Karmaveer Bhaurao Patil University, Satara

Syllabus for

B. Sc. I (Bioinformatics)

Under

Faculty of Science and Technology

(As per NEP 2020)

With effect from Academic Year 2024-2025

Karmaveer Bhaurao Patil University, Satara

Faculty of Science & Technology

B. Sc. (Bioinformatics)

Programme and Credit Structure as per NEP 2020

{Ref. Government of Maharashtra letter no. एनइपी.२०२२/प्र.क.०९/विशि-३शि का ना दिनांक: १३ मार्च २०२४}

The degree shall be titled as 'Bachelor of Science (Bioinformatics) under the faculty of Science and Technology

B. Sc. Sem. I & II from Academic Year 2024-25
B. Sc. Sem. III & IV from Academic Year 2025-26
B. Sc. Sem. V & VI from Academic Year 2026-27
B. Sc. Sem. VII&VIII from Academic Year 2027-28

Programme Outcomes for B. Sc.

PO. No.	**Programme Outcomes** **After completing B. Sc. Programme the students will be able to…….**
PO-1	Provide students with a strong foundation of knowledge in their chosen field of study, including fundamental concepts, theories, and principles.
PO-2	Encourage community engagement and a sense of social responsibility, inspiring students to use their knowledge and skills for the betterment of society.
PO-3	Develop a scientific attitude among the students and to make the students open minded, critical, and curious.
PO-4	Evolve skills in practical work, experiments, and laboratory materials.
PO-5	Equip students with quantitative and analytical skills necessary for data analysis, modelling, and interpretation in their field of study.
PO-6	Enhance students' written and oral communication skills, enabling them to effectively convey scientific concepts, research findings, and ideas to diverse audiences.
PO-7	Equip students with quantitative and analytical skills necessary for data analysis, modelling, and interpretation in their field of study.
PO-8	Develop proficiency in using relevant technology and tools that are essential for their field, including software, laboratory equipment, and data analysis tools.
PO-9	Competent in designing and managing biological databases, ensuring data integrity and accessibility for research purposes.
PO-10	Integrate techniques into biological research, contributing to advancements in fields such as genetics, drug discovery, and disease modelling.
PSO. NO	**Programme Specific Outcomes** **After completing B. Sc. Bioinformatics Programme the students will be able to…….**
PSO-1	Discuss related to the different aspects of bioinformatics.
PSO-2	Perform experiments and projects related to bioinformatics.
PSO-3	Provide students with a strong foundation in bioinformatics principles, algorithms, and methodologies, enabling them to understand and address biological challenges using computational tools.
PSO-4	Teach students how to design and manage biological databases, ensuring efficient data storage and retrieval for research purposes

PSO-5	Enable students to explore protein structure and function, including the prediction of protein structures, protein-protein interactions, and structural analysis.
PSO-6	Present their research findings in research conglomerations like conferences and in research journals in the form of publications.
PSO-7	Graduates will be proficient in analysing and interpreting biological sequences, including DNA, RNA, and protein sequences, using relevant algorithms and tools.
PSO-8	Ability to predict protein structures, analyse protein-ligand interactions, and model three-dimensional structures using computational methods.
PSO-9	Skilled in mining large biological datasets for patterns, associations, and insights, enabling them to make data-driven discoveries.
PSO-10	Skilled in assessing data quality, implementing data validation procedures, and ensuring the reliability of biological data.

Semester, Credit Framework, NSQF Level and Exit Points

Sr. No.	Semester	Year	Year	Credits	Level	Exit Points &Award
1	Sem. I & II	2024-25	1Year	44	4.5	UG Certificate in Bioinformatics
2	Sem. III & IV	2025-26	2Year	88	5.0	UG Diploma in Bioinformatics
3	Sem. V &VI	2026-27	3Year	132	5.5	B. Sc. in Bioinformatics (UG Three Year Degree)
4	Sem. VII & VIII	2027-28	4Year	176	6.0	B. Sc. in Bioinformatics [Honors/Research] (UG Four Year Degree)

Credit Distribution

Sr. No.	Course	3 Year Degree Programme			4 Year Honors Degree Programme			4 Year Honors with Research Degree Programme		
		Courses (3 Yr)	Credits (3 Yr)	%	Courses (4 Yr)	Credits (4 Yr)	%	Courses (4 Yr)	Credits (4 Yr)	%
1	Major	26	52	39.39	34	80	45.45	32	72	40.91
2	Elective	04	08	6.06	08	16	9.09	08	16	9.09
3	IKS	02	04	3.03	02	04	2.27	02	04	2.27
4	VSC	04	08	6.06	04	08	4.55	04	08	4.55
5	FP	01	02	1.52	01	02	1.14	01	02	1.14
6	OJT	01	04	3.03	02	08	4.55	01	04	2.27
7	RP	00	00	0.00	00	00	00	02	12	6.82
8	SEC	03	06	4.55	03	06	3.41	03	06	3.41
9	CEP	01	02	1.52	01	02	1.14	01	02	1.14
Total (Major) (A)		**42**	**86**	**65.15**	**55**	**126**	**71.59**	**54**	**126**	**71.59**
1	Minor & RM	12	24	18.18	13	28	15.91	13	28	15.91
Total (Minor) (B)		**12**	**24**	**18.18**	**12**	**28**	**15.91**	**13**	**28**	**15.91**
1	OE	04	08	6.06	04	08	4.55	04	08	4.55
2	AEC	04	08	6.06	04	08	4.55	04	08	4.55
3	VEC	02	04	3.03	02	04	2.27	02	04	2.27
4	CC	01	02	1.52	01	02	1.14	01	02	1.14
Total (C)		**11**	**22**	**16.67**	**11**	**22**	**12.50**	**11**	**22**	**12.50**
Grand Total (A+B+C)		**65**	**132**	**100**	**79**	**176**	**100**	**78**	**176**	**100**

B.SC. BIOINFORMATICS

Duration:

- The program shall be a full-time program.
- The duration of program shall be three years for Bachelor of Science and four years for Bachelor of Science with Honors or Bachelor of Science with Research.
- Every year students will have exist option with:
- (1^{st} Year: Certificate, 2^{nd} Year: Diploma, 3^{rd} Year: Degree, 4^{th} Year: Honors / Research)
- These students are allowed to re-enter the degree program within three years and complete the degree program within the stipulated maximum period of Seven Years.

Eligibility: 12^{th} Pass with Science, or equivalent.

Medium of Instruction: The medium of instructions shall be in English.

Scheme of Examination & Standard of Passing (CCE and ESE):

As per the decision of the concern Board of Studies

- End Semester Exam (ESE): 30 Marks (Min 12 Marks for Passing)
- Continuous Comprehensive Evaluation (CCE): 20 Marks (Min 08 Marks for Passing)
- Total Marks = 50 Marks
- Minimum 40% Marks Required for Passing and there is separate head of Passing for End Semester Examination (ESE) and Continuous Comprehensive Evaluation (CCE).
- A candidate who acquire 32 credits or more during semester – I & II shall be admitted to B. Sc. II (appear for semester – III & IV examination).
- However the candidate shall not be admitted to B.Sc. III (Semester - V) unless he/she passed in all the subjects at B.Sc. I (Semester - I & Semester - II) and acquire 32 credits or more during semester – III & IV.
- However the candidate shall not be admitted to B. Sc. IV (Semester - VII) unless he/she passed in all the subjects at B. Sc. III (Semester - V & Semester - VI).
- However under the National Education Policy the rules extended by KBP University, time to time regarding ATKT will be applicable.

Eligibility of the Core Faculty:

- As per rules and regulations of Karmaveer Bhaurao Patil University, Satara and Govt. of Maharashtra.

Eligibility for Professor of Practice or Professional Trainer:

- Any other eligibility as per the guidelines and regulations passed by concern board of studies, academic council of the autonomous college and rules & regulations of Karmaveer Bhaurao Patil University, Satara and Government of Maharashtra and UGC norms.

Karmaveer Bhaurao Patil University, Satara

B. Sc. (Bioinformatics) Part-I

Semester I				
Sr. No.	**Components**	**Course code**	**Course Tittle**	**Credits**
1	Course-I (Bioinformatics)	BBiT 111	Fundamental of Bioinformatics	02
		BBiT 112	Bioinformatics for Plant sciences	02
		BBiP 113	Lab I (Based on Paper BBiT 111 & BBiT 112)	02
2	Course-II (Computer Sciences)	BBiT 114	Fundamental of Computer Sciences	02
		BBiT 115	C Programming	02
		BBiP 116	Lab I (Based on Paper BBiT 114 & BBiT 115)	02
3	Course-III (Data Sciences)	BBiT 117	Fundamental of Data Sciences	02
		BBiT 118	Bioinformatics for Data Sciences	02
		BBiP 119	Lab I (Based on Paper BBiT 117 & BBiT 118)	02
4	OE (Music Studies)	BBiT OE I	Indian Musical Instrument	02
5	IKS	BBiT IKS I	Introduction to Indian Knowledge System	02
			Total	**22**

Semester II				
Sr. No.	**Components**	**Course code**	**Course Tittle**	**Credits**
1	Course-I (Bioinformatics)	BBiT 121	Biological Sequence and Protein Sequence Analysis	02
		BBiT 122	Bioinformatics for Animal sciences	02
		BBiP 123	Lab I (Based on Paper BBiT 121 & BBiT 122)	02
2	Course-II (Computer Sciences)	BBiT 124	Database Management System	02
		BBiT 125	R Programming	02
		BBiP 126	Lab I (Based on Paper BBiT 124 & BBiT 125)	02
3	Course-III (Data Sciences)	BBiT 127	Data Mining	02
		BBiT 128	Data Visualization	02
		BBiP 129	Lab I Based on Paper BBiT 127 & BBiT 128)	02
4	OE	BBiT OE II	History of Indian Music	02
5	VEC	BBiT IKS I	Democracy, Election and Indian Constitution	02
			Total	**22**
		EXIT OPTION: Award of UG Certificate in Major **with 44 credits** & an additional 4 credits core NSQF Course/Internship OR Continue with Major & Minor.		

B. Sc. (Bioinformatics) Part-II

Semester III				
Sr. No.	**Components**	**Course Code**	**Course Tittle**	**Credits**
1	Major	BBiT 231	Database Management System	02
		BBiT 232	Genomics,Proteomics & Transcriptomics	02
		BBiP 233	Lab I (Based on Paper BBiT 231 & BBiT 232)	02
2	Minor	BBiT 234	DSC V,DSC VI, DSP III	02
3	OE	BBiTOE III	Music Studies P-III	02
4	VSC	BBiTVSC I	Bioinformatics Method I	02
5	SEC	BBiTSEC I	Web Programming I	02
6	AEC	BBiT AEC I	English for Communication I	02
		BBiT AEC II	English for Communication II	02
			Total	**22**

Semester IV				
Sr. No.	**Components**		**Course**	**Credits**
1	Major	BBiT 241	Python Programming for Bioinformatics	02
		BBiT 242	Biological Sequence and Protein Sequence Analysis	02
		BBiP 243	Lab Based on (BBiT 241 & BBiT 242)	02
2	Minor	BBiT 244	DSC VII, DSC VIII, DSP IV	02
3	OE	BBiTOE IV	Music studies IV	02
4	VSC	BBiTVSC II	Bioinformatics method II	02
5	SEC	BBiTSEC II	Web programming II	02
6	AEC	BBiT AEC III	English for Communication III	02
		BBiT AEC IV	English for Communication IV	02
7	VEC	BBiT VEC I	Environmental Studies	02
			Total	**22**
		EXIT OPTION: Award of UG Diploma in Major and Minor with 88 Credits & an additional 4 credits core NSQF Course/ Internship OR Continue with Major & Minor		

B. Sc. (Bioinformatics) Part-III

Semester V			
Sr. No.	**Components**	**Course**	**Credits**
1	Major	Bioinformatics (P-IX)	02
2	Major	Bioinformatics (P-X)	02
3	Major	Bioinformatics (P-XI)	02
4	Electives	Bioinformatics (P-XIIE1)/ Bioinformatics (P-XIIE2)	02
5	Major Lab	Lab - V	02
6	Elective Lab	Lab - I	02
7	VSC	Bioinformatics Method III	02
8	AEC	English P-III	02
9	OJT	On Job Training in Bioinformatics I	04
10	CEP	Community Engagement Programme in Bioinformatics	02
		Total	**22**
Semester VI			
Sr.	**Components**	**Course**	**Credits**
1	Major	Bioinformatics (P-XIII)	02
2	Major	Bioinformatics (P-XIV)	02
3	Major	Bioinformatics (P-XV)	02
4	Electives	Bioinformatics (P-XVIE1)/ Bioinformatics (P-XVIE2)	02
5	Major Lab	Lab - VI	02
6	Elective Lab	Lab - II	02
7	VSC	Bioinformatics Method III	02
8	SEC	Web Programming III	02
9	FP	Field Project in Bioinformatics	02
10	CC	Co-curricular Course in Bioinformatics	02
11	AEC	English P-IV	02
		Total	**22**
EXIT OPTION: Award of UG Degree in Major with 132 credits OR Continue with Major & Minor.			

B. Sc. (Bioinformatics) Part-IV Honors Degree

Semester VII			
Sr. No.	**Components**	**Course**	**Credits**
1	Major	Bioinformatics (P-XVII)	04
2	Major	Bioinformatics (P-XVIII)	04
3	Major	Bioinformatics (P-XIX)	04
4	Electives	Bioinformatics (P-XXE1)/ Bioinformatics (P-XXE2)	02
5	Major Lab	Lab – VII	02
6	Elective Lab	Lab - III	02
7	Minor	Research Methodology	04
		Total	**22**
Semester VIII			
Sr.	**Components**	**Course**	**Credits**
1	Major	Bioinformatics (P-XXI)	04

2	Major	Bioinformatics (P-XXII)	04
3	Major	Bioinformatics (P-XXIII)	04
4	Electives	Bioinformatics (P-XXIVE1)/ Bioinformatics (P-XXIVE2)	02
5	Major Lab	Lab – VIII	02
6	Elective Lab	Lab - IV	02
7	OJT	On Job Training in Bioinformatics II	04
		Total	**22**
Award of Four year UG Honours Degree in Major and Minor with 176 credits.			

B. Sc. (Bioinformatics) Part-IV Honors with Research Degree

Semester VII			
Sr. No.	**Components**	**Course**	**Credits**
1	Major	Bioinformatics (P-XVII)	04
2	Major	Bioinformatics (P-XVIII)	04
3	Electives	Bioinformatics (P-XIXE1)/ Bioinformatics (P-XIXE2)	04
4	Major Lab	Lab – VII	02
5	Minor	Research Methodology	04
6	RP	Research Project in Bioinformatics I	04
		Total	**22**
Semester VIII			
Sr.	**Components**	**Course**	**Credits**
Sr. No.	**Components**	**Course**	**Credits**
1	Major	Bioinformatics (P-XX)	04
2	Major	Bioinformatics (P-XXI)	04
3	Electives	Bioinformatics (P-XXIIE1)/ Bioinformatics (P-XXIIE2)	04
4	Major Lab	Lab – VIII	02
5	RP	Research Project in Bioinformatics II	08
		Total	**22**
Award of Four year UG Honors Degree in Major and Minor with 176 credits.			

Chairman	Secretary	Chairman
BoS in Bioinformatics	Academic Council	Academic Council

Subject Title according to NEP 2020 (2.0)

Semester	Course	Name of Course	Paper code	Paper title	Credits
I	**Course I** Bioinformatics	DSC - I	BBiT 111	Fundamentals of Bioinformatics	2
		DSC - II	BBiT 112	Bioinformatics for Plant Sciences	2
		DSC(P) - I	BBiT 113	Lab I Based on Paper BBiT 111 & 112	2
I	**Course I** Computer Sciences	DSC - I	BBiT 114	Fundamentals of Computer Sciences	2
		DSC - II	BBiT 115	C Programming	2
		DSC(P) - I	BBiT 116	Lab II Based on Paper BBiT 111 & 112	2
I	**Course I** Data Sciences	DSC - I	BBiT 117	Fundamentals of Data Science	2
		DSC - II	BBiT 118	Bioinformatics for Data Sciences	2
		DSC(P) - I	BBiT 119	Lab I Based on Paper BBiT 111 & 112	2
I	Open Elective	OE- I	BBiT OE I	Music Studies	2
I	IKS – I	IKS – I	BBiT IKS I	Generic	2

Subject Title according to NEP 2020 (2.0)

Semester	Course	Name of Course	Paper code	Paper title	Credits
I	**Course I** Bioinformatics	DSC - III	BBiT 121	Biological Sequence and Protein Sequence Analysis	2
		DSC - IV	BBiT 122	Bioinformatics for Animal Sciences	2
		DSC(P) - II	BBiT 123	Lab I Based on Paper BBiT 121 & 122	2
I	**Course I** Computer Sciences	DSC - III	BBiT 124	Database Management System	2
		DSC - IV	BBiT 125	R Programming	2
		DSC(P) - II	BBiT 126	Lab II Based on Paper BBiT 121 & 122	2
I	**Course I** Data Sciences	DSC - III	BBiT 127	Data Mining	2
		DSC - IV	BBiT 128	Data Visualization	2
		DSC(P) - II	BBiT 129	Lab I Based on Paper BBiT 121 & 122	2
I	Open Elective	OE- 2	BBiT OE I	Music Studies	2
I	VEC-I	DEIC	BBiT VEC-I	Democracy, Election and Indian Constitution	2

Course I: Bioinformatics

BBIT 111- Fundamentals of Bioinformatics

Course Objectives: Students should be able to …

1. Understand the basics of bioinformatics and its scope in the future
2. Recognize the tools necessary for analysis of biological molecules.
3. Conceptualize the types of databases and its online software's.
4. Perceive the knowledge of predicting the structure and function of biomolecules.

Credits (Total Credits 2)	**SEMESTER – I** **BBIT 111- Fundamentals of Bioinformatics**	**No. of hours (30)**
Unit - I	**Basics of bioinformatics** Definition and scope of bioinformatics. Importance of Bioinformatics in modern biology and medicine. Biological molecules, biological databases, bioinformatics tools and software's. Need of bioinformatician, biological & clinical data generation, data visualization. Applications and future trends in bioinformatics.	**(08)**
Unit – II	**Introduction to biological databases and its types** a. Nature, types, and sources of biological data b. File formats of biological data c. NCBI/EBI/EXPASY d. GenBank/EMBL/DDBJ e. UniProtKB f. PDB g. CATH/SCOP/PDBsum 2. Derived databases and data repositories 1. Gene Cards 2. TCGA	**(08)**
Unit – III	**1. Basic sequence analysis and data visualization** a. Dot plots b. Pairwise sequence alignment c. Multiple sequence alignment d. Sequence visualization i. UCSC Genome Browser ii. .NCBI Genome Data Viewer **2. Protein structure prediction & visualization** a. Structure prediction: Homology modelling b. Visualization i. RASMOL/PyMol ii. SPDBV	**(08)**

Unit - IV	1. Biological sequence & structure similarity search a. Pairwise Sequence Similarity Searching BLAST (NCBI), FASTA (EBI), BLAT (UCSC). b. Structure Similarity Searching RCSB (PDB), VAST (NCBI), TM-ALIGN. 2. Functional prediction & analysis 1. Sequence based 1. Prosite 2. Pfam 3. STRING 2. Protein structure based 1. ProFunc 2. DALI 3. CASTp 4. PROSA	**(06)**

Course Outcomes: Students will be able to…

1. Determine the essential scope of bioinformatics.
2. Utilize the tools necessary for analysing the biological molecules
3. Justify the types of nucleic acids and their processes
4. Predict the structure of proteins and analyse it using online databases.

Reference Books:

1. Jean-Michel Claverie and Cedric Notredame (2006) “Bioinformatics for Dummies”, Wiley Publishing Inc., Indianapolis, Indiana. 2nd Edition
2. David W. Mount (2004) “Bioinformatics: Sequence and Genome Analysis" Publisher Cold Spring Harbor Laboratory Press, U.S.2nd Edition.
3. Phillip Compeau and Pavel Pevzner (2014), "Bioinformatics Algorithms: An Active Learning Approach" Active Learning Publishers.
4. Jonathan Pevsner (2015) "Bioinformatics and Functional Genomics" Wiley-Blackwell Publisher.
5. Arthur M. Lesk (2019) "Introduction to Bioinformatics" Publisher OUP Oxford.
6. Marketa Zvelebil and Jeremy O. Baum(2007) "Understanding Bioinformatics" ,Publisher Garland Science,Ist Edition
7. Neil C. Jones and Pavel A. Pevzner(2004) "An Introduction to Bioinformatics Algorithms" ,MIT Press.
8. Jin Xiong (2007)"Essential Bioinformatics" Cambridge University Press,Ist Edition**.**
9. Richard Durbin, Sean R. Eddy, Anders Krogh, and Graeme Mitchison(1998)"Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids" Cambridge University Press .
10. Steven Haddock and Casey Dun (2010) ,"Practical Computing for Biologists" ,Sinauer Associates is an imprint of Oxford University Pres

BBIT- 112: Bioinformatics for Plant sciences

Course Objectives: Students should able to

1. Understand basics of bioinformatics
2. Understand the use of different plant databases
3. Understand sequences, alignments and dynamic programming
4. Recognize gene expression analysis in plants

Credits=2	SEMESTER-I Bioinformatics for plant sciences	No. of lectures per unit
Unit I	**Fundamentals of the plant kingdom**	**07**
	1.1 The general outline of the Five kingdom classification & plant kingdom 1.2 Evolutionary history of Plants, Evolutionary time scale 1.3 . General characters and economic importance of Algae, Fungi, Lichens, Bryophytes, Pteridophytes, 1.4 General characters and economic importance Gymnospersms, Anigosperms	
Unit II	**Structural bioinformatics in plants**	**08**
	2.1 Taxonomy:-Definition, Aims, objectives and functions ,Binomial nomenclature and its significance, Principles of ICBN, Study of outline of Bentham and Hooker's system of Classification of plants. 2.2 Introduction to structure of Plant cell 2.3 Structure of Chloroplast & Chloroplast genome and its importance 2.4 Structure of Plant mitochondria & Plant mitochondrial genome and its importance	
Unit III	**Plant database and its types**	**07**
	3.1 Introduction of Biological Databases- 3.2 Nucleic acid databases: NCBI, DDBJ, and EMBL and GenBank 3.3 Protein databases: Swiss-Prot, UniProt 3.4. Plant database – Phytosomes Plant Garden,plant GDB genome browser	
Unit IV	**Data generation and applications of bioinformatics in plant sciences**	**08**
	4.1.Molecular Identification – 16s RNA, Large scale plant molecular biology data: introduction and generation 4.3 Plant ITS sequence and its importance 4.4 Genome sequencing 4.5 Protein sequencing Applications of Bioinformatics in plant sciences	

Course Outcomes: The students will be able to.

1. Explain different plant groups
2. Identify different molecular structures in plants
3. Use different plant databases
4. Generate different molecular data from plants

References:

1. Claverie J and Notredame C (2011) Bioinformatics for Dummies; John Wiley and Sons
2. Andreas Baxevanis, B. F. Francis Ouellette and B. F. Cuellette (1998) Bioinformatics: A Practical Guide to the analysis of Genes and Proteins, Wiley Publishers, New York
3. Jagota A. (2000) Data Analysis and Classification for Bioinformatics. Published by the Bay Press. University of Michigan, USA
4. Kumar, H. D. (1993). Molecular Biology and Biotechnology, Vikas Publ., New Delhi.
5. Mount D. W. (2001) Bioinformatics Sequence and Genome Analysis. Cold Spring Harbour Laboratory. New York.
6. Eric Lee, (2018) Beginners Guide To Bioinformatics For High Throughput Sequencing
7. R. Amjesh, S.S. Vinodchandra, (2019) Bioinformatics for beginners.
8. Maxwell James, (2021) Bioinformatics for beginners.

BBIP 113: Lab I based on paper BBIT 111 and BBIT 112

Course Objectives: Students should be able to,

1. Operate the online software's for nucleotide and protein analysis.
2. Visualize the sequence and structure of DNA and protein using various software's.
3. Understand the use of different plant databases
4. Recognize gene expression analysis in plants

Credits: 02	**SEMESTER-I BBIP 113 :Lab I based on paper BBIT 111 and BBIT 112**	**No. of hours allotted 60 hrs 4 hrs / Practical**
1	Using online nucleotide databases (NCBI, EBI, EXPASY, Genbank, EMBL, DDBJ), Protein databases(PDB,Uniprot ,SWISS PROT)	4
2	Visit to derived databases and data repositories	4
3	Pairwise sequence alignment using DOTPLOT, Needleman-Wunsch algorithm and Smith-Waterman algorithm.	4
4	Multiple sequence alignment using Clustal Omega	4
5	Sequence visualisation using UCSC Genome browser	4
6	Sequence Similarity search using BLAST	4
7	Functional Prediction based on sequence: Prosite, Pfam, STRING	4
8	Study general characteristics of plants by observing the plant specimens.	4
9	Perform classification of plant species as per Plant taxonomy from Class upto Species level	4
10	Isolation of chloroplasts	4
11	Isolation of plant genomic DNA	4
12	Visit to Phytosomes database	4
13	Retrieval of protein sequence from plant database SWISS PROT	4
14	Retrieval of protein sequence from plant database UNIPROT	4
15	Retrieval of biological sequence from animal databases- NCBI database: gene, mRNA, protein, 16S rRNA sequence retrieval	4
16	Designing the forward and revesrse primers	4

Course outcomes: The students will be able to,

1. Perform Isolation of plant and animals genomic DNA.
2. Use different animal and Plant databases to retrieve and analyze animal biological sequences
3. Use mutation/SNP databases related to animals
4. Comparing genomes of animals to identify similarity and differences between them.

References:

1. Claverie J and Notredame C (2011), Bioinformatics for Dummies; John Wiley and Sons

2. Andreas Baxevanis, B. F. Francis Ouellette and B. F. Cuellette (1998) Bioinformatics: A Practical Guide to the analysis of Genes and Proteins, Wiley Publishers, New York

3. Jagota A. (2000) Data Analysis and Classification for Bioinformatics. Published by the Bay Press. University of Michigan, USA

4. Kumar, H. D. (1993) Molecular Biology and Biotechnology, Vikas Publ., New Delhi.

5. Mount D. W. (2001) Bioinformatics Sequence and Genome Analysis. Cold Spring Harbour Laboratory. New York.

6. Eric Lee(2018), Beginners Guide To Bioinformatics For High Throughput Sequencing

7. R. Amjesh, S.S. Vinodchandra, (2019) Bioinformatics for beginners

8. Maxwell James, (2021) Bioinformatics for beginners.

Course II: Computer Science

B.Sc. I Syllabus for Course in Bioinformatics

SEMESTER – I

Course Code – BBIT 114: Fundamentals of Computer Sciences

Course Objectives: Students should be able to …

1. Define basic concepts and terminology of computers.
2. Understand operate desktop computers to carry out computational tasks.
3. Learn working of Hardware and Software and the importance of operating systems.
4. Identify programming languages, number systems, peripheral devices, and networking, multimedia and internet concepts.

Credits (Total Credits 2)	**SEMESTER – I** **BBIT 114- Fundamentals of Computer Sciences**	**No. of hours (30)**
Unit - I	**Introduction to Computers** Introduction, Characteristics, History & Evolution, Organization of Computers, Applications of Computers in Various Fields, Computer Hardware and Software, Computer Languages – Machine Language, Assembly Language, High-level Language, Language translators: Compiler, Interpreter, Assembler.	**(08)**
Unit – II	**Peripheral Devices** Input Devices – Keyboard, Touch screen, Pointing: Mouse, digitizer, Joystick and scanning devices: Scanner, OMR, OCR, and MICR. Output Devices – Monitors (CRT, TFT, LCD, and Plasma), Screen Image Projector, Printers & its types, Plotters. Memory Devices - Primary Memory & its types (RAM, ROM), Secondary memory & its types (Hard Disk, Flash Drives, Magnetic Tape, Optical Discs- CD, DVD, Blue-Ray)	**(08)**
Unit – III	**Introduction to OS :** Meaning and Definition, Structure of O.S., Types of O.S., Functions of O.S., **Introduction to DOS:** History and versions of DOS. Fundamentals of DOS, Getting Started with DOS: Booting Process (DOS, Windows, Unix/Linux), System Files and Command.com, Internal DOS Files & Directories, Elementary External DOS Commands, Creating a Batch Files, Additional Commands.	**(08)**

	Windows Operating system: Components of window Desktop, windows explorer, control panel, Managing the files and folders, Accessories: Paint, calculator and notepad. **Unix/Linux** : Introduction to Linux, Philosophy and Concepts, Linux Basics and System Startup, User Environment, Graphical Interface, Linux Utilities, Command Line/ Shell, Linux Documentation, File Sharing and Security, Processes, System Configuration and Program Installation, Linux Program Usage	
Unit - IV	**Introduction to MS- Office -** Introduction to software packages, Components of MS-Office, Features of MSOffice. **MS-Word -** Introduction, Menus, Document types, Working with Document, Formating document, Creating table, Tools, Printing document. **MS-Excel -** Introduction, Spread sheet application, Spreadsheet Converting file to different formats, Computation Data- Setting formula, finding total in rows and columns. **MS Power Point -** Introduction, Creating Presentation, Graphics.	**(06)**

Course Outcomes: Students will be able to…

1. Describe basic concepts and terminology of information technology.
2. Evaluate the fundamentals of personal computers and their operations.
3. Operate their small account using the computers and master the world of Information Technology
4. Use the computer for basic purposes of preparing his personnel/business letters, viewing information on internet (the web), sending mails, preparing his business presentations, playing games etc.

Reference Books:

1. P. K. Sinha (1992), Computer Fundamentals, BPB Publications, Sixth Edition.
2. V. Rajaraman (2013), Introduction to Information Technology, PHI, Second Edition.
3. Chetan Shrivastava (2010), Fundamental of Information Technology, Kalyani Publishers..
4. Guy Hart-Davis (2023) "The ABCs of Microsoft Office 97 Professional edition", BPB Publications.
5. Karl Schwartz (1998), "Microsoft Windows 98 Training Guide" BPB Publications.
6. C.S. French(1998) "Data Processing and Information Technology", BPB Publications
7. P.K Sinha (1992)`Computer Fundamentals`, BPB Publications
8. Guy Hart-Davis (2023) "The ABCs of Microsoft Office 97 Professional edition", BPB Publications.
9. Karl Schwartz (1998), "Microsoft Windows 98 Training Guide" BPB Publications.
10. C.S. French(1998) "Data Processing and Information Technology", BPB Publications
11. P.K Sinha (1992)`Computer Fundamentals`, BPB Publications

SEMESTER – I

Course Code – BBIT 115: C Programming

Course Objectives: Students will be able to…

1. Impart adequate knowledge on the need of programming languages.
2. Evaluate the need of problem-solving techniques.
3. Develop programming skills using the fundamentals of C Language.
4. Learn to write algorithms & flowchart of programs in C and to solve the problems.

Credits(total credits 2)	SEMESTER – I BBIT 115 – C Programming	No. of hours(30)
Unit - I	**Introduction to C Programming** • Basics of C Programming • Basic Syntax and Data Types • Operators and Expressions • Control Flow Statements	**(08)**
Unit - II	**Functions and Arrays** • Functions(Definition, Declaration) • Arrays (Definition and initialization of arrays; accessing and modifying array elements.) • String Handling(String literals& arrays of characters, Standard library functions for strings)	**(08)**
Unit - III	**Pointers and Dynamic Memory Allocation** • Introduction to Pointers (Definition, Declaration, Pointer arithmetic and pointer to functions.) • Dynamic Memory Allocation • Structures and Unions (Definition and usage of structures, accessing structure members, nested structures. Understanding unions and their use cases.)	**(08)**
Unit - IV	**File Handling** • File operations (opening, closing, reading, writing) • File pointers and modes • Error handling in file operations • Working with text and binary files	**(08)**

Course Outcomes: - Student will be able to …

1. Illustrate the flowchart and design an algorithm for a given problem and to develop solution
2. Develop conditional and unconditional statements to write C program.
3. Exercise user defined functions to solve real time problems.

Reference Books: –

1) Kanetkar Yashavant , Let Us C: Authentic guide to C programming language 19th Edition , Delhi, BPBpublication Dec 2022,

2) E-Balagurusamy PROGRAMMING IN ANSI C 8th Edition, McGraw Hill, India, 25 March 2019.

3) Thareja Reema , Programming In C,2nd edition, Oxford Higher education publication, India,2016

4) Prinz Peter C in a Nutshell, Relly Media, India December 2015.

5) Kochan G. Stephen, Programming in C, 4th edition, Pearson Education,USA,August 2014.

6) Perry Greg,C programming Absolute Beginner's Guide, British, Que Publishing ,2014.

7) Griffith David, Head First C Abrain friendly guide, Grayscale Indian Edition 2012.

Course code BBIP 116: Practical based on BBIT 114 and 115

Course Objectives: Students should be able to,

1. Understand the use of computer and its working.
2. Execute the commands of the operating systems.
3. Operate the online software's for nucleotide and protein analysis.
4. Visualize the sequence and structure of DNA and protein using various software's.

Credits (Total Credits 2)	**SEMESTER – I** **BBIP 116: Practical based on BBIT 114 and 115** List of Practical (15)	No. of hours (60)
1.	Demonstration of Peripheral Devices	2
2.	MS-DOS Prompt and commands	2
3.	Demonstration of Windows Operating System	2
4.	MS – EXCEL 1) Create a worksheet on students list of 4 faculties and perform following database function on it. a. Sort data by name b. Filter data by Class c. Subtotal of number of students by Class 2) Import or fetch external data from web-sources a. Data cleaning and organizing. b. Data visualization: Prepare a bar chart and pie chart from this data	4
5.	MS – WORD I] Create an article in MS-WORD interpreting the results from any of the above MS-Excel examples. Make use of: Page border and shading, set margins, orientation, size, columns, watermark (as draft).	4
6.	MS-POWER POINT I] Create a MS-PowerPoint presentation to present the results from any of the above MS-Excel examples. Create and work with master slides, make use of transitions and animations, apply themes and layouts, use pictures, graphics, shapes and tables. Given a data and information, summarize them into WORD, EXCEL and POWERPOINT, demonstrating better representation and easy understanding to the audience.	4

7.	Experiments on Linux: A. Checking system status 1. Memory [df, du, free] 2. Task manager [uname, ps, top, kill] 3. Networking [ping, ifconfig, netstat] 4. Workload optimization B. File systems 1. Navigating & Searching files and folders[pwd, ls, cd] 2. Reading Files [cat, more, less, head, tail] 3. Files and folders Creation, deletion, modification [mkdir, rmdir, cp, mv, rm, touch] 4. Permission management & Sharing [chmod, chown] 5. Compression of files & folders [tar, gzip, gunzip] C. Install, Run & Uninstall Linux Programs D. Perform Linux Commands 1. grep 2. find E. Help Doc [man, command_name --help]	8
8.	Write a program to print positive integers from 1 to 10.	4
9.	Write a program to display the following pattern. * * * * * * * * * * * * * * *	4
10.	Write a program to display Fibonacci series.	8
11.	Write a program to add two Matrices; Use Two-Dimensional array.	2
12.	Write a program to insert 5 elements into an array and print the elements of the array.	2
13.	Write a program to calculate factorial of a number using recursion.	4
14.	Write a program to find biggest among three numbers using pointer.	4
15.	Write a C program to create, declare and initialize structure.	2
16.	Write a program to create a file called emp.rec and store information about a person, in terms of his name, age and salary.	4

Course Outcomes: Students will be able to-

1. Demonstrate the different parts with peripheral devices and various commanding systems.

2. create their own files, folders; store, format and present data & information using MS office

3. apply various similarity searches of pairwise, multiple sequence and structural alignments using online software's tools to access and use different bioinformatics databases and repositories

4. Visualize, Analyze and Interpret biological sequence and protein structures.
5. Recognize and understand the syntax and construction of C programming code.
6. Use different basic concepts arrays in C
7. Apply the different concepts of operations on Pointers.

Reference Books:

1. P. K. Sinha (1992), Computer Fundamentals, BPB Publications, Sixth Edition.
2. V. Rajaraman (2013), Introduction to Information Technology, PHI, Second Edition.
3. Chetan Shrivastava (2010), Fundamental of Information Technology, Kalyani Publishers..
4. Guy Hart-Davis (2023) "The ABCs of Microsoft Office 97 Professional edition", BPB Publications.
5. Karl Schwartz (1998), "Microsoft Windows 98 Training Guide" BPB Publications.
6. C.S. French(1998) "Data Processing and Information Technology", BPB Publications
7. P.K Sinha (1992)`Computer Fundamentals`, BPB Publications
8. Jean-Michel Claverie and Cedric Notredame (2006) "Bioinformatics For Dummies",Wiley Publishing Inc.,Indianapolis,Indiana. 2nd Edition
9. David W. Mount (2004) "Bioinformatics: Sequence and Genome Analysis" Publisher Cold Spring Harbor Laboratory Press, U.S.2nd Edition.
10. Kanetkar Yashavant , Let Us C: Authentic guide to C programming language 19th Edition ,
11. E-Balagurusamy PROGRAMMING IN ANSI C 8th Edition, McGraw Hill, India, 25 March.
12. Thareja Reema , Programming In C,2nd edition,Oxford Higher eduction India,2016
13. Prinz Peter C in a Nutshell,ORelly Media,India December 2015.
14. Kochan G. Stephen, Programming in C, 4th edition,Pearson Education,USA,August 2014.
15. Perry Greg,C programming Absolute Beginner's Guide, British, Que Publishing ,2014

Course III: Data Science

Course code BBIP 117: Fundamental of Data Science

Course Objectives: - Student should be able to learn…

1. Fundamental concepts of Data Science.
2. Study basic principles of Data Science.
3. Develop skills for Data Management.
4. Think through the ethics surrounding privacy, data sharing.

Credits=2	SEMESTER-I BDST 111: Fundamental of Data Science	No. of hours per unit/ credits
UNIT I	**Introduction to Data Science**	**(7)**
	Defining data science and big data, Recognizing the different types of data, Gaining insight into the data science process, Data Science Process: Overview, Different steps, Machine Learning Definition and Relation with Data Science.	
UNIT II	**Data management And Analysis**	**(9)**
	Data collection and management: Introduction, Sources of data, Data collection and APIs, Exploring and fixing data, Data storage and management, using multiple data sources. Data analysis: Introduction, Terminology and concepts, Introduction to statistics, Central tendencies and distributions, Variance, Distribution properties and arithmetic, Samples/CLT, Basic machine learning algorithms, Linear regression, SVM, Naive Bayes	
UNIT III	**Data visualization**	**(7)**
	Data visualization: Introduction, Types of data visualization, Data for visualization: Data types, Data encodings, Retinal variables, mapping variables to encodings, Visual encodings.	
UNIT IV	**Applications of Data Science**	**(7)**
	Technologies for visualization, Bokeh (Python), recent trends in various data collection and analysis techniques, various visualization techniques, application development methods of used in data science.	

Course Outcomes: - Student will be able to …

1. Explore the fundamental concepts of data science.
2. Understand data analysis techniques for applications handling large data.
3. Understand various machine learning algorithms used in data science process.
4. Visualize and present the inference using various tools.
5. Learn to think through the ethics surrounding privacy, data sharing and algorithmic decision-making.

Reference Books: –

1. Mittal Gautam, Data Science Simplified: A Hands-on Guide for Beginners, Notion Press, 2021
2. Raghunathan, Foundations of Data Science, CRC Press, 2020
3. Patil Prashant, Big Data Analytics: A Comprehensive Guide, McGraw-Hill Education, 2019
4. Ian H. Witten, Frank, Eibe, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 2016
5. Grus Joel, O'Reilly, Data Science from Scratch: First Principles with Python, 1st edition, 2015.
6. Jeffrey David, Leskovec, Anand Rajaraman, Ullman, Cambridge, Mining of Massive Datasets, J20 University Press, 2nd edition, 2014.
7. Jure Leskovek, Anand Rajaraman, Jeffrey Ullman, Mining of Massive Datasets. v2.1, Cambridge University Press, 2014.
8.O'Neil Cathy, Schutt Rachel, O' Reilly, Doing Data Science, Straight Talk from the Frontline,1st edition, New York, O'Reilly Media,2013.

BBIT 118: Data Science for Bioinformatics

Course Objectives: Student will be able to……

1. Understand the basic principles of Data Science and its applications in Bioinformatics.
2. Develop the essential programming skills by using Python and R and analyze biological data effectively.
3. Learn the fundamental Statistical Methods, Machine Learning and their applications in solving Bioinformatics problems.
4. To provide hands-on experience with real-world Bioinformatics applications.

Credits=2	**SEMESTER-I** **BBIT 118: Data Science for Bioinformatics**	**No. of hours per unit/ credits**
UNIT I	**Basics of Biological Data**	**7**
	Overview of Data Science: Definition and importance of Data Science, Applications of data science in biology and healthcare. Basics of Bioinformatics: Introduction to bioinformatics and its role in modern biology. Key databases and resources in bioinformatics (GenBank, UniProt, PDB) Introduction to Biological Data: Types of biological data: sequences, structures, expression data, Data formats and standards (FASTA, GenBank), Basics of data collection and curation in bioinformatics.	
UNIT II	**Programming and Data Handling**	**8**
	Introduction to Programming for Bioinformatics: Basics of Python programming: syntax, variables, and control structures, Introduction to R programming: data types, functions, and basic operations, Data Manipulation, Handling biological data with Python, Introduction to data preprocessing and cleaning, Working with Biological Databases, Accessing and querying biological databases using Python and R.	
UNIT III	**Statistical Methods and Machine Learning**	**7**
	Fundamentals of Statistics in Bioinformatics: Descriptive statistics: mean, median, mode, variance, standard deviation, Probability distributions and their applications in bioinformatics, Hypothesis testing and p-values in biological research Introduction to Machine Learning: Basic concepts of machine learning: supervised vs. unsupervised learning, Applications of machine learning in Bioinformatics.	

UNIT IV	**Applications of Data Science in Bioinformatics**	**8**
	Genomic Data Analysis: Introduction to genomic data and next-generation sequencing, Overview of genome annotation and variant analysis Structural Bioinformatics: Basics of protein structure: primary, secondary, tertiary, and quaternary structures, Introduction to protein structure prediction methods, Overview of molecular docking and drug discovery, Current trends and advancements in bioinformatics research, Ethical considerations in bioinformatics and data science.	

Course

Outcomes: Student should be able to …

1. Understand the Role of Data Science in Bioinformatics to demonstrate an understanding of how data science techniques can be applied to solve complex biological problems.
2. Perform Data Manipulation and Visualization to utilize Python and R programming languages to manipulate, visualize, and interpret biological data from various sources.
3. Apply Statistical and Machine Learning Techniques to analyze biological datasets, including genomic and proteomic data.
4. Analyze Genomic and Structural Data.
5. Integrate data science methods with biological research, and effectively communicate findings.

Reference Books:

1. Bassi, Sebastian. Python for Bioinformatics. Boca Raton, FL: CRC Press, 2010.
2. Antao, Tiago. Bioinformatics with Python Cookbook. 2nd ed. Birmingham, UK: Packt Publishing, 2018.
3. Gentleman, Robert. R Programming for Bioinformatics. Boca Raton, FL: Chapman & Hall/CRC, 2008.
4. Buffalo, Vince. Bioinformatics Data Skills: Reproducible and Robust Research with Open Source Tools. Sebastopol, CA: O'Reilly Media, 2015.
5. Datta, Somnath, and Dan Nettleton, eds. Statistical Analysis of Next Generation Sequencing Data. Cham: Springer, 2014.
6. Müller, Andreas C., and Sarah Guido. Introduction to Machine Learning with Python: A Guide for Data Scientists. Sebastopol, CA: O'Reilly Media, 2016.
7. Mount, David W. Bioinformatics: Sequence and Genome Analysis. 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 20

Course code BBIP 119: Practical based on BBIT 117 and 118

Course Objectives: Students should be able to,

5. Understand the use of computer and its working.
6. Execute the commands of the operating systems.
7. Operate the online software's for nucleotide and protein analysis.
8. Visualize the sequence and structure of DNA and protein using various software's.

Credits (Total Credits 2)	**SEMESTER – I** **BBIP 119: Practical based on BBIT 117 and 118** List of Practical (15)	No. of hours (60)
1.	Create a simple table	2
2.	Calculate the total and average	2
3.	Fill a sequence of numbers (1 to 20) in a column	2
4.	Highlight employees who earn more than $50,000.	4
5.	Sort the employee data alphabetically	4
6.	Combine first and last names into a full name and extract initials	4
7	Ensure that the Age column only accepts values between 18 and 65.	4
8	Create a drop-down list for the Department column.	4
9	Case Study on Data Center	4
10	Case Study on Data Warehouse.	4
11	Installation of R Programming	
12	Installation of Python Programming.	
13	Basic Python Operation and Syntax	
14	Working with List Python	
15	Working with Dictionary's in Python	
16	Data Import and Export in R	
17	Data Import and Export in Python	
18	Implementation of pandas Lib & Numpy Lib	
19	Statistical analysis using R	
20	Statistical analysis using python	

Course Outcomes: Students will be able to-

Student should be able to …

1. Understand the Role of Data Science in Bioinformatics to demonstrate an understanding of how data science techniques can be applied to solve complex biological problems.
2. Perform Data Manipulation and Visualization to utilize Python and R programming languages to manipulate, visualize, and interpret biological data from various sources.
3. Apply Statistical and Machine Learning Techniques to analyze biological datasets, including genomic and proteomic data.
4. Analyze Genomic and Structural Data.
5. Integrate data science methods with biological research, and effectively communicate findings.

Reference Books:

1. Bassi, Sebastian. Python for Bioinformatics. Boca Raton, FL: CRC Press, 2010.
2. Antao, Tiago. Bioinformatics with Python Cookbook. 2nd ed. Birmingham, UK: Packt Publishing, 2018.
3. Gentleman, Robert. R Programming for Bioinformatics. Boca Raton, FL: Chapman & Hall/CRC, 2008.
4. Buffalo, Vince. Bioinformatics Data Skills: Reproducible and Robust Research with Open Source Tools. Sebastopol, CA: O'Reilly Media, 2015.
5. Datta, Somnath, and Dan Nettleton, eds. Statistical Analysis of Next Generation Sequencing Data. Cham: Springer, 2014.
6. Müller, Andreas C., and Sarah Guido. Introduction to Machine Learning with Python: A Guide for Data Scientists. Sebastopol, CA: O'Reilly Media, 2016.
7. Mount, David W. Bioinformatics: Sequence and Genome Analysis. 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2004.

Course I: Bioinformatics

Semester -II

Course Code – BBiT 121: 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 121 : 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 ofprofile-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.

BBIT -122: Bioinformatics for Animal Sciences

Course Objectives: Students should be able to,

1. Understand basics of animal sciences.
2. Understand the use of different animal databases
3. Understand sequences, alignments and dynamic programming
4. Recognize gene expression analysis in animals

Credits=2	**SEMESTER-I** **(BBIT 122) Bioinformatics for Animal Sciences**	**No. of lectures per unit**
Unit I	**Fundamentals of the Animal kingdom**	**07**
	1. The general outline of the Animal kingdom 2. Evolutionary history of Animal, Evolutionary time scale 3. General characters of Kingdom Animalia and its phylum 4. About Vertebrates and invertebrates	
Unit II	**Bioinformatics data types related to animals**	**08**
	1. **Introduction to structure and properties of biomolecules** a. DNA (Genome) b. RNA (Transcriptome) c. Protein (Proteome) **2. Genetics, Epigenetics and its importance-** Mutations **3. Molecular markers** a. DNA Barcoding - Mitochondrial genome(mtDNA) and its importance b. ITS sequence and its importance	
Unit III	**Bioinformatics databases and analysis related to animals**	**07**
	1. Access to animal data from biological databases i. Animal-Specific databases ii. AnimalTFDB iii. agReg-SNPdb iv. Mouse Genome Informatics v. Goat Genome Variation Database vi. Bgee vii. WormBase viii. FlyAtlas, FlyBase 2. Future perspectives and case studies i. Emerging technologies and bioinformatics methodologies to address biodiversity loss ii. Ethics & data protection related to animals iii. Case study: Bioinformatics in government policy making v. Group presentation: Presentation of published research paper highlighting bioinformatics application in zoology.	

Unit IV	**DNA isolation methods and applications of bioinformatics in animal sciences** 1. Large scale Animal molecular biology data: introduction and generation 2. Gel electrophoresis 3. PCR 4. Genome sequencing 5. Protein sequencing 6. Applications of Bioinformatics in animal sciences	**08**

Course Outcomes: The students will be able to,

1. Explain different animal groups as per taxonomic classification.
2. Use different animal databases to analyse the animal specific data.
3. Understand methods, scope, perspectives, case studies related to animal bioinformatics
4. Generate different molecular data from animals

References:

1. Claverie J and Notredame C (2011), Bioinformatics for Dummies; John Wiley and Sons
2. Andreas Baxevanis, B. F. Francis Ouellette and B. F. Cuellette (1998) Bioinformatics: A Practical Guide to the analysis of Genes and Proteins, Wiley Publishers, New York
3. Jagota A. (2000) Data Analysis and Classification for Bioinformatics. Published by the Bay Press. University of Michigan, USA
4. Kumar, H. D. (1993) Molecular Biology and Biotechnology, Vikas Publ., New Delhi.
5. Mount D. W. (2001) Bioinformatics Sequence and Genome Analysis. Cold Spring Harbour Laboratory. New York.
6. Eric Lee(2018), Beginners Guide To Bioinformatics For High Throughput Sequencing
7. R. Amjesh, S.S. Vinodchandra, (2019) Bioinformatics for beginners
8. Maxwell James, (2021) Bioinformatics for beginners.
9. https://www.animalgenome.org/
10. http://catalog.illinois.edu/graduate/aces/concentration/animal-sciences/bioinformatics/
11. http://www.jakraya.com/journal/pdf/21-vcsArticle_3.pdf
12. https://guides.library.yale.edu/bioinformatics/animal-resources
13. http://cabgrid.res.in/cabin/publication/smfa/Module%20IV/7.%20Bioinformatics%20tools%20for%20classi fication%20and%20prediction_Dinesh%20Kumar.pdf
14. https://academic.oup.com/nar/article/51/D1/D816/6775385

SEMESTER – II

Course code BBIP 123: Practical based on BBIT 121 and 122

Course Objectives: Students should be able to,

1) Recognize the good laboratory practices and maintain biosafety protocols
2) Operate light compound microscope and identify different structures under microscope.
3) Compose isolation methods for DNA/RNA extraction.
4) Perform R programming language with the help of tools.

Credits (Total Credits 2)	**SEMESTER – II** **BBIP 123** **List of Practical (15)**	**No. of hours (60)**
1)	Pair wise alignment using Dot Plot/ Global & Local alignment methods tools	**4**
2)	Multiple sequence alignment using : Clustal, Dialign, Multalign	**4**
3)	Primary and secondary structure prediction methods using GOR Method	**4**
4)	Primary and secondary structure prediction methods using PSI-pred Method	**4**
5)	Primary and secondary structure prediction methods using Chou-Fasman method	**4**
6)	Binding site identification	**4**
7)	Study of Sequence patterns and profiles: Generation of sequence profiles using PSI-BLAST	**4**
8)	Derivation of and searching sequence patterns using MEME/MAST	
10)	Study general characteristics of animals by observing the animal specimens	**4**
11)	Isolation of animal genomic DNA	**4**
12)	Ensembl genome browser: Human, Mouse, Indian Cobra, Zebrafish	**4**
13)	Animal sequence (DNA, RNA, Protein) BLAST search	**4**
14)	Mouse Genome Informatics, Goat Genome Variation Database	4
15)	Bgee, WormBase, FlyAtlas, FlyBase	4

Course II: Computer Science

SEMESTER – II

Course Code–BBIT 124: 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 124: Database Management System**	**No. of** **hours (30)**
Unit - I	• **Introduction to databases and DBMS approach in bioinformatics** • 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	• **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	• **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	• **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 realworlddata

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, 2

SEMESTER – II
Course code – BBIT 125: Introduction to Programming: R language

Course Objectives: Students should be able to...

1. Gain a thorough understanding of R programming fundamentals, including syntax, data types, and basic operations.
2. Learn to import, explore, and clean various types of data from different sources (e.g., CSV, Excel).
3. Develop the ability to manipulate and transform data using fundamental operations and functions
4. Understand and utilize `ggplot2` for creating and customizing a range of visualizations, including histograms, bar charts, and scatter plots.

Credits - 2	**SEMESTER- III** **BBIT 125: Introduction to Programming: R language**	**No. of hours**

UNIT - I	**Introduction to R and Basic Operations**
	• **Getting Started with R**
	  ○ Introduction to R and RStudio
	  ○ Basic syntax and operations: operators, variables, and data types (vectors, lists, matrices, data frames, factors)
	  ○ Basic functions and operations
	• **Data Handling**
	  ○ Importing data from CSV, Excel, and text files
	  ○ Exploring and summarizing data: viewing, summary statistics, structure
	  ○ Cleaning data: handling missing values, data transformation
		(08)

Unit - II	**Data Visualization**
	• **Introduction to Visualization**
	  ○ Basics of ggplot2
	  ○ Creating simple plots: histograms, bar plots, scatter plots
	  ○ Customizing plots: titles, axis labels, legends, colors
	• **Advanced Plotting**
	  ○ Faceting and combining plots
	  ○ Custom themes and annotations
	  ○ Box plots, density plots, and heatmaps
		(08)

Unit - III	**Statistical Analysis**
	• **Descriptive Statistics**
	  ○ Calculating and interpreting measures of central tendency and variability
	  ○ Visualizing distributions
	• **Inferential Statistics**
	  ○ Basics of hypothesis testing: t-tests, chi-square tests
	  ○ Introduction to correlation and regression analysis
		(08)

	○ Simple linear regression models	
Unit - IV	**Advanced Data Handling** • Advanced Data Manipulation ○ Using dplyr for data manipulation: filtering, selecting, summarizing, and grouping ○ Merging and reshaping data with tidyr (e.g., pivoting)	**(06)**

Course Outcomes: Students will be able to…

1. Students will be able to write, debug, and execute R code effectively using RStudio.
2. Students will demonstrate the ability to manipulate and transform data using fundamental functions and operations.
3. Students will be able to import data from various sources, such as CSV and Excel files, and perform essential data cleaning and preprocessing tasks.
4. Students will effectively use visualization techniques to explore and present data insights.

Reference Books:

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

Course code BBIP 126: Lab course based on BBiT 121 & 122

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** BBIT 126 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	Features Data types and objects of R	3
11	Basic operations in R a.Addition, Subtraction, Multiplication & Division b.Row, Column & Other Operations on data frames and matrix c.Repeats, Sorting, Ordering	3
12	2. R functions i. In-built ii. Apply and types iii. Custom	3
13	3. String manipulation i. Expressions: Grep ii. Concatenate, Paste, Splitting functions, Replace functions, Replace	3
14	• Conditional statements o If o If – else • Loops o ‘For’ o ‘While’ • File Handling o Reading: txt, csv, excel files o Writing o Deleting	3
15	Data analysis - Data calculation, Manipulation, Formatting	
16	Data Representation/Visualization - Scatter Plot, Bar Plots, Histogram, Pie chart and their variations	

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).

Course III: Data Science

Semester: II
Course code – BBIT 127: Data Mining

Course Objectives: Student will be able to……

1. Recognize the importance and impact of data mining in Recognize the importance and impact of data. Mining in various industries and research fields.
2. Understand the challenges and ethical issues associated with data mining, including data privacy, security, and the risk of biased outcomes.
3. Understand the advantages of data mining and how its architecture supports efficient data processing.
4. Learn the basics of data scraping, including techniques for extracting data from websites and other digital sources.
5.

Credits=2	**SEMESTER-II** **BBIT: Data Mining**	**No. of hours per unit/ credits**
UNIT I	Fundamental of Data Mining	(8)
	Introduction to Data, Data Mining Concept, Importance of Data Mining, Applications of Data Mining, Key Features for Data Mining, Data Warehouse, Challenges and Issues in Data Mining.	
UNIT II	Data Architecture	(7)
	Data Mining Architecture, Advantages of Data Mining, Data Warehouse Architecture, Top-down approach, Web – Mining, Text Mining, Data Scraping.	
UNIT III	Data Warehouse Concepts	(7)
	Introduction to Data Warehouse, OLTP System, Difference between OLTB System and Data Warehouse, Introduction to Block chain Technology, Applications of Block Chain Technology in Bioinformatics.	
UNIT IV	Advance Data Mining Techniques	(8)
	Introduction to Machine Learning, Supervise Machine Learning, Un-Supervised Machine Learning, Reinforcement Machine Learning, Application of Machine Learning, Tools Used for Machine Learning Algorithms.	

Course Outcomes- Student should be able to …

1. Identify and describe various applications of data mining across different industries, including marketing, healthcare, finance, and bioinformatics.
2. Discuss the advantages of data mining in decision-making and strategic planning for businesses and organizations.
3. Apply basic web mining and text mining techniques to extract valuable information from web content and unstructured text data
4. Differentiate between supervised, unsupervised, and reinforcement learning, and identify appropriate scenarios for applying each type

Reference Book

1. William J. Tastle, Data Mining Applications Using Artificial Adaptive Systems, Springer-Verlag New York Inc, 26 August 2012
2. PANG-NING TAN MICHAEL STEINBACH ANUJ KARPATNE VIPIN KUMAR , Introduction to Data Mining, Pearson, 30 May 2021
3. Parteek Bhatia, Data Mining and Data Warehousing: Principles and Practical Techniques, Cambridge University Press, 27 June 2019
4. Pang-Ning Tan, Introduction to Data Mining, Pearson Education, 10 July 2016
5. DUNHAM, Data Mining: Introductory and Advanced Topics, Pearson Education India, 1 January 2006
6. 6. Han, Data Mining: Concepts and Techniques 3e, Elsevier, 1 January 2007

BBIT 128: Data Visualization Techniques

Course Objectives: Student should be able to …

1. Understand the principles and best practices of data visualization.
2. Utilize various types of charts, graphs, and plots to represent different types of data.
3. Apply critical thinking to choose the most effective visual representations.
4. Use data visualization tools to create interactive and static visualizations.

Credits=2	**SEMESTER-II** **Course IV: Data Visualization Techniques**	**No. of hours per unit/ credits**
UNIT I	**Fundamentals for Data Visualization**	(8)
	Data Visualization, importance of data visualization, Overview of common visualization tools and software, Introduction to datasets: types and sources, Key design principles, Choosing the right type of visualization for the data.	
UNIT II	**Principles of Data Visualization**	(7)
	Line charts, Bar charts, Pie charts, Histograms and Scatter plots, Heatmaps, Treemaps, and bubble charts, Time series and trend analysis, Multivariate data visualization techniques, Key Points used in Data Visualization.	
UNIT III	**Visual Design and Aesthetics in Data Visualization**	(7)
	Importance of labeling, scales, and legends. Aesthetics versus functionality, Use of color, shape, and size in visualizations. Geospatial data visualization: maps, Choropleths.	
UNIT IV	**Advance Visualization Concepts**	(8)
	Visualization of Big Data: techniques and challenges. Identifying and avoiding misleading visualizations, Interactive and Animated Visualizations, Advance Tools for Data Visualization.	

Course Outcomes: - Students will be able to…

1. Apply appropriate visualization techniques to represent different types of data, ensuring clarity, accuracy, and effective communication of insights.
2. Analyze datasets and interpret their underlying patterns, trends, and relationships using a variety of visualization techniques.
3. Demonstrate a solid understanding of the fundamental principles and theories behind data visualization, including visual perception, design principles, and the role of aesthetics.
4. Analyze datasets and interpret their underlying patterns, trends, and relationships using a variety of visualization techniques.

5. Critique and assess the effectiveness of different visualizations, identifying strengths, weaknesses, and potential biases in design and data representation.

Reference Books:

1. Dr. Gaurav Aroraa, Data Analytics: Principles, Tools, and Practices: A Complete Guide for Advanced Data Analytics Using the Latest Trends, Tools, and Technologies, 23 January 2022

2. ajit roy, Applied Big Data Analytics: Evolution, Platforms & Tools, Use cases, Benefits, Impact and Paradox, 18 August 2015

3. Dr Polala Niranjan siripuri Kiran, BIG DATA ANALYTICS USING R-TOOLS REPORT: AICTE ISTE Induction/Refresher Program, 13 December 2019

4. Alberto Ferrari,Introducing Microsoft Power BI, Microsoft Press, 7 July 2016

5 Chandraish Sinha , Mastering Power BI: Build business intelligence applications powered with DAX calculations, insightful visualizations, advanced BI techniques, and loads of data sources, 29 June 2024

6. Mastering Microsoft Power BI - Second Edition: Expert techniques to create interactive insights for effective data analytics and business intelligence, Packt Publishing, 30 June 2022

Course code BBIP 128: Lab course based on BBiT 126 & 127

Course Objectives: Student will be able to…

1. Recognize the importance and impact of data mining in Recognize the importance and impact of data. Mining in various industries and research fields.
2. Understand the challenges and ethical issues associated with data mining, including data privacy, security,
3. and the risk of biased outcomes.
4. Understand fundamental concepts of data.
5. Understand principles of databases.

Credits = 2	SEMESTER-I	No. of hours per unit/ credits
	Group A 1. Load a dataset (e.g., CSV file) and perform basic data exploration. 2. Display top 10 records in the datasets 3. Understand data cleaning. 4. Understand the Removing Null Values in Dataset 5. Perform the Operation and find out the Outilers in the Datasets 6. Perform the Operation to View Min, Median, and Max values in the datasets 7. Remove duplicates and identify outliers using statistical methods 8. Perform exploratory data analysis to understand the dataset. 9. Create different types of plots to visualize data trends and relationships. 10. Analyze and visualize time series data.	
	Group B 1. Installation of Power BI Tool in windows. 2. Learn how to import and connect data from various sources 3. Clean using Power Query Editor 4. Transform data Power Query Editor 5. Create and customize basic visualizations in Power BI 6. Design a dashboard that displays key metrics and insights 7. Apply filters to your visuals to allow for interactive data exploration 8. Apply slicers to your visuals to allow for interactive data	

	exploration 9. Learn how to create calculated columns and measures for custom calculations 10. Publish your Power BI reports to the Power BI service and share them with others	

Course outcomes –Student should be able to …

1. Identify and describe various applications of data mining across different industries, including marketing, healthcare, finance, and bioinformatics.
2. Discuss the advantages of data mining in decision-making and strategic planning for businesses and organizations.
3. Analyze datasets and interpret their underlying patterns, trends, and relationships using a variety of visualization techniques.
4. Critique and assess the effectiveness of different visualizations, identifying strengths, weaknesses, and potential biases in design and data representation.

Reference Books:

1. William J. Tastle, Data Mining Applications Using Artificial Adaptive Systems, Springer-Verlag New York Inc, 26 August 2012
2. PANG-NING TAN MICHAEL STEINBACH ANUJ KARPATNE VIPIN KUMAR , Introduction to Data Mining, Pearson, 30 May 2021
3. DUNHAM, Data Mining: Introductory and Advanced Topics, Pearson Education India, 1 January 2006
4. Han, Data Mining: Concepts and Techniques 3e, Elsevier, 1 January 2007
5. Dr. Gaurav Aroraa, Data Analytics: Principles, Tools, and Practices: A Complete Guide for Advanced Data Analytics Using the Latest Trends, Tools, and Technologies, 23 January 2022
6. Ajit roy, Applied Big Data Analytics: Evolution, Platforms & Tools, Use cases, Benefits, Impact and Paradox, 18 August 2015
7. Dr Polala Niranjan siripuri Kiran, BIG DATA ANALYTICS USING R-TOOLS REPORT: AICTE ISTE Induction/Refresher Program, 13 December 2019