIN 405 (T+P): Modern Drug Design (1T + 1P)

Theory

Objectives:

This course will enable the students to:

- understand the basic concepts of modern techniques of drug discovery and design
- appreciate success stories of drugs developed using modern methods

Syllabus:

- History of drug discovery & modern discovery process (1)
    - Drug vs inhibitor (kinds of inhibitors: competitive, uncompetitive etc)
    - Inhibition constants: IC_{50}, K_i, LD_{50}
- Target identification and Validation, Polypharmacology (1)
- Sources of small molecules: High throughput screening and molecular libraries initiative, natural and synthetic products, commercial libraries (1)
- Ligand based: 3D QSAR – Alignment & Non-alignment based, Pharmacophore modeling, Shape based screening, Bioisosteres (4)
- Structure based: Virtual screening, Ensemble based screening, De novo methods, Pharmacophores, (3)
- Drug repurposing (1)
- ADMET & predictions (1)
- Pharmacogenomics: pharmacokinetics, pharmacodynamics and case studies of variations in drug response (2)
- Success stories of rational design (1)

Practicals

Objectives:

This course will enable the students to

- get hands on experience on tools used in modern drug design

Syllabus

- Data mining databases for drug and drug-like compounds and their associated properties (3)
- Case study on 3D QSAR (3)
- Case study on Ensemble-based virtual screening (3)
- Case study on de-novo methods (3)
- Case study on ligand based pharmacophore screening (3)

References:

- Young David. Computational drug design: A Guide for Computational and Medicinal Chemists. Publisher: Wiley. 2009. ISBN: 9780470126851
- Stroud, Robert et al. Computational and Structural Approaches to Drug Discovery: Ligand-Protein Interactions. Publisher: Royal society of chemistry : Cambridge ISBN: 0854043659
- Gibaldi Milo. Biopharmaceutics and Clinical Pharmacokinetics. Publisher: PharmaMed Press. ISBN: 8188449067
- Abraham, Donald (Ed). Burger's medicinal chemistry and drug discovery. Publisher: John Wiley & Sons, Inc. 2003. ISBN: 0471270903

Elective Courses

BIE 406 (T+P): Translational Bioinformatics (1T+1P)

Theory

Objectives:

- To understand the concept of translational research and translational bioinformatics.
- To learn various biological and clinical data types and computational resources useful for translational bioinformatics.
- To be able to correlate the molecular data at various levels and of various types to the essential biomedical and clinical aspects and entities such as patients, diseases, symptoms, diagnostics, therapeutics etc.)

Syllabus:

- Data resources for translational bioinformatics (6)
    - Large Scale Genome Sequencing Projects
    - EnCode Project
    - The Cancer Genome Atlas
    - Exome Sequencing datasets
- Introduction to precision medicine with a case study (2)
- Identification of biomarkers to diagnose pre-disease states (2)
- Types and varieties of clinical data (1)
- Building probabilistic causal models of diseases – approaches and case studies (2)
- Case Studies in Translational Bioinformatics (2)

Practicals

Syllabus:

- Study of data resources: Exploring & browsing, understanding data formats, searching and retrieval of data (12)
    - Large Scale Genome Sequencing Projects
    - EnCode Project
    - The Cancer Genome Atlas
    - Exome Sequencing datasets
- Case study in translational bioinformatics (3)

References:

- PLOS Computational Biology www.ploscollections.org/translationalbioinformatics
- Readhead B. and Dudley J. Translational Bioinformatics Approaches to Drug Development Advances In Wound Care, 2012; 2(9):470-489.
- Tenenbaum JD. Translational Bioinformatics: Past, Present, and Future; Genomics Proteomics Bioinformatics; 2016; 14:31–41
- 1000 Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491(7422):56-65.
- ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489(7414):57-74.
- Gibson IB, Jiang R, Yu F. The 1000 Genomes Project: paving the way for personalized genomic medicine. Per Med. 2013;10(4):321-324.

BIE 407 (T+P): Advanced Genomics (1T + 1P)

Theory

Objectives:

This course will enable the students to:

- Understand advances in various fields of genomics.
- Understand genomic data archival resources.

Syllabus:

- Genome Projects in High Throughput Sequencing Era (2)
- ChIP-Seq, Bisulfite-treated methylation sequencing (4)
- Introduction to Metagenomics, Epigenomics. (6)
- **Genomic data archival resources:** (3)
    - MEGAN, MGRAST, Integrated Microbial Genomes (IMG/M), EBI-Metagenomics.

References:

- Primrose, S.B. and Twyman, R.M., Principles of Genome Analysis and Genomics (3rd Ed.) Publisher: Blackwell Publishing Company, Oxford, UK. 2003
- Mount, D., Bioinformatics: Sequence and Genome Analysis Publisher: Cold Spring Harbor Laboratory Press, New York. 2004

Practicals

Objectives:

This course will enable the students to

- Browse genomic data archival resources.
- Analyze high through-put sequence data.

Syllabus:

To browse the following genomic data resources

- ChIP-seq (5)
- Bisulfite-Seq (3)
- MEGAN (1)
- MGRAST (1)
- Integrated Microbial Genomes (IMG/M) (1)
- Metagenomics data analysis pipeplines (4)

BIE 408 (T+P): Advanced Molecular Phylogenetics (1T + 1P)

Theory

Objectives:

This course will enable the students to:

- understand the field of Phyloinformatics & phylogeography
- understand various models and role of ***Bayesian*** analysis using MCMC
- understand algorithms and methods for alignment-free phylogeny
- role of phylogeny and clustering in typing viral and bacterial species

Syllabus:

- Assessment of MSA models (2)
    - Need for benchmark datasets
    - Gene and genome based phylogeny
- Substitution models & fitness tests (1)
- Modeling sequence evolution using Bayesian methods (2)
- Molecular clock hypothesis (1)
- Phylogeography: methods & practices (2)
- Phylogeny programs: BEAUTI & BEAST (3)
- Alignment-free phylogeny (3)
- *In-silico* methods for viral and bacterial typing (1)

Practicals

Syllabus:

- Compilation of datasets and benchmarking (2)
- Assessment of alignment accuracy (1)
- Gene-based and Genome-based phylogeny (2)
- Using jmodeltest (1)
- Molecular clock assessment using MEGA & BEAST (1)
- Exploring the packages BEAUTI & BEAST for estimation of divergence (4)
- Phylogeny using RTD and other methods (2)
- Geno- & sero-typing of viruses & bacteria using RTD & MLST (2)

References:

- Holmes, Edward. Evolution and Emergence of RNA Viruses. Publisher: Oxford University Press, 2011. ISBN: 9780199211135
- Pevzner ,Pavel, Bioinformatics for Biologiest. Publisher: Cambridge University Press., 2011. ISBN: 9781107648876
- Vinicius, Lucio, Modular Evolution: How Natural Selection Produces Biological Complexity , Publisher: Cambridge University Press. 2010. ISBN: 9780521728775
- Lemey, Philippe, The Phylogenetic Handbook 2nd Ed: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing, Publisher: Cambridge University Press, 2009. ISBN: 9780521730716
- Bromham,Lindell, Reading the story of DNA: beginners guide to molecular evolution. Publisher: Oxford University Press, 2008 . ISBN: 978 0199220918
- Patthy Laszlo, Protein evolution 2nd ed. Publisher: Blackwell Publishing, 2008. ISBN: 978 1 405151665
- Nielsen, Rasmus. Ed. Statistical methods in Molecular Evolution. Publisher: Spinger, 2005. ISBN: 0387223339
- Isaev, Alexander. Introduction to Mathematical Methods in Bioinformatics. Publisher: Spinger, 2004. ISBN: 3540219730
- Takahata, N. Mechanisms of Molecular Evolution. Publisher: Japan Scientific Societies Press. 1993. ISBN: 0-87893-825-7

BIE 409 (T+P): Advanced Algorithms in Machine Learning (1T + 1P)

Objectives:

The students will learn

- the principles and algorithms of advanced statistical techniques and machine learning paradigm
- applications of the advanced machine learning algorithms in biological systems for predictive analyses and data classification

Theory
Syllabus:

- Support Vector Machines for classification and regression: Soft and hard margins (2)
- Naïve-Bayes classifier (1)
- Logistic regression (1)
- Different attribute selection methods (1)
- Introduction to Image processing and its application in Bioinformatics (2)
- Hidden Markov Models (2)
- Introduction to Time Series and different time series modelling techniques (1)
- Introduction to Deep learning methods (1)
- Introduction to neural networks (NN), convolutional neural network (CNN) and recurrent neural network (RNN) (2)
- Introduction to Large Language Models (LLMs), Transformers, BERT/ BioBERT (2)

Practicals
Syllabus:

- Applications of Support Vector Machines with biological data (4)
- Simultaneous Cancer classification and attribute selection using gene expression data (3)
- Programs for data preprocessing and classification/regression methods using R/ Python (2)
- Applications of ANN, CNN, RNN in biological data analysis: (Sequence and Structure data analysis, biomolecular interactions, function prediction etc.) (4)
- Text mining (2)

Books and references:

- The Elements of Statistical Learning, by Trevor Hastie, Robert Tibshirani, Jerome H. Friedman
- Pattern Recognition and Machine Learning, by Christopher Bishop
- Deep Learning, Goodfellow et al, MIT Press, 20172.

UGC recommended courses (Additional 10 credits)

Cyber security/Information security	**4C (T)**
Skill based credits	**4C (T)**
Human rights education	**2C (T)**