Syllabus for ASRB-NET (BIOINFORMATICS)

UNIT-I: Concepts in Computing: Fundamentals of Computing, Introduction to Operating Systems, WINDOWS, UNIX/Linux operating systems, Batch Processing, Multi-programmed Batch System, Time Sharing System, Distributed System, Real Time System, Dead lock, CPU scheduling, Computer Security, Computer Graphics: Visualization techniques, Interactive Graphics, Viewing in three dimension, Raster algorithms, Rendering, Animation, Image Processing with emphasis on biological systems, Computer Networking, Security of the network, Fire-walls, Network, Network topologyand architecture, Hierarchical networks, TCP/IP family of protocols, Parallel & High performance computing, Cloud computing

UNIT-II: Programming Languages: Programming in C, Arrays and Pointers, Variables, Procedures and functions, Standard Controls, Flow control and loops, string operators, Declaration and definition of user defined functions, Call by value, Call by references, File handling in C, Overview of Object Oriented Programming, Class and Objects, function and operator overloading, inheritance, Introduction to PERL, Overview,Syntax, Data types,Operators and Regular expressions in PERL, BioPerl modules

UNIT-III: Database Systems and Biological Databases: Definition, purpose of database system, Advantages of Database System, Components of Database System, Data Models-Relational, Network, Hierarchical, Three level Architecture for Database System –internal, conceptual and external levels, Data independence, Data Abstraction, Mapping, Data Definition Language, Data Manipulation Language, Role of Schemas, Client/Server architechture, Relational Databases- Relational data models (binary, ternary, quaternary & n-ary relations), Important terms in relational database system, Primary and secondary keys, Structured Query Language, Primary, secondary and derived biological databases, submitting sequence to the Database and retrieval, Data mining & Knowledge discovery in Biological databases, Supervised and unsupervised learning, machine learning techniques, Artificial Neural Network, Support Vector Machine and genetic algorithms

UNIT-IV: Mathematics and Statistics: Introduction to Probability and Statistics, Conditional probability, Bayes' theorem, Random variable, mathematical expectation, Probability Distributions, Binomial, Multinomial, Poisson and Normal Distribution, Introduction and properties of 't', Chisquare and F distribution, Principles of random sampling, Frequency distributionsGraphical representations and Descriptive measures, Standard Probability Distributions, Correlation and regression analysis. Hypothesis testing, Markov Models, Cluster Analysis- Hierarchical and Non-Hierarchical methods, Coordinate geometry with basic concepts of 2D and 3D geometry, Vector algebra– Addition and subtraction of vectors, Dot and cross product, Scalar triple product, Matrix algebra- basic definitions, matrix operations, transpose of a matrix, inverse of matrix, Eigen values, Boolean algebra. Geometric, Arithmetic and Harmonic Progression 2

UNIT-V: Biochemistry and Molecular Biology: Scope and importance of biochemistry in agriculture, hydrophobic, electrostatic and van der Waals forces, General introduction to physical techniques for determination of structure of biopolymers, Historical developments in molecular biology, Nucleic acids as genetic material, Genetic code, Genome organization and regulation in prokaryotes and eukaryotes, DNA replication, transcription and translation, recombinant DNA technology, Site directed mutagenesis, molecular mechanism of mutation, RNA editing, RNA processing, Structure, properties and functions of amino acids, proteins, and nucleic acids, Anatomy of Proteins - Secondary structures, Motifs, Domains, Tertiary and quaternary structures, Classification of Protein topologies, Helices and sheets, Helix-coil theory, Conformation of polypeptides and side chains, Ramachandran plot, Hinge motions in proteins, Algorithms in predicting secondary structure of proteins- Chau-Fasman algorithm, The Protein folding energy landscape, molten globules, enzymes and folding pathways

UNIT-VI: Genomics and Proteomics: Whole genome analysis and comparative genomics, classical ways of genome analysis, large fragment genomic libraries, Physical mapping of genomes, Genome sequencing strategies, Sequence assembly and annotation, exome sequencing, Genome structural and functional annotation, Functional genomics and proteomics, mass spectrometry, protein database and MS data analysis, peptide identification and protein interface, Candidate gene identification, Metabolomics for elucidating metabolic pathways, Linkage analysis, genotyping analysis, Applications of genomics and proteomics in agriculture, Evolution of sequencing technology, Microarrays analysis and applications, Next generation sequencing technologies, Sequencing by synthesis, ligation, single molecular sequencing, emerging NGS technologies

UNIT-VII: Analytical Bioinformatics: Bioinformatics- introduction, origin and history, Applications of Bioinformatics in agriculture, Information retrieval from Biological databases, Analysis of protein and nucleotide sequences, Basic concepts of sequence similarity, Introduction to sequences alignments- local and global alignment, pairwise and multiple sequence alignment, Dynamic programming-Needleman and Wunsch, Smith-Waterman algorithms, Scoring matrices: PAM & BLOSUM, Motifs and Patterns, Phylogenetic analyses, Types of phylogenetic trees, Tree-Building Methods, Character-based and Distance-based methods, Maximum-likelihood, Maximum parsimony, Unweighted Pair Group Method with Arithmetic Mean (UPGMA), Neighbor-Joining (NJ), FitchMargoliash (FM), Minimum Evolution (ME), Tree Evaluation, Bootstrapping, Introduction to systems and network biology

UNIT-VIII: Molecular Modelling: Concepts of Molecular Modelling, Molecular mechanics, Force Fields, Local and global energy minima, Techniques in MD and Monte Carlo Simulation for conformational analysis, DFT and semiempirical methods, Simulated annealing, RNA Secondary Structure prediction techniques, Algorithmic perspective of RNA folding, Protein structure prediction: ab initio, homology modeling and fold recognition methods, Receptor-based and ligand-based drug design, Design of ligands, docking, Classical SAR/QSAR, COMFA & COMSIA, Molecular descriptors, Pharmacophore mapping and applications

UNIT-IX: Algorithms and Data structures in Bioinformatics: Algorithms and complexity, Iterative and recursive algorithms, Fast versus slow algorithms, Big-O Notation, Algorithm design and analysis techniques, Greedy Algorithms, Randomized Algorithms, Divide-and-Conquer approach, Searching and Sorting algorithms, Linear and non-linear data structure, Stack, Queues, Linked list, Trees-Terminologies, Binary trees, Tree traversal (Pre-order, In-order, post-order), AVL trees, Graph theory, Depth-first search and Breadth-first search algorithms, Shortest Superstring Problem