

**MASTER OF SCIENCE
(BIO-INFORMATICS)**



**CENTRE FOR EXCELLENCE IN PHARMACEUTICAL SCIENCES
GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078**



**CENTRE OF EXCELLENCE IN PHARMACEUTICAL SCIENCES
GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078**

INTRODUCTION

Bioinformatics is a rapidly growing field of study that centres on the applications of information science and technology in biology. As such, bioinformatics brings together molecular biology with areas from mathematics, statistics and computer science such as visualization pattern, graph theory, complex systems, and many others. Bioinformatics now entails the creation and advancement of databases, algorithms, computational and statistical techniques and theory to solve formal and practical problems arising from the management and analysis of biological data. Considering the inter-disciplinary nature of bioinformatics, this master programme shall have a major component from basic science like botany, zoology, biochemistry, immunology, molecular biology, chemistry, physics, mathematics, statistics, and computer science.

OBJECTIVES

The core objective of the programme is to provide quality education to the graduates, who want to pursue their career in the emerging area of Bioinformatics or Computational Biology.

- To serve as a nodal point for Bioinformatics programme and its applications.
- To provide industry interface to the students for Project.

TARGET GROUP AND ELIGIBILITY

Candidate desiring for admission to Master of Science in Bioinformatics Programme shall be required to have passed B.Sc. (Life Sciences/ Bioinformatics/ Biotechnology/ Microbiology/ Biochemistry/ Botany/ Zoology/ Physics/ Chemistry/ Mathematics/ Physics/Biophysics/Computer Science as one of the major subjects studied at Graduation level. B.Sc. (Agriculture)/ B.V.Sc./ B.Pharm./ B.E. (Biotech/ Bioinformatics/Computer Science) / B.Tech. (Biotech/Pharmaceutical Sciences Bioinformatics/computer Science/data science/Artificial intelligence/nano sciences) and similar other subjects and must have obtained 55% aggregate at graduation level/CGPA/grade. MBBS or any other course equivalent thereto with 50% marks.

Admission Criteria: Admission shall be on the basis of CET/CUET/DBT BINC/ any recognized National Level examination.

Duration: Two years (Four Semester) – Full Time

Standard of Passing and Award of Division

- a) Minimum marks for passing in each theory/lab course shall be 50% of the marks prescribed for the paper/lab.
- b) Candidate who secures 50% or more marks but less than 60% of the aggregate marks prescribed for four semesters taken together, shall be awarded **Second Class**.

Deena Deswal

- c) Candidate who secures 60% or more of the aggregate marks prescribed for four semesters taken together, shall be awarded **First Class**.
- d) The practical/project shall be assessed by the two examiners, appointed by the University.

Number of Candidates to be admitted

The maximum number of students to be admitted for the Master's Programme in an academic session is 20+2.

Admission

Admission shall be based on merit basis in accordance with the number of applications received/entrance examination conduct on the following criteria:

1.	Duration of Examination shall be three hours.	
2.	Entrance Examination Question Paper shall be in the following Pattern:	
a)	No. of Questions to be covered from broad area of Fundamental and Applied Sciences at Graduation Level.	100 (Compulsory)
3.	There will be no negative marking.	

Examination Question Pattern depending on Credit: Time: 3hours, Maximum Marks: 60

Minimum number of credits required for successful completion of the programme are 83.

Deepa Desmal
Paul
Feder

- c) Candidate who secures 60% or more of the aggregate marks prescribed for four semesters taken together, shall be awarded **First Class**.
- d) The practical/project shall be assessed by the two examiners, appointed by the University.

Number of Candidates to be admitted

The maximum number of students to be admitted for the Master's Programme in an academic session is 20+2.

Admission

Admission shall be based on merit basis in accordance with the number of applications received/entrance examination conduct on the following criteria:

1.	Duration of Examination shall be three hours.	
2.	Entrance Examination Question Paper shall be in the following Pattern:	
a)	No. of Questions to be covered from broad area of Fundamental and Applied Sciences at Graduation Level.	100 (Compulsory)
3.	There will be no negative marking.	

Examination Question Pattern depending on Credit: Time: 3hours, Maximum Marks: 60

Minimum number of credits required for successful completion of the programme are 83.

Grading System of the University

Percentage of marks obtained	Grade	Grade Point
90-100	O	10
75-89	A+	9
65-74	A	8
55-64	B+	7
50-54	B	6
Less than 50 or absent	F	0

Deepa Deswal
Deepa Deswal
Deepa Deswal

- (viii) 'Generic Elective Course' (GE) means an optional course to be selected by a student out of such courses offered in the same or any other School/Centre.
- (ix) 'Open Elective' means an elective course which is available for students of all programmes, including students of same School/Centre. Students of other School/Centre will opt these courses subject to fulfilling of eligibility of criteria as laid down by the School/Centre offering the course.
- (x) 'Credit' means the value assigned to a course which indicates the level of instruction; One-hour lecture per week equals 1 credit, 2 hours practical class per week equals 1 credit. Credit for a practical could be proposed as part of a course or as a separate practical courses.
- (xi) 'SGPA' means Semester Grade Point Average calculated for individual semester.
- (xii) 'CGPA' is Cumulative Grade Points Average calculated for all courses completed by the students at any point of time. CGPA is calculated each year for both the semesters clubbed together.
- (xiii) 'Grand CGPA' is calculated in the last year of the course by clubbing together of CGPA of two years, i.e., four semesters. Grand CGPA is being given in transcript form. To benefit the student a formula for conversation of Grand CGPA into %age marks is given in the transcript.

Deepa Deswal Trichu .
Pavul Beelgal .



**CENTRE OF EXCELLENCE IN PHARMACEUTICAL CHEMISTRY
GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078**

Programme Educational Objectives (PEOs):

- The M.Sc. Bioinformatics program describes accomplishments that graduates are expected to attain within five to seven years after only graduation.
- **PEO1** Should be aware of available bioinformatics resources and information in databases
- **PEO2** Apply critical, analytical and problem solving skills to deal with bioinformatics research
- **PEO3** Gain a deep understanding of the biomolecular systems
- **PEO4** Attain knowledge to identify the research problems in molecular biology
- **PEO5** Develop very efficient algorithms to extract biological knowledge from complex dataset
- **PEO6** Develop programming skills to write their own codes and scripts to resolve research accomplishments
- **PEO7** Possess drafting and writing skills to enhance scientific communication
- **PEO8** Update in accordance with recent scientific advancements and technology
- **PEO9** Gain higher level degrees to pursue a career in academics or scientific organizations
- **PEO10** Should be able to pursue research and also in industry placement in the sectors of Biological Sciences, pharmaceutical, agricultural, environmental science and food industries.

Program Specific Outcomes (PSOs):

After the successful completion of M.Sc. Bioinformatics program, the students are expected to:

- **PSO1** Develop applications to solve biological problems by utilizing the Bioinformatics algorithms and programming languages
- **PSO2** Manipulate scientific datasets and retrieve and access them through biological databases
- **PSO3** Develop knowledge and skills to analyse and interpret biological data for research requirements
- **PSO4** Develop good scientific communication skills
- **PSO5** Gain knowledge of the tools, frameworks, and libraries for bioinformatics applications
- **PSO6** Undergo seminars, conferences and hands on workshops to facilitate domain expertise

Deepa Siswal

Shekhar

Parul



CENTRE OF EXCELLENCE IN PHARMACEUTICAL CHEMISTRY
GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Program Structure:

The Master of Science in Bio-Informatics Course is a Two Year Full-Time Course consisting of four Semester, viz. Semester-I, Semester-II, Semester-III and Semester-IV.

First Year	Part-I	First Semester	Second Semester
Second Year	Part-II	Third Semester	Fourth Semester

Course Credit Scheme at a Glance:

First Semester

Course Code	Nomenclature of the Paper		M.M.	C.E.	E.E.	No. of Hours		Credits
						Th.	Pr.	
BI-801	Fundamental of Biochemistry	CFC	100	40	60	4		4
BI-803	Fundamental of Chemistry	CFC	100	40	60	4		4
BI-805	Essential Mathematics & Biostatistics	CFC	100	40	60	4		4
BI-807	Introduction to Bioinformatics	CC	100	40	60	4		4
BI-809	Fundamental of Computer and Programming	CFC	100	40	60	3		3
BI-851	Lab-I (Bioinformatics Tools)	CFC	100	40	60		6	3
BI-853	Lab-II (C++ & Linux)	CC	100	40	60		6	3
			700			19	12	25

Second Semester

Course Code	Nomenclature of the Paper		M.M.	C.E.	E.E.	No. of Hours		Credits
						Th.	Pr.	
BI-802	Database Management Systems & Data Mining	SEC	100	40	60	3		3
BI-804	Genomics and Proteomics	CC	100	40	60	4		4
BI-806	Programming in Perl	SEC	100	40	60	3		3
BI-808	Structural Bioinformatics and Drug Design & Molecular Imaging	CC	100	40	60	4		4
BI-852	Lab-III (MySQL/Oracle)	SEC	100	40	60		6	3
BI-854	Lab-IV (Systems Biology and Advance Bioinformatics Tools)	AECC	100	40	60		6	3
BI-856	Lab-V (Perl)	SEC	100	40	60		4	2
			700			14	16	22

Deepa Deswal *Techn.* Parul
beehar



CENTRE OF EXCELLENCE IN PHARMACEUTICAL CHEMISTRY

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Third Semester

Course Code	Nomenclature of the Paper		M.M.	C.E.	E.E.	No. of Hours		Credits
						Th.	Pr.	
BI-901	Programming with Python and Biopython	CC	100	40	60	3		3
BI-903	Chemoinformatics and Chemogenomics	CC	100	40	60	3		3
BI-905	Comparative and Functional Genomics	DCE	100	40	60	3		3
BI-907	Algorithms in Bioinformatics	SEC	100	40	60	3		3
BI-909 (A1)	Machine Learning & Artificial Intelligence in Bioinformatics	DCE	100	40	60	3		3
BI-909 (A2)	Cloud Computing							
BI-951	Lab-VI (Drug Design + NGS)	SEC	100	40	60		6	3
BI-953	Lab-VII (Programming with Python and Biopython)	CC	100	40	60		4	2
BI-955	Lab-VIII (Chemoinformatics and Chemogenomics)	CC	100	40	60		4	2
			800			15	14	22

*One course to be selected (909A1/909A2)

Fourth Semester

Course Code	Nomenclature of the Paper		M.M.	C.E.	E.E.	No. of Hours		Credits
						Th.	Pr.	
BI-902	Data Mining In Healthcare	CC	100	50	50	4		4
BI-904 (B1)	Programming in JAVA	GE	100	50	50	4		4
BI-904 (B2)	Immunoinformatics							
BI-906	Developing Entrepreneurial Mindset	NUES	100	50	50	2		2
BI-958	Major Project/Dissertation	CC	100	-	-		24	12
			400			10	24	22

*One course to be selected (906B1/906B2)

Deepa Samal
Jasjit
Sudhanshu
Pamul



CENTRE OF EXCELLENCE IN PHARMACEUTICAL CHEMISTRY
GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Elective Papers

For the Students of M.Sc. (Bio-Informatics)							
Course Code	Course Title	Credits	Teaching Hours per week	Maximum Marks			
				Continuous Evaluation	Mid-Semester Examination	End-semester Examination	Total
909 (A1)	Discipline Centric Elective	3	3	15	25	60	100
909 (A2)	Discipline Centric Elective	3	3	15	25	60	100
Total Credits/Marks		03					100
General Elective							
The Center or any other school offers the following general elective papers to the students.							
Course Code	Course Title	Credits	Teaching Hours per week	Maximum Marks			
				Continuous Assessment	Mid-Semester Examination	End-semester Examination	Total
906(B1)	Programming in JAVA	4	4	15	25	60	100
906(B2)	Immunoinformatics	4	4	15	25	60	100
Total Credits/Marks		4					100

Deepa Deswal
 Behgal
 Tech. Panel



CENTRE OF EXCELLENCE IN PHARMACEUTICAL CHEMISTRY
GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Total Marks of all Four Semesters

SEMESTER	CREDIT	MARKS
I	25	700
II	22	700
III	22	800
IV	22	400
GRAND TOTAL	91	2600

Internal Assessment in theory papers will be made on the basis of one test and continuous evaluation parameters as decided by the University from time to time, while in Laboratory papers it will be decided from continuous assessment in internal viva-voce examination of all the experiments performed. Current guidelines for determining Internal Assessment in theory papers are given as 'Annexure-A'.

Each student will submit a project report at the end of fourth semester duration on the topic to be allotted by the Centre through a constituted committee in 3rd Semester of the M. Sc. Course as per the prescribed schedule. The marks will be awarded by the external examiner and committee on the basis of performance presentation submitted by the student.

Total credits to be earned by the student is 89, leaving behind NUES.

Deepa Dermal
bselga!
Arul . Parul



CENTRE OF EXCELLENCE IN PHARMACEUTICAL CHEMISTRY
GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Annexure-A

Scheme for awarding internal assessment: continuous mode

	Theory (4 Credit/3 Credit/2 Credit)
Criteria	Maximum Marks
Continuous evaluation	15
Mid Term	25
Total	40
	Practical (4 Credit/3 Credit/1 Credit)
Based on Practical Records, Regular viva voce, etc.	40
Total	40

(Objective Learning involves Multiple Choice Test, Matching Test, True / False Test Correct / Incorrect Test, Recall Test, Best Answer Test, Completion Test etc.)

Deepa Siswal

Arjun

Ravi



CENTRE OF EXCELLENCE IN PHARMACEUTICAL CHEMISTRY
GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Letter grades and grade points equivalent to percentage of marks and performances

Marks	Grade Point	Grade
90-100	10	O
75-89	9.0	A+
65-74	8	A
55-64	7	B+
50-54	6	B
45-49	5	C
40-44	4	P
30-39	3	D
20-29	2	E
10-19	1	F
Less than 10 or absent	0	F

A learner who remain absent for any end semester examination shall be assigned a letter grade of F, and a corresponding grade point of zero. He/she should reappear for the said evaluation/examination in due course.

Calculation of Semester Grade Point Average (SGPA) and Cumulative Grade Point Average (CGPA)

1. Performance in a semester will be expressed as Semester Grade Point Average (SGPA) and shall be rounded to two decimal digits.
2. Cumulative performance of all the semesters together will reflect performance in the whole programme and it will be known as Cumulative Grade Point Average (CGPA), and shall be rounded to two decimal digits.
3. The formula for calculation for SGPA and CGPA is given below:

$$SGPA = \frac{\sum_i C_i G_i}{\sum_i C_i}$$

$$CGPA = \frac{\sum_n \sum_i C_{ni} G_{ni}}{\sum_n \sum_i C_{ni}}$$

Where;

C_i – number of credits for the ith course.

G_i – grade point obtained in the ith course.

C_{ni} – number of credits of the ith course of the nth semester.

Deepa Dharwal

Dr. Anil Kumar

Ravi

Cni – number of credits of the i^{th} course of the n^{th} semester.

Mni – marks of the i^{th} course of the n^{th} semester.

Gni – grade points of the i^{th} course of the n^{th} semester.

4. The successful candidates as per clause 11.6 and having an overall CGPA higher than or equal to the minimum CGPA specified in the Syllabi and Scheme of Teaching and Examination for the award of the degree, shall be awarded the degree and shall be placed in Divisions as below:

- **CGPA of 4.00 – 4.99** shall be placed in the Third Division.
- **CGPA of 5.00 – 6.49** shall be placed in the Second Division.
- **CGPA of 6.50 or above** shall be placed in the First Division.
- **CGPA of 10** shall be placed in the Exemplary Performance. Exemplary Performance shall be awarded, if and only if, every course of the programme offered to the student is passed in the first chance of appearing in the paper that is offered to the student. A student with an academic break shall not be awarded the exemplary performance.
- The $\text{CGPA} \times 10$ shall be deemed equivalent to percentage of marks obtained by the student for the purpose of equivalence to percentage of marks.

Deepa D. S. S. Teacher, Principal
B. S. S. S.

Scheme of M.Sc. (Bioinformatics)

Semester-I	Semester-II	Semester-III	Semester-IV
BI-801: Fundamental of Biochemistry ④	BI-802: Database Management Systems & Data Mining ③	BI-901: Programming with Python and Biopython ③	BI-902: Data Mining In Healthcare ④
BI-803: Fundamental of Chemistry ④	BI-804: Genomics and Proteomics ④	BI-903: Chemoinformatics and Chemogenomics ③	BI-904 (B1): Programming in JAVA* ④
BI-805: Essential Mathematics & Biostatistics ④	BI-806: Programming in Perl ③	BI-905: Comparative and Functional Genomics ③	BI-904 (B2): Immunoinformatics ④
BI-807: Introduction to Bioinformatics ④	BI-808: Structural Bioinformatics and Drug Design & Molecular Imaging ④	BI-907: Algorithms in Bioinformatics ③	BI-906: Developing Entrepreneurial Mindset ②
BI-809: Fundamental of Computer and Programming ③	BI-852: Lab-III (MySQL/Oracle) ③	BI-909 (A1): Machine Learning & Artificial Intelligence in Bioinformatics ③	BI-958: Major Project/Dissertation ⑫
BI-851: Lab-I (Bioinformatics Tools) ③	BI-854: Lab-IV (Systems Biology and Advance Bioinformatics Tools) ③	BI-909 (A2): Cloud Computing ③	
BI-853: Lab-II (C++ & Linux) ③	BI-856: Lab-V (Perl) ②	BI-951: Lab-VI (Drug Design + NGS) ③	
		BI-953: Lab-VII (Programming with Python and Biopython) ②	
		BI-955: Lab-VIII (Chemoinformatics and Chemogenomics) ②	



CC: Core Course

DCE: Discipline Centric Elective

SEC: Skill Enhancement Course

CFC/AECC: Compulsory Foundation Course/Ability Enhancement Compulsory Course

GE: Generic Elective Course/OE: Open Elective Course

NUES: Non University Examinational Subject (Entitled for credit and not to be considered for the purpose of declaration of Result)

***Note: Student would have to opt one course each from 909A1/909A2 and 904B1/906B2.**



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-801	FUNDAMENTALS OF BIOCHEMISTRY	L	T	P	C
		Compulsory Foundation Course	4	-	-	4
Pre-requisite		<u>Basic Knowledge in Biomolecules</u>	Syllabus		2025	

The main objectives of this course are to:

1. The course aims to provide students with an understanding of biomolecules, the basic building blocks of living organisms, focusing on their structural underpinnings, unique properties, biological roles and functions and inter relations.
2. Emphasis will be on the association between structure and function of various biomolecules at a chemical level with a biological perspective.
3. It provides comprehensive understanding of DNA replication.
4. The objective of the course is to provide detailed knowledge about enzymes, the biological catalysts with remarkable properties that sustain life, so as to develop an understanding of enzyme kinetics.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

- Comprehend the structure, function and acid base properties of amino acids.
- Introduced to the structure, properties and roles of carbohydrates, lipids and nucleic acids.
- Learn about the molecular basis of processes like DNA replication.
- Learn the nature and importance of enzymes in living systems.
- Gain insight into molecular basis of catalysis by enzymes and the underlying basis of their specificity.

Unit-I

Carbohydrates and Glycobiology: Monosaccharides - structure of aldoses and ketoses; Ring structure of sugars, conformations of sugars, mutarotation, anomers, epimers and enantiomers; Structure of biologically important sugar derivatives, oxidation and reduction of sugars; Formation of disaccharides, reducing and non-reducing disaccharides; Polysaccharides-homo and heteropolysaccharides, structural and storage polysaccharides; Structure and role of glycoconjugates - proteoglycans, glycoproteins and glycolipids (gangliosides and lipopolysaccharides); Carbohydrates as informational molecules.

Building blocks of lipids - fatty acids, glycerol, ceramide; Storage lipids - triacyl glycerol and waxes; Structural lipids in membranes - glycerophospholipids; Galactolipids and sulpholipids, etherlipids, sphingolipids and sterols, structure, distribution and role of membrane lipids.

Unit-II

Nucleic Acids: Nucleotides - structure and properties of bases, pentoses, nucleosides; Nucleic acid structure-Watson-Crick model of DNA, forms of DNA; Structure of major species of RNA - mRNA, tRNA and rRNA.

General features of replication, the chemistry of DNA synthesis, DNA polymerase, the replication fork, enzymes and proteins in DNA replication.

Deepa Deswal *Teaching Assistant* - *Beelgaal* - *Paul*

Unit-III

Introduction to amino acids, peptides and proteins: Amino acids and their properties - hydrophobic, polar and charged. Organization of protein structure into primary, secondary, tertiary and quaternary structures. N-terminal and C-terminal amino acid analysis. Sequencing techniques - Edman degradation. Generation of overlap peptides using different enzymes and chemical reagents. Disulfide bonds and their location. Forces stabilizing the protein structure - covalent and non-covalent. Importance of primary structure in protein folding. The peptide bond, dihedral angles psi and phi, helices, sheets and turns, Ramachandran map. Motifs and domains.

Unit-IV

Introduction to enzymes: General characteristics of enzymes; nature of enzymes - protein and non-protein. Co-factor and prosthetic group, apoenzyme, holoenzyme. Classification and nomenclature of enzymes. Enzyme activity, specific activity, units to express enzyme activity. Features of enzyme catalysis, factors affecting the rate of chemical reactions, collision theory, activation energy and transition state theory. Catalytic power and specificity of enzymes (concept of active site), Fischer's lock and key hypothesis, Koshland's induced fit hypothesis.

Enzyme kinetics: Relationship between initial velocity and substrate concentration, equilibrium constant, steady state kinetics, mono-substrate reactions. Michaelis-Menten equation, Lineweaver-Burk plot, Eadie-Hofstee and Hanes plot. Determination of K_M and V_{max} , K_{cat} , specificity constant. Effect of pH and temperature on the activity of enzymes. Enzyme inhibition, Reversible inhibition (competitive, uncompetitive, non-competitive and mixed) and irreversible inhibition.

Suggested Readings:

1. Lehninger: Principles of Biochemistry (6th Ed.), Nelson D.L. and Cox M.M., W.H. Freeman & Company (New York), 2013 (ISBN:13: 978-1-4641-0962-1/ISBN: 10-14641-0962-1).
2. Biochemistry (4th Ed.), Voet Donald, Voet Judith & Pratt Charlotte, Wiley & Sons, Inc. (New Jersey), 2013 (ISBN:978-1-11809244-6).
3. Biochemistry (5th Edition), Jeremy M Berg, John L Tymoczko, and Lubert Stryer, New York: W H Freeman, 2002 (ISBN-10: 0-7167-3051-0).

Reference Books:

1. Lewin's Genes XII, Krebs Jocelyn E., Goldstein Aachorage Elliott S. and Kilpatrick Stephen T., Jones & Bartlett, 2018.
2. Enzymes: Biochemistry, Biotechnology, Clinical Chemistry, Palmer Trevor and Bonner Philip L, Wood Head Publishing, 2007.

Deepa Desmal
Tushar
Bachar

Paul



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-803	FUNDAMENTALS OF CHEMISTRY	L	T	P	C
		Compulsory Foundation Course	4	-	-	4
Pre-requisite		<u>Basic Knowledge in Chemistry</u>	Syllabus		2025	

The main objectives of this course are:

1. To provide basic knowledge of Physical Chemistry and Organic Chemistry
2. Students will become familiar with background knowledge in basic thermodynamics, statistical mechanics, Chemical Kinetics, Quantum Chemistry and Organic Chemistry needed for this course.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

- Comprehend applicability of Physical Chemistry in Bioinformatics.
- Relate the usefulness of knowledge of stereochemistry and weak chemical interactions of organic reactions in Bioinformatics.
- Learn the basic principles of statistical mechanics, which correlate the microscopic properties of systems with the macroscopic observables.
- Appreciate microscopic connection between classical mechanics and thermodynamics.
- Learn theories of chemical kinetics, and catalysis.

Unit-I

Chemical Kinetics

The concept of reaction rates. Effect of temperature, pressure, catalyst and other factors on reaction rates. Order and molecularity of a reaction. Derivation of integrated rate equations for zero and first order reactions. Half-life of a reaction. Concept of activation energy and its calculation from Arrhenius equation. Simultaneous reactions - opposing, parallel and consecutive reactions, the steady state approximation. Theories of Reaction Rates: Collision theory and Activated Complex theory of bimolecular reactions-enthalpy, entropy and free energy of activation, potential energy surfaces, reaction coordinates, kinetic isotope effects, factors determining reaction rates in solution.

Unit-II

Quantum Chemistry

Atomic Structure-Recapitulation of: Bohr's theory and its limitations, dual behaviour of matter and radiation, deBroglie's relation, Heisenberg Uncertainty principle. Basic concepts and postulates of quantum mechanics, Hydrogen atom, Quantization of angular momentum, Many electron atoms, Variation theorem, Perturbation theory, Molecular orbital and valence bond theories, Introductory treatment of semi-empirical and ab initio calculations on molecular systems, Density functional theory.

Unit-III

Organic Chemistry

Basics of Organic Chemistry Organic Compounds: Classification, and Nomenclature, Hybridization, Shapes of molecules, Influence of hybridization on bond properties. Electronic Displacements: Inductive, electromeric, resonance effects, hyperconjugation and their

Deepa Deswal *Ruchit* - *Pav*

applications; Dipole moment; Organic acids and bases; their relative strength. Concepts of aromaticity, antiaromaticity and non-aromaticity.

Weak Chemical Forces: van der Waals forces, ion-dipole forces, dipole-dipole interactions, induced dipole interactions, Instantaneous dipole-induced dipole interactions. Repulsive forces, Hydrogen bonding (theories of hydrogen bonding, valence bond treatment)

Unit-IV

Stereochemistry: Introduction and terminology. Elements of symmetry, Concept of chirality (upto two carbon atoms). Configuration: Geometrical and Optical isomerism; Enantiomerism, Diastereomerism and Meso compounds). Threo and erythro; D and L; cis – trans nomenclature; CIP Rules: R/ S (for upto 2 chiral carbon atoms) and E / Z Nomenclature (for upto two C=C systems). **Topicity** in molecules Homotopic, stereoheterotopic (enantiotopic and diastereotopic) groups and faces – symmetry, substitution and addition criteria. Prochirality nomenclature: Pro-R, Pro-S, Re and Si faces.

Suggested Readings:

1. Stereochemistry of Carbon Compounds, E. J. Eliel, John Wiley, 1997.
2. Organic Chemistry, Pearson Education, P. Y. Bruice, 3rd edition, 2006.
3. Organic Chemistry, Morrison R. N. & Boyd R. N., Dorling Kindersley (India) Pvt. Ltd. (Pearson Education).
4. Stereochemistry Conformation and Mechanism, Kalsi P. S., New Age International, 2005.
5. Advanced Organic Chemistry, 6th Ed., March J., John Wiley & Sons, 2006.
6. Statistical Mechanics, McQuarrie D.A., Viva Books Pvt. Ltd.: New Delhi, 2003.
7. Chemical Kinetics, Laidler K.J., 3rd Ed., Benjamin Cummings, 1987.
8. Physical Chemistry, Castellan, G. W., 4th Ed., Narosa, 2004.
9. Physical Chemistry, Levine, I.N., 6th Ed., Tata McGraw Hill, 2010.
10. Physical Chemistry, Glasstone Samuel S., Affiliated East-West.
11. Molecular Thermodynamics, McQuarrie, D. A. & Simon, J. D., Viva Books Pvt. Ltd., New Delhi, 2004.
12. Stereochemistry of Organic Compounds: Principles and Applications, D. Nasipuri.

Reference Books:

1. Atkin's Physical Chemistry, Atkins P.W. & Paula J. De, 10th Ed., Oxford University Press, 2013.
2. Elements of Statistical Thermodynamics, Nash L.K., 2nd Ed., Addison Wesley, 2006 (Reprint).
3. Organic Chemistry (Volume 1), Finar I. L., Dorling Kindersley (India) Pvt. Ltd. (Pearson Education).
4. Organic Chemistry (Volume 2: Stereochemistry and the Chemistry of Natural Products), Finar I. L., Dorling Kindersley (India) Pvt. Ltd. (Pearson Education).
5. Statistical Mechanics: Principle & Selected Applications, Hill, T.L., Dower Publication, New York, 1987.
6. Physical Chemistry, Engel, T. & Reid P., 3rd Ed., Prentice-Hall, 2012.

Deepa Deswal Teacher
Bechgal

Parul



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-805	ESSENTIALS OF MATHEMATICS & BIostatISTICS	L	T	P	C
		Compulsory Foundation Course	4	-	-	4
Pre-requisite	<u>Basic Knowledge of Mathematics & Statistics</u>		Syllabus		2025	

The main objectives of this course are to:

1. Provide knowledge of probability and statistical concepts.
2. Teach statistical methodologies in analyzing and interpreting various real world data.
3. Formulation of sampling hypothesis and testing.
4. Explain fitting regression and correlation models.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

- Formulate as well as analyze mathematical and statistical problems, precisely define the key terms, and draw clear and reasonable conclusions.
- Apply mathematical and statistical concepts to solve and develop Bioinformatics tools applied in life sciences research.
- Apply Biostatistics in major areas of drug design and discovery for example to evaluate the different hypotheses using ANOVA, t-test, correlation, and regression generated during the exercise of computational technique.

Unit-I

Basic Mathematics: Vectors: Vector algebra, Dot and Cross products. Matrices: Algebra of Matrices, Transpose and inverse, Diagonalization of Matrices and Characteristic roots. Basic Differentiation and Integration Techniques.

Central Tendency and Dispersion: Terminologies: Population, Sample, Variable, Parameter, Primary and Secondary Data; Screening, Representation and Classification of Data; Frequency Distribution, Tabulation and Graphical Representations; Mean, Geometric Mean, Harmonic Mean, Median and Mode; Quartiles and Percentiles; Measures of Dispersion: Range, Variance, Standard Deviation, and Coefficient of Variation.

Unit-II

Correlation & Regression: Methods of studying simple correlation: Scatter Diagram, Karl Pearson's Co-efficient of Correlation, Spearman's Rank Correlation, Multiple and Partial Correlation, Correlation ratio and intra-class correlation; Regression Analysis: Introduction: Meaning and Purpose of Regression. Simple and Multiple Linear Regressions using Least Square Principle, Coefficient of Determination.

Deepa Siswal
behgat
Fudini
Pamf

Unit-III

Probability and Distributions: Sample space, events, Equally likely events. Probability and types; Different Approaches, Independent Events, Addition and multiplication rules, Conditional probability, Rules for Calculating Probabilities.

Unit-IV

Analysis of Variance and Testing Hypothesis: Assumptions for Use of the ANOVA, One-Way ANOVA, F-test, Two-way ANOVA for the Randomized Block Design, Test of Significance, Testing Hypotheses about the Variance, Parametric and non-parametric test, The Goodness-of-Fit Test; Chi Square Test.

Suggested Readings:

1. Bio Statistical Methods, Zar J.H., Prentice Hall International Edition, USA, 1984.
2. An Introduction to Biostatistics, Gurumani N., MJP Publisher, 2nd Edition, 2015.
3. A text book of Vector Calculus, Shanti Narayan, S Chand & company, New Delhi.
4. Introduction to Matrices and linear transformations, D T Finbeiner, CBS publishers, Delhi.
5. Differential Calculus, Prasad G., Pothisala Publication, 2016.
6. Integral Calculus, Prasad G., Pothisala Publication, 2015.

Reference Books:

1. Medical Biostatistics, Indrriyan A, 3rd Edition Star Publisher, 2012.
2. Fundamentals of Biostatistics, Rosner B., Duxbury Press, 2005.
3. Numerical Methods, Balaguruswamy, Tata Megra Hill.
4. Statistical Methods in Bioinformatics: An Introduction, Ewens W.J. and Grant, Springer-Verlag, 2001.
5. Probability and Statistics, Devore, J.L., 5th edition, Thomson Asia, 2002.
6. An Introduction to Biostatistics, Sundar Rao & Richard, PHI -3rd Ed.
7. Theory & Problem of Probability, Lipschutz-Shaum's Outline Series, Tata McGraw Hill.
8. Statistical Methods in Bioinformatics: An Introduction, Ewens W.J. and Grant, Springer-Verlag, 2001.
9. Biostatistics-Basic Concepts and Methodology for the Health Sciences, Daniel & Cross, Wiley.
10. Methods in Biostatistics, Mahajan, Jaypee.

Deepa Deswal
Teacher
Delhi

Pand



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-807	INTRODUCTION TO BIOINFORMATICS	L	T	P	C
	Core Course		4	-	-	4
Pre-requisite	<u>Basic Knowledge in Bioinformatics</u>		Syllabus	2025		

The main objectives of this course are to:

1. Make students understand what bioinformatics is and why it is important.
2. Applications of Bioinformatics, with a focus on the topics that will be covered in the course.
3. How bioinformatics help us to store and organize data.
4. Describe the different types of data that can be extracted from various databases such as, NCBI and EBI resources.
5. Explain how to locate and extract data from key bioinformatics databases and resources.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

- Learn to locate and use the main databases at the NCBI and EBI resources.
- Learn to differentiate between various databases, tools, repositories and be able to use each one to extract specific information.
- Extract data from specific databases using accessions numbers, gene names etc.
- Learn comfortable using selected tools at NCBI and EBI to run simple analyses on genomic sequences, learning phylogeny of sequence and gene annotation.

Unit-I

Introductory Concepts: Bioinformatics as an Emerging Discipline, Applications of Bioinformatics in Various Areas, Overview of Available Bioinformatics Resources on the Web, Protein and Genome; Information Resources and Analysis Tools; Established Techniques and Methods; Sequence File Formats FASTA, GenBank, FASTQ and Structured File Formats.

Biological Databases: Protein Sequence and Structural Databases, Nucleotide Sequence Databases; NCBI, PubMed, Protein Data Bank(PDB), PIR, UniProt, EMBL, GenBank, DDBJ, SRA, UniGene; Specialized Databases: Pfam, SCOP, GO, Metabolic Pathways.

Unit-II

Methods of Sequence Analysis: Pairwise sequence alignment methods; Heuristic Methods; BLAST and its variants, Statistics of Sequence Alignment Score; E-Value, P-Value, Scoring matrix, PAM, BLOSUM and Gap Penalty; Multiple Sequence Alignments; ClustalW, Hidden Markov Models, HMM Based Multiple-Sequence Alignment.

Unit-III

Phylogenetic Analysis: Distance and Character Based Methods and Software, Computing Tools for Phylogenetic Analysis, Distances, GROWTREE, PAUP, PHYLIP and MEGA; Construction and Visualization Phylogenetic Tree; and Application of Phylogenetic Analysis.

Deepa Deswal *Tejas* *Seehgal* *Parekh*

Unit-IV

Protein Structure Analysis Tools: Sequence Motif Databases, Pfam, PROSITE, Protein Structure Classification; SCOP, CATH, Other Relevant Databases, KEGG, Protein Structure Alignments; Structure Superposition, RMSD, Different Structure Alignment Algorithms, DALI, and TM-align.

Suggested Readings:

1. Bioinformatics Sequence and Genome Analysis, David. W. Mount, Cold spring Harbor Lab. NY.USA, 2001.
2. Comparative Genomics, Ann Gibbons (Science), 1998.
3. Genomes 4, T.A Brown, Taylor and Francis Group, 2018.

Reference Books:

1. Understanding Bioinformatics, Jeremy O. Baum, Marketa J. Zvelebil. 2007, Garland Science, USA.
2. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, Andreas D. Baxevanis, B. F. Francis Ouellette, Wiley Publishers, 1998.
3. Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]
4. Bio-Informatics: Algorithms and Applications- SWAYAM
5. <https://nptel.ac.in/courses/102/103/102103044/>

Deepa Deswal ^{Teacher}
Beehgaal Paul



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-809	FUNDAMENTALS OF COMPUTER AND PROGRAMMING	L T P C
		Compulsory Foundation Course	3 - - 3
Pre-requisite		<u>Basic Computer Knowledge</u>	Syllabus 2025

The main objectives of this course are to:

1. Make the students understand the basic aspects of programming.
2. Practice the use of conditional and looping statements.
3. Gain skills to handle strings and files.
4. Explain and be able to use fundamental programming constructs such as sequencing, decisions and iteration.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

- Explain and make use of data types, variables, constants, assignment statements, and arithmetic and Boolean expressions in writing programs.
- Understand object-oriented programming and are able to use arrays and structures.
- Explain and make use of data types, variables, constants, assignment statements, and arithmetic and Boolean expressions in writing programs.

Unit-I

Overview of Computing system: Overview and organization of a computer system, storage, devices, memory, etc., Types of Processing: Batch, Real-Time, Online, Offline, Types of modern computers: The workstation, The Minicomputer, Mainframe Computers, Parallel Processing Computer, The Super Computer, etc.

Unit-II

Basic Programming, Operating system and softwares: Software concepts: Concepts of flowcharting, Algorithm development, Relationship between hardware and software, Types of software: System software and Application software. Introduction to Operating Systems, Process management, Memory management, File management, Device management and Security. Introduction to Windows/Unix/Linux, C++ and Perl & Bioperl.

Unit-III

Computer Networking: OSI Reference Model, topologies and protocols, designing networks, Networking gadgets (Router, Switch, etc.); Data Communication (ISDN, VPN, DSL, cable modem, cellular modem, etc.); Communication Links (Wire pairs, Coaxial cables, Fiber optics, Microwave, Satellite, etc.).

Unit-IV

Internet: The Internet and its Resources, Internet protocols, services, and related terminologies. Web browsers, customizing browsers, Blocking popup windows, Internet programming languages.

Deepa Deswal . *Kuldeep* - *Parul*
Sevag

Gates and Circuits: Computers and Electricity; Logic Gates – AND, OR, NOT, XOR, NAND and NOR Gates. Gate Processing; Gates with More Inputs; Constructing Gates; Transistors; Circuits – Combinatorial Circuits: Adders and Multiplexers, Circuit as Memory; Integrated Circuits; CPU Chips.

Suggested Readings:

1. C++ How to Program: Introducing the New C++, 14 Standard, P.J. Deitel & H.M. Deitel, 10th Edition, Pearson Education, 2016.
2. Computer Science Illuminated, Dale & Lewis, Narosa Publishing House, Keddall & Systems Analysis and Design, Kendall, Prentice Hall India, ITL Esl: Introduction to Computer Science, Pearson Education.
3. Beginning Perl for Bioinformatics, James Tisdall, O'Reilly Media, Inc., 2001.
4. Mastering Perl for Bioinformatics: Perl Programming for Bioinformatics, James D Tisdall, O'Reilly, 2003.
5. Networking for Beginners, Russell Scott, 2019.
6. Logic Gates, Circuits, Processes, Compilers & Computers, Jan Friso Groote, Rolf Morel, Julien Schmaltz & Adam Watkins, 2021.
7. Digital Circuits & Logic Design, Lee S.C., 1980
8. Digital Design, M. M. Mano, 2001

Reference Books:

1. Object Oriented Programming with C++, E Balagurusamy, 8th Edition, The McGraw Hill companies, 2020.
2. The Complete Reference C++, Herbert Schildt, 4th Edition, McGraw Hill Education, 2017.

Deepa Desmal Tulsi. Parul
Deekha



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-851	BIOINFORMATICS TOOLS (LAB-I)	L	T	P	C
		Compulsory Foundation Course	-	-	6	3
Pre-requisite		Basic Computer Knowledge	Syllabus 2025			

The main objectives of this course are to:

1. Make the students understand the basic aspects and applications of Bioinformatics.
2. Know the computational methods for Sequence Alignment and the related scoring algorithms.
3. Understand the steps in Evolutionary analysis.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

- Describe the file formats, nomenclature of macromolecules and have an idea of Databases.
- Interpret the algorithms, scoring functions involved in the sequence alignment.
- Evaluate the phylogenetic relationship of an organism and identify genes, repeats, domains from sequences using bioinformatics tools.
- Model 2D and 3D structure of a target from the sequence.
- Exploit the information in the sequences for further research.

Experiments:

1. Retrieving, viewing and printing of the specific protein sequence (by accession no. or name) using a public database site.
2. Exploring the NCBI, ExPASy, www.ebi.ac.uk/Tools etc. websites for information and tools available there.
3. Pairwise alignment of Protein and DNA sequences & data interpretation.
4. Local and global alignment of sequence data and comparing both results.
5. Retrieving DNA and/or protein sequences of a given item (by name or accession number) from GENBANK. Performing a sequence similarity search using the BLAST.
6. Retrieving this protein sequence of a given organism and downloading the structure of this protein from existing database. Short-listing protein sequences of highest similarity from the list of BLAST search result and doing a multiple sequence alignment (Using CLUSTALW). Finding out the regions of exact/good match in the protein sequences of these sequences.
7. Aligning nucleotide sequences; designing a degenerate primer of 20 bases from nucleotide alignment data, and calculate the level of degeneracy of this primer.
8. Learning about the Phylip/MEGA program and its uses for the construction of phylogenetic trees.
9. Searching and downloading protein structure data using Entrez. Viewing the structure using public domain software.
10. Protein structures: Visualizing and analysis of inter atomic distances, H-bond calculations, secondary structure analysis and salt bridge analysis of protein structures using different software. Prediction of 3D structure of protein.

Deepa Deswal
D. Deswal
T. Deswal
S. Deswal

Few Specific Bioinformatics tools/webservers:

Pymol, VMD, GROMACS, SWISS MODEL, Expasy web server, SAVES server, ERRAT, PROSA, PROSA 2, VERIFY 3D, AMOEBA, ModLoop for Loop Modeling, Rampage for Ramachandran plot assessment, WHAT CHECK, PROCHECK, I-TASSER, CYTOSCAPE, DAVID, STRING database, PATHOGEN BOX, PANDEMIC BOX databases, AUTODOCK, QSAR-Co, QSARToolBox.

Reference web Links for Practical Sessions:

<https://pymol.org/2/>

<https://blast.ncbi.nlm.nih.gov/Blast.cgi>

<https://swissmodel.expasy.org/>

<https://www.gromacs.org/>

<https://www.ks.uiuc.edu/Research/vmd/>

<https://www.expasy.org/>

<https://saves.mbi.ucla.edu/>

<https://www.doe-mbi.ucla.edu/errata/>

<https://prosa.services.came.sbg.ac.at/prosa.php>

<https://www.doe-mbi.ucla.edu/verify3d/>

<https://zlab.umassmed.edu/bu/rama/>

<https://modbase.compbio.ucsf.edu/modloop/>

<https://cytoscape.org/>

<https://david.ncifcrf.gov/>

<https://string-db.org/>

<https://swift.cmbi.umcn.nl/gv/whatcheck/>

<https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/>

<https://zhanggroup.org/I-TASSER/>

<https://autodock.scripps.edu/>

<https://qsartoolbox.org/>

Deepa Deswal *Deepa* . *Deswal* . *Deepa*



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-853	C++ & LINUX (LAB-II)	L	T	P	C
	Core Course		-	-	6	3
Pre-requisite	<u>Basic Computer Operating Skills</u>		Syllabus 2025			

The main objectives of this course are to:

1. To make the students understand the basic aspects of programming.
2. Develop an in-depth understanding of functional, logic, and object-oriented programming paradigms.
3. Explain and be able to use fundamental programming constructs such as sequencing, decisions and iteration.
4. To enhance problem solving and programming skills in C & C++.
5. To make them familiar with Linux environment.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

- Have a good understanding about the concept of C & C++.
- Explain and make use of data types, variables, constants, assignment statements, and arithmetic and Boolean expressions in writing programs;
- Understand object-oriented programming using C++
- Explain and are able to use arrays and structures;
- Write and read basic codes in C & C++.
- Use Linux and its basic commands.

Experiments:

1. Basic Linux commands
2. Basic C++ I/O functions/operators
3. Conditional statements: if-else, and switch
4. Looping and unconditional branching statements
5. Arrays and matrix operations
6. Defining classes and creating objects
7. Defining constructors, destructions, Array of objects
8. Inheritance & Virtual Functions
9. Template – Implementing genetic searching and sorting algorithms
10. File and file operations

Suggested Readings:

1. Programming in ANSI C, E. Balagurusamy, Tata McGrawHill Publishing Company Limited, 2007.
2. Object Oriented Programming using C++, Lafore R., Galgotia Publishers, 2006.

Reference Books:

1. Teach Yourself C++ in 24 hours (5th edition), Jesse Liberty and Rogers Candanhead, Pearson Education Inc., 2012.
2. Head First C, David Griffiths and Dawn Griffiths, O'Reilly, 2013.

Deepa Deswal Teacher
D. Deswal



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-802	DATABASE MANAGEMENT SYSTEMS & DATA MINING	L	T	P	C
		Skill Enhancement Course	3	-	-	3
Pre-requisite	Math Foundation and Data Structures		Syllabus		2025	

The main objectives of this course are to:

1. Present an introduction to database management systems.
2. An emphasis on how to organize, maintain and retrieve - efficiently, and effectively.
3. Use and implementation/design.
4. Experience with SQL.
5. Increased proficiency with the programming language.

Expected Course Outcomes:

Upon successful completion of this course, students should be able to:

- Describe the fundamental elements of relational database management systems.
- Explain the basic concepts of relational data model, entity-relationship model, relational database design, relational algebra and SQL.
- Design ER-models to represent simple database application scenarios.
- Convert the ER-model to relational tables, populate relational database and formulate SQL queries on data.
- Improve the database design by normalization.
- Gain familiarity with basic database storage structures and access techniques: file and page organizations, indexing methods.

Unit-I

Basic Concepts & ER Diagram: Data, Database and Database Management System; Database vs Traditional File Systems; Three-Tier Architecture of DBMS, Data Independence; Categories of DBMS: Hierarchical, Network and Relational Database, Data and Database Models, Categories of Database Models, Entity relationship (ER) Model: Basic Concepts & their Representations, Entity, Entity Type, Entity Set, Attributes and Keys, Relationships, Relationship Types, Structural Constraints, Weak Entity, Naming Conventions, Design Issues in ER Model. ER Diagrams.

Relational Database Model: Structure of Relational Model; Domains, Attributes, Tuples, and Relations; Characteristics of Relations; Relational Constraints – Domain Constraints, Key Constraints, Entity Integrity, and Referential Integrity Constraints; Relational Database Schema.

Unit-II

Functional Dependencies and Normalization: Informal Design Guidelines for Relation Schemas; Functional Dependencies; Inference Rules for Functional Dependencies; Normalization using Functional Dependencies – First Normal Form (1NF), Second Normal Form (2NF), Third Normal Form (3NF), and Boyce-Codd Normal Form (BCNF).

Unit-III

Relational Algebra and SQL: Relational Algebra Operations – Select, Project, Rename, Union, Intersection, Set Difference, Join, and Division Operations; Aggregate Functions and Groupings. SQL Concepts, Schema and Table Deletion; Table Modification; Insert, Delete, and Update Statements; SELECT-FROM-WHERE Structure; Renaming Attributes; Nested Queries and Set Comparisons; Aggregate Functions.

Unit-IV

Advanced Database Concepts: Emerging database technologies and applications: Spatial databases, Mobile databases, Multimedia databases, Genome data management, Overview of Data Warehouse, OLAP, Big data, biological big data, and big data analytics.

Suggested Readings:

1. Introduction to Database Systems (Vol. I & II), Date C.J., 8th Edition. Addison Wesley, 2004.
2. Principles of Database and knowledge base Systems (Vol. I & II), Ullman J.D., Computer Science Press New York, 1989.
3. Database Design, Gio Wiederhold, McGraw Hill, 1997.
4. Fundamentals of Database Systems, Elmasri R. and Navathe S.B., Fifth Edition, Pearson, 2007.
5. Database Systems- Concepts, Designs and Application. Singh S.K., 2nd Edition. Pearson, 2011.
6. Database System Concepts, Silberschatz A. Korth H. F. Sudarshan S., Sixth Edition. McGraw-Hill, 2010.
7. An Introduction to Database Systems, Date K., Swamynathan S., Eighth Edition, Pearson, 2012.
8. Big Data Analytics in Bioinformatics & Healthcare, Baoying Wong, Ruowang Li and William Perrizo, 1st edition, 2014.
9. Multimedia Database Management Systems, Guojun Lu, 1999.
10. Applied Optimization for Wireless, Machine Learning, Big Data, Prof. Aditya K. Jagannatham, IIT Kanpur, SWAYAM.

Reference Books:

1. Database System Concepts, Silberschatz Korth, 4th ed., McGraw Hill, 2006.
2. Database Management Systems (3rd Ed.), Raghu Ramakrishnan and Johannes Gehrke, McGraw Hill, 2003.
3. Database System-Design, Implementation and Management (7th Ed.), Peter Rob and Carlos Coronel, Cengage Learning, 2007.
4. Genomes, Browsers and Databases: Data mining tools in integral Genomic Databases, Peter Schattner, 2008.

Deepa Desai Tutor.
Parul Bhangal.



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY

SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-804	GENOMICS AND PROTEOMICS	L	T	P	C
Pre-requisite	Core Course		4	-	-	4
	<u>Concepts of Genome and Proteome</u>	Syllabus			2025	

The main objectives of this course are to:

1. Understand the genome architecture with gene function and regulation.
2. Provide students with the skills of genomic data analysis.
3. Get clear idea about the Computational transcriptomics and epigenomics.
4. Use different techniques for generating biological data like genome sequences, gene expression data & proteomics data.
5. Have knowledge of Proteomics and advance clinical Proteomics.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

- Study the variation in DNA and genome helps in understanding biological processes.
- Explain genomic technologies and the ways in which genomic data are stored.
- Have hands-on experience on various bioinformatics tools available for analyzing genes and genomes.
- Handle big biological data generated by the sequencing projects and its analysis.
- Understand clinical transcriptomics and epigenomics evaluation.

Unit-I

Introduction to Genomics: Genomics and its diversifications, structural organization of prokaryotic and eukaryotic genomes, C value paradox, types and significance of repeats in the genome, Organelle genomes; Evolution of genes and genomes; Epigenomics and gene silencing; Application of Genomics.

Unit-II

Sequencing Techniques and Comparative genomics: DNA sequencing techniques, Structural features of Genome sequences, Tools for the prediction of genes and regulatory elements; Comparative genomics, Comparing by gene content, mobile elements and noncoding elements; Minimal gene set, Signature genomes COG, Gene Ontology. NGS concepts and applications; DNA alignment algorithms and assessment; Data QC and Filtering

Unit-III

Transcriptomics: Transcriptome and analysis: ESTs, Digital Northern, DEG's, SAGE, Microarrays, qRT PCR, RNA-seq and algorithm, Online sequence sampling resources, Transcriptome data mining, heatmaps Global profiling of gene expression, Applications of expression profiling in human disease, Metatranscriptomics.

Deepa Deswal, Teaching Assistant, and Dr. [Signature]

Unit-IV

Proteomics: Web and internet sites for comparing and identifying protein, comparing physical properties of proteins, motifs and domains; Techniques: 2D gel electrophoresis; Mass spectrometry and analysis, LC/MS-MS; Quantitative proteomics - isotope-coded affinity tag (ICAT), isobaric tagging for relative and absolute quantitation (iTRAQ), Structural proteomics

Interaction Proteomics: Interactomes and Proteomic interactions, Bioinformatics tools and databases for studying PPI, protein microarrays, Nucleic Acid Programmable Protein Array (NAPPA), Label-free nanotechnologies in proteomics, Surface Plasmon Resonance (SPR); Understanding post-translational modifications; Proteome- wide interaction maps

Suggested Readings:

1. Genome, Brown T.A., John Wiley Press, US, 2002.
2. Discovering Genomics, Proteomics and Bioinformatics, Campbell A.M. & Heyer L. J., Benjamin/Cummings, 2002.
3. Next-Generation DNA Sequencing Informatics, Stuart M. Brown, Second Edition, New York University School of Medicine (ISBN-13: 978-1621921236).
4. Next Generation Sequencing Data Analysis, Xinkun Wang, CRC Press, (ISBN13: 9781482217889).
5. Biotechnology and Genomics, Gupta P.K., First Edition, Rastogi Publications, Meerut, 2004.
6. Plant Biotechnology, Gupta P.K., Rastogi Publications, Meerut, 2009.
7. Handbook of Plant Genome Mapping: Genetic and Physical, Meksem K. and Gahl K., 233-256, Wiley-VCH, 2005.
8. Advanced Genetics. Miglani G.S., Narosa Publishing House, New Delhi, 2007.
9. Principles of Gene Manipulation and Genomics, Primrose S.B. and Twyman R.M., Blackwell Publishing, Australia, 2006.
10. Biotechnology: Expanding Horizons, Singh B.D., Second Edition, Kalyani Publishers, Ludhiana, 2009.
11. Plant Biotechnology, Singh B.D., Kalyani Publishers, Ludhiana, 2009.
12. Functional Proteomics: Methods and Protocols, Thompson J.D., Schaeffer-Reiss C., and Ueffing M., Humana Press, New York, 2008.
13. Principles of Proteomics, Twyman R.M., Taylor & Francis, 2004.

Reference Books:

1. Principles of Genome Analysis & Genomics, Primrose and Twyman, Blackwell, 2003.
2. An Introduction to Molecular Human Genetics, Pasternak, Fitzgerald, 2000.
3. Genes to Genomes, Dale and Schartz, Humana, 2003.
4. Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]
5. Big Data Computing, Prof. Rajiv Misra | IIT Patna | SWAYAM
6. <https://nptel.ac.in/courses/102/104/102104056/>
7. <https://nptel.ac.in/courses/102/103/102103017/>

Deepa Deswal Kuchel.
Rahul Bsehgal.



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-806	PROGRAMMING IN PERL	L	T	P	C
		Skill Enhancement Course	3	-	-	3
Pre-requisite		Perl Software, Notepad, NCBI-Blast	Syllabus		2025	

The main objectives of this course are to:

1. Familiarize the students with programming skills in Perl
2. Make them learn to read and write perl scripts for bioinformatics applications

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

- Write perl scripts for sequence manipulation & accessing data from various databases.
- Parse the results of various molecular biology programs including Blast, clustalw, TCOffee, Genscan etc.
- Write and execute a script in Perl.
- Enable routine and module calls and their implementation using Bioperl.

Unit-I

Introduction: History of Perl, Availability, Support, Versions, Installation. Basic Concepts, Significance of Perl in Bioinformatics.

Unit-II

Basics Constructs: Scalar Data, Numbers, Strings, Scalar Operators, Scalar Variables, Scalar Operators and Functions. Arrays and List Data: Literal Representation, Variables, Array Operators and Functions, Scalar and List Context.

Hashes: Hash Variables, Literal Representation of a Hash, Hash Functions, Hash Slices. Control Statement blocks, Loops and Conditions.

Unit-III

Advance Constructs and Features: Regular Expressions: Concepts of Regular Expressions, Simple Usage of Regular Expressions, Patterns, Matching Operators, Substitutions, Split and Join functions, Subroutines: System and User Functions, Local Operator, Length, Parameter Lists, Lexical Variables, File Handles and File Tests: Opening and Closing a File handle, Using Pathnames and Filenames, Die, Using File handles.

Object-Oriented Perl: Introduction to Modules, Creating Objects and References.

CGI Programming: The CGI.pm Module, CGI Program in Context, Simple CGI Programs, Passing Parameters via CGI, Perl and the Web.

Unit-IV

Application of Perl in Bioinformatics: To align the sequence using Local Blast, Write a script to search for genes from Genscan, Protein Sequence Generation, To count start and stop codons in a sequence, To calculating the reverse complement of DNA Sequence, To concatenating DNA.

Fragments Transcription: DNA to RNA, Program to simulate DNA Mutation, Write a code to check file extensions using Perl, Write a code to search the given string in an entered sequence, to compute total and average length of the proteins in these files, Reading a protein sequence from a file, Write a sequence in a file.

Suggested Reading:

1. Developing Bioinformatics Computer Skills, Cynthia Gibas & Per Jambeck, O'Reilly & Associates, 2000.
2. Genomic Perl, Rex A. Dawyer, Cambridge University Press.
3. Learning Perl, 3rd Edition, Randal L. Schwartz and Tom Phoenix, O'Reilly.
4. Genomic Perl, Rex A. Dwyer, Cambridge University Press, 2003 (ISBN: 978-1449303587).
5. <https://www.perltutorial.org/>

Reference Books:

1. Beginning Perl for Bioinformatics, James Tisdall, O'Reilly & Associates, 2001.
2. Mastering Perl for Bioinformatics, James Tisdall, O'Reilly, 2003.
3. Online tutorials Sakai web site

Seepa Saxena

Tejinder
Bachgal

Paul



The main objectives of this course are to:

1. Provide concepts of bond angle, bond stretching, bond distance and different types of protein-protein & protein ligand interactions.
2. Provide a theoretical background to the various methods of molecular modelling, mechanics.
3. Develop and understand the mechanism of drug design using computational methods.
4. Provide practical experience in the analysis of protein sequences and structures.
5. Understand the use of informatics in drug design and development, finding new targets to treat disease.
6. Gain insights on protein-ligand docking and knowledge-based scoring functions.

Expected Course Outcomes:

On successful completion of the course, the student will be able to:

- Perform protein structure prediction and understand energy minimization methods and its importance in drug action.
- Gain insight on the molecular dynamics simulation and Monte Carlo simulation methods.
- Understand the concept of Pharmacophore and QSAR.
- Apply the knowledge gained to find new targets and design drug to treat disease conditions.

Unit-I

Fundamentals of protein structure: Amino acid chemistry and the peptide bond; Protein folding; Primary, secondary, tertiary and quaternary structure; Enzymes, receptors and transport proteins; Experimental methods for structure determination; Structural quality assurance; Motion and disorder in proteins. Molecular structure databases and visualization

Structure, function and Evolution: THE PDB and mmCIF formats; Structure classification databases: SCOP and CATH; Structure comparison and alignment; Structure and functional assignment; Secondary structure assignment; identifying structural domains in proteins; Prediction of protein-protein interactions; protein design and engineering

Unit-II

Proteins as drug targets: attributes of drug targets; candidate gene prioritization; experimental validation; practical aspects and case studies; Structural bioinformatics in drug discovery; Structure prediction: Homology modelling; fold recognition methods, ab initio methods, Casp and CAFASP experiments and their findings

Unit-III

Concepts in Molecular modeling: Introduction – coordinate systems – potential energy surfaces – introduction to quantum mechanics – postulates – Schrodinger wave equation – hydrogen molecule– Born-Oppenheimer approximation, Molecular mechanics and energy minimization: Empirical force field models – Bond stretching – angle bending – torsional term –nonbonding interactions – thermodynamics properties using a force field – derived and non-derived energy minimization method – simplex – sequential univariate method – steepest descent method – conjugate gradient method- Newton-Rapson method, Molecular Dynamics and Monte Carlo simulation : Introduction – Using single Model – time steps – Multiple steps – Setting up MD – energy conservation in MD Simulation Examples – Monte Carlo – Random number generation – Difference in MD & MC.

Unit-IV

Principles and methods of docking and ligand design: Ligand and structure based design, The drug discovery pipeline; Library design; Binding site prediction; virtual screening; docking and scoring; structure based and de novo drug design; molecular affinities and assemblies; design against protein-protein interactions.

Quantitative Structure Activity Relationship (QSAR): SAR versus QSAR, QSAR methodology, Quantum-chemical descriptors, COMFA and COMSIA. Molecular mechanisms of diseases and drug action. Pharmacokinetics: Models and Applications, Drug–receptor interactions. Pharmacodynamics, Biomolecular interactions and binding thermodynamics, Biochemical and cell-based assays, Molecular recognition and specificity, Future developments.

Suggested Readings:

1. Molecular Modelling: Principles and applications (2nd Edition), Leech Andrew, Prentice Hall, 2001.
2. Molecular Modeling for Beginners, Hinchliffe Alan, Wiley, 2003.
3. Computational Chemistry and Molecular Modeling-Principles and Applications, Ramachandran, Deepa and Namboori, Springer Verlag, 2008.

Reference Books:

1. Molecular Modelling for Beginners, (2nd Edition), Alan Hinchliffe., John Wiley & Sons Ltd., 2008.
2. Molecular Modeling and Simulation – An Interdisciplinary Guide, Tamar Schlick, Springer Verlag 2000.
3. Computational Medicinal Chemistry for Drug Discovery, Patrick Bultinck, Marcel Dekker Inc., 2000.
4. Drug Design: Structure- and Ligand-Based Approaches, Kenneth M. Merz, Jr. Cambridge University Press.

Deepa Deswal

Parul

Tushar
Beergal



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-852	MySql/ORACLE (LAB-III)	L	T	P	C
		Skill Enhancement Course	-	-	6	3
Pre-requisite		Basic Knowledge of DEMS	Syllabus		2025	

The main objectives of this course are to:

1. Enhance the knowledge and understanding of Database analysis and design.
2. Enhance the knowledge of the processes of Database Development and Administration using SQL.
3. Enhance Programming and Software Engineering skills and techniques using SQL.

Expected Course Outcomes:

On successful completion of the course, the student will be able to:

- Develop Oracle Database Objects
- Develop Advanced package concepts
- Produce readable output using SQL*Plus
- Use Oracle supplied packages
- Create database triggers
- Manage subprograms and triggers and object dependencies
- Manipulate and create Oracle large objects.

Experiments:

1. SQL statements to create, update, and delete databases and tables.
2. SQL statements to insert, update, and delete records from tables.
3. SQL statements to create, update, and delete views.
4. Simple SQL queries to retrieve information from a database.
5. Nested SQL queries to handle complex information retrieval requirements.
6. SQL queries using aggregate functions like count, average, sum, etc.
7. PL/SQL blocks using basic data types and operators.
8. PL/SQL blocks using branching and looping constructs.
9. Database triggers using PL/SQL.
10. Database functions/procedures using PL/SQL.

Suggested Reading:

1. Database system concepts. 5th Ed., Abraham Silberschatz, Henry F. Korth and S. Sudhashan, McGraw Hill Publications, 2005..
2. Fundamentals of Database Systems, ElmasriRamez and NovatheShamkant, Benjamin Cummings Publishing Company, 2007 (ISBN-10: 0321369572).
3. Database Management System, 3rd Ed., P. Ramakrishnan Rao, McGraw Hill Publications, 2003 (978007123057).
4. Beginning MySQL, Robert Sheldon, Wiley Publishing
5. jQuery: <http://www.w3schools.com/jquery/>
6. HTML: <http://www.w3schools.com/html/>
7. CSS: <http://www.w3schools.com/css/>
8. PHP: <http://www.w3schools.com/php/>

Reference Books:

1. Database Management system, V.K. Jain, Dreamtech Press, 2002 (ISBN 8177222279).
2. Introduction to database management, Vol. 1, Vol. 2, Vol. 3, Date C.J., Addison Wesley, 2009.
3. Principles of Database systems, Ullman, JD, Galgottia Publication, 1992.
4. Principles of Database Management systems, James Martin, PHI, 1985.

Deepa Saran Student
Sachin Parv



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-854	SYSTEM BIOLOGY AND ADVANCE BIOINFORMATICS TOOLS (LAB-IV)	L	T	P	C
		Ability Enhancement Compulsory Course	-	-	6	3
Pre-requisite		<u>Knowledge of Basic Informatics Tools</u>			Syllabus	2025

The main objectives of this course are to:

1. Understand the general principles of gene organization and expression.
2. Explore the various levels of gene regulation and protein function.
3. Analyse the various genetic and molecular changes occur in a normal cell.

Expected Course Outcomes:

On successful completion of the course, the student will be able to:

- Represent and illustrate the structural organization of genes and the control of gene expression.
- Explore the prokaryotic and eukaryotic protein synthesis mechanism.
- Conceptualize mechanisms of signal transduction, cell cycle and cell death.
- Link the concepts of cell and molecular biology to a better understanding of diseases, including cancer.

Experiments:

1. Visualization of biological networks (graph) using Cytoscape/other tools.
2. Topological analysis of biological networks using Cytoscape/other tools.
3. Statistical analysis of high-throughput gene expression data.
4. Implementation of gene regulatory network reconstruction algorithm.
5. Implementation of modelling methods for metabolic network.
6. Working with Cell Designer or other biological network editing tools.
7. GO enrichment and DAVID analysis.
8. Searching of adjacent and patterned repeats of nucleotide sequences using Tandem Repeat Finder or Repeat Masker.
9. Open Reading Frames using ORF Finder. Finding probable restriction sites for different enzymes using Neb Cutter tool.
10. RNA structure prediction using M-Fold, RNA-Fold, RADAR and P-Fold server.

Suggested Reading:

1. A software environment for integrated models of biomolecular interaction networks
Genome Research, Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape, 13(11):2498-504, Nov. 2003.

Reference Web links:

<https://cytoscape.org/>

<https://www.celldesigner.org/>

<https://tandem.bu.edu/trf/trf.html>

<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-856	PERL (LAB-V)	L	T	P	C
	Skill Enhancement Course		-	-	4	2
Pre-requisite	<u>Perl Software, Notepad, NCBI-Blast</u>	Syllabus	2025			

The main objectives of this course are to:

1. Make the students familiar with Perl programming environment.
2. Make them familiar with the installation of various systems components required for Perl
3. Able to formulate stepwise implementation of a Perl script for a given problem in Bioinformatics.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

1. Align sequence in local BLAST.
2. Write a script to search for genes from GENSCAN.
3. Generate protein sequence.
4. Write a program to count start and stop codons in a sequence.

Experiments:

- Align the Sequence using Local Blast.
- Gene from Genscan.
- Protein Sequence Generation.
- To Count Start and Stop Codons in a Sequence.
- Reverse Complement of DNA.
- DNA Fragments Transcription.
- DNA Mutation.
- To Check File Extensions.
- To Search String in Sequence.
- Total and Average length of Proteins in files.
- Reading Protein Sequence from files.
- Sequence into a file.

Suggested Readings:

1. Beginning Perl for Bioinformatics, James Tisdall, O'Reilly Media, 2001.
2. Mastering Perl for Bioinformatics, James Tisdall, O'Reilly Media, 2003.

Reference Web Links:

<https://blast.ncbi.nlm.nih.gov/Blast.cgi>
<https://bio.tools/genscan>
<https://www.expasy.org/>
<https://www.uniprot.org/>
<https://www.bioinformatics.org/sms2/>

Deepa Deswal Tuli. Paul.



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-901	PROGRAMMING WITH PYTHON AND BIOPYTHON	L	T	P	C
		Core Course	3	-	-	3
Pre-requisite		<u>Basic Knowledge of Bioinformatics</u>	Syllabus		2025	

The main objectives of this course are to:

1. Introduce the effectiveness of scripting language concepts with regard to biological applications to solve computational biology/Bioinformatics problems.
2. Acquire programming skills in core Python.
3. Acquire Object Oriented Skills in Python.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

- Read a script and understand its pros and cons on scripting languages vs classical programming languages.
- Write and execute a script in Python.
- Devise stepwise execution of a Python for a given problem in Bioinformatics.

Unit-I

An Introduction to Python Programming: Working with Python, An interpreter for python, Relational operators, Logical operators, Bitwise operators, Variables and assignment Statements, Keywords, Script mode.

Basic Concepts: Control structures, if-else conditional statement, Looping statements, Nested loops, break, continue and pass, Debugging, Scope of variables, Strings, String manipulations, Regular Expressions, Built-in Functions, I/O Functions, Function Definition and Call, Importing user defined modules, Command-line arguments, Mutable and Immutable objects, Recursion.

Unit-II

Advanced Concepts: Lists, accessing lists, working with lists, operations, related functions and methods, Tuples, Accessing tuples, working with tuples, Operations, related Functions and Methods, Dictionary, Working with dictionary, Accessing values in dictionaries, working with dictionaries, Operations, related Functions and Methods. Files and Exceptions: File Handling, Writing structures to a file, Errors and Exceptions, Handling exceptions using try-except, File processing examples.

Unit-III

OOP concepts: OOPs concepts, Classes and objects, Constructor, Destructor, Attributes, Encapsulation, Data Hiding and Data Abstraction, Inheritance, Polymorphism, Overloading, overriding, Inbuilt Object-Oriented functions and modules. Managing Databases using SQL.

Dr. Anil Kumar

Unit-IV

Biopython: Introduction to Biopython, Installation, Inbuilt modules related to sequence objects, sequence annotation objects, sequence analysis, sequence input/output, sequence alignment objects and tools, Applications of Biopython. Overview of Scikit module.

Suggested Readings:

1. Python Programming: A Modular Approach, Taneja & Kumar, Pearson India, 2017.
2. Fundamental of Python, Kenneth & Lambert Cengage Learning India Pvt. Ltd., 2019.
3. Biopython Tutorial and Cookbook (ebook), Jeff Chang, et al., 2021,
<http://biopython.org/DIST/docs/tutorial/Tutorial.html/>

Reference Books:

1. Python for Bioinformatics, Sebastian Bassi, Chapman and Hall, 2010.

Deepa Deswal *Kuchi!* *Pareel*
Ischgal



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-903	CHEMOINFORMATICS AND CHEMOGENOMICS	L	T	P	C
		Core Course	3	-	-	3
Pre-requisite	<u>Basic background in Chemistry, Life Science</u> <u>Chemistry Syllabus</u>		Syllabus	2025		

The main objectives of this course are to:

1. Introduce the basic concepts of using chemical structure databases.
2. Apply the concepts and learn the use of Cheminformatics tools.
3. Understand the applications of Cheminformatics in drug design.

Expected Course Outcomes:

Upon successful completion of this course, students should be able to:

- Gain skills to analyse the properties of small molecules.
- Design the biological targets and properties of the small molecule under investigation.
- Better understanding of the drug discovery and development process.
- Apply the concepts to create novel leads.

Unit-I

Chemoinformatics: Basic concepts of Cheminformatics, Application of and modern prospective of Cheminformatics. Manipulations in 2D and 3D structures of chemical compounds, representation of chemical reactions, molecular descriptors, calculations of physical and chemical data, calculation of structural deciphers, In silico representation of chemical information (SMILES, IchI, SDF, MOL, PDB, PDBQT, etc), Computer-oriented chemical compounds nomenclature, Computer-Assisted structure elucidation, Molecular drawing and interactive visualization, Chemical name and formula searching.

Unit-II

Molecular Structures: CheminformaticsDatabases, Chemical Database Design, Bio Catalysts Database, The MOS Database, The Failed Reaction Database, Protecting Groups Database, Solid-Phase Synthesis Database, Compound Databases (ZINC, PubChem, ChEMBL, DrugBank, etc.). Chemistry & Information technology, chemical data collation, retrieval, analysis & interpretation, hypothesis generation & validation, development of structure activity/property relationships, artificial intelligence techniques in chemistry, Building molecules on a computer, quantum and molecular mechanics methods for geometry optimization, simulation methods for molecules and material

Unit-III

Personalized Medicine: Computational Chemistry in drug design, Role of Pharmacogenomics in Drug Discovery & Therapeutics. Pharmacogenetics; Biomarkers, Promise of Personalized

Medicine. Genetic drug response profiles, Effect of drug on gene expression, Pharmacogenetics, Drug Metabolism. Pharmacogenomics of drug transporters-OATP, OAT, OCT, OCTN, PepT, MRP, MDR, ABC transporter, Interethnic Differences in drug response, Drug metabolism.

Unit-IV

Combinatorial Chemistry: Combinatorial chemistry technologies & libraries, Solution phase synthesis, High-Throughput Synthesis and Screening, Target-oriented synthesis: Designing organic synthesis, Retrosynthetic analysis, disconnection approach, linear and convergent synthesis. Diversity-oriented synthesis: concept of forward-synthetic analysis, appendage diversity, skeletal diversity, stereo-chemical diversity, complexity and diversity. Asymmetric Synthesis: Use of chiral auxiliaries, chiron approach. Principles and use of enzymes in the synthesis of industrially important sugar / fatty acid esters, sugar nucleotide derivatives.

Drug Design and Discovery: Strategies for drug development. Chemical & physiochemical parameters in drug design, Design of enzyme inhibitors, Target selection and validation, Structure based drug designing, Computational techniques in drug discovery: Molecular modelling tools, structural homology modelling tools, docking tools and screening tools. Computational assessment of activity and toxicity and drugability.

Suggested Readings:

1. Chemoinformatics: A Textbook, Johann Gasteiger, Thomas Engel Publisher: Wiley publication, 2004 (ISBN: 978-35-273-0681-7).
2. An Introduction to Chemoinformatics, Leach AR and Gillet VJ, Springer, 2010 (ISBN – 10 8184892551).

Reference Books:

1. Chemoinformatics: Advanced Control and Computational Techniques, Hossein G. Gilani, Katia G. Samper, Reza KhodaparastHaghi, Apple Academic Press, First edition, 2012.
2. Chemoinformatics for Drug Discovery, Bajorath J, John Wiley & Sons, 2010 (ISBN-978-1-118-13910-3).
3. Chemoinformatics: Advanced Control and Computational Techniques, Gilani HG, Samper KG and Haghi RK, CRC Press, 2012 (ISBN-9781466559332).

Deepa Deswal Truchi
 Ischgal Parul



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-905	COMPARATIVE AND FUNCTIONAL GENOMICS	L	T	P	C
		Discipline Centric Elective	3	-	-	3
Pre-requisite		<u>Basic Bioinformatics Knowledge</u>	Syllabus	2025		

The main objectives of this course are to:

1. Provide an insight into the complete genome sequences of a few organisms as well as the Human genome through Comparative and Functional genomics.
2. Acquaint knowledge on functional genomics techniques such as microarrays, EST, SAGE and interpret data obtained through high throughput expression studies.
3. Develop an understanding of the entire protein complement of a cell through analytical approaches, Data mining and other software tools.

Expected Course Outcomes:

On successful completion of the course, the student will be able to:

- Gain an insight of the basic and advanced concepts and applications of sequencing technologies.
- Understand the mechanism of genomics and proteomics and exploit the same in the growing field of omics.
- Apply functional genomics techniques to analyse data for biological system.
- Implement techniques and database search to analyze complex protein samples.

Unit-I

Comparative genomics: Databases for genomics resources, Technologic principles of currently utilised methods within the field of functional genomics. Sequence alignment: global versus local; Inferring gene function from relatedness to other genes, Finding Open Reading Frames, overall structure and organisation of the genome in microorganisms, plants and animals. Extra-chromosomal elements: Mitochondrial genome, mitochondrial plasmids. Organization of nuclear and organellar genomes; Repetitive DNA-satellite DNAs and interspersed repeated DNAs, Transposable elements, LINES, SINES, Alu family and their application in genome Mapping.

Molecular Phylogenetics: Fine structure of gene, split genes, pseudogenes, non-coding genes, overlapping genes and multi-gene families. Genome sequencing methods, Genome annotation at different levels, Comparative genome sequencing. Molecular Phylogenetics Origins, Applications of Molecular Phylogenetics. Phylogenomic comparisons, introduction to programs, Phylogenetic analyses: tree terminology and parsimony, Phylogenetic analyses to determine relationships and interpret character evolution. Single gene disorders- conventional and contemporary methods. Candidate gene identification; Genetic polymorphism and disease susceptibility.

Handwritten signature

Unit-II

Databases and Software for Genome Biology: Gene Expression Analysis using Microarrays and RNA-Seq, Application of DNA microarrays for the analysis of gene expression, protein-DNA binding, chromatin structure, chromatin modifying complexes, and RNA polymerase occupancy. Error models and data normalization techniques for high-resolution array technologies. Clustering genes into sets and discovering gene set features that can be used for diagnostic purposes. Importance of chromatin structure in contemporary modelling, relationship between chromatin structure and transcriptional regulation.

Unit-III

Functional Genomics-I: Determining the Functions of Individual Genes, Pathway and GO annotation systems, Non-coding DNA, gene prediction methods and tools. Genome Evolution, Genome and Transcriptome Assembly Algorithms, Profile HMMs for Protein Family Modelling, Human Genome Project, Genomics and Personalized medicine, Sex-linked inheritance: Conceptual basis, sex influenced traits, mechanism of sex determination. Mutation variability across the genome. Mutation detection, diagnosis and therapy. Genes and Environment - heritability, penetrance and expressivity.

Unit-IV

Functional Genomics-II: Integration of genomic, epigenetic and external effects in the context of understanding gene and genome function, Genome regulation: Introduction, levels of regulation, evidences and experimental designs/methodologies, role of genetic analysis in understanding gene function and regulation. Transcriptional control - DNA sequence elements and transcription factors. Mutation - Classification, mechanism, repair, role in genetic analysis and evolution. Post-transcriptional regulation - Alternative RNA splicing, RNA editing, RNA transport and localization, RNA stability, Regulation of translation - RNA structure, control at initiation, codon usage, Post-translational modifications. RNA-mediated control of gene regulation.

Suggested Readings:

1. Discovering Genomics, Proteomics and Bioinformatics 2nd edition, A. Malcolm Campbell and Laurie J. Heyer., Cold Spring Harbor Laboratory Press, 2006.
2. Analysis of Genes and Genomes, Richard J.R., Wiley Publications, 2003.
3. Comparative genomics: empirical and analytical approaches to gene order dynamics, map alignment and the evolution of gene families, Sankoff D. & Nadeau J.H., Kluwer Academic Publishers, Netherlands, 2000.

Reference Books:

1. Principles of Genome Analysis and Genomics (3rd Ed.), Primrose S.B. and Twyman R.M., Blackwell Publishing Company, Oxford, UK. 2003.
2. Bioinformatics and Functional Genomics, Pevsner J., John Wiley and Sons, New Jersey, USA, 2003.
3. Bioinformatics: Sequence and Genome Analysis, Mount D., Cold Spring Harbor Laboratory Press, New York, 2004.
4. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, Baxevanis A. and Ouellette F.B.F (Editors), John Wiley and Sons, New York, 1998.

Seema Deswal Tuteja, Seehgal, Parul



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-907	ALGORITHMS IN BIOINFORMATICS	L	T	P	C
		Skill Enhancement Course	3	-	-	3
Pre-requisite	Knowledge of Algorithms and Proteomics		Syllabus 2025			

The main objectives of this course are to:

1. Develop a quantitative understanding of how living things are built.
2. Raise the awareness of the impact of algorithms on the efficiency of the system.
3. Develop skills to analyse algorithms related to Bioinformatics

Expected Course Outcomes:

Upon successful completion of this course, students should be able to:

- Know the main problems in the field of bioinformatics and computational molecular biology.
- Analyse the correctness of algorithms and how it works.
- Describe the divide-and-conquer paradigm and explain when an algorithmic design situation calls for it.
- Apply the algorithms and design techniques to solve problems.
- Analyse the complexities of various problems in different domains.

Unit I

Introduction to molecular biology: basic concepts in biochemistry and bioinformatics, what are proteins and why they are important, Database search and sequence alignment: BLAST and FASTA 2 algorithms, the BLAST web server, multiple sequence alignment, Introduction to protein structure: Protein structure, the protein folding problem, protein structure visualization and classification.

Unit-II

Protein folding: Folding methods homology modeling, threading, ab-initio folding, folding models, folding vs. protein structure determination, Biomolecular simulations: Molecular dynamics, montecarlo. Applications of MD, force fields and energy models.

Geometry based methods: Transformations, distance measurement, introduction to geometric hashing, Geometric hashing in docking, Geometric hashing and its applications in protein structural alignment.

Unit III

Introduction: Introduction to Robotics applications in bioinformatics, Robotics-based methods and their applications in folding, flexible docking and conformational search. Other topics in bioinformatics: Systems biology, networks.

Other topics in bioinformatics: Evolutionary biology, data mining. Simple Algorithms, Analyzing Algorithms.

Unit-IV

Asymptotic Notation Design Methods: General Consideration, Algorithm Design Paradigms and Representative Problems. Divide and Conquer (Binary search, Merge Sort), Greedy Method (Minimal Spanning Tree). Dynamic Programming (Chained Matrix Multiplication), Longest common subsequence, Backtracking (8-queens problem), Branch and Bound (0/1 Knapsack Problem). String Matching Problem, Brute Force Method, KMP Algorithm, Boyer-Moore Algorithm, Approximate String matching.

Suggested Readings:

1. Algorithms in Bioinformatics, A Practical Introduction, Wing-Kin Sung, 2009.
2. An Introduction to Bioinformatics Algorithms (Computational Molecular Biology), Neil C. Jones, Pavel A. Pevzner.

Reference Books:

1. Data Structure (Outline Schaum Series), Lipshutz and Pai, McGraw Hill, Indian edition, 2010.
2. Data Structure Techniques, Standish T. A., Addison-Wesley, 1980.
3. File Structures, Salzberg, B., Prentice-Hall, 1988.
4. File Organization and Processing, Tharp, A.L., John Wiley and Sons, 1988.
5. The Design and Analysis of Computer Algorithms, Aho, A., Hopcroft, J. and Ullman, J., Addison Wesley.
6. Fundamentals of Computer Algorithms, Horowitz E., Sahani, V S. Rajasekaran S., Galgotia, New Delhi, 1993.
7. Introduction to the Design and Analysis of Algorithms, Goodman S.E. and Hedetniemi S.T., McGraw Hill.
8. Algorithmics, Brassard G. and Bratley P., Prentice-Hall India, 1988.
9. Design Methods and Analysis of Algorithms, Basu, S.K., Prentice Hall of India (Pvt.) Ltd, New Delhi, 2005.

Deepa Desmal *Fuller.* *Pavul*
beelgal.



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-909 (A1)	MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE IN BIOINFORMATICS	L	T	P	C
		Discipline Centric Elective	3	-	-	3
Pre-requisite	<u>Basic statistics, probability, Python, Bash</u>		Syllabus	2025		
Terminal						

The main objectives of this course are to:

1. Introduce the neural networks for classification and regression.
2. Give design methodologies for artificial neural networks.
3. Provide knowledge for network tuning and over fitting avoidance.
4. Demonstrate neural network applications on real-world tasks.

Expected Course Outcomes:

Upon successful completion of this course, students should be able to:

- Demonstrate fundamental understanding of artificial intelligence (AI) and expert systems.
- Apply basic principles of AI in solutions that require problem solving, inference, perception, knowledge representation, and learning.
- Demonstrate proficiency in applying scientific method to models of machine learning. Discuss the awareness of ANN and different optimizations techniques.
- Apply basic principles of AI in solutions that require problem solving, inference, perception, knowledge representation, and learning.

Unit-I

Introduction to Machine Learning: Idea of Machines learning from data, Classification of problem – Regression and Classification, Supervised and Unsupervised learning. k-Means Algorithm, Optimization objective, Random Initialization, Choosing number of clusters.

Linear and Logistic Regression: Model representation for single variable, Single variable Cost Function, Gradient Decent for Linear Regression, Multivariable model representation, Multivariable costfunction, Gradient Decent in practice, Normal Equation and non-invertibility. Logistic Regression-Classification, Hypothesis Representation, Decision Boundary, Cost function, Advanced Optimization, Multi-classification (One vs All), Problem of Overfitting, Regularization.

Unit-II

Neural Networks & Support Vector Machine: Non-linear Hypothesis, Biological Neurons, Model representation, Intuition for Neural Networks, Multiclass classification, Cost Function, Back Propagation Algorithm, Back Propagation Intuition, Weights initialization, Neural Network Training. SVM-Optimization Objective, Large Margin Classifiers, Kernels, SVM practical considerations.

Unit-III

AI Fundamentals: Defining Artificial Intelligence, Defining AI techniques. State Space Search and Heuristic Search Techniques: Defining problems as State Space search, Production systems and characteristics, Hill Climbing, Breadth first and depth first search, Best first search.

Knowledge Representation Issues: Representations and Mappings, Approaches to knowledge representation. Representing simple facts in logic, Computable functions and predicates, Procedural vs Declarative knowledge, Logic Programming, Forward vs backward reasoning.

Unit-IV

Symbolic Logic: Non-monotonic Reasoning, Logics for non-monotonic reasoning. Statistical Reasoning: Probability and Bayes Theorem, Certainty factors, Probabilistic Graphical Models, Bayesian Networks, Fuzzy Logic, Markov processes and Hidden Markov models.

Important Applications of AI: Introduction to Natural Language Processing, Hopfield Networks, Neural Networks, Recurrent Networks, Symbolic AI

Suggested Readings:

1. Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids, R. Durbin, S. Eddy, A. Krogh, and G. Mitchison, Cambridge University Press, 1998.
2. Intelligent Bioinformatics: The Application of Artificial Intelligence Techniques to Bioinformatics Problems, Edward Keedwell and Ajit Narayanan, Wiley, 2005.
3. Bioinformatics: The Machine Learning Approach, P. Baldi and S. Brunak.

Reference Books:

1. Data-driven advice for applying machine learning to bioinformatics problems, Olson et al., 2018.
2. Probabilistic Modeling in Bioinformatics and Medical Informatics, Husmeier D, Dybowski R, and Roberts S, Springer, 2005.
3. Printing proteins as microarrays for high-throughput function determination. Science. MacBeath G, Schreiber SL, Sep. 8-289 (5485), 1760-3. PubMed, 2000 (PMID: 10976071).
4. Influence of 17 β -Estradiol on Gene Expression of *Paracoccidioides* during Mycelia-to-Yeast Transition, Shankar J, Wu TD, Clemons KV, Monteiro JP, Mirels LF, et al., PLoS ONE 6(12), 2011 (e28402. doi:10.1371/journal.pone.0028402)

Other useful Web links:

1. Link to NPTEL course contents: <https://nptel.ac.in/courses/106104019/>
2. Link to topics related to course: <https://www.advancedsciencenews.com/machine-learning-for-bioinformatics-and-neuroimaging/>
https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_neural_networks.htm
<https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/>

Deepa Deswal
Tulsi
Behgal
Parul



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-909 (A2)	CLOUD COMPUTING	L	T	P	C
		Discipline Centric Elective	3	-	-	3
Pre-requisite	<u>Knowledge of Computer Algorithm</u>		Syllabus	2025		

The main objectives of this course are to:

1. Identify the technical foundations of cloud systems architectures.
2. Analyze the problems and solutions to cloud application problems.
3. Apply principles of best practice in cloud application design and management.
4. Identify and define technical challenges for cloud applications and assess their importance.

Course Outcomes:

Upon successful completion of this course, students should be able to:

- Understand the fundamental principles of distributed computing.
- Understand how the distributed computing environments known as Grids can be built from lower level services.
- Understand the importance of virtualization in distributed computing and how this has enabled the development of Cloud Computing.
- Analyze the performance of Cloud Computing.
- Understand the concept of Cloud Security.
- Learn the concept of Cloud Infrastructure Model

Unit-I

Basic Concepts of Cloud Computing: Introduction to computer networks, evolution of computer networks and its uses, Advantages and Disadvantages of Computer Network, reference models: OSI reference Models, TCP/IP Protocol Suit Networking fundamentals: Internet, Circuit switching vs Packet switching, ISPs, Delay and Loss in Packet, Switched Networks: Local Area Network LAN Architecture, LAN topologies- Bus/ Tree LAN, Ring LAN, Star LAN, Wireless LAN, Ethernet and Fast Ethernet, Token Ring, application layer and data link layer Application Layer Protocols: HTTP, FTP, SMTP, DNS Data link layer design issues, Flow Control- Stop and Wait, Error Detection, Error Control, error detection and correction, data link layer protocols, sliding window protocols, example of data link protocol- HDLC, Medium access layer Channel allocation problem, multiple access protocols, Introduction to ALOHA, CSMA/CD, CSMA/CA, The network layer Introduction, Routers, shortest path routing, flooding, distance vector routing, link state routing (without algorithms), congestion control and quality of service, internetworking, IP, Ipv4 Addressing vs Ipv6, The transport layer services, elements of transport protocols, TCP and UDP, Brief introduction to presentation and session layer, E-mail

Unit-II

Distributed Systems: Introduction, Goals and Examples, trends and resources of distributed systems, types of System Models, Physical Architectural and Fundamental models, Networking and Internetworking, Internet protocols, Basics of Inter-Process communication, Multicast

Communication, Network virtualization: overlay networks, Remote invocation, Remote Procedure Call, Remote method invocation, Operating System Support, Web Services and Security Introduction, Operating System layer and Architecture, Virtualization at Operating System level, Introduction to web services, Service descriptions and IDL for web services, Applications of web services, Overview of Security Techniques, Cryptographic algorithms, Digital signatures, Needham-Schroeder, Kerberos techniques, Distributed File Systems and Distributed Transactions Introduction (URI, URL's), Name services and the Domain Name System (Name Space, Name Resolution, DNS), Directory services, Transactions, Nested transactions, Locks, Optimistic concurrency control, Flat and nested distributed transactions, Atomic commit protocols, Concurrency control in distributed transactions, Distributed deadlocks, Distributed Multimedia System and Distributed Algorithms Characteristics of multimedia data, Quality of service management, Resource management, Stream adaptation, Introduction to communication protocols, Balanced sliding window protocol, Routing algorithms, Destination based routing, APP problem, Deadlock free Packet switching.

Unit-III

High Performance Computing: Parallel and Distributed Programming Models Introduction to high performance computing, basic definitions: cluster, grid, meta-computing, middleware etc., examples of representative applications. Programming models: shared memory, message passing, peer-to-peer. Development of parallel and distributed applications, Design phases, Common parallel patterns, Performance metrics and profiling, Overview of Cluster Computing The Role of Clusters, Definition and Taxonomy, Distributed Computing, Limitations, Architecture of cluster-based systems, Design Decisions, Network Hardware, Network Software, Protocols Distributed File Systems, Virtualization technologies, Issues in cluster design: performance, single system-image, fault tolerance, manageability, programmability, load balancing, security, storage, Introduction of Grid Computing Introduction, Evolution of the Grid, Definitions of Grid Computing, Infrastructure of hardware and software, Grid models, Applications, Examples of usage, Research possibilities / scope in Grid Computing, HPC and Grids, Scheduling HPC applications in Grids, Grid Monitoring Architecture (GMA) – An Overview of Grid Monitoring Systems, Integrating task parallelism with data parallelism Introduction and motivation, A model for integrating task parallelism into data parallel programming platforms, Integration of the model into ARC, Design and implementation applications, performance analysis, guidelines for composing user programs, related work. Anonymous remote computing and communication model: Introduction, Location in dependent inter task communication with DP, DP model of iterative grid computations, Design and implementation of distributed pipes.

Unit-IV

Interfacing with virtualization and cloud computing tools and techniques: Virtualization and cloud computing, Understanding Server Virtualization, types of server virtualization, Virtual machine basics, types of virtual machines, hypervisor concepts and types, : Microsoft's Infrastructure Optimization Model, Virtualization and the Infrastructure Optimization Model, Benefits of Virtualization, Achieving the Benefits of Datacenter Virtualization, Achieving the Benefits of Client Virtualization, Achieving the Benefits of Cloud Virtualization, Challenges while migrating to Cloud, Broad approaches to migrating into the cloud why migrate -deciding on cloud migration, the Seven-step model of migration into a cloud, Migration Risks and Mitigation, Enterprise cloud computing paradigm, relevant Deployment Models for Enterprise

Deepa Deswal *Teju-
Behgal.* *Pamul*

Cloud Computing, Adoption and Consumption Strategies, issues for enterprise applications on the cloud.

Suggested Readings:

1. Distributed System: Concepts and Design, Coulouris, Dollimore, Kindberg, Blair, Fifth Edition, Pearson Ed.
2. Advanced Concept in Operating Systems, Singhal&Shivaratri, McGraw Hill.
3. Grid Computing a Research Monograph, D. Janakiram, Tata McGraw hill publications.
4. Information Storage and Management, Wiley Publication (ISBN: 978-81-265-2147-0).
5. Advanced Server Virtualization: VMware and Microsoft Platform in the Virtual Data Center, David Marshall, Wade A. Reynolds.
6. Auerbach Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online - Michael Miller – Que, 2008.

Reference Books:

1. Computer Networks: Tanenbaum, Andrew S, Prentice Hall.
2. Computer Networks, 4th ed., Andrew S. Tanenbaum, Prentice Hall Gerald Tel.
3. Distributed Algorithms, Cambridge University Press Joshy Joseph & Craig Fellenstein.
4. Grid Computing, Pearson Education "Grid Computing: A Practical Guide to technology and Applications, Ahmar Abbas.
5. Building Storage Networks, Charles River media, Marc Farley Osborne, Tata McGraw Hill.
6. Storage Networks: The Complete Reference, Robert Spalding, Tata McGraw Hill, Publications, 2006.
7. Cloud Computing (Principles and Paradigms), RajkumarBuyya, James Broberg, AndrzejGoscinski, John Wiley & Sons, Inc. 2011.
8. Cloud computing a practical approach - Anthony T. Velte, Toby J. Velte Robert Elsenpeter, TATA McGraw- Hill, New Delhi – 2010.
9. Cloud computing for dummies- Judith Hurwitz, Robin Bloor, Marcia Kaufman, Fern Halper, Wiley Publishing, Inc, 2010.
10. Cloud Computing (Principles and Paradigms), RajkumarBuyya, James Broberg, AndrzejGoscinski, John Wiley & Sons, Inc. 2011.

Deepa Deswal *Teacher*
budget: *Paul*



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-951	DRUG DESIGN AND NGS (LAB-VI)	L	T	P	C
		Skill Enhancement Course	-	-	6	3
Pre-requisite		<u>Knowledge of PDB</u>	Syllabus		2025	

The main objectives of this course are to:

1. Provide practical experience in the analysis of protein sequences.
2. Understand the use of informatics in drug design and development, finding new targets to treat disease; mechanism of drug designing.
3. Gain insights on protein-ligand docking and knowledge-based scoring functions.

Expected Course Outcomes:

Upon successful completion of this course, students should be able to:

- Perform, understand, and interpret the results of the calculations and bring them in a publication ready form.
- Understand the importance of drug-like properties and their prediction.
- Describe the use of lead candidates and database representations.
- Apply Molecular modeling and molecular dynamics methods to study structure from sequence.

Experiments:

1. Generating drug library using blaster tool and analysing drug properties using Drug Classification tool.
2. Predicting cytochromes P450 inhibition using Which Cyp tool & analysing ADMET properties using ALOGPS2.
3. Predicting binding site in the protein molecule using FINDSITE tool & estimating binding free energy using BAPPLserver.
4. Designing ligand molecule using eDesign tool, Ligand-Protein Docking using AUTODOCK VINA & SwissDock tool.
5. Searching docking resources for structure-based drug design using SPLINTER.
6. Working with NGS databases (NCBI-SRA, etc.), NGS file formats, File format conversion.
7. Quality checking and trimming using freely available software (e.g. Fast QC).
8. Read mapping using BWA or Bowtie or any other freely available software, Interpretation and Visualization of Sequence Alignment Map (SAM) file.
9. Identification of SNPs using GATK Pipeline.
10. RNA-Seq and Chip-seq analysis using cloud-based server (e.g. Galaxy).

Suggested Readings:

1. Textbook of Drug Design and Discovery, Kristian Stromgaard, Povl Krogsgaard Larsen, Ulf Madsen, CRC Press, 2009.

Reference Books:

1. Drug Design and Discovery: Methods and Protocols, Volume 716, Seetharama D. Satyanarayanajois, Humana Press, 2011.

Deepa Deswal Kuching Paul

2. Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]
<https://nptel.ac.in/courses/102/106/102106070/>

Deepa Desmal ~~Teacher~~ Parul
Incharge



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-953	PROGRAMMING WITH PYTHON AND BIOPYTHON (Lab-VII)	L	T	P	C
		Skill Enhancement Course	-	-	4	2
Pre-requisite		<u>Knowledge of Python and Biopython</u>				2025
Syllabus						

The main objectives of this course are to:

1. Provide Python and Biopython Programming and applications to Bioinformatics.
2. To familiarize and master the programming skills in Perl and Python.
3. Implement problem-solving algorithms efficiently in the Python and Biopython programming language
4. Evaluate the rationale behind these problem-solving algorithms
5. Evaluate published research that employs algorithmic problem-solving strategies

Expected Course Outcomes:

Upon successful completion of this course, students should be able to:

- Analyze the syntax and semantics of diverse coding elements of Python and Biopython program
- Select an appropriate problem-solving algorithm for a given bioinformatics problem
- Select an appropriate data structure to store and efficiently manipulate data

Experiments:

1. Python File handling-Read Write Fasta.
2. Gene Expression Analysis with Python.
3. Using BioPython- Part I
4. Using BioPython- Part II
5. Sequence analysis- Reading DNA/Protein sequences, sequence length, GC%, handle string search.
6. Programs using string handling function substring search, count of nucleotides, check RNA/DNA, ORF finding, Transcription, Translation.
7. Program with regular expression.
8. Define RE for a set of sequences, search for subsequences/patterns, locations.
9. Programs for processing FASTA files.
10. Analyze Biological/Bioinformatics data: Epidemic Modeling of COVID dataset.

Suggested Readings:

1. Tim J. S., Wayne B. (2015), Python Programming for Biology Bioinformatics and Beyond, CUP
2. Biopython Tutorial and Cookbook (ebook), Jeff Chang, et al., 2021, <http://biopython.org/DIST/docs/tutorial/Tutorial.html/>
3. Martin Jones. 2013. Python for Biologist – A programming course for complete beginners. <http://pythonforbiologists.com>.

Reference Books:

1. Python for Bioinformatics, Sebastian Bassi, Chapman and Hall, 2010.

Deepa Deswal *Ruchi Behgal* *Pavul*



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-955	CHEMOINFORMATICS AND CHEMOGENOMICS (LAB-VIII)	L	T	P	C
		Skill Enhancement Course	-	-	4	2
Pre-requisite		<u>Knowledge of PDB</u>	Syllabus	2025		

The main objectives of this course are to:

1. Should be aware of available Cheminformatics resources and information in Databases.
2. Apply critical, analytical and problem solving skills to deal with Cheminformatics research.
3. Attain knowledge to identify the research problems in Cheminformatics.
4. Develop programming skills to write their own codes and scripts to resolve research accomplishments.

Expected Course Outcomes:

Upon successful completion of this course, students should be able to:

- Develop applications to solve biological problems by utilizing the Cheminformatics algorithms and programming languages.
- Able to apply the knowledge in Cheminformatics such as computational biology, chemical principles that underlie biochemistry, molecular biology and genomics.
- Develop and implement software effectively to retrieve information from biological databases and use this information for computation.

Experiments:

1. Chemical database techniques and data mining methods and handling.
2. Structure-Activity relationships (SAR) to predict the behaviors of unknown compounds.
3. XML and CML data format to store molecular information.
4. Examples with regards to modified lead substructures relating to absorption in GI tract and ability to cross Blood Brain Barrier (BBB).
5. Experiments on interactions between chemical compounds and living systems in terms of induced genomic response.
6. Generation of SMILES (simplified molecular input line entry specification) strings for various compounds.

Suggested Readings:

1. Computational Chemogenomics, J.B Brown, Springer protocols, Humana Press, 2019.
2. Practical Chemoinformatics, Springer, MuthukumarasamyKarthikeyan, Renu Vyas, 2021.

Reference Books:

1. Tutorials in Chemoinformatics, Alexandre Varnek (Editor), ISBN: 978-1-119-13796-2, August 2017

Deepa Desmal Teacher - Deepa Desmal



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-902	DATA MINING IN HEALTHCARE	L	T	P	C
		Core Course	4	-	-	4
Pre-requisite	Programming/Statistics Language: R, Python, C++, Java		Syllabus 2025			

Instruction to Paper Setters:
Attempt five questions

Time: 3 hours
Maximum Marks: 60

Question Paper shall contain **Five Sections**

- The student has to attempt **five questions** from five sections.
- All sections are of 12 marks each.
- Section **I** is compulsory, and it should have objective or short answer type questions and cover the entire syllabus.
- Section **II** to **V** shall have two questions in each section, and the student needs to attempt only one question from each section. (Section II to V shall reflect Unit I to IV of the syllabus respectively).

The main objectives of this course are to:

1. Make students implement the data mining techniques in healthcare, to solve health-related problems.

Course Outcomes:

On the successful completion of the course, student will be able to:

- Learn techniques for managing and mining big data.
- Extract, manipulate, manage and analyze medical and biological data.
- Use techniques such as Python and R.

Unit I

Health care Analytics: Introduction to Healthcare Data Analytics, Electronic Health Records, Privacy-Preserving Data Publishing Methods in Healthcare, Clinical Decision Support Systems.

Unit II

Healthcare Data Analytics Natural Language Processing and Data Mining for Clinical Text: Core NLP Components, Information Extraction and Named Entity Recognition, Social Media Analytics for Healthcare: Tracking of Infectious Disease Outbreaks, Readmission risk Prediction.

Unit III

Genomic Data Analytics: Microarray Data, Microarray Data Analysis.

Deepa Deswal

*Shuchi
Bergal*

Paul

Unit IV

Genomic Data Analysis for Personalized Medicine, Patient Survival Prediction from Gene Expression Data, Genome Sequence Analysis.

Suggested Readings:

1. Data Science and Predictive Analytics: Biomedical and Health Applications using R, Ivo D. Dinov, Springer, 2018.
2. Data Handling in Biomedical Sciences, Peter White, Cambridge University Press, 2012.

Reference books:

1. Healthcare Data Analytics, Chandan K. Reddy, Charu C. Aggarwal CRC Press, 2015.

Deepa Desmal Tucker. Rehgal. Parul



GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-904 (B1)	PROGRAMMING IN JAVA	L	T	P	C
		Generic Elective Course	4	-	-	4
Pre-requisite		<u>Knowledge of Computers</u>	Syllabus		2025	

The main objectives of this course are to:

1. Understand and implement OOPs concept through JAVA programming language.
2. Learn applet programming.
3. Understand how JAVA helps in solving different complex problem in biology or data analysis.
4. Learn the fundamentals of Java Programming in order to apply it in Bioinformatics.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

- Understand the operators of Java and its applications.
- Learn the control statements, classes and methods of Java.
- Know packages, applets and string handling in Java.
- Learn event handling and AWT controls in Java.
- Understand, manipulate software tools and distributed packages for various sequence manipulation and analysis according to a particular need.

Unit-I

Introduction: Environment and Programming Structure: Java White Paper Buzzwords, History of Java, Choosing a Development Environment: Command-Line Tools, Running a Graphical Application, A Simple Java Program, Comments, Data Types, Variables, Operators, Input and Output, Control Flow, Big Numbers, Arrays. 10

Unit-II

Class, Objects and Inheritance: Introduction to OOP, Predefined Classes, User Defined Classes, Static Fields and Methods, Method Parameters, Object Construction, Packages, Class Path, Class Design; Inheritance: Super-classes and Subclasses, Types of Inheritance, Polymorphism, Abstract class, Object: The Cosmic Super class, Generic Array Lists, Inheritance Guidelines.

Unit-III

Interfaces, String Handling and Exceptions: Interfaces, Object Cloning, Interfaces and Callbacks, Inner Classes, Proxies, String Handling APIs: String, Immutable String, String Comparison, String Concatenation, Substring, Methods of String Class, String Buffer Class, String Builder Class, Creating Immutable Class, String Tokenizer Class. Exception Handling.

Unit-IV

Generic Programming & Collection: Definition of Generic Class, Generic Methods, Bounds for Type Variables, Generic Code and the Virtual Machine, Restrictions and Limitations,

Deepa Deswal
Rohit
Ravi

Inheritance Rules for Generic Types, Wildcard Types, Reflection and Generics. Collection Interfaces, Concrete Collections, The Collections Framework, Algorithms, Legacy Collections

Java GUI Programming & JDBC: Introduction to Swing, Creating a Frame, Positioning a Frame, Displaying Information in a Component, Displaying Images, Event Handling, Basics of Event Handling, Actions, Mouse Events, The AWT Event Hierarchy; JDBC: Basic JDBC Programming Concepts, Installing JDBC, Statements, Executing Queries, Result Sets.

Suggested Reading:

1. Core Java Volume I: Fundamentals: 1, 11th Edition, Horstmann C, Pearson Education, 2020.
2. Core Java Volume II: Advanced Features, Horstmann C, Pearson Education, 2020.
3. Java How to Program, 11th Edition, Dietel&Dietel, Pearson Education, 2018.

Reference Books:

1. JavaScript: The Complete Reference, 2nd Edition, A P Thomas and F Schneider, TMH, 2004.

Deepa Deswal Fuchi . Iselga . Parul



GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code BI-904 (B2)

IMMUNOINFORMATICS

L T P C

Generic Elective Course

4 - - 4

Pre-requisite

Basic Concepts of Immunology

Syllabus 2025

The main objectives of this course are to:

1. Enable the students to acquire skill set for in-silico immunological studies.
2. Make students familiarize and use the web based tools and computer aided vaccine design.
3. Gain experience in application of sequence analysis in immunology.
4. Study viral escape.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

- Apply knowledge of immune responses to various pathogens by integrating omics with bioinformatics strategies.
- Analyze computer aided vaccine design.
- Apply the computational techniques in solving immunological problem.
- Use Clustering techniques and understand the role of Microarray in immunology.
- Use Methods to predict B cell epitopes.
- Handle different tools and servers for vaccine designing.

Unit-I

Immunoglobulins: Monoclonal antibodies: structure and function. B Cell generation and differentiation: BCR--Antibody diversity: Genetic basis--T- dependent activation of B cells-B lymphocyte signal transduction. Cytokines. Complement system. Antigen- antibody interaction: antibody affinity and activity- Isolation of lymphoid cells from blood and lymphoid organs--precipitation reaction, agglutination reaction --Radioimmunoassay, ELISA, Western Blot, Immunoprecipitation, Immun-fluorescence, flow cytometry. Cell cultures and Experimental animal models. Analysis of gene expressions.

Unit-II

Introduction to Immunoinformatics and Immunological Databases: Introduction to immunology & Bioinformatics, immunoinformatics, the immune system, cellular immunity, antibody mediated immunity. Immunological databases-db MHC-MHC database at NCBI, MHCBN-comprehensive database of MHC binding and non-binding peptides, T-cell epitope databases, B-cell epitope databases. SYFPEITHI MHC presented epitopes, Antigen recognition by B-cells vaccine design, The IMGT® Immunoinformatics page, Exploring Databases with Immunoglobulins (or Antibodies) (IG), T cell receptors (TR), Major histocompatibility (MH), Antigens, Allergens, Peptides binding to MH etc.

Unit-III

Deepa Deswal Teacher
Bechgal
Pam

Immunological Tools: Experimental and theoretical description of peptide-MHC binding, selection of epitopes using bioinformatics, prediction of proteasome processing, and TAP binding, Predictions of Class I and Class II MHC Epitopes, IEDB analysis L T P/S SW/F W
TOTAL CREDIT UNITS 2 0 0 0 02 Resource, CTLPred, Population Coverage analysis, Epitope conservancy analysis.

Unit-IV

Computational Vaccinology: Introduction to vaccines, Different generations of Vaccines, Concepts of reverse vaccinology, case study of Reverse Vaccinology with Meningococcus B, Comparison of Traditional Vaccinology and Reverse Vaccinology, Tools & servers for computational Vaccine design-from Genome to Vaccine.

Suggested Reading:

1. Immunoinformatics: Bioinformatic Strategies for Better Understanding of Immune, Wiley Publications, 2008.
2. Immunology and Immunotechnology, K. Chakarvarthy, Oxford University Press, 2006.
3. Immunoinformatics: Predictive Immunogenicity insilico, D.R. Flower, Humana Press, 2007.

Reference Books:

1. Immunological Bioinformatics, Ole Lund, MIT press, September 2005.
2. Immunoinformatics: Bioinformatics Strategies for Better Understanding of Immune Function, Rammensee, No. 254 [E-Book], Novartis Foundation.

Deepa Suresh Kumar
Ischgal
Pammy



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-906	DEVELOPING ENTREPRENEURIAL MINDSET	L	T	P	C
		Non University Examinational Subject	2	-	-	2
Pre-requisite		<u>Nil</u>				2025

The main objectives of this course are to:

1. The course aims at developing an entrepreneur's attitude in the student by helping them to understand the steps involved in becoming an entrepreneur and developing a mindset of entrepreneurship.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

- Quickly sense, take action, and get organized under certain conditions.
- Acquire self-leadership that utilizes behavior strategies, reward strategies, and constructive thought patterns.
- Become more creative to produce new ideas and inventions that are exclusive and of value to others.

Unit-I

Introduction to Entrepreneurship: Who is an Entrepreneur? Advantage of becoming entrepreneur, characteristics of entrepreneur, competencies, and skills possessed by Entrepreneur, Myths about the entrepreneur, etc. Difference between entrepreneur and manager, entrepreneur, and entrepreneurship. Case studies on Indian entrepreneurs.

Unit-II

Steps involving in starting enterprise: Deciding the type of organization to start a business, deciding the enterprise's name, registration formalities, identification of opportunities, sources of finance, arranging the finance, and managing the enterprise.

Unit-III

Definition of MSME & Institutional Support: Definition as per MSMED Act, 2016 revised guideline 2020, incentive available to MSME by Govt. of India, Institutional setups available at the centre and state level supporting MSME. Case study on MSME enterprises in India.

Unit-IV

Developing entrepreneurship attitude: Practical training on developing creativity and innovation in the students, entrepreneur attitude using behavioural scales, entrepreneurship scorecard for the students, improving public speaking and negotiation skills, doing a live project.

Suggested Readings:

1. Entrepreneurship Development and Small Business Enterprise, Charantimath, Pearson Education, 2013.

Deepa Deswal *[Signature]* *[Signature]* *[Signature]*

2. Essentials of Entrepreneurship and Small Business Management, 8th Ed., Scarborough N.M. and Comwall H.R., Pearson Education, 2016.
3. Entrepreneurship Theory and Practice, TaingKalpana, Anmol Publication Pvt. Ltd., Delhi, 2014.

Reference Books:

1. Entrepreneurship Development and Small Scale Industries, NathSuryakant, Neha Publishers & Distributors, Delhi, 2012.
2. Entrepreneurship New Venture Creation, Holt D.H., Pearson Education, 2016.

Deepa Desmal Teacher:

Iselgal

Ramul



GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY
SECTOR-16C, DWARKA, NEW DELHI-110078

Course Code	BI-958	MAJOR PROJECT/DISSERTATION	L	T	P	C
			-	-	-	12
Maximum Marks: 100						

Dissertation work would comprise of research work carried out by each student during semester IV under the supervision of a particular faculty member. The student would carry out the review of literature on the topic of research and formulate the plan of work in consultation and in the supervision of the mentor. The student would then conduct the research experiments for the proposed work. Towards the end of semester IV, the student will compile the research work including review of literature, aims and objectives, methodology and results and discussion in the form of a dissertation in the supervision of the mentor. At the end of semester 4, students would make presentations in the presence of all faculty members and would be collectively judged by the faculty members. Marks will be assigned to each student collectively by the faculty based on his/her performance, work and continuous assessment throughout the year by the mentor.

Deepa Deswal *Teacher*

beehgal

Pamul