

Course: Advanced Diploma In Bioinformatics

Scheme of Examination and Evaluation

Semester	Subject	Marks		Total Marks (External+ Internal))
		External	Internal	
Sem.-I	DBI 101T (Theory)	60	40	100
	DBI 102T (Theory)	60	40	100
	DBI 103T (Theory)	60	40	100
	DBI 104T (Theory)	60	40	100
	DBI 105P (Practical)	50	25	75
	DBI 106P (Practical)	50	25	75
	Total	340	210	550

Semester	Subject	Marks		Total Marks (External+ Internal)
		External	Internal	
Sem.-II	DBI 201T (Theory)	60	40	100
	DBI 202T (Theory)	60	40	100
	DBI 203T (Theory)	60	40	100
	DBI 204T (Theory)	60	40	100
	DBI 205P (Practical)	50	25	75
	DBI 206P (Practical)	50	25	75
	Project	300	00	300
		Total	640	210

Total Marks:

Sr. No.	Academic Session	Marks
1	Semester – I	550
2	Semester – II	850
	Total (Semester I + Semester II)	1400

Abbreviations:

DBI= Diploma in Bio-Informatics, T = Theory, P = Practical, S = Seminar, Sem. = Semester.

DETAILED SYLLABUS

SEMESTER - I

THEORY:

DBI 101 T: Fundamentals of Computer & Computing.

Unit – I	<p>Computer: Types of computers, Computer Architecture, Input output devices: Peripheral devices, input devices, asynchronous data transfer, mode of data transfer, DMA, I/O processors serial communication. Storage: primary storage & secondary storage (Memory organization): Basic memory cell, RAM, ROM, DDRAM, SDRAM, PROM, EPROM, main memory, auxiliary memory, associative memory, cache memory, virtual memory.</p>
Unit – II	<p>Number systems: Binary, Decimal, Octal & Hexadecimal number system & there inter-conversion, Computer codes, BCD, EBCDDIC & ASCII. Binary arithmetic. Boolean Algebra and Logic Circuits: fundamentals concepts, Boolean functions, and complements of Boolean function, Logic gates, Logic Circuits.</p>
Unit – III	<p>Operating System: Introduction, concept of operating system, Windows & Linux.</p>
Unit – IV	<p>Basics of Bioinformatics: Biology & Bioinformatics, Computers in Biology & Medicine, Macromolecules, and (Protein –Structure, Folding & Function, and DNA & RNA – Structure & Functions). Different fields in Bioinformatics, Contribution of Bioinformatics in Biological Sciences, different Field and Limitation of Bioinformatics.</p>
Unit - V	<p>Information Retrieval: LAN, WAN, Introduction to Internet, WWW, NICNET, ERNET, VSNL, ISDN, INTERNET (Basic, Connectivity), E-Mail, Publications on WWW – Online publishing ventures eg. E-Biomed On line International database access.</p>
	<p>Reference:</p>
1.	Sinha P. K; “Fundamentals of Computer”, BPB publication.
2.	Hooman H. Rashdi, Lucas K. Buehler, “Bioinformatics Basics”
3.	Cynthia Cribas & Per Jam Beck, “Developing Bioinformatics Computer skills”, O’Reilly associates.

DBI 102 T: Mathematics and Statistic

	Mathematics: (25 hrs)
Unit I	Number System: Integer, Real, Rational & irrational Set, Relation & function (Polynomials, Trigonometric, Exponential) Graphical. Representations of functions. Trigonometric function: Sin, Cos, Tan, Cot, Series expansion of these functions & other related functions.
Unit - II	Limits: Limits of Sequences, Series, Limits of functions. Calculus: Differentiation: First & Second order derivatives, Derivatives of Composite functions. Integration: Definite Integral; Indefinite Integral, Integration by parts Application (area, volume). Differential Equations: First order Differential equations, Growth equations.
Unit - III	Vector: Addition, Subtraction, dot & cross product, Triple scalar product, Divergence, curl of a vector, equation of normal. 2 D Coordinate Geometry: Equation of a line, circle, ellipse, parabola, hyperbola. 3 D Geometry: Equations of sphere & cone. Matrix Algebra: Addition, Subtraction, multiplication, transpose inverse & conjugate of matrix
	Statistics (20 hrs)
Unit - IV	Statistics: Descriptive Statistics: Numerical Description of Data. Measure of central Tendency, The Arithmetic Mean, median & mode. Measuring the variation in DATA standard Deviation, Population variance, sample variance, and significance of standard deviation. Percentile, Quartiles & Box Plots. Correlation and regression: Scatter diagram, correlation between, two variables, rank correlation, linear regression, Principles of least squares, time series data.
Unit - V	Probability: Sample space & events. Axioms of probability. Sample spaces having equally likely outcomes. Conditional Probability. Independent events. Bayes formula. Random Variables: Discrete Random Variables, Continuous Random Variables, Expected Value, and Variance. Binomial Distribution, Normal Distribution, Poisson Distribution, Geometric Distribution, Uniform Distribution, Chi-squared, Distribution, Student's t-distribution. Gumball Distribution. Limit Theorems, Central Limit Theorem, and Strong law of large numbers.
	References
1.	Calculus and Analytic Geometry 9 th Edition, G.B. Thomas & R. L. Fenney.
2.	Contemporary statistics A computer. Approach, Mc Graw Hill International edition-1994.
3.	Introduction to mathematics for life scientists 3 rd edition, Springer-verlay-1992.
4.	A first concept in probability, Sheldon Ross- Sixth Edition, Reason education Asia 2002.
5.	Contemporary statistics A computer. Approach, Mc Graw Hill International edition-1994.

DBU 103 T: Database Management System

Unit – I	<p>Database System: Basic of Databases: Database system-Definition, Database models-Instances & Schemes, Database language, Storage management, Database Administration, Database language,</p> <p>Database system Architecture: Introduction to centralized system, client server system, and parallel system and distributed system.</p>
Unit – II	<p>Entities Relationship model: Entity and entity sets, Relation and relationship sets, E-R diagrams, Reducing diagram to table trees.</p> <p>Relational Algebra and Calculus: Structure of Relational Databases: Relational Algebraic operations- relational calculus, domain relational calculus.</p> <p>Relational Databases: Relational Database design-Decomposition, importance of normalization, functional dependencies, normal form-first, second, third & fourth normal form.</p>
Unit – III	<p>Query Language an Query Optimization: SQL Basic – Schema definition in SQL, SQL commands, SQL operator, tables, views, indexes, aggregates functions, updates operations in SQL, Queries, Sub-queries, and Equivalence of queries.</p>
Unit – IV	<p>ASN.1 and NCBI data Model: Data model for Biological sequences, data types, - ASN.1 – Flat files, Inter database connectivity – Entrez and SRS Search Engine.</p>
Unit - V	<p>Networking Technology: Network Basics – TCP/IP, WWW, FTP – Web server, CGI programs, Perl/CGI, Developing dynamics application, CGI programming with Perl, connecting database with ODBC – Hyper Text Markup Language (HTML) – XML.</p>
	References
1.	“Database system concepts” Hanery Korth & Abraham. Silber Schatz. Tata Mac-Graw Hill Publications.
2.	“Parallel & Distributed Databases”, Wilteach et.al.
3.	“Introduction to Database system” C.J. Date.
4.	“Database system organization”, J.M. Martin, Princeton-Hall
5.	“Oracle: Power Object Handbook”, Bruce Kolste & David peterson.
6.	“Database system concepts”, Hanery Korth & Abraham. Silber Schatz. Tata Mac-Graw Hill Publications.

DBI 104 T: Programming with C & C++

Unit – I	Introduction to C: Control statements – Loops, Decision control structures – fundamentals of arrays, Function, Strings, Structure – Unions – Command line arguments.
Unit - II	Introduction to Data Structures: Storage and representation of Arrays, Stacks and Queues – Handling pointers with structures, Arrays, Functions and Strings, single linked list, Doubly linked list.
Unit – III	General Concept: OOP as a problem solving model – Objects – Classes – Abstraction – Encapsulation – Subclasses – Inheritance & Polymorphism – Extensibility – Reusability, Maintainability and Reliability of s/w system – Introduction to object and Classes. Enhanced features of C++ over C: Drawback of structured programming – object oriented languages, characteristics and fundamentals – differences between C and C++.
Unit - IV	Data Encapsulation and Data Hiding: Object and classes in C++, self reference in classes – constructor, destructor – overloading of operators. Introduction to concept on Inheritance, Virtual, Functions. Polymorphism: Derived class construct, derived classes with parent class – Late Binding, virtual base classes – abstract classes – friend classes.
Unit - V	File handling in C/C++, Streams.
	References
1.	“Let C++ 4 th Ed”, Yashwant Kanetkar, BPB Publication
2.	“Let C”, Yeshwant Kanitkar, BPB publication.
3.	“C – Compete Reference”, Tata McGrew Hill publication.

Practical - Semester – I

DBI 105 P: Practical – I (30 Practical)

A.	Practical related with windows 0.5 & Linux Utilities like menu bar, submenu bar, Taskbar, Desktops, Recycle bin, Help menu, Control panel, Directory & file management, Backup etc.
B.	Unix practical: Login Directory & file management, Help menu & Unix Operation commands Unix editor.
C.	Programming in C: Minimum 15 programs on following topics
D.	Programming in C++: Writing Minimum 15 programs in C++.

DBI 106 P: Practical - II (30 Practical)

A.	Database and Data Retrieval- Goggle search NCBI, PIR, EMBL, DDBJ, SWISS-Prot, PDB, CATH, SCOP, COG, Pubmed, Enzyme database, Genome Database. Entrez & SRS data retrieval: Subject search, Boolean logic, Truncation, Author Search, Limit Search, Joining Query, History. Sequence format & format Inter Conversion.
B.	SQL: Data Definition Statement, Data Manipulation Statement, Data Control Statements, And Simple Query & Joined Query.
C.	Designing 5 small databases, normalization & implementation. Designing 2 Biological databases, normalization & implementation.
D.	WebPages Designing using HTML, XML & DHTML in FrontPage & Notepad

Semester - II

THEORY:

DBI 201T: Computational Biology

	Introductory Bioinformatics: (3 hrs)
Unit - I	Prokaryotic and Eukaryotic genomes, structure, organization and function, Molecular Evolution, Gene Structure, Genetic Codes and Mutations. Biological Databases: Primary source of sequence and structure database. Secondary database. Databases in Bioinformatics: Nucleic acid sequence database: Gen Bank, EMBL, DDBJ, and submission of Data. Protein sequence database: SwissProt, TrEMBL, PIR, Enzyme, G-Protein. Protein Structure & Classification database: MMDB, PDB, COG, CATH, SCOP. Genome Databases: EnEMBL Enzyme database: ExPaSy Enzyme, Branda's Enzyme database. Metabolic Pathway database: KEGG, EMP database.
Unit - II	Sequence Analysis: Primer selection, ORF, Intron – Exon, Promoter, Repeat regions, Gene finding, restriction mapping, Pairwise sequence alignment: Dot Matrix Method, Scoring Matrices (PAM & BLOSUM), Gap penalty, Dynamic programming: Needleman & Wuncsh and Smith waterman algorithm. Heuristic methods: FASTA: Algorithm & Searching. BLAST: Algorithm & Searching. Gapped BLAST, Statistical Analysis.
Unit - III	Multiple Sequence Alignment: Simple method, Progressive method: ClustalW, PILEUP, Iterative Methods: Genetic Algorithms and HMM. & PSI-BLAST. Patterns, Profiles & Motifs – Regular expression and Positions specific scoring Matrices, Repeats finding and pattern recognition. Derived databases of pattern, motifs.
Unit - IV	Phylogenetics Analysis: Evolutionary Changes in nucleotides sequences, rates and patterns of nucleotides substitution, Models for nucleotides substitution, Methods for Phylogenetics estimation, Maximum parsimony, Distance Matrix, Methods and Maximum Likelihood Methods.
Unit - V	Functional Perl programming: String processing, Regular Expression, Object- Oriented programming in Perl. Modules – CGI, DBI and GD – Perl. Bioperl.
	References:
1.	“Introduction Bioinformatics”, T. K. Attwood, Pearson Publication Asia.
2	“Bioinformatics: a practical guide to the analysis of gene & proteins “, Baxevanis, A. D. & Quellettee, Wiley-Sons Interscience- 2003- Asia
3.	“Bioinformatics - sequence and genome Analysis”, Mount D.W Cold spring Harbor CSHL press-2003 Asia.

DBI 202T: Structural Bioinformatics

Unit – I	Protein structure determination: Isolation and purification of proteins, Crystallization of Proteins. Basic principles of X-ray diffraction studies, phase determination, calculation of electron density Map, Refinement of structures. Electron crystallography of protein. NMR techniques for protein structure determination.
Unit – II	Protein structure anatomy: Physical & chemical properties of amino acid & polypeptide. Physical nature of non-covalent interaction, conformational properties of protein - Ramchandran plot, secondary, super secondary, tertiary & quaternary structure of proteins, protein structure stability, protein folding pathway, factor affecting protein stability. Anatomy of DNA: A, B, Z DNA, DNA bending, RNA structure, Structure of Ribosome
Unit – III	Protein structure prediction: Secondary structure prediction: Statistical method: Chou-Fasman, GOR. Neural Network method: PHD, PSIPRED, JNET. Tertiary structure prediction: Homology Modeling, Threading, fold recognition. Prediction of Structural classes, motif, folds and domains. Classification of three-dimensional structures in Brook haven Protein Data Bank (HSSP, SCOP, FSSP, CATH).
Unit – IV	Molecular dynamics & Molecular Simulation studies: Molecular mechanics & molecular dynamics of oligopeptides, proteins, nucleotides and small molecules, Empirical forcefields for biomolecular simulations, Energy minimization technique, Molecular Dynamics simulations, Monte Carlo Simulation, Techniques of efficient conformational search: Simulated Annealing, Calculation of relative Free energy using simulation techniques.
Unit - V	Methods for comparison of 3D protein structures, Methods of 3D structure prediction of nucleic acid, Molecular Interaction: Protein – Protein, Protein – DNA, Protein – Carbohydrate, Protein - small molecules, DNA – Small molecules etc, Molecular Docking.
	Reference
1.	Protein Structure Prediction: a practical approach, q1996 Sternberg. M.J.E.,
2.	Mechanism of Protein Folding, 1994, Pain, R.G
3.	Protein structure Prediction: Methods and Protocol, David M. Webster Ed, Humana Press Totowa, New Jersey
4.	Molecular Modeling: Principle and applications, Leach A.R.
5.	Introduction to Protein structure(1991), Carl Branden and John Tooze Garland Publishing, Inc.; New York and London, 1991
6.	Molecular Modeling, Claude Cohen, Academic Press, New York.
7.	Protein Structure Stability and Function, Keneth Murphy
8.	Protein Folding, Creighton, T.: 1992

DBI 203T: - Advanced Bioinformatics

Unit – I	Genomic: genome databases, assembly of full genome from sequence fragments, gene prediction algorithms, operon prediction methods genome alignment methods gene order comparison tools, Gene identification, Fourier Analysis of DNA sequences. Comparative genomic, gene annotations. Pharmacogenomics & DNA based drug designing.
Unit – II	Proteomics: definition identification and analysis, 2D gel data (image analysis), mass spectroscopic data.(principle, analysis, information integration, data validation and peptide sequence determination)
Unit – III	Drug designing: QSAR for drug design, De novo protein design and engineering, High throughput searching, virtual screening, chemoinformatics-2D/3D chemical or drug databases. Structure based / ligand based & de novo drug design
Unit – IV	Metabolic pathway engineering: enzyme database, metabolic pathway databases
Unit - V	Gene expression profile databases & analysis. Methods for analyzing gene expression and micro array data. Techniques: clustering, SVMs basics of designing and micro assay image analysis. Normalization. Variability and replication. The interpretation of large-scale data. Clustering algorithms.
	Reference:
1	QSAR and computer aided drug design: Structural Bioinformatics Philip E Bourne, Helge Weissig Ced
2	Molecular modeling in drug design, Editor: N. ClandeCohen Academic press, New York
3	“Bioinformatics - sequence and genome Analysis”, Mount D.W Cold spring Harbor CSHL press-2003 Asia

DBI 204T: Computer Programming – II

A	<p>Programming in Java: An introduction to JAVA Programming Object –Orientated Programming and JAVA JAVA Basics Working with objects Conditionals, hoops and Arrays Creating classes and application in JAVA More about methods JAVA applets basics Graphics, fonts, & color Simple animation & threads Advanced animation, images & sound Creating user interfaces with AWT Windows, networking Packages and interfaces Exception Multithreading Streams and 7 /0 Using Native methods and libraries</p>
B	<p>Programming with visual Basics Introduction to Application Development using Visual Basic Working with code & forms Variables, procedures & controlling program executors Standard control Data access using Data control Connecting to ORACLE Database using visual Basics</p>
	<p>References:</p>
1	<p>“Visula Basics –6- complete Reference”, Tata Mcgraw Hills , New delhi</p>
2	<p>“JAVA: How to program Java 2 (2nd edition)”, Dietel & Dietel (2003) Pearson Education.</p>
3	<p>“JAVA the complete Reference (2003)”, H. Schildt , Tata Mcgraw Hills , New delhi</p>

PRACTICALS:

DBI 205 P: Practical - I

A.	Sequence Analysis: - (Standalone & Web based Software tools) Sequence Comparison using. Dot Matrix method. Dynamic Programming. Jemboss & Emboss. Similarity search using BLAST & FASTA, Statistical analysis of search. Multiple sequences Alignment. Phylogenetic Analysis.
B.	Programming in Java Minimum 20 programs in Java
C.	Programming in visual Basics Minimum 5 programs on Application development Minimum 2 programs on Database connectivity.

DBI 206P: Practical – II (30 Practical)

A	Structural Database: PDB, CATH, SCOP, MMDB, Retrieval of Data.
B	Introduction to RasMol, VMD, spdbv & Hyperchem software tools. Using different visualization & rendering options with RasMol & spdbv, Manipulation of protein molecule, energy calculation, Electrostatic surface generation & animation using spdbv
C	Protein structure prediction: Secondary structure prediction. Tertiary structure prediction, 3 D structure analysis
D	3D structure prediction of nucleic acid
E	Molecule Designing & energy minimization, protein homology modeling, protein structure prediction using ab-invitro method
F	Docking of molecules.

PROJECT: Duration 2 months

Project will be a part of their curriculum for the partial fulfillment of Diploma in Bioinformatics. Student has to undergo training at industry or research institute for two months. After completion of training they have to submit their project report. For the final project valuation student has to deliver presentation and has to pass through viva.

The project valuation will be done on the basis of project report submitted, presentation and performance in viva.