



Rayat Shikshan Sanstha's

Yashwantrao Chavan Institute of Science, Satara (Autonomous)

Lead College, Karmaveer Bhaurao Patil University, Satara

Reaccredited by NAAC (3rdCycle) with 'A+' grade (CGPA 3.57).

ISO 9001:2015 Certified

Bachelor of Science Part – II

Bioinformatics

Syllabus

to be implemented w.e. f. June, 2024 NEP 2020

Rayat Shikshan Sanstha's
Yashavantrao Chavan Institute of Science, Satara
Department of Bioinformatics
Syllabus for Bachelor of Science - Part II

1. TITLE: Bioinformatics

2. YEAR OF IMPLEMENTATION: New Syllabi for the B.Sc. II Bioinformatics will be implemented from June 2024 onwards.

3. PREAMBLE:

Bioinformatics is a rapidly evolving interdisciplinary field at the intersection of biology, computer science, and data analysis. It plays a pivotal role in unraveling the complexities of life sciences, enabling us to decode the secrets of genomics, proteomics, and structural biology. The Bachelor of Science (B.Sc.) in Bioinformatics program is designed to equip students with the knowledge and skills necessary to excel in this dynamic and groundbreaking field.

4. Programme Objectives :

- 1.To provide students with a strong foundation of knowledge in their chosen field of study, including fundamental concepts, theories, and principles.
2. To cultivate research skills, including the ability to design experiments, gather and analyze data, and draw evidence-based conclusions.
3. To enhance students' written and oral communication skills, enabling them to effectively convey scientific concepts, research findings, and ideas to diverse audiences.
4. To equip students with quantitative and analytical skills necessary for data analysis, modelling, and interpretation in their field of study.
5. To develop proficiency in using relevant technology and tools that are essential for their field, including software, laboratory equipment, and data analysis tools.
6. To encourage community engagement and a sense of social responsibility, inspiring students to use their knowledge and skills for the betterment of society.

5. Programme Outcomes:

1. Students will have a deep understanding of the principles and concepts of science.
2. Students will be competent in designing and managing biological databases, ensuring data integrity and accessibility for research purposes.
3. Graduates will be able to integrate techniques into biological research, contributing to advancements in fields such as genetics, drug discovery, and disease modeling.
4. Graduates will effectively communicate scientific findings and research results through written reports, presentations, and publications.
5. Graduates will collaborate with researchers and professionals from diverse fields, including biology, computer science, and medicine, to solve complex biological problems.
6. Graduates will demonstrate problem-solving skills by applying techniques to real-world biological challenges and proposing innovative solutions

6. Programme specific objectives:

1. To provide students with a strong foundation in bioinformatics principles, algorithms, and methodologies, enabling them to understand and address biological challenges using computational tools.
2. To equip students with in-depth knowledge of genomics, including DNA sequencing technologies, genome assembly, annotation, and analysis.
3. To enable students to explore protein structure and function, including the prediction of protein structures, protein-protein interactions, and structural analysis.
4. To develop students' skills in processing, analyzing, and interpreting biological data, including sequence alignment, phylogenetics, and functional annotation.
5. To teach students how to design and manage biological databases, ensuring efficient data storage and retrieval for research purposes.
6. To familiarize students with bioinformatics software and databases commonly used in research and industry, enabling them to leverage existing resources effectively.
7. To provide opportunities for students to work on real-world bioinformatics projects, allowing them to apply their skills to practical research problems.

8. Program specific outcomes:

1. Graduates will be proficient in analyzing and interpreting biological sequences, including DNA, RNA, and protein sequences, using relevant algorithms and tools.
2. Graduates will have the ability to predict protein structures, analyze protein-ligand interactions, and model three-dimensional structures using computational methods.
3. Graduates will be skilled in mining large biological datasets for patterns, associations, and insights, enabling them to make data-driven discoveries.
4. Graduates will be able to create clear and informative data visualizations that facilitate the communication of complex biological findings to both technical and non-technical audiences.
5. Graduates will be skilled in assessing data quality, implementing data validation procedures, and ensuring the reliability of biological data.
6. Graduates will be prepared for careers in bioinformatics research, pharmaceuticals, healthcare, and related industries, with the skills and knowledge needed to excel in their chosen roles.

Rayat Shikshan Sanstha's
Yashwantrao Chavan Institute of Science, Satara (Autonomous)
Department of Bioinformatics
“NEP Implementation 2020”

B.Sc. II SEM - III Bioinformatics

Sr.no	Course Category	Course code	Name of Course
1	Major -I	BBiT231	Database Management System
2	Major-II	BBiT232	Genomics, Proteomics & Transcriptomics
3	Major Practical III	BBiT233	Based on BBiT231
4	Major Practical IV	BBiT234	Based on BBiT232
5	Minor-I	BBiT235	Statistical Analysis methods
6	Minor practical III	BBiP 236	Lab course Based on BBiT235
7	VSC	BBiTVSC 1	Vocational Skill Course : Bioinformatics methods I
8	SEC	BBiTSEC 2	Skill Enhancement Course – Web-based Programming
9	AEC	BBiTAEC 1	English I
10		BBiTAEC 2	English II
11	VEC	BBiTVEC 2	Value Education Course – Environmental awareness from Bioinformatics

B.Sc. II SEM - IV Bioinformatics

Sr.no	Course Category	Course code	Name of Course
1	Major -I	BBiT241	Python programming for Bioinformatics
2	Major-II	BBiT242	Biological Sequence and Protein structure Analysis
3	Major Practical V	BBiP 243	Based on BBiT 241
4	Major Practical VI	BBiP244	Based on BBiT 242
5	Minor-I	BBiT245	Data Visualization using R
6	Minor- practical IV	BBiP 246	Lab course based on BBiT 245
7	VSC	BBiTVSC 2	Vocational Skill Course : Bioinformatics methods II
8	SEC	BBiTSEC 3	Skill Enhancement Course – Programming in Perl & Bioperl
9	AEC	BBiTAEC 1	English III
10		BBiTAEC 2	English IV
11	CC	BBiT CC 2	CC 2 from CC Board

Course Structure for B.Sc. II. (Semester- III)

Theory				Practical				
Course Title	Course Code	Lecture per week	Credits	Course	Course Title	Course Code	Lecture per week	Credits
Database Management System	BBiT- 231	4	2	Practical -1	Practical based on Paper BBiT 231 and 232 Lab- II	BBiP- 233 & BBiP 234	4	2
Genomics, Proteomics & Transcriptomics	BBiT- 232		2					

Course Structure for B.Sc. II. (Semester-IV)

Theory				Practical				
Paper Title	Paper Code	Lecture per week	Credits	Course	Paper Title	Paper Code	Lecture per week	Credits
Python programming for Bioinformatics	BBiT- 241	4	2	Practical 1 and 2	Practical based on Paper BBiT 241 and 242 Lab- II	BBIP -243 and BBIP 244	4	2
Biological Sequence and Protein structure Analysis	BBiT 242		2					

Note: B=B.Sc., T=Theory and P=Practical.

Structure and Title of Courses of B.Sc. II**B.Sc. II Semester III**

Course Number	Course Code	Course Name
I	BBIT-231	Database Management System
II	BBIT-232	Genomics, Proteomics & Transcriptomics
Lab-III	BBIP-233	Practical based on Paper BBiT 231
Lab – IV	BBIP-234	Practical based on Paper BBiT 232

B.Sc.II Semester IV

Course Number	CourseCode	CourseName
III	BBIT241	Python programming for Bioinformatics
IV	BBIT 242	Biological Sequence and Protein structure Analysis
Lab-V	BBIP243	Practical based on PaperBBIT 241
Lab -VI	BBIP 244	Practical based on Paper BBIT 242

Rayat Shikshan Sanstha's
Yashavantrao Chavan Institute of Science, Satara(Autonomous)
Lead college of Karmaveer BhauraoPatil University, Satara
Department of Bioinformatics

B.Sc.II Syllabus for Course in Bioinformatics

SEMESTER – III

CourseCode–BBIT231:Database Management System

Course Objectives: Student should be able to ...

1. Understand the fundamental concepts of data.
2. Imbibe principles of databases.
3. Identify the database management operation.
4. Discuss the concept of procedure oriented, and object-oriented programming languages.

Credits (Total Credits 2)	SEMESTER – III BBIT 231 - Database Management System	No. of hours (30)
Unit - I	<ul style="list-style-type: none"> • <u>Introduction to databases and DBMS approach in bioinformatics</u> • Database system concepts and architecture in bioinformatics • Three-schema architecture and data independence • Centralized and client-server architectures for DBMS in bioinformatics • Introduction – Database System Versus File Systems, Characteristics of Database, Database Concepts, Schemas & Instances, DBMS architecture and Data Independence, Database Languages & Interfaces, View of Data, Database users and Administrators, Database System Structure, Database System Applications 	(08)
Unit – II	<ul style="list-style-type: none"> • Data Models – • ER Model: Keys, Constraints, Design Issues, Extended ER features, Reductions of ER • Schema to Tables. Relational Model: Structure, Relational Algebra; Hierarchical Model, Network Model, Object Oriented Model 	(08)

Unit – III	<ul style="list-style-type: none"> • Basics of Structured Query Language (SQL) – • Basic Structure, Set Operations, Aggregate Functions, Null Values, Nested Sub queries, Views, Integrity: Domain constraints, Joined Relations, Data-Definition ,Language 	(06)
Unit - IV	<ul style="list-style-type: none"> • Relational Database Design for Bioinformatics • Relational Database and Storage – Pitfalls in Relational Design Database, Functional dependencies, Decomposition Normal Forms – 1NF, 2NF, 3NF & Boyce-Codd NF, Data Storage – Ordered indices, Hashing concepts - Security and Authorization. • Concurrency control techniques & Information retrieval – Transactions: Properties of transactions: Concurrency problems, Serialisability and Locking techniques 	(08)

Course Outcomes: Students should be able to...

1. Demonstrate the basics of data, information, system and Database.
2. Evaluate basics of different database models for software development.
3. Design the basics of Relational algebra operations and Relational Calculus.
4. Apply SQL basics and write queries to perform different operations on realworld data

Reference Books:

1. Ramez Elmasri and Shamkant B. Navathe. Fundamentals of Database Systems (6th Edition). Pearson Education, 2010.
2. Raghu Ramakrishnan and Johannes Gehrke. Database Management Systems (3rd Edition). McGraw-Hill, 2002.
3. Abraham Silberschatz, Henry F. Korth, and S. Sudarshan. Database System Concepts (6th Edition). McGraw Hill, 2010.
4. Abraham Silberschatz. Database System Concepts. McGraw Hill, 2021.
5. Ben Chan. SQL Programming: Learn the Ultimate Coding, Basic Rules of the Structure Query Language. Notion Press, 2020.
6. Toby Teorey. Database Modeling and Design: Logical Design. Morgan Kaufmann, 2010.
7. Ivan Bayross. SQL, PL/SQL The Programming Language of ORACLE. BPB publication, 2021.

SEMESTER – III**Course code BBIP 233: Lab course based on BBIT 231 Database Management system****Course Objectives: Students will be able to,**

1. understand the fundamental concepts of databases and their application in bioinformatics.
2. explore various data models and their relevance in representing bioinformatics data.
3. develop skills in SQL for effective data manipulation and retrieval in bioinformatics.
4. gain expertise in relational database design and management specific to bioinformatics applications.

Credits (Total Credits 2)	SEMESTER – III BBIP 233 List of Practical (20)	No. of hours (60)
1	Introduction to databases, data acquisition, curation and formatting ,DBMS in bioinformatics, comparison with file systems.	3
2	Understanding three-schema architecture, data independence, implementing centralized architecture.	3
3	Constructing ER diagrams, converting to relational tables, performing relational algebra	3
4	Identifying keys, constraints, addressing redundancy, normalization to 3NF.	3
5	Learning SQL syntax, executing queries, handling NULL values, nested subqueries.	3
6	Creating, manipulating views, implementing domain constraints, exploring joined relations.	3
7	Identifying pitfalls in database design, understanding functional dependencies, normalizing to BCNF.	3
8	Exploring data storage methods, implementing indexing techniques, discussing security.	3
9	Investigating concurrency problems, understanding transaction properties, implementing locking techniques.	3
10	Studying information retrieval, analyzing transaction properties, designing and executing transactions	3
11	SQL statements to create, update, and delete databases and tables	3
12	SQL statements to insert, update, and delete records from tables	3
13	SQL statements to create, update, and delete views	3
14	Simple SQL queries to retrieve information from a database	3
15	Nested SQL queries to handle complex information retrieval requirements.	3
16	SQLqueries using aggregate functions like count, average, sum, etc.	3
17	PL/SQL blocks using basic data types and operators, branching and looping constructs	3

18	Data file to SQL database using R, Connecting accessing, updating SQL database using R	3
19	Database triggers using PL/SQL	3
20	Database functions/procedures using PL/SQL	3

Course Outcomes: Students should be able to-

1. Explain database system architecture in bioinformatics.
2. Design and implement bioinformatics databases.
3. Proficiency in SQL for bioinformatics data handling.
4. Examine data integrity and concurrency control in bioinformatics databases.

Reference Books:

1. Coronado, S., Zavala, A., López, F., & López, M. Database Systems: Concepts, Design, and Applications (2019).
2. Elmasri, R., & Navathe, S. B. Fundamentals of Database Systems (2019).
3. Ramakrishnan, R., & Gehrke, J. Database Management Systems (2003).
4. Silberschatz, A., Korth, H. F., & Sudarshan, S. Database System Concepts (2010).
5. Date, C. J. An Introduction to Database Systems (2003).
6. Garcia-Molina, H., Ullman, J. D., & Widom, J. Database Systems: The Complete Book (2008).
7. Rob, P., & Coronel, C. Database Systems: Design, Implementation, and Management (2009).
8. Connolly, T., & Begg, C. Database Systems: A Practical Approach to Design, Implementation, and Management (2014).
9. Navathe, S. B. Fundamentals of Database Systems (2015).
10. Ramakrishnan, R., Gehrke, J., & Johannes, G. Database Management Systems (2006).

SEMESTER III**Course Code – BBIT 232:Genomics, Proteomics & Transcriptomics****Course Objectives: Students should be able to ...**

1. Describe the fundamental of genome organization, genomics and the techniques used for the analysis of genomic data.
2. Compare DNA and RNA microarray techniques with the applications in basic research.
3. Demonstrate the protein and protein purification strategies.
4. illustrate the protein identification techniques like 1D and 2D gel electrophoresis & other Protein analysis techniques

Credits (Total Credits 2)	SEMESTER – III BBIT 232 : Genomics, Proteomics & Transcriptomics	No. of hours (30)
Unit - I	Genomics A) Genomics overview, omes and omics, Concepts and applications B) Genome overview at the level of Chromosome (with model organisms example); Strategies for large scale DNA sequencing, Whole genome analysis techniques, Next generation sequencing methods; Organization, structure and mapping of genomes (with model organisms example), Genome markers and mapping , Functional genomics C) Comparative genomics - Goals, genome annotation, methods and limitations	(06)
UNIT II	Transcriptomics A) Introduction to transcriptomics and expression profiling B) RNA sequencing techniques : Microarray, RNAseq-Sanger,Illumina,Oxford nanopore ,Concept working and analysis databases and bioinformatic tools Microarray databases and bioinformatics tools C) Investigative techniques –EST, SAGE, SNP D) Applications in basic research and medical genetics, Metagenomics, Toxicogenomics, Pharmacogenomics, Gene disease association.	(08)
Unit – III	Proteomics A) Proteomics overview, Concept, application, advantages and limitations of Expressional, Functional and Structural Proteomics,types of proteins, protein folding and misfolding and diseases, Anfinsen’s experiment B) Proteomics-with at least one explanatory example for each. C) Applications	(08)

Unit - IV	Techniques in Proteomics A) Strategies in protein identification, Polyacrylamide gel electrophoresis (PAGE), 2D Gel electrophoresis, Isoelectric Focusing (IEF), staining methods for 1D and 2D PAGE. B) Mass spectrometry in proteomics – C) Principle, techniques, components and variations (HPLC, ESI, MALDITOF, FT-MS, MS/MS, Quadrupole) and analysis, applications D) Protein- Protein interactions- experimental and computational two hybrid, Phage display; E) Protein Microarray- Preparation, working and analysis. Proteomics and Microarray databases and allied bioinformatics tools	(08)
------------------	--	-------------

Course Outcomes: Students should be able to ..

1. Recognize recent advances in genomics and proteomics.
2. Acquire knowledge about tools and techniques used in genomics and proteomics research.
3. Apply concept and applications of gene expression studies
4. Utilize the techniques of proteomics for their future studies.

REFERENCES:

1. Langauer T. Bioinformatics - From Genomes to Drugs (editor) Wiley- VCH; 1st edition (2001)
2. Mount D. W. Bioinformatics-Sequence and Genome Analysis Cold Spring Harbor Laboratory Press; 2nd edition ((2004)
3. Graham D. R. M. Broad-based Proteomics strategies: a practical guide to proteomics and functional screening J. Physiol 563: 1, 1-9, (2005).
4. Miller W. et al, Comparative Genomics Annu. Rev. Genomics Hum. Genet, 5: 15-56 (2004)
5. Campbell A. M., Heyer L. J., Cummings B. Discovering genomics, Proteomics and Bioinformatics 2nd edition, (2006).

SEMESTER – III**Course code BBIP 234: Lab course based on BBIT 232****Course Objectives: Students will be able to,**

1. provide an overview and hands-on experience in genomics, transcriptomics, and proteomics techniques.
2. familiarize students with the experimental methods and bioinformatics tools used in genomic and proteomic analysis.
3. develop skills in protein purification, identification, and characterization techniques.
4. enable students to critically analyze and interpret experimental data and apply them to biological research.

Credits (Total Credits 2)	SEMESTER – I BBIP 234 List of Practical (20)	No. of hours (60)
1	Study of different types of DNA using micrographs and model / schematic Representations .	3
2	Study of different types of RNA using micrographs and model / schematic Representations .	3
3	Study of semi-conservative replication of DNA through micrographs / schematic representations	3
4	Isolation of genomic DNA from E. coli	3
5	Isolation and purification of plasmid DNA	3
6	Restriction digestion of plasmid DNA	3
7	Estimation of DNA using UV spectrophotometer.	3
8	Estimation of RNA using UV spectrophotometer	3
9	Resolution and visualization of DNA by Agarose Gel Electrophoresis	3
10	Resolution and visualization of proteins by Polyacrylamide Gel Electrophoresis (SDS – PAGE).	3
11	Protein separation using HPLC (demo)	3

12	Instrumentation in molecular biology - Ultra centrifuge, Transilluminator, PCR	3
13	Performing practicals of genomic sequences using National Centre for Biotechnology Information (NCBI).	3
14	Using biological search engine- Entrez	3
15	Introduction to literature database at NCBI and querying the PUBMED central database using the ENTREZ search engine	3
16	Analysis of 3D structure of protein using RASMOL through command line.	3
17	Pair-wise and Multiple sequence alignment by using Clustal omega	3
18	Introduction of BioEdit. Effect of insertion INDEL from given amino acid	3
19	Pairwise and multiple sequence alignment of DNA using BioEdit.	3
20	Using online DNA editing tools	3

Course Outcomes: Students should be able to-

1. demonstrate proficiency in conducting genomic and proteomic experiments, including DNA sequencing, microarray analysis, and protein purification.
2. utilize bioinformatics tools for genome annotation, transcriptome analysis, and protein identification.
3. gain insights into the applications of genomics and proteomics in basic research and medical genetics.
4. effectively communicate experimental procedures, results, and interpretations through written reports and presentations.

References :

1. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. Molecular Biology of the Cell (6th ed.). Garland Science, 2014.
2. Brown, T. A. Genomes (3rd ed.). Oxford University Press, 2017.
3. Hood, L., & Taron, C. H. "Proteomics: The Path to the Future." Journal of Proteome Research 1, no. 1 (2002): 5-6.
4. Korf, B. R., Yandell, M. D., & Eichler, E. E. Introduction to Genomics (3rd ed.). W. H. Freeman, 2016.
5. Liebler, D. C. Introduction to Proteomics: Tools for the New Biology (2nd ed.). Humana Press, 2014.

6. Pawlowski, M., & Wojcik, J. "Proteomics in Laboratory Medicine." *Clinical Biochemistry* 50, no. 9 (2017): 555-564.
7. Schena, M. *Microarray Analysis* . Wiley-Liss, 2000.
8. Tisdale, E. J. "Protein Engineering through In Vivo Incorporation of Phenylalanine Analogues." *Methods in Enzymology* 510 (2012): 157-174.
9. Valenzuela, P., Martinez, V., & Santisteban, M. *DNA Sequencing: Principles and Methods* (2nd ed.). Humana Press, 2013.
10. Wilkins, M. R., & Williams, K. L. *Proteome Research: New Frontiers in Functional Genomics* . Springer, 2000.

Semester III
Minor Paper –I Syllabus

Semester III**BBIT-235 Statistical Analysis Methods**

Course Objective: Student should be able to....

1. Understand the techniques of Statistical data analysis.
2. Develop the skills of understanding of the data.
3. Learn the methods of various measures of Central Tendencies and Dispersion.
4. Study the theories and methods of sampling techniques.

Credits = 2	SEMESTER-III BBIT 235 :Statistical Analysis Methods	No. of hours per unit/credits
UNIT I	Meaning and scope of Statistical Analysis in Bioinformatics	(07)
	Definition and use of statistics in Bioinformatics , Scope and limitations of Statistics. Concept of variables and its types. Concept of data and its type. Concept of tabulation and its types. Concept of classification.	
UNIT II	Concept of Central Tendency	(08)
	Concept of central tendency, criteria for good measures of central tendency. Arithmetic Mean (A.M.), Weighted Mean, Combined Mean. Concept of Median, Mode and their properties. Computations of median and mode for ungrouped and grouped data.	
UNIT III	Measures of Dispersion	(06)
	Concept of Dispersion, Requirements of a good measure of dispersion. Absolute measures of dispersion, Relative measures of dispersion. Range, Mean Deviation, Standard Deviation (S.D.) and their relative measures. Variance, coefficient of variation and its use.	
UNITIV	Methods of Data Collection	(09)
	Methods of data collection. Concept of sample, sample size, sampling frame, population, population size, type of population. Concept of census method, sampling method, advantages of sampling over census method. Concept of simple random sampling, SRSWR and SRSWOR, Stratified sampling, Systematic sampling. Analysis methods: T-test, ANOVA, Chi-square test, Pearson Correlation, Regression analysis, Man-Whitney U Test, Kruskal-Wallis Test, Wilcoxon Signed Rank Test, McNemar's Test, Fischer's Exact test	

Course Outcomes: Students will be able to

1. Interpret the data and its techniques of statistical analysis.
2. Summarize the scope of data analysis in Industry and various sectors.
3. Execute the concept of Mathematical Averages.
4. Develop the skill to interpret the results.

Reference books :

1. Agarwal, B. L. Basic Statistics . 6th ed. Delhi: New Age International Private Ltd., 2013.
2. Daniel, W. W. Biostatistics: A Foundation for Analysis in the Health Sciences . 10th ed. Wiley, 2010.
3. Freund, J. E., & Perles, B. M. Modern Elementary Statistics . 12th ed. Pearson, 2010.
4. Goon, A. M., Gupta, M. K., & Das Gupta, B. Fundamentals of Statistics . 4th ed. Calcutta: The World Press Private Ltd., 1968.
5. Gupta, S. C. Fundamentals of Statistics . 7th ed. Mumbai: Himalaya Publishing House, 2018.
6. Gupta, S. C., & Kapoor, V. K. Fundamentals of Mathematical Statistics . 10th ed. New Delhi: Sultan Chand and Sons, 2000.
7. Gupta, S. P. Statistical Methods . 29th ed. Sultan Chand & Sons, 2009.
8. Johnson, R. A., & Wichern, D. W. Applied Multivariate Statistical Analysis . 6th ed. Pearson, 2007.
9. Keller, G. Statistics for Management and Economics . 10th ed. Cengage Learning, 2012.
10. Kenney, J. F., & Keeping, E. S. Mathematics of Statistics . 3rd ed. D. Van Nostrand Company, 1962.
11. Levin, R. I., Fox, J. A., & Foran, M. Elementary Statistics in Social Research . 12th ed. Pearson, 2016.
12. Mann, P. Introductory Statistics . 8th ed. Wiley, 2011.
13. Ott, R. L., & Longnecker, M. An Introduction to Statistical Methods and Data Analysis . 7th ed. Cengage Learning, 2015.
14. Sharma, J. K. Business Statistics . 3rd ed. Pearson, 2014.

Semester III**BBIP 236 : Lab course –III based on Statistics Analysis Methods**

Course Objectives : Student should be able to ..

1. familiarize with fundamental concepts and types of data in statistical analysis.
2. develop skills in organizing and presenting data through tabulation techniques.
3. compute and interpret various measures of central tendency accurately.
4. provide the ability to perform advanced statistical computations.

Credits = 2	SEMESTER-III BBIP 236 : Lab course –III based on Statistics Analysis Methods	No. of hours per unit/credits
	1. Introduction to Data Types and Classification	3
	2. Tabulation Techniques and Methods	3
	3. Calculating Measures of Central Tendency	3
	4. Computing Weighted Mean and Combined Mean Calculation	3
	5. Measures of Dispersion Computation	3
	6. Variance and Coefficient of Variation Analysis	3
	7. Exploring Methods of Data Collection	3
	8. Simple Random Sampling Implementation	3
	9. Determining Sample Size	3
	10. Census vs. Sampling Methods Discussion	3
	11. Stratified Sampling Procedure and Systematic Sampling Technique	3
	12. Analyzing Grouped and Ungrouped Data	3
	13. Interpretation of Statistical Measures	3
	14. Applying Central Tendency and Dispersion Measures	3
	15. Evaluating Sampling Methods	3
	16. Research Design and Data Collection Planning	3
	17. T – test, ANOVA, Chi square test	
	18. Pearson Correlation, Regression analysis, Man Whitney U test	

	19. Kruskal Wallis test , Wilcoxon Signed Rank test	
	20. Mc Nemars Test, Fishecrs Exact test	

Course Outcomes: Students will able to....

1. Understand the different types of data and their classification methods.
2. Apply tabulation techniques effectively to organize and summarize data.
3. Compute measures of central tendency accurately and interpret their significance.
4. Calculate weighted mean and combined mean to analyze complex datasets.

Reference Books :

1. Newbold, Paul, William L. Carlson, and Betty Thorne. *Statistics for Business and Economics* . Pearson, 2017.
2. Freedman, David, Robert Pisani, and Roger Purves. *Statistics* . W.W. Norton & Company, 2007.
3. Bulmer, M.G. *Principles of Statistics* . Dover Publications, 1979.
4. Cochran, William G. *Sampling Techniques* . Wiley, 1977.
5. Creswell, John W., and J. David Creswell. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* . SAGE Publications, 2017.
6. Devore, J. L., & Peck, R. (2015). *Statistics: The Exploration & Analysis of Data* (7th ed.). Cengage Learning.
7. Moore, D. S., & McCabe, G. P. (2018). *Introduction to the Practice of Statistics* (9th ed.). W. H. Freeman.
8. Montgomery, D. C., Runger, G. C., & Hubele, N. F. (2017). *Engineering Statistics* (5th ed.). Wiley.
9. Rosner, B. (2015). *Fundamentals of Biostatistics* (8th ed.). Cengage Learning.
10. Witte, R. S., & Witte, J. S. (2019). *Statistics* (11th ed.). Wiley.

Semester –IV

B.Sc.II Semester -IV**Course Code – BBIT 241: Python programming for Bioinformatics**

Course Objectives : Student should be able to,

1. Acquire solid foundation in Python programming language and basic concepts.
2. Introduce control structures and functions in Python, enabling them to write efficient and structured code.
3. Familiarize with file handling techniques and Python libraries such as NumPy, Pandas, and Matplotlib for data manipulation and visualization.
4. Practice the necessary skills to perform sequence analysis and bioinformatics tasks using the Biopython library.

Credits = 2	SEMESTER-IV BBIT 241 : Python Programming for Bioinformatics	No. of hours per unit/credits
UNIT I	Introduction to Python Programming	(07)
	Introduction to programming concepts, Overview of Python programming language, Setting up Python environment (interpreter, IDEs), Writing and executing Python scripts, Basic datatypes: integers, floats, strings, and Booleans, Variables and data assignment, Basic input/output operations.	
UNIT II	Control Structures and Functions	(08)
	Conditional statements: if, else, if else, Iterative statements: for loops, while loops, Nested loops and loop control statements: break, continue, Introduction to functions: Function definition and invocation, Parameters and return values, Scope and lifetime of variables	
UNIT III	File Handling and Introduction to Libraries	(06)
	Reading from and writing to files, Opening, closing, and handling file objects, File modes: read mode, write mode, append mode, Error handling with try-except blocks, Introduction to Python libraries: NumPy, Pandas, Matplotlib, Importing and using libraries in Python scripts	
UNIT IV	Biopython	(09)
	Sequence Analysis with Biopython: Introduction to Biopython library for bioinformatics. Reading, writing, and parsing sequence files (FASTA, FASTQ), Performing sequence alignment and similarity searches, Extracting sequence features and annotations, Calculating sequence statistics and analyzing genomic data	

Course Outcomes: Students will be able to

1. Understand and apply fundamental programming concepts in Python, including variables, data types, and input/output operations.
2. Write Python scripts utilizing control structures such as conditional statements and

- loops for efficient program execution.
3. Demonstrate proficiency in file handling techniques and error handling mechanisms in Python programming.
 4. Perform sequence analysis tasks using the Biopython library, including reading, writing, and parsing sequence files, sequence alignment, and annotation extraction.

Reference books :

1. Chang, Jeff, Brad Chapman, Iddo Friedberg, Thomas Hamelryck, Michiel de Hoon, and Peter Cock. Biopython Tutorial and Cookbook . O'Reilly Media, 2020.
2. Jones, Martin. Python for Biologists: A complete programming course for beginners (2nd ed.). CreateSpace Independent Publishing Platform, 2013.
3. Lutz, Mark. Learning Python (5th ed.). O'Reilly Media, 2013.
4. Matthes, Eric. Python Crash Course . No Starch Press, 2019.
5. Model, Mitchell L. Bioinformatics Programming Using Python: Practical Programming for Biological Data . O'Reilly Media, 2009.
6. Stevens, Tim J., and Wayne Boucher. Python Programming for Biology: Bioinformatics and Beyond (2nd ed.). Cambridge University Press, 2015.
7. Sweigart, Al. Automate the Boring Stuff with Python: Practical Programming for Total Beginners . No Starch Press, 2015.
8. Tisdall, James. Beginning Python for Bioinformatics (2nd ed.). O'Reilly Media, 2009. Antao, Tiago. Bioinformatics with Python Cookbook (2nd ed.). Packt Publishing, 2015.
9. Zelle, John. Python Programming: An Introduction to Computer Science (3rd ed.). Franklin, Beedle & Associates Inc., 2016.

B.Sc.II Semester -IV**Course Code – BBIT 242 :Biological Sequence and Protein Sequence Analysis****Course Objectives: Students should be able to,**

1. Learn fundamental concepts of sequence similarity, identity, and homology, and gain proficiency in using scoring matrices for nucleic acid and protein sequences.
2. Develop skills in pairwise and multiple sequence alignments using algorithms like Needleman-Wunsch, Smith-Waterman, CLUSTALW, and PileUp.
3. Gain hands-on experience with tools for filtering repetitive sequences, gene identification, promoter prediction, and database searching by sequence.
4. Know the sequence patterns, motifs, and profiles, and conduct profile-based database searches using various algorithms.

Credits (Total Credits 2)	SEMESTER – III BBIT 242 : Biological Sequence and Protein Sequence Analysis	No. of hours per unit
Unit - I	Sequence Analysis: Basic concepts of sequence similarity, identity and homology, definitions of homologues, orthologues, paralogues and xenologues Scoring matrices: basic concept of a scoring matrix, Matrices for nucleic acid and proteins sequences, PAM and BLOSUM series, matrix derivation methods and principles. Repeats: Tandem and Interspersed repeat finding, Motifs, consensus, position weight matrices	(06)
Unit – II	Pairwise sequence alignment -Basic concepts of sequence alignment, gap penalties, Needleman and Wunsch, Smith and Waterman algorithms for pairwise alignments and application in Nucleic acid and protein sequences alignments. Multiple sequence alignments (MSA) The need for MSA, basic concepts of various approaches for MSA (e.g. progressive, hierarchical etc.). Algorithm of CLUSTALW and PileUp and application, concept of dendrogram and its interpretation, Use of HMM- based Algorithm for MSA (eg. SAM method)	(08)
Unit – III	Genomic Analysis. Filtering of repetitive sequences using Repeatmasker. Exon and gene identification: Genscan. Promoter identification. Ppnn web-site. The Transfac databases. Identifying Transfac Profiles with Match. Probabilities of TFBSs with RSAT	(08)

	cDNA- Genomic DNA alignment. The Santa Cruz Human and Mouse Genome Map web-site. The NCBI Genomic web-sites. Database searching by sequence- Filtering of low-complexity and repetitive sequences – Seg and Dust.	
Unit - IV	Sequence patterns and profiles: Basic concept and definition of sequence patterns, motifs and profiles, various types of pattern representations viz, consensus, regular expression (Prosite-type) and sequence profiles, profile-based database searches using PSL-BLAST. analysis and interpretation of profile-based searches. Algorithms for derivation and searching sequence patterns MEME, PHL-BLAST, SCanProsite and PRATT. Algorithms for generation of sequence profiles Profile Analysis method of Gribskov, HMMer, PSI-BLAST	(08)

Course Outcomes: Students should be able to,

1. Understand and differentiate between sequence-related terms, and apply scoring matrices effectively in sequence analysis tasks.
2. Perform accurate pairwise and multiple sequence alignments, and interpret alignment results in biological contexts.
3. Utilize genomic analysis tools for sequence filtering, gene and promoter identification, and efficient database searching.
4. Analyze sequence patterns and profiles, conduct profile-based searches, and interpret results for practical applications.

Reference Books:

1. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. "Basic local alignment search tool." *Journal of Molecular Biology* 215, no. 3, 1990: 403–410.
2. Attwood, T. K., & Parry-Smith, D. J. *Introduction to Bioinformatics*. Pearson Education, 2001.
3. Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., ... & Noble, W. S. "MEME SUITE: Tools for motif discovery and searching." *Nucleic Acids Research* 37, suppl. 2, 2009: W202–W208.
4. Durbin, R., Eddy, S. R., Krogh, A., & Mitchison, G. *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. Cambridge University Press, 1998.

5. Gribskov, M., McLachlan, A. D., & Eisenberg, D. "Profile analysis: Detection of distantly related proteins." *Proceedings of the National Academy of Sciences of the United States of America* 84, no. 13, 1987: 4355–4358.
6. Krogh, A., Brown, M., Mian, I. S., Sjolander, K., & Haussler, D. "Hidden Markov models in computational biology: Applications to protein modeling." *Journal of Molecular Biology* 235, no. 5, 1994: 1501–1531.
7. Lesk, A. M. *Introduction to Bioinformatics*. Oxford University Press, 2008.
8. Mount, D. W. *Bioinformatics: Sequence and Genome Analysis*. Cold Spring Harbor Laboratory Press, 2004.
9. Pearson, W. R. "Flexible sequence similarity searching with the FASTA3 program package." *Methods in Molecular Biology* (Clifton, N.J.) 132, 2000: 185–219.
10. Thompson, J. D., Higgins, D. G., & Gibson, T. J. "CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." *Nucleic Acids Research* 22, no. 22, 1994: 4673–4680.

SEMESTER – IV**Course code BBIP 243 : Lab course based on BBIT 241 – Python programming for Bioinformatics**

Course Objective: Student will be able to

1. Introduce students to the fundamental concepts of Python programming language and its applications in bioinformatics.
2. Provide hands-on experience in utilizing Python libraries such as NumPy, Pandas, Matplotlib, and Biopython for data analysis and manipulation.
3. Familiarize students with file handling techniques and error handling mechanisms in Python programming.
4. Equip students with the necessary skills to perform sequence analysis tasks and genomic data analysis using Biopython library.

Credits = 2	SEMESTER-IV BBIP 243: Python programming for Bioinformatics	No. of hours per unit/credits
	1. Setting Up Python Environment	3
	2. Exploring Python Basics	3
	3. Basic Input/Output Operations	3
	4. Understanding Variables and Data Types	3
	5. Introduction to Programming Concepts	3
	6. Implementing Conditional Statements	3
	7. Using Iterative Statements	3
	8. Understanding Functions	3
	9. Working with Nested Loops	3
	10. Scope and Lifetime of Variables	3
	11. Reading from and Writing to Files	3
	12. Handling File Objects	3
	13. Introduction to NumPy	3
	14. Introduction to Pandas	3
	15. Introduction to Matplotlib	3
	16. Introduction to Biopython	3
	17. Performing Sequence Alignment	3
	18. Extracting Sequence Features	3
	19. Calculating Sequence Statistics	3
	20. Analyzing Genomic Data	3

Course Outcomes: Students should be able to..

1. Understand the syntax and basic principles of Python programming language.
2. Apply Python programming concepts to solve bioinformatics-related problems effectively.

3. Utilize Python libraries such as NumPy, Pandas, and Matplotlib for data manipulation, analysis, and visualization.
4. Perform sequence analysis tasks and genomic data analysis using Biopython library.

Reference Books:

1. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. "Basic local alignment search tool." *Journal of Molecular Biology* 215, no. 3 (1990): 403–410.
2. Attwood, T. K., & Parry-Smith, D. J. *Introduction to Bioinformatics*. Pearson Education, 2001.
3. Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. "MEME SUITE: Tools for motif discovery and searching." *Nucleic Acids Research* 37, suppl. 2 (2009): W202–W208.
4. Durbin, R., Eddy, S. R., Krogh, A., & Mitchison, G. *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. Cambridge University Press, 1998.
5. Gribskov, M., McLachlan, A. D., & Eisenberg, D. "Profile analysis: Detection of distantly related proteins." *Proceedings of the National Academy of Sciences of the United States of America* 84, no. 13 (1987): 4355–4358.
6. Krogh, A., Brown, M., Mian, I. S., Sjolander, K., & Haussler, D. "Hidden Markov models in computational biology: Applications to protein modeling." *Journal of Molecular Biology* 235, no. 5 (1994): 1501–1531.
7. Lesk, A. M. *Introduction to Bioinformatics*. Oxford University Press, 2008.
8. Mount, D. W. *Bioinformatics: Sequence and Genome Analysis*. Cold Spring Harbor Laboratory Press, 2004.
9. Pearson, W. R. "Flexible sequence similarity searching with the FASTA3 program package." *Methods in Molecular Biology (Clifton, N.J.)* 132 (2000): 185–219.

10. Thompson, J. D., Higgins, D. G., & Gibson, T. J. "CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." *Nucleic Acids Research* 22, no. 22 (1994): 4673–4680.

SEMESTER – IV**Course code BBIP 243 : Lab course based on BBIT 242 Biological Sequence and Protein Structure analysis****Course Objectives: Students will be able to ..**

1. Understand the fundamental principles and techniques involved in biological sequence analysis.
2. Develop proficiency in utilizing various bioinformatics tools and databases for sequence and protein structure analysis.
3. Gain practical experience in performing pairwise and multiple sequence alignments to identify similarities and differences among sequences.
4. Acquire knowledge and skills in predicting protein structures, motifs, domains, and binding sites using computational methods.

Credits (Total Credits 2)	SEMESTER – IV BBiP 243: Biological Sequence and Protein Structure analysis List of Practical (20)	No. of hours (60)
1	Sequence Databases: EMBOSS, NCBI ToolKit, Expamy tools	3
2	Search for Advanced tools against Databases: i. BLAST ii. FASTA	3
3	Pair wise alignment using Dot Plot	3
4	Pair wise alignment using Global alignment methods tools	3
	Pair wise alignment using Local alignment methods tools	3
5	Multiple sequence alignment using : Clustal, Dialign, Multalign	3
6	Primary and secondary structure prediction methods using GOR Method	3
7	Primary and secondary structure prediction methods using PSI-pred Method	3
8	Primary and secondary structure prediction methods using Chou-Fasman method	3
9	Binding site identification	3
10	Study of Sequence patterns and profiles: Generation of sequence profiles using PSI-BLAST	3
11	Derivation of and searching sequence patterns using MEME/MAST	3
12	Derivation of and searching sequence patterns using PHI-BLAST	3
13	Derivation of and searching sequence patterns using	3

	SCanProsite	
14	Derivation of and searching sequence patterns using PRATT	3
15	Protein motif and domain analysis using : MEME/MAST, eMotif, InterproScan	3
16	Protein motif and domain analysis using ProSite, ProDom, Pfam	3
17	Protein motif and domain analysis using ProSite	3
18	Protein motif and domain analysis using ProDom,	3
19	Protein motif and domain analysis using Pfam	3
20	Phylogentic analysis Mega/ Paup/phylip	3

Course Outcomes: Students should be able to,

1. Efficient utilization of sequence databases and bioinformatics tools for sequence analysis.
2. Proficiency in pairwise and multiple sequence alignments to identify similarities and conserved motifs.
3. Ability to predict protein structures and identify binding sites using computational methods.
4. Skill in deriving sequence patterns and motifs for functional annotation and understanding protein structure-function relationships.

Reference Books :

1. Attwood, Teresa K., and David J. Parry-Smith. Introduction to Bioinformatics . 1st ed. Pearson, (2001).
2. Baxevanis, Andreas D., and B. F. Francis Ouellette, eds. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins . 3rd ed. Wiley-Blackwell, (2005).
3. Durbin, Richard, et al. Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids . Cambridge University Press, (1998).
4. Gusfield, Dan. Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology . Cambridge University Press, (1997).
5. Lesk, Arthur M. Introduction to Bioinformatics . 4th ed. Oxford University Press, (2014).
6. Mount, David W. Bioinformatics: Sequence and Genome Analysis . 2nd ed. Cold Spring Harbor Laboratory Press, (2004).
7. Ranganathan, Shoba. Introduction to Bioinformatics . CRC Press, (2018).

8. Setubal, João C., and JoãoMeidanis. Introduction to Computational Molecular Biology . PWS Publishing Company, (1997).
9. Thompson, Julie D., et al. Bioinformatics: Sequence and Genome Analysis . Cold Spring Harbor Laboratory Press, (2003).
10. Tisdall, James D. Beginning Perl for Bioinformatics . O'Reilly Media, (2001).

Semester IV

Minor paper – II syllabus

Semester IV**BBIT 245: Data Visualization using R**

Course Objective: Student will be able to

1. Understand the importance of data visualization in data analysis and decision-making processes.
2. Learn the principles of effective data visualization design.
3. Explore different types of data visualizations and their applications.
4. Gain hands-on experience with popular data visualization tools and libraries.

Credits = 2	SEMESTER-IV BBIT 245:Data Visualization using R	No. of hours per unit/credits
UNIT I	Introduction to Data Visualization	(09)
	Introduction to Data Visualization, Importance of data visualization, Historical overview of data visualization, Role of data visualization in various fields (Bioinformatics, Physical Science, Business)	
UNIT II	Principles of Data Visualization	(08)
	Principles and visual perception, Design principles for effective visualizations, Choosing the right chart types for different types of data. Data cleaning and preprocessing: Data aggregation and summarization.	
UNIT III	Basic Visualization Tools	(07)
	Scatter plots, Line charts, Bar charts, Histograms, Pie charts, Donut charts, Box plots and Violin plots.	
UNIT IV	Specialized Visualization Tools	(06)
	Word Clouds, Radar Charts, Waffle Charts, Heat maps and correlation plots, Geospatial visualizations.	

Course Outcomes: Students should be able to..

1. Understand to how to analyze the data visually
2. Illustrate the scope of data visualization in various sectors.
3. Apply the concept of appropriate visualization techniques.
4. Develop the skill of interpreting the results and compelling visualization to communicate insights effectively

Reference books :

1. Tufte, Edward. *The Visual Display of Quantitative Information*. Graphics Press, 1983.
2. Yau, Nathan. *Data Points: Visualization That Means Something*. Wiley, 2013.
3. Nussbaumer Knaflic, Cole. *Storytelling with Data: A Data Visualization Guide for Business Professionals*. Wiley, 2015.
4. Murray, Scott. *Interactive Data Visualization for the Web*. O'Reilly Media, 2013.
5. Wickham, Hadley. *ggplot2: Elegant Graphics for Data Analysis*. Springer, 2016.
6. Chang, Winston. *R Graphics Cookbook*. O'Reilly Media, 2013.
7. Cook, Dianne, and Deborah F. Swayne. *Interactive and Dynamic Graphics for Data Analysis: With R and Ggobi*. Springer, 2007.
8. Wickham, Hadley, Winston Chang, Joe Cheng, and Barbara Borges Ribeiro. *Mastering Shiny*. O'Reilly Media, 2020.
9. Healy, Kieran. *Data Visualization: A Practical Introduction*. Princeton University Press, 2018.
11. Wickham, Hadley, and Garrett Grolemund. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media, 2017.
12. Khandekar, Narayan. *Data Visualization with R: 100 Examples*. Packt Publishing, 2016.
13. Jain, Dipanwita, and Panchajanya Paul. *Data Visualization with R*. Apress, 2019.
14. Bagh, Ramnath Vaidyanathan. *Interactive and Animated Graphics for Data Analysis: With R and Ggobi*. CRC Press, 2017.
15. Kaur, Ankita, and Sharanjit Singh. *R Data Visualization Recipes: A Cookbook with Hands-On Recipes to Create Interactive, Stunning, and Informative Visualizations*. Packt Publishing, 2018.
16. Raj, Abhik. *Learning Data Visualization with R*. Packt Publishing, 2016.
17. Kulkarni, Hrishikesh D. *Data Visualization using R*. CRC Press, 2019.

Semester IV**BBIP 246 – Lab Course based on BBIT 245 Data Visualization using R**

CourseObjective: Student will be able to..

1. Introduce the fundamental concepts and principles of data visualization.
2. Creating a variety of visualizations to represent different types of data effectively.
3. Familiarize with various visualization techniques and tools available in R.
4. Apply data visualization techniques to real-world datasets and communicate insights effectively.

Credits = 2	SEMESTER-IV BBIP 246: Data Visualization using R	No. of hours per unit/credits
	1. Introduction to Data Visualization	3
	2. Visual Perception and Design Principles	3
	3. Data Cleaning and Preprocessing	3
	4. Creating Scatter Plots in R	3
	5. Line Charts and Time Series Visualization in R	3
	6. Bar Charts and Histograms in R	3
	7. Pie Charts and Donut Charts in R	3
	8. Box Plots and Violin Plots in R	3
	9. Creating Word Clouds in R	3
	10. Radar Charts in R	3
	11. Waffle Charts in R	3
	12. Heat Maps and Correlation Plots in R	3
	13. Geospatial Visualizations in R	3
	14. Exploring R Packages for Data Visualization	3
	15. Customizing Visualizations in R	3
	16. Interactive Visualizations in R	3
	17. Data Visualization Best Practices in R	3
	18. Data Storytelling with R Visualizations	3
	19. Dashboard Design in R	3
	20. Data Visualization Case study	3

Course Outcomes: Students should be able to..

1. Understand the importance and role of data visualization across various domains.
2. Demonstrate proficiency in using R for creating different types of visualizations such as scatter plots, line charts, histograms, etc.
3. Apply principles of visual perception and design to develop effective visualizations.
4. Design and present interactive dashboards and narratives using R visualizations.

Reference Books :

1. Tufte, Edward. The Visual Display of Quantitative Information. Graphics Press, 1983.
2. Yau, Nathan. Data Points: Visualization That Means Something. Wiley, 2013.
3. Nussbaumer Knaflic, Cole. Storytelling with Data: A Data Visualization Guide for Business Professionals. Wiley, 2015.
4. Murray, Scott. Interactive Data Visualization for the Web. O'Reilly Media, 2013.
5. Wickham, Hadley. ggplot2: Elegant Graphics for Data Analysis. Springer, 2016.
6. Chang, Winston. R Graphics Cookbook. O'Reilly Media, 2013.
7. Cook, Dianne, and Deborah F. Swayne. Interactive and Dynamic Graphics for Data Analysis: With R and Ggobi. Springer, 2007.
8. Wickham, Hadley, Winston Chang, Joe Cheng, and Barbara Borges Ribeiro. Mastering Shiny. O'Reilly Media, 2020.
10. Healy, Kieran. Data Visualization: A Practical Introduction. Princeton University Press, 2018.
11. Wickham, Hadley, and Garrett Grolemund. R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. O'Reilly Media, 2017.
12. Khandekar, Narayan. Data Visualization with R: 100 Examples. Packt Publishing, 2016.
13. Jain, Dipanwita, and Panchajanya Paul. Data Visualization with R. Apress, 2019.
14. Bagh, Ramnath Vaidyanathan. Interactive and Animated Graphics for Data Analysis: With R and Ggobi. CRC Press, 2017.
15. Kaur, Ankita, and Sharanjit Singh. R Data Visualization Recipes: A Cookbook with Hands-On Recipes to Create Interactive, Stunning, and Informative Visualizations. Packt Publishing, 2018.
16. Raj, Abhik. Learning Data Visualization with R. Packt Publishing, 2016.
17. Kulkarni, Hrishikesh D. Data Visualization using R. CRC Press, 2019.

Syllabus of Vocational Skill Course (VSC) for B.Sc.II

BBiTVSC1: Bioinformatic Methods I**Course Objectives: Student will be able to**

1. Understand the basic principles of bioinformatics
2. Apply Bioinformatic methods to analyze biological data
3. Interpret and critically evaluate bioinformatics research literature
4. Utilize bioinformatics tools and databases for biological research

Credits (Total Credits 2)	SEMESTER-III BBiTVSC1: Bioinformatic methods I	No. of hours per unit/credits
1.	Introduction to NCBI Database: Familiarize students with the NCBI website, its navigation, and various resources available.	
2.	Exploring Biological Data Sources: Discuss different types of biological data available in NCBI Database and their sources.	
3.	File Formats in NCBI Database: Understand and work with common file formats like FASTA, GenBank, and PDB used in NCBI Database.	
4.	BlastP Search: Perform a BlastP search to find homologous protein sequences and analyze the results	
5.	PSI-Blast: Conduct a PSI-Blast search to identify distantly related protein sequences and understand its iterative process.	
6.	6. Translated Blast: Perform a Translated Blast search to find nucleotide sequences that translate into a given protein sequence	
7.	Basic Sequence Analysis: Introduce fundamental concepts such as sequence similarity, homology, and sequence alignment.	
8.	Needleman-Wunsch Algorithm: Implement the Needleman-Wunsch algorithm for global sequence alignment.	
9.	Smith-Waterman Algorithm: Implement the Smith-Waterman algorithm for local sequence alignment.	
10.	Pairwise Sequence Alignment: Practice pairwise sequence alignment using both global and local alignment algorithms	
11.	Multiple Sequence Alignment: Perform multiple sequence alignment using tools like ClustalW or MUSCLE and analyze the results	
12.	Sequence Similarity and Homology: Understand the relationship between sequence similarity and evolutionary homology.	

13.	Phylogenetic Analysis: Construct phylogenetic trees using distance-based or character-based methods and interpret the evolutionary relationships.	
14.	Phylogenetic Trees Visualization: Visualize phylogenetic trees using software like MEGA or FigTree and interpret the tree topology	
15.	Sequence Visualization with UCSC Genome Browser: Explore genomic data using the UCSC Genome Browser and understand its features.	
16.	. Functional Prediction with Prosite: Predict protein function using Prosite patterns and motifs.	
17.	Functional Prediction with Pfam: Use Pfam database to identify conserved protein domains and infer function.	
18.	Functional Prediction with STRING: Analyze protein-protein interaction networks using STRING database and predict functional associations.	
19.	NGS Technologies Overview: Discuss the significance of Next Generation Sequencing (NGS) technologies in genomics research and medical diagnostics.	
20.	NGS Applications: Perform hands-on exercises on various NGS applications such as whole-genome sequencing, RNA-seq, ChIP-seq, and metagenomics, and analyze the generated data.	

Course Outcomes: Students will be able to..

1. Navigate and utilize various bioinformatics tools and databases effectively.
2. Proficient in performing pairwise and multiple sequence alignments using appropriate algorithms.
3. Interpret the results to infer sequence homology and evolutionary relationships.
4. Apply their knowledge to know techniques in NGS

References:

1. Arthur M. Lesk "Introduction to Bioinformatics" Publisher OUP Oxford. 2019
2. David W. Mount "Bioinformatics: Sequence and Genome Analysis" Publisher Cold Spring Harbor Laboratory Press, U.S.2 nd Edition. 2004
3. Lesk, Arthur M. Introduction to Bioinformatics. Oxford University Press, 2019.
4. Lesk, Arthur M. Introduction to Genomics. Oxford University Press, 2017.

5. Marketa Zvelebil and Jeremy O. Baum(2007)"Understanding Bioinformatics" ,Publisher Garland Science, Ist Edition
- 6.Mount, David W. Bioinformatics: Sequence and Genome Analysis. 2nd ed. Cold Spring Harbor Laboratory Press, 2004.
7. Neil C. Jones and Pavel A. Pevzner(2004) "An Introduction to Bioinformatics Algorithms" ,MIT Press. 18. Jin Xiong (2007)"Essential Bioinformatics" Cambridge University Press,Ist Edition
- 8.Pevsner, Jonathan. Bioinformatics and Functional Genomics. Wiley, 2015.
- 9.Xiong, Jin. Essential Bioinformatics. Cambridge University Press, 2007.
- 10.Zvelebil, Marketa, and Jeremy O. Baum. Understanding Bioinformatics. Garland Science, 2007.

BBiTVSC1: Bioinformatics methods II**Course Objectives: Student should be able to**

1. Navigate and utilize major protein databases effectively
2. Analyze protein structures and sequences using bioinformatic tools
3. Interpret and extract meaningful information from protein database resources
4. Apply computational methods to study protein-ligand interactions

Credits (Total Credits 2)	SEMESTER-III BBiT VSC1: Bioinformatics methods I	No. of hours per unit/credits
1.	Familiarize with PDB, UniProt, Pfam; explore interfaces and search functions.	3
2.	Investigate the importance of protein databases through case studies.	3
3.	Analyze data formats used in protein databases; convert data between formats.	3
4.	Navigate PDB, download protein structures; use PyMOL, Chimera for analysis.	3
5.	Annotate protein functions with Gene Ontology terms.	3
6.	Predict functional domains in proteins using Pfam or InterPro.	3
7.	Explore protein interaction databases like STRING and BioGRID.	3
8.	Retrieve interaction data; analyze and visualize protein-protein interaction networks.	3
9.	Review recent research utilizing protein databases and bioinformatics tools	3
10.	Integrate protein structure and interaction data from multiple databases.	3
11.	Implement a basic algorithm for predicting protein-protein interactions or functional domains.	3
12.	Compare features and data content of different protein databases.	3
13.	Create a tutorial demonstrating how to use a protein database or bioinformatics tool.	3
14.	Organize an interactive workshop for hands-on practice.	3
15.	Design a quiz or test assessing understanding of protein databases and analysis.	3
16.	Host a discussion forum for sharing insights and discussing challenges.	3
17.	Invite a guest speaker to discuss their work in protein database analysis.	3

18.	Assign a project where students address a biological question using protein databases.	3
19.	Develop practical skills in searching, retrieving, and analyzing protein data.	3
20.	Apply knowledge gained to real-world scenarios through case studies and projects.	3

Course Outcomes: Students will be able to..

1. demonstrate proficiency in navigating major protein databases
2. Develop a solid understanding of protein structure and function, including protein domains, motifs, folds, and their relationship to biological function.
3. Gain proficiency in annotating protein functions using bioinformatic tools and resources such as Gene Ontology (GO), InterPro.
4. Analyze protein-protein interaction networks using databases like STRING and BioGRID, and apply network analysis techniques to understand biological pathways and interactions.

References:

1. Attwood, T. K., and Parry-Smith, D. J. Introduction to Bioinformatics. Pearson Education, 2001.
2. Baxevanis, Andreas D., and B. F. Francis Ouellette. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. Wiley, 2001.
3. Durbin, Richard, Sean R. Eddy, Anders Krogh, and Graeme Mitchison. Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids. Cambridge University Press, 1998.
4. Jones, Neil C., and Pavel A. Pevzner. An Introduction to Bioinformatics Algorithms. MIT Press, 2004.
5. Lesk, Arthur M. Introduction to Bioinformatics. Oxford University Press, 2019.
6. Lesk, Arthur M. Introduction to Genomics. Oxford University Press, 2017.
7. Mount, David W. Bioinformatics: Sequence and Genome Analysis. 2nd ed. Cold Spring Harbor Laboratory Press, 2004.
8. Pevsner, Jonathan. Bioinformatics and Functional Genomics. Wiley, 2015.
9. Xiong, Jin. Essential Bioinformatics. Cambridge University Press, 2007.
10. Zvelebil, Marketa, and Jeremy O. Baum. Understanding Bioinformatics. Garland Science, 2007.

Syllabus for Skill Enhancement Course SEC

B.Sc.II Semester III**Syllabus for Skill Enhancement Course****Course Code- BBitSEC 2 : Web-based Programming****Course Objectives: Students will be able to...**

1. Understand the foundational concepts of the Internet and web programming.
2. Develop proficiency in HTML and HTML5 for creating structured web documents.
3. Master Cascading Style Sheets (CSS) and JavaScript for enhancing web page styling and interactivity.
4. Gain practical experience in server-side scripting using PHP and database connectivity with MySQL.

Credits (Total Credits 2)	SEMESTER – III BBitSEC 2 : Web-based Programming	No. of hours per unit
1.	Explore the history of the Internet and differentiate between Internet and Intranet.	3
2.	Set up and configure a local web server using WAMP/XAMPP.	3
3.	Create a simple web page using HTML and understand the basic structure of a HTML document.	3
4.	Use web authoring tools to design and build a basic website.	3
5.	Create a webpage with HTML5 semantic elements (header, footer, article, section).	3
6.	Design a webpage layout using tables and lists.	3
7.	Implement hyperlinks and navigation menus on a webpage.	3
8.	Incorporate multimedia elements such as audio and video using HTML5.	3
9.	Style a webpage using CSS properties for background, text formatting, and font control.	3
10.	Implement JavaScript for client-side form validation and event handling.	3
11.	Create dialog boxes and interactive elements using JavaScript.	3
12.	Develop a responsive webpage layout using CSS for different screen sizes	3

13.	Set up a PHP server and establish connectivity with MySQL database	3
14.	Create a PHP script to capture form data and perform basic database operations (insert, delete, update, select).	3
15.	Implement state management techniques such as cookies and sessions in a web application.	3
16.	Develop a PHP script to handle string manipulation and pattern matching using regular expressions.	3
17.	Understand PHP server variables and query strings for URL rewriting.	3
18.	Implement form handling with PHP using hidden fields and query parameters.	3
19.	Execute SQL queries with PHP for database operations and handle query results.	3
20.	Implement authentication and authorization mechanisms in a web application using PHP and MySQL.	3

Course Outcomes: Students should be able to..

1. Explain the history of the Internet, understand web system architecture, and distinguish between Internet and Intranet.
2. Demonstrate proficiency in creating well-structured HTML/HTML5 documents, including the use of semantic elements, lists, hyperlinks, tables, forms, multimedia elements, and graphics.
3. Showcase proficiency in CSS styling techniques, JavaScript programming for client-side interactivity, event handling, and form validations.
4. Develop dynamic web applications using PHP, including handling form data, working with variables, arrays, functions, and database operations with MySQL.

References :

1. "HTML and CSS: Design and Build Websites" by Jon Duckett
2. "JavaScript and JQuery: Interactive Front-End Web Development" by Jon Duckett
3. "PHP and MySQL Web Development" by Luke Welling and Laura Thomson
4. "Learning PHP, MySQL & JavaScript: With jQuery, CSS & HTML5" by Robin Nixon
5. "Web Design with HTML, CSS, JavaScript and jQuery Set" by Jon Duckett
6. "Beginning HTML, XHTML, CSS, and JavaScript" by Jon Duckett
7. "JavaScript: The Good Parts" by Douglas Crockford
8. "PHP Solutions: Dynamic Web Design Made Easy" by David Powers
9. "MySQL Crash Course" by Ben Forta
10. "Web Development and Design Foundations with HTML5" by Terry Felke-Morris

B.Sc.II Semester III**Syllabus for Skill Enhancement Course****Course Code- BBitSEC 2 : Programming in Perl & Bioperl****Course Objectives: Students will be able to...**

1. Familiarize students with the fundamentals of Perl programming language and its significance in Bioinformatics.
2. Acquaint with basic constructs and data types in Perl for effective programming.
3. Enable students to utilize advanced features of Perl such as regular expressions, subroutines, and file handling.
4. Introduce Object-Oriented Perl and its application in CGI programming for web-based bioinformatics applications.

Credits (Total Credits 2)	SEMESTER – III BBitSEC 3 : Programming in Perl & Bioperl	No. of hours per unit
1.	Install Perl: Guide students through the installation process of Perl and Bioperl.	3
2.	Basic Scripting: Write a Perl script to print "Hello, World!" to the console	3
3.	Scalar Data: Create variables for integers, floats, and strings, and perform arithmetic operations on them.	3
4.	Arrays: Declare arrays, add elements, access elements by index, and print array contents.	3
5.	Hashes: Initialize hashes, add key-value pairs, access values by keys, and print hash contents.	3
6.	Control Statements: Implement conditional statements (if-else) and loops (for, while) in Perl.	3
7.	Basic I/O: Write scripts to read input from users and print output to the console.	3
8.	Regular Expressions: Use regular expressions to match patterns in strings and perform substitutions.	3
9.	Subroutines: Define and call subroutines for reusable code blocks.	3
10.	File Handling: Open, read, write, and close files using file handles	3

11.	Object-Oriented Perl: Create classes and objects, define methods, and access object properties.	3
12.	CGI Programming: Develop a basic CGI script to process form input and display results	3
13.	DNA Sequence Storage: Write a program to store DNA sequences in variables.	3
14.	Transcription: Convert DNA sequences to RNA using Perl.	3
15.	Protein File Parsing: Read protein data from a file and display relevant information.	3
16.	Motif Identification: Implement a script to find specific motifs in DNA sequences.	3
17.	Nucleotide Counting: Count occurrences of each nucleotide in a DNA sequence.	3
18.	String Manipulation: Perform string operations like concatenation, splitting, and joining.	3
19.	Random DNA Generation: Generate random DNA sequences of specified lengths.	3
20.	BioPerl Integration: Utilize BioPerl modules to parse and analyze biological data files.	3

Course Outcomes: Students should be able to..

1. Demonstrate proficiency in basic Perl programming constructs and understand their relevance in Bioinformatics.
2. Apply regular expressions and advanced features of Perl for data manipulation and analysis in bioinformatics.
3. Develop Perl scripts for handling biological data, including file parsing and manipulation.
4. Create object-oriented Perl programs and utilize CGI programming for building web-based bioinformatics applications.

References :

1. Schwartz, Randal L., Tom Phoenix, and Brian D. Foy. Learning Perl. .
2. Lee, James. Beginning Perl.
3. Wall, Larry, Tom Christiansen, and Jon Orwant., Programming Perl..
4. Model, Mitchell L. Bioinformatics Programming Using Python: Practical Programming for Biological Data..

5. Buffalo, Vince. Bioinformatics Data Skills: Reproducible and Robust Research with Open Source Tools..
6. Birney, Ewan, Brian D. Joseph, and Damian Smedley. Bioperl: Perl Tools for Bioinformatics..
7. Tisdall, James D. Mastering Perl for Bioinformatics..
8. Christiansen, Tom, and Nathan Torkington. Perl Cookbook..
9. Tisdall, James. Beginning Perl for Bioinformatics..
10. Schwartz, Randal L., Brian D. Foy, and Tom Phoenix. Learning Perl, 7th Edition..

B.Sc.II Bioinformatics

Semester III

Syllabus for Value Education Course

(VEC)

Semester III

Syllabus For Value Education Course (VEC)

BBiT-VEC-II: Environmental Awareness for Bioinformatics

Course Objectives: Student will be able...

1. Understand the Environmental Issues
2. Discuss the role of Computer in creation of environmental issues
3. Implement the Environmental Laws
4. Identify the Computational Sustainability

Credits (Total Credits 2)	BBiT-VEC-II Environmental Awareness for Bioinformatics	No. of hoursperu nit/credits
Unit I:	Environmental issues	(8)
	Pollution (Air, water and Land), Fresh-water overuse, Natural disasters, Fuel and Energy shortage due to overuse, Increase in wasteland, Biodiversity loss, Global warming and climate change(Causes and intensity of the problem), role of Computer in creation of environmental issues.	
Unit II	Environmental laws and ethics for Bioinformatics	(8)
	Environmental Protection Act, Wildlife Protection Act, Forest Conservation Act, Prevention and Control of Pollution Act (Air, Water, Land), from unsustainable to sustainable development, Responsibilities of an Environmentally aware citizen.	
Unit III	Sustainable Development Goals for Bioinformatics	(7)
	Implementation, History, Structure of goals, targets and indicators, Goals with their targets and indicators, Challenges in sustainable development of goals in Environmental Awareness of Bioinformatics	
Unit IV	Role of Bioinformatics in Computational sustainability	(7)
	Computational sustainability, Computational Research in Sustainability, Balancing environmental and socioeconomic needs, Biodiversity and conservation, Computational Synergies, Green IT	

Course outcomes: The student should be able to...

1. Relate the Environmental issues.
2. Identify how Environmental Laws and Ethics are work.
3. Devise the Challenges in sustainable development of goals in Environmental Awareness of Bioinformatics.
4. Take a part in Balancing environmental and socioeconomic needs.

Reference books :

1. Carson, Rachel. *Silent Spring*. Mariner Books, 2002.
2. Miller, G. Tyler, and Scott Spoolman. *Environmental Science*. Cengage Learning, 2014.
3. Rajagopalan, R. *Environmental Studies: From Crisis to Cure*. Oxford University Press, 2010.
4. Divan, Shyam, and Armin Rosencranz. *Environmental Law and Policy in India: Cases and Materials*. Oxford University Press, 2022.
5. Kharchenko, Vyacheslav. *Green IT Engineering: Social, Business and Industrial Applications*. CRC Press, 2019.
6. United Nations Development Programme. *Sustainable Development Goals Report*. United Nations, 2020.
7. Sterling, Stephen, John Harte, and John J. Parson. *Computer Models for Environmental Management: Case Studies in Environmental Planning*. Routledge, 2018.
8. Allen, Thomas F., Thomas W. Baugh, and Jennifer C. Smith. *Sustainability in Environmental Management: Designing Development Projects*. CRC Press, 2016.
9. Prakash, Shri. *Environmental Studies*. S. Chand Publishing, 2018.
10. Jorgenson, Andrew K., and Juliet D. Jorgenson. *Introduction to Computational Sustainability*. Cambridge University Press, 2016.