DEPARTMENT OF MICROBIOLOGY

CHAUDHARY CHARAN SINGH UNIVERSITY, MEERUT

OVERVIEW

1. Minimum Eligibility for admission

A three/ four-year Bachelor's degree or equivalent in Science (Biology group/Bioinformatics/Microbiology/Math Group/ Biotech./ Computer Science/ BCA/B.Sc. Ag/BMLT/Statistics) awarded by a University or Institute established as per law and recognized as equivalent by this University withminimum 45% percentage marks or equivalent grade, shall constitute the minimum requirement for admission to the Master in Bioinformatics.

2. Number of seats and fee structure

Initially there should be only 20 seats which may be altered depending upon the facilities available in the Department/College/Institute. Reservation shall apply as per the policy of the University for other Courses in the University. This course is approved under self finance scheme of the University/State Govt. and annual tuition fee shall be decided by the concerned Department/College/Institute depending upon the resources available.

3. Appointment of Examiners

Course Coordinator is authorized to make a proposal of the examiners (both for theory and practical examination) in consultation with the members of Board of Students either through telephonic conversation or through electronic media. Alternately, a meeting of Board of Studies may be convened

4. Programme Objectives

The M.Sc. **(Bioinformatics)** programme explores the interface between two of the most dynamic subject areas- Bioinformatics and Information technology. The sequencing of the first genome of a free-living organism moved biology irrevocably into the new data-rich era of bioinformatics. The availability of the blueprint of simple organisms and of humans has not of itself answered all our questions. Beyond the detail of connecting gene to function, what is missing is a clear model of how a single gene functions within a complex network to govern expression of the blueprint, an understanding of how the protein folds, a vision of how metabolic pathways and, in their turn, cells are integrated to deliver the features we commonly recognize in living cells. While much remains to be learned, the fundamental shift identified above would not have been possible without the power of modern information technology. The storage, retrieval and analysis of vast quantities of sequence and structural data pose leading edge computational challenges. The course has a strong focus on hands-on experience of in-demand skills of the industry.

Surender Singh

hemant gautam

5. Programme Outcomes

Program Outcomes After successful completion of M.Sc. program, the students would be able

PO1: To get opportunities in higher education. They are also developed on the professional front. It also provides opportunities for career advancement in teaching, research, and industries.

PO2: An ability to acquire in-depth theoretical and practical knowledge of Biology, Computer Science, and Mathematics. Students will acquire knowledge of computer languages- PERL, R programming, Python, MySQL, PHP and JAVA, C++ and Bio-computing programming to write programs to solve biological problems.

PO3: To get introduced to the basic concepts of Bioinformatics and its significance in Biological data analysis. Students will become familiar with a wide variety of bioinformatics tools and software's and apply these to conduct basic bioinformatics research and thus develop platform for molecular biology experiments

PO4: An ability to acquire proficiency in tackling problems related to Biology using software or by development of new algorithms that help in addressing the biological challenges faced by humankind.

PO5: An ability to get an innovative perspective on Biology by providing support in terms of hardware, software and Big Data-handling.

P06: An understanding of the intersection of life and information sciences, the core of shared concepts, language and skills the ability to speak the language of structure-function relationships, information theory, gene expression, and database queries.

6. Program Specific Outcomes

PSO1: After successful completion of this program, the students would be able to apply knowledge of Bioinformatics in all the fields of learning, including higher research and its extensions. Gain multi-disciplinary knowledge and practical skills in computational, mathematical and biological sciences for challenging careers in academics, research, and biotechnology, pharmaceutical and health care industries.

PSO2: Enhance knowledge in the scientific domains of Programming languages, Structural bioinformatics, Genomics, Proteomics, Systems biology, Bioethics, Bio-safety, IPR, Molecular modeling and Drug discovery

PSO3: Apply the appropriate programming and analytical skills in big data analysis and make meaningful predictions.

PSO4: Design new algorithms and in-silico interventions to solve industrial and societal problems. Develop entrepreneurial skills and become professionals in various fields.

PSO5: Students will have analytical and scientific facet in the field of research Students will be able to identify societal problems and recognize the importance of designing scientifically sound and ethical research to solve societal problems.

PSO6: It makes them equipped with knowledge to crack lectureship and fellowship exams like UGC/ CSIR – NET, SET/ ISRO/DRDO and other competitive exams. Students will qualify for higher education and industry needs.

Surender Singh hemant gautam

Syllabus of M.Sc. Bioinformatics
(Under Self Finance Scheme)
Applicable from academic session 2024-25

COURSE STRUCTURE

Following course structure is approved

Year	Sem.	Course Code	Course Title	Core Compulsory / Value added	Internal Marks	External Marks	Credits
Year I	Sem. I	BI -CCSU– 101	Bioinformatics and Biological Database	Core Compulsory	30	70	4
		BI -CCSU – 102	R Programming and Biostatistics	Core Compulsory	30	70	4
		BI -CCSU – 103	Biology - I	Core Compulsory	30	70	4
		BI -CCSU – 104	Computer and operating system **or** Introduction to AI with Machine Learning	Elective	30	70	4
		BI -CCSU - 105	Practical 101-104	Core Compulsory		100	8
		Total Credits					**24**
	Sem. 2	BI -CCSU - 201	Bio-Python Programmimg	Core Compulsory	30	70	4
		BI -CCSU - 202	Sequence Analysis	Core Compulsory	30	70	4
		BI -CCSU - 203	Biology - II	Core Compulsory	30	70	4
		BI -CCSU - 204	Data Mining **or** Block Chain Technology	Elective	30	70	4
		BI -CCSU - 205	Practical 201-204	Core Compulsory		100	8
		Total Credits					**24**
Year II	Semester III	BI -CCSU - 301	Biocomputing Programming	Core Compulsory	30	70	4
		BI -CCSU - 302	Protegenomics, System Biology and NGS Analysis	Core Compulsory	30	70	4
		BI -CCSU - 303	Molecular Modeling & Drug Design	Core Compulsory	30	70	4
		BI -CCSU - 304	Fundamental of Data science **or** Multimedia Applications with Web Technology	Elective	30	70	4
		BI -CCSU - 305	Practical 301-304	Core Compulsory		100	8
		Total Credits					**24**
	Semester IV	BI -CCSU – 401	Industrial Training/ Research Project	Core Compulsory	500		**24**
		Credits					**96**

Surender Singh hemant gautam

YEAR FIRST

SEMESTER I

BI -CCSU- 101: Introduction to Bioinformatics and Biological Database

Course Outcomes:

- Students will become familiar with a variety of currently available primary, secondary and composite databases.
- Students will be able to search and retrieve information from genomic and proteomic databases (e.g. GenBank, Swiss-Prot).
- Use computers to record, visualize and store biological database
- Describe the method of heuristic approach and study of protein visualization tool.

Unit I: Introduction to Bioinformatics: Definition and History of Bioinformatics, Problems and Applications of Bioinformatics, Bioinformatics resourses: NCBI, EBI, SIB, Open access bibliographic resources and literature databases: PubMed, BioMed Central, Public Library of Sciences (PLoS), CiteXplore.

Unit II: Biological Database- Primary Database & Secondary Database., Primary Nucleotide sequence database-GenBank, EMBL, DDBJ.,Primary Protein sequence database-PIR,SWISS PROT.,Primary Molecular structure database-PDB,CSD PubChem, ChemBank., Secondary database of nucleotide sequence-TrEMBL, NDB, Secondary database of protein sequence-PROSITE, Pfam, PRINTS, UniProt., Secondary structure database-CATH, SCOP. DSSP, FSSP, RNAbase. NLM database. Indian medicine database-DataRequisite

Unit III:-Composite and other databases: Compositedatabase-OWL, SPTR (SWISS PROT+TrEMBL), NCBI. Genome database-OMIM.,Metabolic pathway database-KEGG., Specialized database-PubMed.Model organism databse- Thale cress (Arabidopsis thaliana). , Domain Databases- SMART. Antibody database, Enzyme database

Unit IV: Introducion of file Format: Various file formats for bio-molecular sequences: GenBank, FASTA, GCG, MSF, NBRF-PIR format and utilities etc. Protein and nucleic acid tools: Proteomics tools at the ExPASy server, and EMBOSS, Conversion between formats, Introduction about Annotation and Archival.

Unit V: Heuristic Method: BLAT, FASTA, BLAST (BLASTP, Quick BLASTP, BLASTN, BLASTX, TBLASTN, TBLASTX, DELTA-BLAST). Statistical analysis of BLAST results, Database Searching techniques - ENTREZ, Introduction to RASMOL, SWISS-PDB Viewer.

Suggested Readings (Latest Editions):

1. Bioinformatics: Sequence, Structure and Databanks: A Practical Approach (The Practical Approach Series, 436), Des Higgins (Editor), Willie Taylor.
2. Bioinformatics: Sequence and Genome Analysis, David W. Mount.
3. Introduction to Bioinformatics, Teresa Attwood, David Parry-Smith
4. Bioinformatics: Databases and Systems, Stanley Letovsky, Bioinformatics for Dummies
5. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, Second Edition, Andreas D. Baxevanis, B. F. Francis Ouellette.
6. Introduction to Bioinformatics, Arthur M. Lesk, Oxford University Press

Surender Singh hemant gautam

BI -CCSU - 102: R Programming and Biostatistics

Course Outcomes:

- Explain R programming concepts for statistical analysis
- Study, how to install and configure R-Studio
- Understand OOP concepts in R programming
- Explain the use of data structure and loop functions
- Analyse data and generate reports based on the data
- Apply various concepts to write programs in R
- Understand the concept of a random, representative sample from a databases
- Explain how statistical techniques we have studied are incorporated in the analysis of medical research data and its presentation.

Unit I: Introduction to R programming: R command Prompt, R script file, comments, Installing a R Package, R DataTypes: Vectors, Factors, Data Frame, R Variables: Variable assignment, Data types of Variable, Finding Variable, Deleting Variables, R Operators: Arithmetic Operators, Relational Operators, Logical Operator, Assignment Operators, Miscellaneous Operators, R Decision Making: if statement, if else statement, switch statement, R Loops: repeat loop, while loop, for loop, break statement, next statement.

Unit II: R-Function : function definition, Built in functions, user defined function, R Strings, R Vectors : Sequence vector, vector access, vector names, vector math, vector recycling, vector element sorting, R List: Creating a List, List manipulation Tags and Values, R Arrays, R Matrices: Accessing Elements of a Matrix, Addition, subtraction, Multiplication, R Factors: creating factors, generating factor.

Unit III: Introduction to Biostatistics: Definition, Types of statistics, Applications and uses of Biostatistics, Identification and types of variable, Tabulation of data, Graphical presentation (categorical and metric data), charting of data using MS-Excel; sampling and techniques, Data Collection and presentation: Types of data, methods of collection and graphic representation of data, Frequency distribution; Measures of central tendency (mean, median and mode); Measures of dispersion: mean deviation and standard deviation; Correlation and regression, Curve fitting..

Unit IV: Basic concepts related to probability theory, classical probability. Probability distributions: Introduction and simple properties of Binomial, Poisson, Normal and skewed distribution and their applications in biology. Sampling: Concept of sampling and sampling techniques. Test of significance: Chi-square test, t-test, Z-test, f-test and standard error.

Unit V: Formulation of Hypothesis (One-tailed & Two-tailed), Type I and Type II errors, Analysis of Variance, One-way ANOVA, Two-way ANOVA and Three-way ANOVAwith their layout and statistical analysis. Applications of statistical methods using statistical software, SAS.

Suggested Readings (Latest Editions):

1. A Handbook of Statistical analysis using R, Brain Everitt and Torsten Hothorn
2. The art of R programming, Norman Matloff
3. Data Analysis and Graphics using R, W. John Braun
4. R for Data Science, Garrett Grolemund and Hadley Wickham
5. Linear Models with R, Julian J. Faraway
6. Biostatistics: A foundation for analysis in the Health Sciences, W.W Daniel. Publisher: John Wiley and Sons.
7. Biostatistics, P.N Arora and P.K Malhan. Publisher: Himalaya Publishing House.
8. Introduction to Biostatistics, Ronald N. Forthfer and Eun Sun Lee .Publisher: Elsevier.
9. Biostatistics: A foundation for analysis in the Health Sciences, W.W Daniel. Publisher: John Wiley and Sons.
10. Statistical Methodology, S.P Gupta. Publisher: S.Chand& Co.

hemant gautam

Surender Singh

Bl -CCSU -103 Biology - I

Course Outcomes:

- Students will learn basic introduction and division of microbes through classification of five and eight kingdom.
- The students will learn about cells and their structure and properties of cells
- The students will learn virus and their structure and organization of virus.
- The students will learn immune system and its type and differentiate between B and T cells immunity.
- The students will learn antigen-antibody reaction also learn structure of immunoglobulins.

Unit 1: Discovery of microbial world, controversy over spontaneous generation; origin of life. Microbialevolution and diversity; five kingdom and eight-kingdom classification.

Unit 2: General account of prokaryotes: structural organization of Eubacteria and Archae (cell membrane, cytoplasmic matrix, inclusion bodies, nucleoid, flagella, pilli and endospore); General characters of eukaryotic microbes, structure and organization of a typical eukaryotic cell.

Unit 3: History and discovery of viruses; nature of viruses; General characters of viruses; Nomenclature and Classification of viruses; Bacteriophage: Structure and life cycle pattern of T-even phage; Genome organization of viruses

Unit 4: Introduction to the immune system: Innate immunity; anatomic, physiological, phagocytic & inflammatory barriers. Adaptive immunity; natural & artificial immunity. Cells involved in immune response: lymphoid lineage (producing B & T lymphocytes) & Myeloid lineage (phagocytes: macrophages, neutrophils & eosinophils. And auxillary cells; basophils, mast cells & platelets). Organs involved in immune system: primary & secondary lymphoid organs.

.Unit 5: Types of antigens; Structure & types of Immunoglobulins, Cytokines, major histocompatibility complex molecules; vaccines & their types, Hybridoma technology, applications of monoclonal antibodies. Antigen-antibody reactions *in vitro*.

Suggested Readings (Latest Edition): -

1. Prescott - Harley - Klein -2007- Microbiology - IV Edition - International edition - McGraw Hill - ISBNO-07-115830-8.
2. D. J. Taylor - N. P. O. Green - G. W. Stout – Biological Sciences - III Edition - Ed. - R. Soper - Cambridge University Press - ISBNO - 521 - 639239 (Low Price Paperback)
3. K. P. Talaro & A. Talaro,. Foundations in Microbiology, HI - International Edition. WCB / McGraw Hill –ISBNO, 697, 35452 . 0
4. Guntram Seltmann, Otto Holst Bacterial cell walls. Springer ISBN:3540426086
5. Nelson D. L. and Cox M. M. (2005) Lehninger's Principles of Biochemistry, Fourth edition, W. H. Freeman & Co. New York.
6. White David Physiology and Biochemistry of Prokaryotes. 2nd Ed. Oxford University Press, New York.
7. Berg Jeremy, Tymoczko John, Stryer Lubert Biochemistry 4th Ed, W. H. Freeman, New York.

hemant gautam Surender Singh

BI -CCSU -104 (Elective I): Introduction to Computer and Operating System

Course Outcomes:

- Recognizing and analyzing of image acquisition storage, processing, communication & display.
- Able to understand the formation of image model & basics enhancements techniques.
- Learn the image segmentation processing in detail.
- Able to understand the basic applications of image processing in medical systems..

Unit 1: Definition and characteristics, Components (Input/Output unit, Control Unit, Primary Storage Unit, Arithmetic and Logic Unit), Types of Memory, Communication Pathways (Control Bus, Address Bus, Data Bus), Classification of Computers (according to logic & size),Software.

Unit 2: Networking- Definition, Client/Server, Types (LAN, WAN, MAN); Network Connecting Devices; Topologies (Bus Network, Ring Network, Star Network, Mesh Network, Tree Network) and their advantages & disadvantages. Elements of Networking (Network Services such as File Services, Database Services, Print Services, Application Services); Transmission Medias (Coaxial Cable, Fiber Optics, Twisted Pair), Internet and Internet tools.

Unit 3: Introduction: Definition and Concepts of operating system; Function of Operating System; Batch Processing; Multiprogrammed Batch System; Time Sharing System; Parallel System; Distributed System; Real Time System.

Unit 4:. Process & Memory Management: Process; Process State(New, Running, Waiting, Ready, Termination); Process Control Block; Process Scheduling (Round Robin Scheduling, Priority Scheduling, Multiple Queues, Shortest Job Scheduling); Operations on Process; Basic Management of Memory; Swapping Virtual Memory; Paging, Dead Lock.

Unit 5: UNIX/LINUX Operating Systems: Introduction; Concepts; Layers of UNIX; Role of System Administrator and Ordinary User; Tree Structure of UNIX; Root File System; /bin Directory; /dev Directory; /bin Directory; /etc Directory; /lib Directory; /proc Directory; /mnt Directory; /root Directory; /sbin Directory; /tmp Directory; /var Directory; Relative Path; Absolute Path; Creation of Directory; Creating file; removing file; Listing Files and Directories copying file; renaming file; Changing File Permission; Changing Directory Permission; Changing Group; Changing Owner; Pipe; Filters; pwd command; date command; head command; tail command less command; more command; grep command; VI Editor (Creating a new File; Inserting Text in File; Deleting Text in File; Copy , Cut & Paste Text; Save File).

Suggested Readings (Latest Editions):

- Ramesh Bangia. 2015. Learning Unix. BPB Publication.
- Peter Baer Galvin. 2016. Operating System Concepts. BPB Publication
- Kenneth H. Roshan. 2007. The complete reference Unix Tata Mac Graw Hill..
- D. M. Dhamethire. , 2011. System Programming and Operating Systems. Tata Mac Graw Hill.
- Kirrgcox. 2001. Red Hat Linux by. Printice Hill India.
- Sumetabha Das. 2013. Unix (Concept and Application). Tata Mac Graw Hill.
- Raja Raman V. 1999Computer Programming In C, Prentice Hall of India Pvt. Ltd. New Delhi .

hemant gautam Surender Singh

BI -CCSU -104 (Elective II): INTRODUCTION TO AI WITH MACHINE LEARNING

Course Outcomes:

- Understand the features of machine learning to apply on real world problems
- Understand the informed and uninformed problem types and apply search strategies to solve them.
- Understand the informed and uninformed problem types and apply search strategies to solve them.
- Examine the issues involved in knowledge bases, reasoning systems and planning

UNIT 1: An Introduction to AI & ML:
Overview of artificial intelligence, History and foundations of AI, Introduction to machine learning, Types of ML, Applications used for ML and AI, Essential for ML and AI.

UNIT 2: Techniques and Supervised of Machine Learning:
Introduction to Supervised, unsupervised, semi-supervised, and reinforced machine learning techniques, Training, validation, test data, and Over fitting and complexity

UNIT 3: Managing data and Data Processing:
Data description, Data processing, Dimension Reduction, Comprehend the meaning, process, and importance of data preparation, feature engineering and scaling of datasets.

UNIT 4: Evaluating model performance:
Measures performance for the classification problem. Discuss different performance measures.

UNIT 5: Resampling methods:
Discuss different cross validation techniques. Leave-One-Out Cross-Validation, k-Fold Cross Validation
Examination Scheme:

Suggested Readings (Latest Editions):

- Programming Collective Intelligence by Toby Segaran
- Machine Learning for Hackers by Drew Conway and John Myles
- Machine Learning by Tom M. Mitchell
- Pattern Recognition and Machine Learning by Christopher M. Bishop (Author)
- Machine Learning Yearning by Andrew NG
- The Elements of Statistical Learning by Trevor Hastie , Robert Tibshirani , Jerome Friedman

BI -CCSU -105 -Practical in First Semester is based on, BI-CCSU -101 to BI-CCSU -104

hemant gautam Surender Singh

YEAR FIRST

SEMESTER II

BI -CCSU -201: Bio-Python programmimg

Course Outcomes:

- Understand basic principles of computers
- Understand basics of binary computation
- Understand the programming basics (operations, control structures, data types, etc.)
- Readily use the Python programming language
- Apply various data types and control structure
- Understand class inheritance and polymorphism
- Understand the object-oriented program design and development
- Understand and begin to implement code

Unit 1: Overview of Python, applications, usage and comparative study with other softwares, Syntax, Data Types, Variables, Operators, Input/output, Flow of Control, If, If else, Nested if-else, Looping, For, While, Nested loops, Break, Continue.

Unit 2: Introduction and Accessing list, working with Lists, Declare, assign and retrieve values from Lists Operations, Function and Methods.

Unit 3: Strings and Tuples, Accessing Strings, Basic Operations, String Manipulation Functions, Introducing Tuples, Accessing tuples, Files, Modules, Dictionaries,

Unit 4: Advance Python: Object Oriented, OOPs concept, Class and object, Attributes, Inheritance, Overloading, Overriding, Data hiding, Operations Exception.

Unit 5: Introduction to Bio.Seq, Using Seq class, Sequences reading and writing, Bio classes for sequences, Bio.SwissProt. SProt and Bio.WWW.ExPASy, Reading entries, Regular expressions in Python, Prosite - Bio.GenBank, Reading entries and Running Blast and Clustalw.

Suggested reading (Latest Edition):

1. Think Python by Allen B. Downey
2. Introducing Python by Bill Lubanovic
3. Hello World by Warner Sande and Carter Sande
4. Learning Python , 5th Edition , Mark Lutz
5. Python For Data Analysis by W Mckinney
6. Mitchell L. Model, "Bioinformatics Programming using Python: Practical
7. Programming for Biological Data" , O'Reilly Media",.
8. Mark Lutz and David Ascher, "Learning Python", O'Reilly Media, 5th Edition,
9. Jeff Chang et al, Biopython – "Biopython Tutorial and Cookbook", Link: http://biopython.org/DIST/docs/tutorial/Tutorial.html.
10. Python Essential Reference 4th edition by David M. Beazley

hemant gautam Surender Singh

BI -CCSU -202: Sequence Analysis

Course Outcomes:

- After completion of the Course, the student will learn how to use sequence analysis in different way.
- The students will learn the different scoring matrix and its use in sequence alignment.
- Students also aware of different between pairwise and multiple sequence alignment.
- The students will also learn about phylogenetic tree and its different terminology.
- Students also making different tree by learn different methods of tree and how to view of that's tree.
- Students would be able in differentiate between motif and domain and also learn EST process.

Unit 1: Concept of Scoring matrix (PAM and BLOSUM), local alignment, global alignment, Dynamic programming-Dot matrix method, Needleman-Wunsch algorithm, Smith-Waterman algorithm. Weight matrix as Identify scoring, genetic code scoring scheme chemical scoring, observed substation matrix and Gap penalties.

Unit 2: What is Sequence Analysis, Sequence patterns & representation- consensus sequence, regular expression, profiles, PSSMs, Pattern, contigs. PSI- & PHI BLAST. , Identity and homology, definitions of homologous, orthologous, paralogues, Analogous ,Xenologous. Importance of sequence alignment. Multiple sequence alignment method- Progressive alignment method, Iterative method, , hierarchical method, Consensus methods and Phylogeny-aware methods, Clustal W, T-COFFEE

Unit 3: Phylogenetic prediction: Relationship of phylogenetic analysis to sequence alignment,Genome complexity and phylogenetic analysis, Introduction to Phylogenetic: Phylogenetic Basics, Tree Styles, Terminologies, types of phylogenetic tree, Binary trees, Tree traversal (Pre-order, In-order, post-order). Gene tree & Species tree, Newick format of tree representation, True tree & Inferred tree.

Unit 4: Phylogenetic Tree construction Methods: UPGMA, Neighbor-Joining, Maximum Parsimony, Maximum likelihood. Phylogenetic. Tree Evaluation method: Bootstrapping, Randomized jack-knifing method. Molecular Clock, Rate heterogeneity, LBA problem, Phylogenomics, Supertree & Super matrix method. Phylogenetics software and tools like MEGA, PHYLIP, PhyML and PHYLODENEDRON.

Unit 5: Functional genomics Strategies for generating EST and full length insert, EST clustering and assembly, statistical analysis of EST and EST data, TIGER indices, Motif and Domain, sequence assembly, global assembly.

Suggested Readings (Latest Editions):

1..Mount, D. "Bioinformatics: Sequence and Genome Analysis"; Cold Spring Harbor Laboratory Press, New York.
2.Inna Dubchak et al. Active Conversation of Non-coding Sequences revealed by three way species comparisons. Genome Research. 10, 1305–1306.
3.Heijne, Gunnar Von. Sequence Analysis in Molecular biology: treasure troureor trivial Pursuit
4 Koski, T. Dordrecht Kluwer. 2013. Hidden Markou Modles for Bioinformatics. Academic Publishers.
5.Darbin, R. Eddy, S. Krogh, A. & Mithchison G.. 2010. Biological Sequence Analysis: Probabilistic models of Proteins and Nuclic acid Cambridge University Press, Cambridge
6. Sharma, Munjal and Shankar. A Text book of bioinformatics.2009. Rastogi Publications
7Advances in Bioinformatics by M. S. Krishna Kumar.
8.Rastogi, S.C., Mendiratta, N. and Rastogi, P. 2007. Bioinformatics: concept skill and applications

hemant gautam Surender Singh

BI -CCSU -203: Biology - II

Course Outcomes:

- Students will demonstrate a core knowledge base in the theory and practice of Molecular Biology.
- Students will function successfully in the laboratory and use safe laboratory practices.
- To describe the structure of nucleic acids & proteins and their interactions.
- To study the molecular mechanisms of gene regulation in prokaryotes and eukaryotes.
- To familiarize the student with emerging field of biotechnology i.e. Recombinant DNA
- Technology as well as to create understanding and expertise in wet lab techniques in genetic engineering.
- To study the different molecular techniques and their applications.

Unit 1: Nucleic acids as genetic information carriers, DNA structure, types of DNA. DNA replication in prokaryotes & eukaryotes. Structural features of RNA (mRNA, tRNA, rRNA). Transcription in prokaryotes & eukaryotes.

Unit 2: Regulation of gene expression. Basic features of the genetic code. Protein synthesis in prokaryotes and eukaryotes. Recombination: general principles.Gene transfer mechanisms: transformation, transduction, and conjugation.

Unit 3: Mutations: spontaneous mutation, Induced mutagenesis- mutagens (physical mutagens: non ionizing & ionizing radiations; chemical mutagens: Base analogues, alkylating agents, deaminating agents, intercalating agents & others), DNA repair mechanism: repair by direct reversal, excision repair, recombinational repair & SOS repair.

Unit 4: Basic steps of r-DNA technology. Restriction endonucleases. Cloning vectors: general properties, plasmids, bacteriophages, cosmids, shuttle vectors, bacterial artificial chromosomes. Gene libraries: genomic library, c DNA library

Unit 5: Molecular Techniques; Principles, methods & their applications in medical diagnosis -such as PCR, Southern Blotting, Northern Blotting, Western Blotting, DNA finger printing.

Suggested Readings (Latest Editions):

1. Nelson D and Cox MM. Lehninger's Principles of Biochemistry. W.H. Freeman an Company, New York.
2. Voet D and Voet JG.Principle's of Biochemistry. John Wiley and sons New York.
3. J.L. Jain.Fundamentals of Biochemistry, S. Chand and Co.
4. Cooper, G.M. and Hausman, R.E. The Cell- A molecular approach (4th eds.). A S M Press, Sinauer
5. Associate Inc.
6. Karp, G. Cell and Molecular Biology, Concepts and Experiments, John Wiley and Sons.
7. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. 2008. Molecular Biology of Cell (5th eds.). Garland Sciences. Robert J. Brooker . Genetics, Analysis and principles, Mc Graw Hill.
8. Brown, T.A. 2008. Gene Cloning and DNA analysis (5th eds.). Blackwell Sciences LTD.
9. Gupta, P.K. 2008. Biotechnology and Genomics (1st ed.). Rastogi Publication.

hemant gautam Surender Singh

BI -CCSU -204 (Elective I) : Data Mining

Course Outcomes:

- To introduce students to the basic concepts and techniques of Data Mining.
- To develop skills of using recent data mining software for solving practical problems.
- To gain experience of doing independent study and research.
- Use the basic concepts of Database Systems in Database design
- Design a Database using ER Modelling
- Apply normalization on database design and its relationship

Unit 1: Definition; Purpose of Database System; Advantages of Database System; Components of Database System (., Hardware, Software & Users), Database Administrator; Data Administrator; Data Models (Relational, Network, Hierarchical); Three Level Architecture for Database System (Internal Level; Conceptual Level; External Level); Data Independence; Data Abstraction; Mapping; Data Definition Language; Data Manipulation Language; Data Sub Language; Role of Schemas in Three Level Architecture; Client/Server architecture; Distributed Processing; Database Technologies (Flat Files, Relational & Object).

Unit 2: Introduction of data mining, Data Mining Goals, Data Mining Techniques, Data warehouse, Need for data warehouse, definition, goals of data warehouse, Data Mart, Extract and load process, clean and transform data, star, snowflake and galaxy schemas for multidimensional databases, fact and dimension data, Designing fact tables.

Unit 3: Data preprocessing, data integration and transformation, data reduction, Discretization and concept Hierarchy Generation, Data mining primitives, Types of Data Mining, Data Mining query language.

Unit 4: Architectures of data mining. data generation & Summarization based characterization, Analytical characterization, Mining class comparisons.

Unit 5: Definition; Relational Data Model (Binary, Ternary, Quaternary & n-ary Relation); Important terms in Relational database system (Tuple, Records, Fields, Domain, Degrees, Cardinality); Keys(Primary Key, Candidate Key, Composite Key, Foreign Key& Alternate Key), Creating Table; Applying Column & Table Constraints; Inserting Values in Table; Deletion(of Rows & Table); Updating Values; Altering Table; Retrieving Values from Table; Revoke Command; Drop Command; Grant Command; Commit Command; Rollback Command

Suggested Readings (Latest Editions):

1. Building Data Ware House by W.H.Inmon, John Wiley & Sons
2. Data warehousing by S . Anahory and D.Murray, Pearson Education, ASIA
3. Data Mining Concepts & Techniques by Jiawei Han & Micheline Kamber; Harco
4. SumeetDua, PradeepChowriappa, Data Mining for Bioinformatics, CRC press
5. Pang-NingTan , Michael Steinbach, Vipin Kumar, Introduction to Data Mining, Addison-Wesley
6. Jiawei Han, Jian Pei,MichelineKamber, Data Mining: Concepts and Techniques, Morgan Kaufmann; 3 edition
7. Ivan Bayross. 2010. SQL, PL/SQL the programming language of Oracle., BPB Publication
8. C.J. Date. 2010. An indroduction to database systems. Addison Wesely
9. Orpipa Bosu and Simminder Kaur, Thukaral. 2009. Bioinformatics: Database, tools and algorithms. Oxford Publications

hemant gautam

BI -CCSU -204 (Elective II) : Block chain Technology

Course Outcomes:

- To introduce students to the benefits, and the limitations of blockchain technologies.
- Be able to state the key differentiators for blockchain from other technology systems.
- Apply various blockchain concepts to analyze examples, proposals, case studies, and preliminary blockchain system design discussions.
- Determine real world challenges that blockchain technologies may assist in solving.

Unit 1: Overview of Blockchain Technology- Blockchain concepts, decentralized systems, consensus, Blockchain types: public, private, and consortium, Cryptography fundamentals for blockchain.

Unit 2: Development Platforms and Tools- Blockchain development frameworks (e.g., Ethereum, Hyperledger), Setting up the blockchain development environment, Smart contracts.

Unit 3: Security and Privacy- Blockchain security challenges and attacks, Cryptographic techniques for securing blockchain transactions, Privacy and anonymity considerations in blockchain systems.

Unit 4 Decentralized Applications (DApps)- Smart contract development, Interacting with smart contracts using web interfaces and APIs, Building and deploying decentralized applications (DApps).

Unit 5 Blockchain Applications uses in Industry- Blockchain applications in finance, supply chain, healthcare, and other domains, Regulatory and legal considerations for blockchain adoption, Impact of blockchain on various industries.

Suggested Readings (Latest Editions):

1. "Mastering Blockchain: Unlocking the Power of Cryptocurrencies, Smart Contracts, and Decentralized Applications" by Imran Bashir.
2. Andreas Antonopoulos, "Mastering Bitcoin: Unlocking Digital Cryptocurrencies", O'Reilly
3. "Blockchain Basics: A Non-Technical Introduction in 25 Steps" by Daniel Drescher.
4. Melanie Swa, "Blockchain", O'Reilly
5. "Hyperledger Fabric", https://www.hyperledger.org/projects/fabric
6. Bob Dill, David Smits, "Zero to Blockchain - An IBM Redbooks course", https://www.redbooks.ibm.com/Redbooks.nsf/RedbookAbstracts/crse0401.html

BI -CCSU -205 -Practical in Second Semester is based on **BI-CCSU** -201 to **BI-CCSU** -204

hemant gautam Surender Singh

YEAR SECOND

SEMESTER III

BI -CCSU-301: Biocomputing Programming

Course Outcomes:

- Concepts of computer programming languages like C, JAVA helps in solving different complex problem in biology or data
- Analysis.
- Knowledge of Traditional HTML interfaces for input to and output from Bioinformatics analysis on the Web are highly variable in style, content and data formats.
- Knowledge of PHP There is an extremely rich variety of programming languages, web development frameworks, tools and libraries for "general" web development that are perfectly fit to handle Bioinformatics data.
- Arithmetic and logical operators; Conditional statement and Loops; Regular Expressions; Function and subroutines Application of PERL in Bioinformatics; concatenating DNA fragments; DNA to RNA; Reading protein Files; Finding motifs; ORFs; DNA to protein.
- Introduction to PERL as scripting language; variables; Array; Initialization and manipulation

Unit 1: Hyper Text Programming Language: Structure of HTML program(<HTML>, <HEAD>, <TITLE>, <BODY>); Titles & Footer; Text Formatting(Paragraph break<P>, line break
); Text Styles(Bold<B>, Italics<I>, Underline<U>); Font Style, Color & Size; Image Tag(<IMG>); Table (<TABLE>, <TR>, <TH>, <TD>) & Attributes(Border, Width, Align); Frames; Forms (Text Box, Check Box, Command Button, List Box); Anchors; CSS (Cascade Style Sheet).

Unit 2: Java Script- Data Types; Literals; Variables; Arrays; Operators (Arithmetic, Comparison, Logical, String, Assignment); Condition Check(if-then-else); looping(for, while); Functions(Built-in, user defined); scope of functions; Dialog Boxes(Alert Dialog Box, Prompt Dialog Box; Confirmed Dialog Box).

Unit 3: PERL- Scalar Data; Scalar Variable; List Data & Variable; Operators, <STDIN>; print & printf; Indexed Arrays; Associative Array; Hash elements; Conditionals (if & elseif); Loops; Manipulating Lists and Strings.

Unit 4: Basic PHP - PHP Syntax, Variables, Strings, Constants, Operators, Echo / Print, Conditionals (if, if…else, if…elseif); for loops, Foreach, while loop.

Unit 5: PHP with MySQL - Functions, string functions, user defined functions. MySQL data types, DML and DDL, Aggregate functions, Sub query and join.

Suggested Readings (Latest Editions):

- *Dick Oliver. HTML 4 in 24 hrs. Techmedia Publication*
- *Dick Oliver. JavaScript 4 in 24 hrs. Techmedia Publication*
- *Ivan Bayross. Web Enabled commercial application development using HTML, DHTML, JavaScript and PERL CGI. BPB Publication*
- *James Tisdall. Beginning perl for Bioinformatics. O-Reilly*
- *Cynthia Gibos, Per Jambeck. Developing Bioinformatics Computer Skills. O-Reilly*
- *Randal L. Schwartz, Tom Phoenix. Learning Perl. O-Reilly*
- *Techmedia Aldo. Perl programming, Wiley Publication.*
- *Eric C. Hermamm. Mastering Perl 5, BPB Publication.*
- *PHP for Absolute Beginner's, Apress Publication.*

hemant gautam

Surender Singh

BI -CCSU -302: Proteogenomics, System Biology and NGS Analysis

Course Outcomes:

- Identify the different NGS technologies in the market
- Describe the different workflows in the analysis of NGS data
- Interpret the results from the different NGS applications
- Understand the proteogenomics and system biology apllication
- Knowledge about the genome editing tools and process

Unit 1: Introduction and history of proteogenomics, Application, uscand goal of proteogenomics. Database of protegenomics- integrated proteogenomics database (iPtgxDB), VarDB2.0, dbPepVar. Proteogenomics engine-Spritz. Tool sand software use in proteogenomics- proteogenomics browser, Peppy, PRODIGY (PROtein binDIng enerGY prediction) and PgxSAVy.

Unit 2: Introduction to Genomics, Epigenome, cBioPortal, Genotype, Gene expression and Phenotype. C-value paradox, structural and comparative genomics, transcriptomics, proteomics and metabolomics. SNPs BLAST2, VISTA.Comparative Genomics Databases- COG, VirGen, Gramene. Introduction to metabolomics: metabolome, metabonomics, metabolite profiling, metabolome fingerprinting, tools of metabolome studies: NMR, MS, GC, LC, IR and its application,

Unit 3: Aims, strategies and challenges in proteomics, Quantitative proteomics, Principles of protein structure; Hierarchical organization of protein structure – Primary.Secondary, Super secondary, Tertiary and Quaternary structure; Structural elements and terminology: Helix, Sheet, Strand, Loop and coil, Active site, Architecture, Blocks, Class and Domains, Fold, Motif, leucine zipper, coiled coil, trans-membrane, signal peptide, cleavage site. Protein Structure Visualization Tools. BIND, MINT, STRING database.

Unit 4: System biology, System-level-Understanding of Biological Systems, System Behavior Analysis: Simulation, Analysis Methods, and Robustness of Biological Systems. System Control; Redundancy, Modular Design, Control, Structural Stability, Impacts of systems Biology.

Unit 5: Introduction to NGS: Overview of generation to DNA sequencing technologies, typical, Different NGS Platforms – Illumina, Ion Torrent, Nanopore Sequencing, Drawbacks of NGS, NGS File formats (FASTQ File Format), & applications. NGS Data sources – NCBI SRA, EBI-ENA, DDBJ-SRA, GEO; retrieving data from data sources - SRA toolkit; Aspera connect. Introduction of Genome editing and tools

Suggested Readings (Latest Editions):

1. Ali Masoudi-Nejad, Zahra Narimani, Nazanin Hosseinkhan; "Next Generation Sequencing and Sequence Assembly", Methodologies and Algorithms, Springer.
2. Stuart M. Brown, "Next-Generation DNA Sequencing Informatics", Cold Spring Harbor Laboratory Press.
3. Hecker, M. & Mullner, S., Berlin. Proteomics of microorganisms. Springer-Verlag.
4. Liebler, D.C. & Yates, J.R. III.. Introduction to proteomics: tools for the new biolog. New York. Humana Press.
5. David W Mount, "Bioinformatics sequence and Genome analysis", Second Edition, Cold Spring Harbor Laboratory Press,.
6. Introduction to Genomics, Arthur M. Lesk, Oxford University Press.

hemant gautam Surender Singh

BI-CCSU – 303: Molecular Modeling and Drug Design

Course Outcomes:

- Have knowledge of macromolecular structures and their properties
- Get to know the interactions of drugs and proteins
- Acquired knowledge of the in silico tools and their applications
- To aware the energy minimization process and different between QSAR and QSPR method.
- Students will learn different immunoinformatics method use in computational biology.

Unit 1- Concepts in molecular modelling: Coordinate systems, potential energy, molecular graphics, SMILE format, Mol2 forma, PDBQTt fomat etc. Force fields, Bond stretching, angle bending, introduction to non-bonded interactions, electrostatic interactions, Van der Waals interaction, hydrogen bonding in molecular mechanics, Description of various force fields for the simulation (AMBER,CHARMM MM3,MM4 etc.)

Unit 2- Simulations: Monte Carlo Simulations, Techniques for efficient conformational search: Simulated Annealing, Calculation of Free energy using simulation techniques. Energy minimization techniques:Concept of local and global minima, energy minimization protocol, energy minimization algorithms (steepest descent, conjugate gradient, Newton Raphson)

Unit 3- Chemoinformatics: Introduction, representing 2D & 3D structures, 2D chemical database applications & molecular descriptors and their classifications, database searching and applications in CADD. Lipinski rule of five, Drug target identification & computer-aided drug design processes, structure-based (receptor fitting) and ligand-based (receptor mapping) molecule design, Binding-site predictions, Docking Introduction, search algorithms, scoring (MM, Grid, etc.) Databases of small molecules (PubChem, ZINC etc.)

Unit 4- QSAR methodology, biological and physicochemical parameters, feature selection (PLS, PCA, MLR, etc.), model building and validation, QSAR applications in drug design, Quantitative structure-property relationships (QSPR), CoMFA and CoMSIA, 3D and nD-QSAR methods. ADMET: Oral bioavailability, BBB permeability, toxicity.

Unit 5- Methods for 3D Structure Prediction: Homology modeling of protein 3D structures. ab initio approach, Threading approach. A Comparison of protein structure prediction methods: CASP. Basic principles of modeling, steps in homology modeling , tools and databases, Structure validation - generation and analysis of Ramachandran plot using PROCHECK , WHATCHECK, VERIFY 3D etc.

Suggested Readings (Latest Editions):

1. Philip E. Bourne (Editor), Helge Weissig (Editor). 2003. Structural Bioinformatics – Methods of biochemical Analysis V. 44. New Jersey. Wiley-Liss.
2. Jan Drenth. Principles of protein X-ray Crystallography Springer-Verlag.
3. Branden, Carl & Tooze, John , Introduction to Protein Structure. Garland Publishing.
4. Friesner, R.A. Ed., Prigogine, L. Ed. & Rice, S.A.New York. John wiley & sons.. Computational methods for protein folding : advances in chemical physics vol. 120. Inc. publication.
5. Hans Dieter and Didier Rognan.. Molecular Modeling: Basic Principles and application. Wiley VeH Gmbh and Co. KGA.
6. Heilmeyer, L. & Friedrich, P. Amsterdam . Protein modules in cellular signaling edited. IOS Press.
7. Chemoinformatics: A Textbook, Johnann Gasteiger, Wiley-VCH, ,

hemant gautam Surender Singh

BI-CCSU – 304 (Elective I): Fundamental of Data Science

Course Outcomes:

- Understand the fundamental concepts and principles of data science.
- Apply data manipulation techniques using appropriate tools and libraries.
- Analyse and visualize data to extract insights and make data-driven decisions.
- Apply statistical methods and machine learning algorithms to analyse data.
- Communicate data findings effectively through data visualization and storytelling.

Unit 1 Introduction to Data Science- Overview of data science and its significance, Data science process and lifecycle, Ethical considerations in data science

Unit 2 Data Preparation and Manipulation- Data acquisition and data cleaning techniques, Data pre-processing: handling missing data, outliers, and data normalization, Exploratory data analysis and data profiling

Unit 3 Data Science Tools- Tools used in Data storage, Databases, Data Analysis, Data Visualization.

Unit 4 Machine Learning for Data Science- Introduction to machine learning, Supervised and unsupervised learning techniques, Model evaluation and selection.

Unit 5 Data Visualization and Communication Principles of data visualization, Data visualization techniques and tools, Communicating data findings and storytelling

Suggested Readings (Latest Editions):

1. "Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython" by Wes McKinney.
2. An Introduction to Data Science, Jeffrey Stanton, Syracuse University A Simple Introduction to DATA SCIENCE, Lars Nielsen, Noreen Burlingame
3. Introduction to Data Science, DAN POTTER,CARSTEN BINNING,ELI UPFAL
4. "Data Science for Business: What You Need to Know about Data Mining and DataAnalytic Thinking" by Foster Provost and Tom Fawcett.

hemant gautam Surender Singh

BI-CCSU – 304 (Elective II): Multimedia Applications with Web Technology

Course Outcome:

- Understand the concepts and components of multimedia technology.
- Design and develop multimedia content using appropriate tools and techniques.
- Integrate various multimedia elements such as text, images, audio, and video.
- Apply multimedia technologies in interactive applications and presentations.
- Evaluate and optimize multimedia content for different platforms and devices.

Unit 1 Multimedia Technology- Overview of multimedia technology and its components, Multimedia elements: text, images, audio, video, Multimedia file formats and compression techniques.

Unit 2 Tools and Techniques- Multimedia authoring software and tools, Design principles for multimedia content, Multimedia scripting and programming languages.

Unit 3 Image and Video Processing- Image and video acquisition and editing, Image and video compression techniques, Image and video enhancement and effects.

Unit 4 Audio and Animation in Multimedia- Digital audio concepts and formats, Audio editing and processing techniques, Animation principles and techniques.

Unit 5 Integration and Application Development- Integration of multimedia elements in interactive applications, Multimedia in web design and development, Optimization and delivery of multimedia content.

Suggested Readings (Latest Editions):

1. "Multimedia: Making It Work" by Tay Vaughan.
2. "Multimedia Systems: Algorithms, Standards, and Industry Practices" by Parag Havaldar and Gerard Medioni.
3. The effect of Multimedia annotation modes on L2 vocabulary acquisition: A comparative study. Language learning & technology, by A1-Seghayer, K., 5, 202-232.
4. The Application of multimedia information resources to carry out multimedia network teaching system, www.bazhouedu.coml.
5. Guo Jianmin, On the multimedia technology in teaching, http://www.govyi.coml.
6. On the application of the computer multimedia In teaching, http://www.ttdxs.coml.2009.5.25.

BI-CCSU -305 -Practical in Second Semester is based on, BI-CCSU -301 to BI-CCSU -304

hemant gautam Surender Singh

YEAR SECOND
SEMESTER IV