M.Tech Curriculum 2023-24

Department of Applied Science

(NEP based Curriculum)

Guidelines for M. Tech & Duel degree M. Tech-PhD 2023-24

PC- Programme Core, PE- Programme Elective, ES- Engineering Science, BS- Basic Science, OE- Open Elective, HM- Humanities and Management, MR- Minor, HR- Honor, AC- Audit Course

L-T-P (hr)

Lecture (L): 1 Credit = 1 hr/ week, Tutorial (T): 1 Credit = 1hr/week, Practical (P): 1 credit = 2 hr/week

The credits for all courses are given in LTP format. L stands for Lecture, T for Tutorial, P for Practice/Practical. L=3 signifies 3 hours of Lectures per week, T=1 signifies 1 hour of Tutorial per week, an P=1 corresponds to 2 hours Practice/Practical per week. For M. Tech. students, T and P need not be explicitly mentioned in the Time Table. The Practice/Practical component of the course could also be a Term paper or Term Paper along with the lab. The discretion is left to the Teacher teaching the course.

The dual degree M.Tech.-Ph.D. Programme has been kept reserved. This is in tune with the spirit of NEP 2020. It is envisaged with a target towards research leading to a PhD programme with a minimum duration of 04 years, and with a possible exit for the award of MTech degree and/or PG Diploma after securing required credits.

Please refer to the M.Tech. and Ph.D. Ordinance for the other details of the M. Tech. and Ph.D. programs that is available at the following link: https://aaa.iiita.ac.in/

The AS Department is offering the following specializations in each group:

Bioinformatics (M.Tech & Dual degree M.Tech-PhD)

Bioinformatics is one of the interdisciplinary courses which involve study of biology, computer science, and information technology. Bioinformatics uses both computation and assessment tools to collect and translate biological data. This is suitable for development of algorithms and tools suitable for large biological databases, design data frameworks, and develop and adjust algorithms. M. Tech Bioinformatics at the Indian Institute of Information Technology – Allahabad, is a four semester's program designed to develop the skills in Information Technology and Data science applied on Biological Data. M. Tech in Bioinformatics is an excellent avenue to build your career, which provides an interdisciplinary platform through the amalgamation of Information Technology, Statistics, Mathematics and Biology. The coursework is designed to include a balance of functional knowledge as well as practical learning spread over four semesters covering, Computer programming, Data structure and Algorithms, Database Management System/Biological Information System and Management(Applied on Large Scale Biological Data), Mathematics, Statistics, Machine Learning and emerging topics such as Big Data, Deep Learning, Pattern Recognition as electives.

Total Credit Distribution

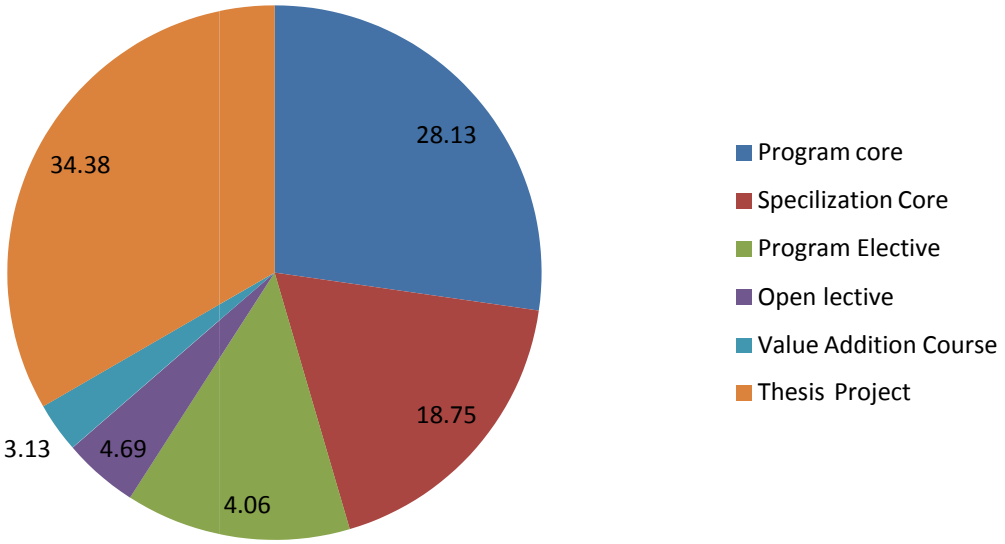
Semester	Semester I	Semester II	Summer Semester	Semester III	Semester IV	Total Credits
M. Tech	18	19	0	17	12	66
PG Diploma	18	19	5	-	-	42

Dual degree M. Tech-PhD*	20	19	***After Second semester, Dual degree M. Tech-PhD students will follow PhD ordinance/Curriculum**

M.Tech BI, Credit Distribution

M. Tech BI		I	II	III	IV	Total Credits	
	Program core	6	12			18	28.13 %
	Specialization Core	12				12	18.75%
	Program Elective		3	6		9	14.06%
	Open Elective			3		3	4.69%
	Value Addition Course		2			2	3.13%
	Thesis/Project		2	8	12	22	34.38%

1



Dual Degree M. Tech-PhD BI, Credit Distribution

M. Tech BI		I	II	III	IV	V	VI	VII	VIII	Total Credits	
1.	Program core	6	12	8						26	20%
2.	Specialization Core	12								12	10%
3.	Program Elective		3							3	2.5%
4.	Open Elective			3						3	2.5%
5.	Value Addition Course		2							2	1.6%
6.	Thesis/Project/Summer Project		2	8						10	8%
7.	Research Credit/ Colloquium			8	12	12	12	12	12	68	57%

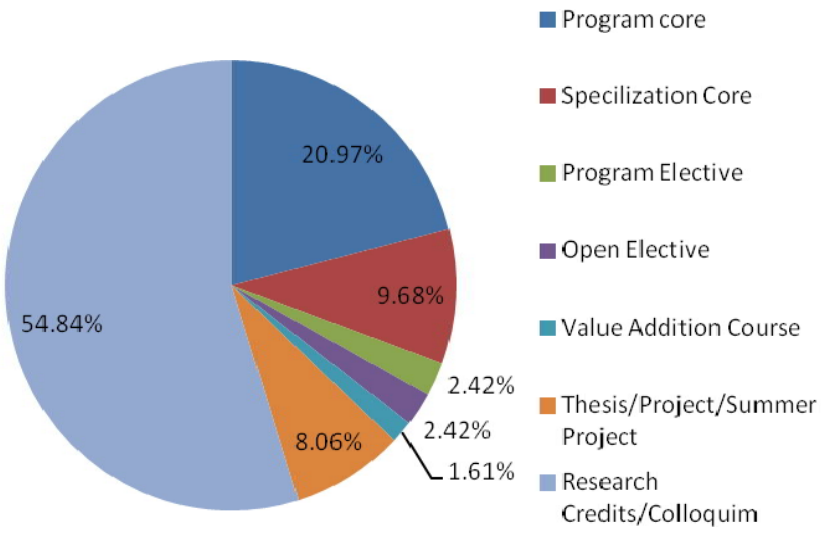


Table 1: First Semester (M. Tech and Dual Degree M. Tech-PhD)

Sl. No.	Course Name	Code	Type	Credit	Hours L-T-P-S
1.	OMICS	BS-AS-OMI501	PCC	4	3–0–2-0
2.	Biological Data Analytics	PC-AS-BDA501	PCC	4	3–0–2-0
3.	Scripting and Computer Environments	ES-AS-SCE501	PCC	4	2–0–4-0
4.	Data Structure and Algorithms	ES-AS-DSA502	PCC	4	2–0–4-0
5.	Research Methodology & IPR (**for M. Tech and M. Tech-PhD**)	PC-AS-RMA504	ELC	2	2–0–0-0
6.	Research Methodology (**for M. Tech-PhD**)	PC-AS-RMA505	ELC	2	2–0–0-0
Total Credit:				18/20	24Hrs./week 26Hrs./week

Table 2: Second Semester (M. Tech and Dual Degree M. Tech-PhD)

Sl. No.	Course Name	Code	Type	Credit	Hours L-T-P-S
1.	Next Generation Sequencing Tools and Algorithms	PC-AS-NGS502	PCC	4	3–0–2-0
2.	Molecular Structure Prediction and isualization	PC-AS-MPV502	PCC	4	3–0–2-0
3.	Biological Information System and Management	ES-AS-BIM503	PCC	4	3–0–2-0
4.	HSMC	HM-AS-RMI501	VEC	2	2–0–0-0
5.	Elective-1	PE-AS-XXX	PEC	3	3–0–0-0
6.	Mini Project	PC-AS-MPJ510	ELC	2	0–0–4-0
Total Credi :				**19**	24 Hrs./week

EXIT: *after the end of second Semester, after clearing all the papers, the M. Tech student may be eligible for PG Diploma in Bioinformatics. However, these students have to secure* ***additional 5 credits*** *from summer semester for awarding PG diploma, which is mandatory.*

Dual degree M. Tech-Ph.D.: *After the end of II Semester the M. Tech-PhD student moves into the Ph. D Section of the Degree and will be henceforth governed by the Ph.D. ordinance. Please note that there is a* ***Summer Semester*** *which will be between II and III Semester. For regular M. Tech students there is no need of summer semester.*

Table 3: Summer Semester

Sl. No.	Course Name	Code	Type	Credit	Hours L-T-P-S
1.	Summer Project (for **PG Diploma**)	PC-AS-SPJ503	ELC	2	0–0–4-0
2.	Systems Biology/ Cheminformatics/NPTEL (for **PG Diploma**)	PC-AS-SCN509	PCC	3	3–0–0-0
3.	Summer Project (for M. Tech-PhD)	PC-AS-SPJ504	ELC	8	0–0–16-0
		Total Credit:		**5/8**	7 Hrs./week 16 Hrs./week

Table 4: Third Semester (M. Tech)

Sl. No.	Course Name	Code	Type	Credit	Hours L-T-P-S
1.	Elective II	PE-AS-XXX	PEC	3	3–0–0-0
2.	Elective III	PE-AS-XXX	PEC	3	3–0–0-0
3.	Open Electives	OE-EC/IT	OEC	3	3–0–0-0
4.	Major Project	PC-AS-MRP501	ELC	8	0–0–16-0
		Total Credi :		**17**	25 Hrs./week

Table 5: Third Semester (M. Tech-PhD)

Sl. No.	Course Name	Code	Type	Credit	Hours L-T-P-S
1.	Research specific breadth course/Self study	PC-AS-RSB501	PCC	4	0–0–0-8
2.	Research specific depth course/Self study	PC-AS-RSD502	PCC	4	0–0–0-8
3.	Open Electives	OE-EC/IT	OEC	3	3–0–0-0
4.	Research Credit	PC-AS-RGC 601	ELC	8	0–0–16-0
		Total Credi :		**19**	35 Hrs./week

Table 6: Fourth Semester (M. Tech)

Sl. No.	Course Name	Code	Type	Credit	Hours L-T-P-S
1.	M. Tech Thesis	PC-MTH502	ELC	12	0–0–24-0
			Total Credit:	**12**	24 Hrs./week

Table 7: Fourth Semester (M. Tech-PhD)

Sl. No.	Course Name	Code	Type	Credit	Hours L-T-P-S
1.	Research Credit	PC-AS-RPC601		12	0–0–24-0
			Total Credi :	**12**	24 Hrs./week

Table 8: List of Electives (for elective I, II and III)

Sl. No.	Course Name	Code	Type	Credit	Hours L-T-P-S
1.	Machine Learning for Biological Systems	PE-AS-MLB501	PEC	3	2-0-1-0
2.	Cheminformatics	PE-AS-CHI502	PEC	3	3-0-0-0
3.	Molecular Medicine	PE-AS-MOM503	PEC	3	3-0-0-0
4.	Cognitive Modeling	PE-AS-CGM504	PEC	3	3-0-0-0
5.	Advance Data Analytics Deep Learning	PE-AS-ADD505	PEC	3	2-0-1-0
6.	Systems Biology	PE-AS-SYS506	PEC	3	3-0-0-0
7.	Parallel Computing	PE-AS-PLP507	PEC	3	3-0-0-0
8.	Pattern Recognition	PE-AS-PRE508	PEC	3	2-0-1-0
9.	Numerical Methods	PE-AS-NUM509	PEC	3	3-0-0-0
1	All the electives provided in M. Tech BME basket are available for M. Tech BI students also				
1	**OPEN ELECTIVES**				
1	Open elective from IT	OE-IT		3	
1	Open elective ECE	OE-ECE		3	

Syllabus for M. Tech BI

OMICS

1. Name of the Course and Code: Omics (BS-AS-OMI501)

2. LTP structure of the course: 3-0-1

3. Objective of the course: The aim is to provide the fundamental knowledge of Molecular biology, Biochemistry, Genomics and Proteomics to the 1st semester M. Tech. BI and 7th semester B. Tech (IT)-M. Tech. Dual Degree BI students.

4. Outcome of the course: Since a mixed population of students from Biology, Computer Science and Information Technology is taking this course, it is very necessary to provide them the basic understanding about the Advanced Biology, especially in the areas of Molecular biology and Biochemistry. The students will be endowed with the knowledge about the different biological processes and the biomolecules involved. In addition, the students will learn the principles of different laboratory techniques from Proteomics and Genomics which will be implemented in the practical classes. They will also learn how to handle different wet laboratory instruments. The overall goal is to inculcate the passion for biological research among the engineering students to improve the interdisciplinary research.

5. Course Plan:

Component	Unit	Topics for Coverage
C1	Unit1	DNA as a genetic material, Nucleic acid (DNA & RNA) structure, Central dogma of molecular biology, Gene structure (Regulator, Operator, Promoter, and Structural genes, Exons, Intron, and ORF) and its expression, Genetic codon, Operon, Restriction enzymes and mapping, Site directed Mutagenesis.
	Unit 2	DNA Replication, Transcription and Translation in prokaryotes, Regulation of gene expression.

C2	Unit 3	Cloning and expression plasmid, Recombinant DNA technology, Construction and screening of Genomic DNA library.
	Unit 4	Principles of Gel electrophoresis, 2D-PAGE, Blotting Techniques (Southern, Northern, Western), Immunoprecipitation, Chip-seq, EMSA, PCR, RT-PCR, ELISA, Mass spectrometry, DNA Microarray, Protein microarray.

6. Text Books and References:

i. Molecular Biology of the Gene by James D. Watson, Tania A. Baker, Stephen P. Bell, Alexander Gann, Michael Levine, and Richard Losick.

ii. Biochemistry by Jeremy M. Berg, John L. Tymoczko, and Lubert Stryer.

iii. Kuby Immunology by Thomas J. Kindt, Richard A. Goldsby, Barbara A. Osborne, Janis Kuby.

iv. Lehninger Principles of Biochemistry by Albert L. Lehninger, David L. Nelson and Michael M. Cox

****

Biological Data Analytics

1. Name of the Course and Code: Biological Data Analytics **(PC-AS-BDA501)**

2. LTP structure of the course: 3-0-1

3. Objective of the course: To expose the M. Tech. BI students with the advanced techniques of Probability and statistics applied to biological data.

4. Outcome of the course: The students will learn modern statistical techniques (sampling, hypothesis tests, correlation and regression analysis etc.) to analyse different kinds of data associated to Health Sciences. As the examples would be taken directly from the health sciences literature instead of contrived examples, we believe that this course should appear more interesting for the students, and would provide very good platform for them to build their skill to become practicing health professionals.

5. Course Plan:

Component	Unit	Topics for Coverage
C 1	Unit 1	Review of the basic concepts of Probability (up to Bayes Theorem) and Statistics (Central tendencies and standard deviations)
	Unit 2	Probability Distribution functions: Binomial, Poisson and Normal distributions, Central Limit Theorem and it's applications.
C2	Unit 3	Sampling distribution, Estimation, Interval estimation, Confidence interval, Test of hypotheses, Z-test, t-test, the chi-square test, F-test and ANOVA test.
	Unit 4	Correlation and Regression analyses, Correlation Coefficients, Least square method and curve fittings, Single and multi variable regression.

6. Text Book:'Biostatistics -A Foundation for Analysis in the Health Sciences' by Wayne E. Daniel and Chad L. Gross.

7. References: 'Fundamental of Biostatistics' by Bernard Rosner.]

****

Scripting and Computer Environments

1. **Name of the Course and Code:** Scripting and Computer Environments (ES-AS-SCE501)

2. **LTP structure** of the course: 2-0-2

3. **Objective of the course**: To expose students to scripting languages of relevance to Bioinformatics.

4. **Outcome of the course:** Bioinformatics is a discipline that requires expertise in using various scripting languages and tools. At the end of course, the students will be able to use scripting languages Perl, Python and R to accomplish tasks required for general purposes in Bioinformatics.

5. Course Plan:

Component	Unit	Topics for Coverage
C1	Unit 1	Introduction to programming, statements, numeric, string literals, variables, arrays and hashes, control statements, subroutines, file handling, regular expression
	Unit 2	References, Advanced Programming in Perl Packages, Object Oriented Programming, BioPerl
C2	Unit 3	Introductory Python and R, Local & Global Alignment Algorithms, Dynamic Programming: Smith & Waterman, Needleman & Wunsch Algorithm.
	Unit 4	Multiple Sequence Alignment, Concepts & Implementations, Amino Acid Substitution Matrices PAM & BLOSUM Derivation of Dayhoff Matrices, Profiles & Motifs General Tools, Techniques & Resources Clustal W, BLAST and FASTA.

6. Text Book: Mandatory for UG core courses

7. References:

- Learning Perl Randal Schwartz, Tom Phoenix, drian d foy (O'Reilly)
- Molecular Modeling: Principles and Applications (2nd Edition) Andrew R. Leach (Prentice Hall)
- Proteins: Structures and Molecular Properties Thomas E. Creighton (Freeman)
- Fast Lane to Python http://heather.cs.ucdavis.edu/~matloff//Python/PLN/FastLanePython.pdf

****

Data Structure and Algorithms

1. **Name of the Course and Code:** Data Structure and Algorithms (ES-AS-DSA502)
2. **LTP structure** of the course: 2-0-2
3. **Objective of the course:** To impart the knowledge of basic programming and elementary data structure to non-IT/CS students
4. **Outcome of the course:**
5. **Course Plan:**

Component	Unit	Topics for Coverage
C1	Unit 1	Computer basics Flowcharts, Algorithms, Data representation Characters, Integers, Fractions, Hexadecimal & Binary Conversions
	Unit 2	C programming, Arrays and Strings, Conditions, Loops, File Handling, Data Types and Pointers
C2	Unit 3	Linked lists, Sorting Algorithms
	Unit 4	Graphs, Binary Search Trees, Shortest Path Algorithms

6. **References:**

- Think Python (2e), Allen Downey, O'Reilly
- Introduction to Algorithms, Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein
- Data Structures Using C and C++, Yedidyah Langsam, Moshe J. Augenstein, and Aaron M. Tenenbaum

****

Biological Information System and Management

1. **Name of the Course and Code:** Biological Information System and Management (ES-ASBIM503)

2. **LTP structure** of the course: 3-0-1

3. **Objective of the course:** To expose the M. Tech. BI students with the techniques of Database Management System along with different available flat files and database handling techniques

4. **Outcome of the course:** The students will learn DBMS role, usage, creation, designing and implementation database to create in-house database. Further they will learn the usage of different flat file formats along with their parsing and information retrieval. This will be sufficed with few case studies using available data mining techniques and data clustering and classification.

5. **Course Plan:**

Component	Unit	Topics for Coverage
C1	Unit 1	Introduction to the database, Database models, Flat model, Hierarchical model, & Network model, Profile & Block, Secondary and Tertiary sequence databases, Relational model, Codd's rule with explanation, 1st, 2nd and 3rd level normalisation, Relational operations, Dimensional Model and Object database models.
	Unit 2	Applications of Databases, Introduction to MySQL, Introduction to Database interfacing language, PHP, Introduction to PHP manipulating APACHE server.
C 2	Unit 3	Database internals, Indexing and triggers, Binary tree and Hash-key based indexing, Transactions, concurrency, and Replication. Data Mining & Warehousing, Association, Clustering & Classification.

	Unit 4	Applications: Introduction to Protein and Nucleic Acid Databases (PDB, NCBI etc.), PDB and NCBI database formalism, Further manipulation of PDB and NCBI data with the help of already acquired RDBMS knowledge based on PHP-MySQL manipulation, Pearl based data mining and linkage with databases (DBI).

6. Text Book:

7. References:

i. Biological Databases by Attwood.

ii. Programming the Perl DBI by O'Reilly

iii. Essential of MATLAB for Scientist and Engineer by Hahn Brian D

iv. Beginning PHP and MySQL 5: from novice to professional by W.Jason Gilmore

****

Next Generation Sequencing Tools and Algorithms

1. Name of the Course and Code: Next Generation Sequencing Tools and Algorithms (PC-AS-NGS502)

2. LTP structure of the course: 3- 0-1

3. Objective of the course: To get acquainted with the high throughput sequencing data and its processing. Since these data often pose a problem of big data domain, the existing algorithm to tackle such problems will be discussed with the limits and lacunas of each such existing technique. This will enable the students to ponder more about the string processing techniques and to come with novel approach of genomic strings processing.

4. Outcome of the course: Trained individuals with basic knowhow of the string processing techniques and a good understanding about the tools for such data analytic.

5. Course Plan:

Component	Unit	Topics for Coverage

C1	Unit 1	DNA sequencing, strings, and matching: DNA sequencers and working principle, DNA as a string. Parsing and manipulating real genome sequences and real DNA sequencing data. Naive exact matching, homology detection; optimal pair-wise sequence alignment, alignment score statistics, efficient database searches (BLAST), Data science of metabolomics, pathway models.
	Unit 2	Preprocessing, indexing and approximate matching: Improving on naive exact matching with Boyer-Moore. Preprocessing and indexing. Indexing through grouping and ordering, k-mers and k-mer indexes. Approximate matching and the pigeonhole principle. Edit distance, assembly, overlaps: Hamming and edit distance. Algorithms for computing edit distance. Dynamic programming. Global and local alignment. De novo assembly. Overlaps and overlap graphs.
C2	Unit 3	Algorithms for assembly: Shortest common superstring and the greedy version. How repetitive DNA makes assembly difficult. De Bruijn graphs and Eulerian walks. How real assemblers work. The future of assembly.
	Unit 4	Data variability and replication, Data transforms, Clustering, Dimension reduction, Pre-processing and normalization, Linear models with categorical covariates, Logistic regression, Null and alternative hypotheses analysis, false discovery rate, permutation and bootstrapping, Gene expression repository (GEO).

7. References:

i. Analytical Techniques In DNA Sequencing by Veena Kumari

ii. DNA Sequencing From Experimental Methods To Bioinformatics by Alphey, Luke Next-Generation Sequencing Data Analysis by Xinkun Wang

****

Molecular Structure Prediction and Visualization

1. **Name of the Course and Code:** Molecular Structure Prediction and Visualization (PC-AS-MPV502)

2. **LTP structure of the course:** 3-0-1

3. **Objective of the course:** To provide M.Tech. IT (spln. In Bioinformatics) students knowledge and exposure in hands-on experience on Molecular Structure Prediction for use in control of disease (patho-physiological condition of health) through application

4. **Outcome of the course:** student will learn state of the art and specialized algorithm to tackle the non-linear and complex problem of prediction of molecular structure that will be useful as part of target specific drug design and to explore structure-influenced patho-physiological conditions.

5. **Course Plan:**

Component	Unit	Topics for Coverage
C1	Unit 1	Basic structural principles: Building blocks of life, Chemical properties of polypeptides & PDB Database, Intermolecular forces: Types of intermolecular forces, Entropy and temperature, Protein folding & Levinthal Paradox.Levels of protein structure: Primary structure, Secondary structure, Tertiary structure & Quaternary structure, Motifs of protein structure: Hydrophobic and hydrophilic regions, Ramachandran plot Alpha-helix, Beta sheets, Loops, Topology diagrams &various structural motifs.
	Unit 2	**Structure determination primer:** X-Ray crystallography and NMR: Structure determination methods & Structure evaluation methods. Structure prediction primers: Protein structure prediction: Impediments, Secondary/fold recognition, Threading/teritiary structures, Sequence considerations, Structural considerations, Energy consideration, Energy landscape

		&Validation.
C2	Unit 3	Structure prediction of small proteins using ab initio stochastic models: Lattice simulation, Randomwalk model, Self-avoiding model & HP-models, Structure prediction of small proteins using ab initio deterministic models Ergodic hypothesis, Use of Newtonian equations of motion , Optimization techniques: Steepest descent, GA, simulated annealing & Force fields (Amber, CHARMM)
	Unit 4	Nucleic acid structures: DNA structures, RNA structures & Secondary structure prediction in RNA, Useful Tools: Visualization using VMD, PROCHECK, WHATIF & Simulation using Amber.

Text/Reference Books:

1. Introduction to Protein Structure: Carl Branden, John Tooze (Garland)

2. Proteins: Structures and Molecular Properties: Thomas E. Creighton (Freeman)

Guide lines for practical's: A two credit lab is to be conducted by covering the most relevant and useful topics from afore mentioned syllabus.

References:

1. Molecular Modelling: Principles and Applications (2nd Edition): Andrew R. Leach (Prentice Hall)

2. Principles of Nucleic Acid Structure Stephen Neidle (Academic Press)

****

****

Elective Course Syllabus: (Electives-I, II and III)

1. Machine Learning for Biological Systems

1. **Name of the Course and Code:** Machine Learning in Biological System (PE-AS-MLB501)

2. LTP structure of the course: 2-0-2

3. Objective of the course: To provide M. Tech. IT (spln. In Bioinformatics) students knowledge and exposure in hands-on experience on solution of "exon-intron sequence identification", "gene identification", "secondary structure prediction" and various solution of biological problems through existing and state of the art machine learning methods.

4. Outcome of the course: Intelligent solutions of non-linear and complex biomolecular sub-systems are needed to get holistic information of pathobiological system (health condition) from the angle of Systems Biology. Students will have exposure of such intelligent solutions through this course.

5. Course Plan:

Component	Unit	Topics for Coverage
C 1	Unit 1	Foundation of Machine learning: Turing Machine, Concepts of John von Neumann, computation of amount of learning of a machine, Concept of supervised and unsupervised learning, concept of clusters and classes, concept of training and testing. Statistical Machine Learning: Design of rule based expert system, knowledge engineering, forward chaining and backward chaining inference techniques, Application to discriminate intron from exon within eukaryotic DNA, Application of rule based system to discover knowledge from data, concept of clustering, condition to find best clusters. Various clustering techniques.
	Unit 2	Important components of a classifier, Probabilistic classifier, Bayesian classifier, Hidden Markov Models and applications, Applications to discriminate Exon from Intron, to predict secondary structures of proteins, discover group of genes having similar up-regulation or down-regulation pattern from micro-array data.
C 2	Unit 3	Nearest Neighbor Classifier, Discriminant Function Analysis (Linear and

		non-linear) as precursor to Artificial Neural Network. Application for protein secondary structure prediction
	Unit 4	Soft computing method based machine learning: Artificial Neural Network for clustering and classification, local optimization of ANN weights, Back propagation network, Hopfield network, Genetic algorithm for optimizing parameters of classifiers, Support Vector Machine foundation, constrained local optimization using Lagrange, Multiplier, application and Cross validation.

Text:

1. Pattern recognition and image analysis by Earl Gose.

2. Pattern Classification by Duda, Richard and David Stork

3. Machine Learning by Mitchell and Tom

Reference Books: 1. Artificial Intelligence and Molecular Biology: Lawrence Hunter (MIT Press) (freely available ebook)

****

2. Cheminformatics

1. **Name of the Course and Code:** Cheminformatics and Molecular modeling (PE-AS-CHI502)

2. **LTP structure of the course: 3**-0-0

3. **Objective of the course: Students should be able to build up QSAR models**

4. **Outcome of the course:** Efficiency in drug design

5. **Course Plan:**

Component	**Unit**	**Topics for Coverage**

C1	Unit 1	Detail Introduction to Cheminformatics in Drug Discovery. 2D Databases and Database searching: Substructure search, Virtual Screening, property searching, similarity searching, Representation and manipulation of 2D Molecular Structures, 3D Databases: experimental data sources, database searching, Representation and manipulation of 3D Molecular Structures, Selecting Diverse Sets of Compounds.
	Unit 2	QSAR: Quantitative Structure and Activity Relationship, Historical Development of QSAR, Hammett Equation, Hansch Equation, Kubinyi bilinear model, Tools and Techniques of QSAR: Biological Parameters, Statistical Methods: Linear Regression Analysis. Parameters used in QSAR: Electronic Parameters, Hydrophobicity Parameters, Steric Parameters, Molecular Structure Descriptors, Quantitative Models: Linear Models, Nonlinear Models, Free- Wilson Approach, Applications of QSAR: Isolated Receptor Interactions, Interactions at the Cellular Level Interactions in-Vivo, Comparative QSAR: Database Development, Software: GRID, CoMFA.
C2	Unit 3	QSPR : Quantitative Structure and Property Relationship, Octonal Water Partition Coefficient, Quantum Chemical Descriptor, HUMO/LUMO, Predictive Quantitative Structure –Activity Relationships Modeling: Data Preparation and General Modeling Workflow, Reaction network Generation, Open Source Chemoinformatics Software and Database Technologies, Machine Learning based Bioinformatics Algorithms-Applications to Chemicals.
	Unit 4	Combinatorial Library Designing: Diverse and Focussed Libraries, Monomer Selection, Product. based Library Design, Structure Based Library design. High Throughput / Virtual screening Screening, Introduction, Basic Steps, Important Drug Databases, Database Designing Lipinski's Rule of Five, ADMET screening

6. Text Book: Mandatory for UG core courses

7. References:

i. Chemoinformatics, Concepts, Methods & Tools for Drug Discovery; Ed. Jurgen Bajorath (Humana Press)

ii. Chemoinformatics Ed by Johann Gasteigen, Thomas Engel, Wiley-VCH

iii. Molecular Modeling, Principles & Applications, Andrew R. Leach

iv. Bioinformatics from Genomes to Drugs ; Vol I & 2

v. An Introduction to Chemoinformatics, Andrew R. Leach, Valerie J. Gillet. R Eisberg and R. Resnick.

****

3. Molecular Medicine

1. Name of the Course and Code: Molecular Medicine (PE-AS-MOM503)

2. LTP structure of the course: 3-0-0

3. Objective of the course: Objective was to make students aware aboutphysical, chemical, biological, bioinformatics and medical techniques used to describe molecular structures and mechanisms to identify fundamental molecular and genetic errors of disease, and to develop molecular interventions to correct them.

4. Outcome of the course: Students were able to emphasize cellular and molecular phenomena and interventions rather than the previous conceptual and observational focus on patients and their organs.

5. Course Plan:

Component	Unit	Topics for Coverage
C1	Unit 1	Concept and perspective of molecular medicine,History of Drug Development, Basic pharmacodynamics and pharmacokinetics, Principles of Chemotherapy, Human genome : implication and applications..

	Unit 2	Bioinformatics, Genomics, Proteomics and Metabolomics in Biomedical research, Molecular basis of diseases(Mixed system) , Molecular basis of inherited diseases, Pathological basis of neurological diseases, Molecular diagnostics of Infectious diseases..
C2	Unit 3	Drug targets at the molecular level, bonding forces, classification of drugs, Receptors-Role, activation, Active site, Strategies of drug design.
	Unit 4	Stem cells and regenerative therapy, Gene therapy for human diseases, personalized medicine, Structure Based Drug Designing (SBDD), Ligand Based Drug Designing (LBDD), Pharmacophore Generation,Docking and scoring methods for proteins-ligands, protein- protein, protein-DNA, DNA-ligand,Identifying Cavities and Surface Matching, Shape Complementarity, Solvent-Accessible Surface.Targeted drug delivery, Nanotechnology in medicine, Biomaterials in tissue engineering, Enzymes in clinical diagnosis.

6. Text book/Reference:

1. Drug Design: Structure and ligand-based approaches: Kenneth M.Merz, Dagmar Ringe, Charles H.Reynolds.

2. Bioinformatics-from genomes to drugs (Vol.2- Applications Lengauer, Thomas (ed.).

3. Burger`s medicinal chemistry & drug discovery; Vol.-2(Drug discovery and drug development) Abraham, Donald J. (ed.)

4. Drug design : structure and ligand-based approaches: edited by Kenneth M. Merz, Dagmar Ringe, Charles H. Reynolds

5. Chemoinformatics; (Vol.-275 - Methods in molecular biology) : concepts, methods and tools for drug discovery : Bajorath, Jurgen (ed.)

****

4. Cognition and Cognitive Processes Modeling

1. **Name of the Course and Code:** Cognition and Cognitive Processes Modeling (PE-AS-CGM504)

2. **LTP structure of the course: 3**-0-0

3. **Objective of the course:**

   a. To provide an overview of cognition in human brain.

   b. To introduce students about several AI debates and pro and against arguments of realization of true AI.

   c. To provide comprehensive details about the cutting-edge approaches and recent developments of cognitive systems.

   d. Introducing students about several cognitive architectures and hand-on working in these architectures.

4. **Outcome of the course:**

   a. Students will get the understanding of how human cognition works as per the explanations till date.

   b. Students will get new side of AI development (Using cognitive architectures).

   c. Students will get to know the challenges which have been accomplished and which are yet to be addressed to make true AI systems.

5. **Course Plan:**

Component	**Unit**	**Topics for Coverage**
C 1	Unit 1:	Introduction: Human Brain: Introduction, cognitive faculties: memory, attention, vision and language, What is cognition, introduction about approaches to cognition, theories of mind: mind - body dualism, materialist theory of mind, identity theory of mind, computational theory of mind.
	Unit 2:	Consciousness: First person approach , third person approach, Chalmers view of consciousness, problem of third person approach, Pattern-Information duality, Free Will: Sloman view, free will as continuous dimension,

		design distinctions for agent modeling.
C2	Unit 3:	First AI Debate: Is AI possible? Pro: Roger Penrose, moravec, Herbert Simon. Artificial mind via symbolic AI, Turing test of AI. Against: Dreyfus five stages of learning, Searle's chinese room thought experiment, Degrees of understanding, godel's incompleteness theorem Second AI Debate: Connectionist Model, Objectives of Connectionist model, Feldman's hundred step rules, Brain vs computer model of mind, Lloyd's cautions, Fodor's attack, Chamlmers' defense, Rule based AI.
	Unit 4:	Cognitive Architectures: ACT-R, CLARION, SOAR, Reinforcement Learning, Distributed Cognition, Learning and Memory Architectures.
	Projects	1. Hands-on on cognitive architectures. 2. Analysis of cognition of brain using complex networks.

Books:

1. Artificial Mind by Stan Franklin
2. Siegelbaum, Steven A., and A. J. Hudspeth. Principles of neural science. Eds. Eric R. Kandel, James H. Schwartz, and Thomas M. Jessell. Vol. 4. New York: McGraw-hill, 2000.
3. Research papers for brain modeling.

****

5. Advance Data Analytics

1. Name of the Course and Code: Advance Data Analytics Deep Learning (PE-AS-ADD505)

2. **LTP structure of the course:** 2-0-2

3. **Objective of the course:** Talks about domain specific mining issues and methods. Large data mining

4. **Outcome of the course:** Students will get exposure of various methods of performing data mining.

5. **Course Plan:**

Component	Unit	Topics for Coverage
C1	Unit 1	Association mining, Classification and Clustering: Revision.Data Streams mining, Social Network Analysis, Graph mining.
	Unit 2	Mining algorithms for large data, Mining Big Data, Hadoop, Map- Reduce, HDFS, Spark + seminars.
	Unit 3	Mining Sequence pattern in TD, Mining, Time-series data Mining WWW + seminars
C2	Unit 4	Advanced Machine Learning: Deep Learning, probabilistic learning + seminars Detail + seminars

6. **Text Book:** Mandatory for UG core courses

7. **References Books:**

i. Ian H. Witten and Eibe Frank, Data Mining: Practical Machine Learning Tools and Techniques (Second Edition), Morgan Kaufmann, 2005, ISBN: 0-12-088407-0.

ii. Hadzic F., Tan H. & Dillon T. S. “Mining data with Complex Structures” Springer, 2011 Yates R. B. and Neto B. R. “Modern Information Retrieval ” Pearson Education, 2005

****

6. Systems Biology

1. **Name of the Course and Code:** Systems Biology (PE-AS-SYS506)
2. **LTP structure of the course**: 3-0-0
3. **Objective of the course**: Expose students to interdisciplinary branch, Systems Biology
4. **Outcome of the course:** Upon completion of the course, the student will be able to appreciate and gain insights into multi-disciplinary approaches made necessary by advancing technology in the field of Systems Biology. Familiarity into various aspects of integrated approaches to understand organisms/organs/tissues/cells at various levels will be obtained.
5. **Course Plan:**

Component	Unit	Topics for Coverage
C1	Unit 1	Biological Systems, Introduction to Mathematical Modeling, Introduction to Graph Theory
	Unit 2	Static Network Models, The Mathematics of Biological Systems, Parameter Estimation
C2	Unit 3	Gene Systems, Protein Systems, Metabolic Systems
	Unit 4	Signaling Systems, Design of Biological Systems

6. Text Book: A first Course in System Biology, *Eberhard Voit*, Mar 2012.
7. References:
    i. Systems Biology: A Text Book, *Edda Klipp et al.* Aug 2009.

****

7. Parallel Computing

1. **Name of the Course and Code:** Parallel Computing (PE-AS-PLP507)

2. LTP structure of the course: 3-0-0

3. Objective of the course: To introduce concepts of Parallel Computing.

4. Outcome of the course: Upon successful completion, the students will be able to approach designing of parallel computation based better. They shall have not only the theoretical concepts but also practical skill to implement the solutions.

5. Course Plan:

Component	Unit	Topics for Coverage
C1	Unit 1	Introduction, Motivation, Scope of parallel computing, Basics of Parallelization, Mutual exclusion
	Unit 2	Concurrent objects, Principles of parallel algorithm design Scheduling and Work Distribution, Foundations of Shared Memory, Primitive Synchronization Operations
C2	Unit 3	Tools and Platforms: C++11 threads, Intel Threading Building Blocks, Open CL and CUDA, Introduction to LAM/MIPCH, Issues of Multicore Programming, Basic Communication Operations, Analytical Modelling of Parallel Programs
	Unit 4	Universality of Consensus, Spin Locks and Contention Monitors and Blocking Synchronization, Parallel Algorithms & Data Structures, Decomposition Techniques, Characteristics of Tasks and Interactions, Mapping Techniques for Load Balancing, Methods for Containing Interaction Overheads Parallel Algorithm Models

6. Text Book: The Art of Multiprocessor Programming by Maurice Herlihy and Nir Shavit, Morgan Kaufmann Publishers

7. References:

i. The Art of Concurrency by Clay Breshears, O Reilly

ii. Introduction to Parallel Computing (2 Ed) by Ananth Grama, Anshul Gupta, George Karypis,

Vipin Kumar, Addison Wesley

iii. Professional C++ by M Gregoire, NA Solter, SJ Kleper (2Ed)

****

8. Pattern Recognition

1. Name of the Course and Code: Pattern Recognition (PE-AS-PRE508)

2. LTP structure of the course: 2-0-2

3. Objective of the course: This course deals with pattern recognition which has several important applications. For example, multimedia document recognition (MDR) and automatic medical diagnosis are two such.

4. Outcome of the course: Students will learn Pattern Recognition techniques and its applications.

5. Course Plan:

Component	Unit	Topics for Coverage
C1	Unit 1	Preliminary concepts and pre-processing phases, coding, normalization, filtering, linear prediction, Feature extraction and representation thresholding, contours, regions, textures, template matching
	Unit 2	Data structure for pattern recognition, statistical pattern recognition, clustering Technique and application. Study of pattern classifiers: Supervised and unsupervised.
C2	Unit 3	Pattern Classifiers: Naïve Bayes, Linear Discriminant Analysis, k-nearest neighbour (K-NN), Artificial Neural Network etc. and Case studies
	Unit 4	Application: Finance, Multimedia.

6. References:

1. R.O.Duda, P.E.Hart and D.G.Stork, Pattern Classification, John Wiley, 2001.

2. K. Fukunaga, Statistical pattern Recognition; Academic Press, 2000.

3. Devi V.S.; Murty, M.N., Pattern Recognition: An Introduction, Universities Press, Hyderabad, 2011

Numerical Methods

1. Name of the Course and Code: Numerical Methods (PE-AS-NUM509)

2. LTP structure of the course: 3-0-0

3. Objective of the course: To expose students to numerical methods of relevance to computational Biology and Bioinformatics

4. Outcome of the course: Students will be able to convert a qualitative problem into computer based solutions

5. Course Plan:

Component	Unit	Topics for Coverage
C1	Unit 1	Introduction, Curve fitting, Taylor series, Roots of equations, Linear algebraic equations
	Unit 2	Optimization and minimization, Numerical differentiation and integration
C2	Unit 3	Ordinary differential equations, Eigen Values and Eigen Vectors
	Unit 4	Molecular dynamics, Monte Carlo algorithm

7. Text book: Numerical Methods for Engineers by Chapra and Can