



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Accredited A++ Grade by NAAC | 12B Status by UGC | Approved by AICTE

www.sathyabama.ac.in

School of Bio & Chemical Engineering

Department of Biotechnology

1.1.3 Courses having focus on Employability / Entrepreneurship / Skill Development

CONTENTS

B.Sc Bioinformatics and Data science

S. No.	Course Code	Course Name	Page No.
Employability related courses			
1.	SBIA1101	Introduction to Bioinformatics	5
2.	SBIA1102	Cell Biology and Biochemistry	6
3.	SBIA1201	Sequence Analysis	8
4.	SMTA1207	Basic Statistics	9
5.	SBIA1304	Programming With C And C++	10
6.	SBIA1301	Molecular Biology And Genomics	12
7.	SBIA1303	Proteomics And Interactomics	13
8.	SBIA1304	Perl For Bioinformatics	14
9.	SCSA1102	Fundamentals Of Python Programming	16
10.	SCSA1301	Database Management Systems	17
11.	SBIA1401	Structural Biology	18
12.	SBIA1402	Immunology And Immunoinformatics	19
13.	SBIA1501	Modelling Of Molecules	21
14.	SBIA1502	Drug Design And Development	22
15.	SBIA1503	Systems Biology	23
16.	SBIA1504	Cheminformatics	24
17.	SBIA1601	Datamining And Machine Learning	26
18.	SBIA1602	Microarray Data Analysis	27
19.	SCSA3016	Data Science	28



20.	SBIA3003	R Programming for Biologists	29
21.	SITA3008	Internet Of Things	30
22.	SBIA3004	Next Generation Sequencing	31
23.	SBIA3005	Quantitative Modelling in Biology	32
24.	SCSA3015	Deep Learning	33
25.	SCSA1501	Operating Systems	34
26.	SBIA3007	Bigdata In Healthcare	35
27.	SBIA3010	Health Informatics	36



S.No	Course Code	Course name	Page no
Skill development related courses			
1.	SBIA1101	Introduction to Bioinformatics	4
2.	SBIA1102	Cell Biology and Biochemistry	6
3.	SBIA1201	Sequence Analysis	8
4.	SMTA1207	Basic Statistics	9
5.	SCSA1202	Programming With C And C++	10
6.	SBIA1301	Molecular Biology and Genomics	12
7.	SBIA1303	Proteomics And Interactomics	13
8.	SBIA1304	Perl For Bioinformatics	14
9.	SCSA1102	Fundamentals Of Python Programming	16
10.	SCSA1301	Database Management Systems	17
11.	SBIA1401	Structural Biology	18
12.	SBIA1402	Immunology And Immunoinformatics	19
13.	SBIA1501	Modelling Of Molecules	21
14.	SBIA1502	Drug Design and Development	22
15.	SBIA1503	Systems Biology	23
16.	SBIA1504	Cheminformatics	24
17.	SBIA1601	Datamining And Machine Learning	26
18.	SBIA1602	Microarray Data Analysis	27
19.	SCSA3016	Data Science	28
20.	SBIA3003	R Programming for Biologists	29
21.	SITA3008	Internet Of Things	30
22.	SBIA3004	Next Generation Sequencing	31
23.	SBIA3005	Quantitative Modelling in Biology	32
24.	SCSA3015	Deep Learning	33
25.	SCSA1501	Operating Systems	34
26.	SBIA3007	Bigdata In Healthcare	35
27.	SBIA3010	Health Informatics	36



S.No	Course Code	Course name	Page no
Entrepreneurship			
1.	SMTA1207	Basic Statistics	9
2.	SBIA1402	Immunology and Immunoinformatics	19
3.	SBIA1504	Cheminformatics	24
4.	SCSA3016	Data Science	28
5.	SBIA3003	R Programming for biologists	29
6.	SITA3008	Internet of things	30
7.	SBIA3004	Next generation sequencing	31



SBIA1101	INTRODUCTION TO BIOINFORMATICS	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

The basic objective is to give students an introduction to bioinformatics and emphasis will be given to the application of bioinformatics in various fields of science with preliminary knowledge on biological databases.

Unit I Basics of Biomolecules

9 hrs

Cell Structures and Organelles. Introduction to Macromolecules like DNA, RNA and Proteins. Representation of biomolecules – Sequences and structures

Unit II Introduction to Bioinformatics

9 hrs

Bioinformatics -Definition, History, Scope and Applications. Fields related to Bioinformatics. Emerging areas of Bioinformatics

Unit III Biological Databases

9 hrs

Biological Databases: introduction, classification. Nucleotide sequence databases, Protein sequence databases, Protein structure database. Bibliographic databases – OMIM, PUBMED, and Medline.

Unit IV Formats and Visualisation

9 hrs

Sequence file formats. Sequence submission – Bank it and Sequin. Visualization of biomolecules-RASMOL, SWISS PDB viewer, Pymol.

Unit V Introduction to Computers

9 hrs

Overview and functions of a computer system, Input and output devices, Storage devices, Memory Introduction to Computers and Programming Languages. Internet, World Wide Web, Browsers

Course outcomes

CO1 - Describe the basic structures of biomolecules.

CO2 - Relate the role of bioinformatics in biology

CO3 - Summarize on the databases that are crucial in storing biological data

CO4 - Correlate few file formats and also illustrate the ways to submit sequences to a database

CO5 - Visualize the biomolecules in three dimension and understand the geometrical properties

CO6 - Comprehend the basics of computers and their role in handling biological data

Text / Reference books

1. Arthur Lesk Introduction to Bioinformatics 3rd Edition, Oxford University Press 2008
2. Jin Xiong Essential bioinformatics Cambridge University Press, New York 2006
3. Dan E. Krane Fundamental Concepts of Bioinformatics 1st Edition by Pearson 2002
4. Dr Darrell W Hajek Introduction To Computers CreateSpace Independent Publishing Platform 2017



SBIA1102	CELL BIOLOGY AND BIOCHEMISTRY	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

This course primarily aims to educate the students on the basics of cell biology and the biomolecules that are responsible for the existence and functioning of cells

UNIT I Basics of cell structure

9hrs

Basics of Cell Biology (structure & function) – Discovery of cell and Cell Theory; Comparison between plant and animal cells; Structure of Prokaryotic and Eukaryotic cells. Cell wall, plasma Membrane, Cytoplasm

Unit II Cell organelles

9hrs

Mitochondria; Chloroplast; ER; Golgi complex; Lysosome, endosome and microbodies; Ribosome; Centriole; Nucleus; Chemical components of a cell

UNIT III Cell Division and Transport

9hrs

Cell division - Mitosis & Meiosis, Cell cycle, Cell cycle Check points. Transport across cell membranes: Passive and active transports, Permeases. Basics of cell signaling- autocrine, paracrine and endocrine signaling.

UNIT IV Carbohydrates and Proteins

9hrs

Classification, Properties and Biological importance. Isomers, epimers, enantiomers, mutarotation, open chain and closed chain structures of glucose. Amino acids: classification- essential and non-essential amino acids, protein and non-protein amino acids, Zwitter ions. Proteins: Classification Structure of proteins: primary, secondary, tertiary and quaternary, biological significance.

UNIT V Lipids and Vitamins

9hrs

Lipids: Introduction, Classification, Properties and Biological importance. Fatty acid nomenclature and structure. Vitamins: fat soluble and water-soluble vitamins. Minerals: Micro and Macro minerals. Biological importance of vitamin and minerals, deficiency symptoms

Course outcomes

- CO1 – Describe the structures and purposes of basic components of prokaryotic and eukaryotic cells
- CO2 – Exhibit knowledge on the cell organelles and their functions
- CO3 – demonstrate the knowledge on the various aspects of cell division
- CO4 – Understand the basic structure of carbohydrates and proteins and their role in the function of cells
- CO5 – Comprehend the structure and function of lipids and vitamins
- CO6 – appreciate the basis of life and gain an understanding on the very existence of life

Text / Reference books

1. Darnell J Lodish, H. Baltimore, Molecular cell biology, Free Man, 1990.
2. Stryer, L., Biochemistry, 4th Edition, W.H. Freeman & Co., 2000.
3. Lehninger, Nelson and Cox, Principles of Biochemistry, W. H. Freeman, 4th Edition, 2005
4. Donald Voet, Judith Voet and Charlotte Pratt, Principles of Biochemistry, John Wiley and Sons, 2008



SBIA2101	CELL BIOLOGY AND BIOCHEMISTRY LAB	L	T	P	Credits	Total Marks
		0	0	4	2	100

LIST OF EXPERIMENTS

1. Isolation and viable cell count of bacteria by Pour plant and spread plate method
2. Study of chromosomal abnormalities (Permanent slides)
3. Preparation of buccal smear for identification of Barr bodies
4. Units, Volume and weight measurements, concentration units
5. pH measurement & Principle of pH meter
6. Principle and working of colorimeter and centrifuge
7. Qualitative analysis of Carbohydrates
8. Qualitative analysis of amino acids
9. Quantitative estimation of carbohydrates by anthrone method
10. Quantitative estimation of Protein by Lowry's method

SBIA2102	BIOLOGICAL DATABASES LAB	L	T	P	Credits	Total Marks
		0	0	4	2	100

LIST OF EXPERIMENTS

1. NCBI – GENBANK
2. EMBL
3. SWISSPROT
4. PDB
5. SEQUENCE SUBMISSION TOOLS – BANKIT AND SEQUIN
6. OMI, PUBMED
7. RASMOL



SBIA1201	SEQUENCE ANALYSIS	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

The students will be able to describe the contents and properties of the most important bioinformatics tools in sequence analysis for DNA and proteins, perform alignments and analyse and discuss the results in light of molecular biological knowledge

Unit 1 Sequence comparison

9hrs

Sequence comparison and alignment. Global and local alignments. Pairwise and multiple sequence alignment. Tools used – NEEDLE, WATER, CLUSTAL. Scoring matrices

Unit 2 Similarity search

9hrs

Basics of database similarity search. Homology search: BLAST- Algorithm and modules. FASTA - Algorithm and modules.

Unit 3 Phylogenetic analysis

9hrs

Terms in phylogenetic analysis. Relation with multiple sequence alignment. Phylogenetic trees. Steps in phylogenetic tree construction.

Unit 4 DNA analysis

9hrs

Feature identification of DNA: Gene finding. Restriction site analysis, ORF prediction, SNP identification, primer designing

Unit 5 Protein analysis

9hrs

Introduction to protein structure. Secondary structure prediction, Tertiary structure prediction. Domain identification using PROSITE.

Course outcomes

The students will be able to

CO1 – Describe the need for sequence comparison and understand its significance

CO2 – Infer the different kinds of alignment and rightly decide on their necessity

CO3 – Conclude on the most similar sequences for a given query

CO4 – Derive a relationship between sequences under comparison

CO5 – Formulate an approach in analysing a given DNA sequence

CO6 – Designing a protocol to study about an unknown protein

Text / Reference books

1. David W. Mount Bioinformatics: Sequence and Genome Analysis Cold Spring Harbor Laboratory Press 2004
2. Marketa Zvelebil, Jeremy O. Baum Understanding Bioinformatics Garland Science 2007
3. Richard C. Deonier, Simon Tavaré, Michael S. Waterman Computational Genome Analysis: An Introduction Springer 2010
4. Arthur Lesk Introduction to Bioinformatics 3rd Edition, Oxford University Press 2008



SMTA1207	BASIC STATISTICS	L	T	P	CREDITS	TOTAL MARKS
		3	0	0	3	100

COURSE OBJECTIVES

To be familiar with basic concepts sample space; To be familiar with discrete and continuous Random variables; To be familiar with correlation and regression analyses; To understand the standard distribution and Time Series.

UNIT I DESCRIPTIVE STATISTICS

9 Hrs.

Collection, Classification and Tabulation of data, Bar diagrams and Pie diagrams, Histogram, Frequency curve and frequency polygon, Ogives. Mean, median, mode, Standard deviation.

UNIT II CURVE FITTING, CORRELATION AND REGRESSION

9 Hrs.

Fitting a straight line and second-degree parabola by Least Squares. Correlation – Scatter diagram – Limits of correlation coefficient – Spearman's Rank correlation coefficient – Simple problems – Regression – Properties of Regression coefficients and regression lines.

UNIT III PROBABILITY THEORY& RANDOM VARIABLES

9 Hrs.

Probability theory, Random experiments, Sample Space, Conditional Probability, Baye's theorem, Random variable, (discrete and continuous), Probability density function (discrete and continuous), Distribution function (discrete and continuous random variable).

UNIT IV STANDARD DISTRIBUTION

9 Hrs.

Binomial distribution, Poisson distribution, Geometric Distribution, Uniform distribution, Exponential distribution, Normal distributions.

UNIT V TIME SERIES

9 Hrs.

Component of time series – Measurement of trend – Graphical Method – Semi average method - Moving average method – Measurement of seasonal variation – Method of simple average method – Ratio Moving average method .

TEXT / REFERENCE BOOKS

- David S. Moore, George P. McCabe, Bruce A. Craig. Introduction to the practice of statistics, 7th edition, W.H. Freeman and Company, New York 2011.
- PSS Sunder Rao and J. Richard, an introduction to Biostatistics and Research Methods, 5th edition PHI Learning Private Ltd, New Delhi, 2012.
- P. R. Vittal, Mathematical Statistics, Margham Publications, Chennai, 2002.
- J. N. Kapur and H. C. Saxena, Treatment and Content as in Mathematical Statistics, 20th Edition, S. Chand & Co. Ltd., New Delhi, 2010.
- Hogg, R.V. & Craig. A. T. (1998): Introduction to Mathematical Statistics, Macmillan.

COURSE OUTCOMES

- CO1 Identify, classify and tabulate the data. Construct various frequency charts. Evaluate the various measures of central tendency and measures of variation.
- CO2 Evaluate the Karl Pearson's correlation coefficient, Spearman's correlation coefficient and Regression Equations. Explore the relationship between correlation and regression. Fitting of straight line by least squares.
- CO3 Understand the basic concepts of probability and enable the students to apply Baye's theorem using conditional probability.
- CO4 Explain the standard discrete and continuous distributions and enable the students to use these distributions to solve real time problems.
- CO5 It empathizes the students about the importance of Time series and methods of Time series.
- CO6 Explore the various problems involving probability and to evaluate real time problems using the concepts of probability.



SCSA1202	PROGRAMMING WITH C AND C++	L	T	P	CREDITS	TOTAL MARKS
		3	0	0	3	100

COURSE OBJECTIVES

To learn the fundamental programming concepts and methodologies which are essential to building good C/C++ program.

UNIT I BASICS OF C PROGRAMMING

9Hrs

Introduction to C: Features of C - Structure of C program-Data Types-'C' Tokens-Input/output statements-Control Statement, Functions: – Types of Functions –Recursion.

UNIT II Arrays, Strings And Structures

9Hrs

Arrays: Single and Multidimensional Arrays— Array as Function Arguments, Strings: String Handling Functions, Structure: Nested Structures – Array of Structures – Structure as Function Argument–Function that Returns Structure, Union.

UNIT III Pointers And File Processing

9Hrs

Pointers: Introduction, Arrays Using Pointers – Structures Using Pointers – Functions Using Pointer, Dynamic Memory Allocation, Storage Classes, File Handling in 'C'.

UNIT IV Object Oriented Programming Concepts

9Hrs

Introduction-Procedure vs. object-oriented programming-Concepts: Classes and Objects-Operator & Function Overloading-Inheritance-Polymorphism and Virtual Functions.

UNIT V Templates and exception handling

9Hrs

Function Templates and Class Templates – Name spaces – Standard Template Library - Casting – Exception Handling – case study.

COURSE OUTCOMES

On completion of the course the student will be able to

- CO1. Develop simple applications in C using basic constructs.
- CO2. Design and Implement applications using arrays and strings.
- CO3. Develop and Implement applications using memory allocation and File concepts.
- CO4. Use proper class protection to provide security.
- CO5. Describe the reusability of code through Inheritance.
- CO6. Demonstrate the use of virtual functions to implement polymorphism.

TEXT / REFERENCE BOOKS

- Deitel and Deitel, C how to Program, 7th Edition, Pearson Education, 2013.
- Brian W.Kernighan and Dennis M.Ritchie, The C Programming Language, Pearson Education, 2015.
- Yashavant Kanetkar, Understanding Pointers in C, 4th Revised & Updated Edition, Bbp Publications, 2008.
- E Balagurusamy, Object Oriented Programming with C++, 3rd edition, Tata McGraw Hill, 2006.
- Bhave, Object Oriented Programming with C++, Pearson Education, 2004.
- John R Hubbard, "Programming with C++", Schaums Outline Series, McGraw Hill, 2nd edition, 2009.
- Bjarne Stroustrup, Programming: Principles and Practice using C++, 1st Edition, Addison Wesley Publications, 2008.



SBIA2201	SEQUENCE ANALYSIS LAB - I	L	T	P	Credits	Total Marks
		0	0	4	2	100

SUGGESTED LIST OF EXPERIMENTS

1. Pairwise Global Alignment
2. Pairwise Local Alignment
3. Homology search using BLAST and FASTA
4. MSA using CLUSTAL OMEGA
5. GENSCAN
6. GOR
7. SWISSMODEL

SCSA2105	PROBLEM SOLVING TECHNIQUES LAB	L	T	P	Credits	Total Marks
		0	0	4	2	100

COURSE OBJECTIVES

Identify the problem.
 To analyse the various steps in program development.
 Evaluate and select the best algorithm to solve the problem.
 Deploy suitable methods to get the desired output.
 Create the solutions for various Real-World Problems

SUGGESTED LIST OF EXPERIMENTS:

1. Program to find GCD.
2. Program to find the max and min from the three numbers.
3. Program to find Exponentiation.
4. Program to find sum of an array of numbers.
5. Program to implement Sine function computation.
6. Program to Generate the Fibonacci sequence.
7. Program to find the roots of a Quadratic equation.
8. Program for reversing the digits of an integer.
9. Program to find the smallest divisor of an integer.
10. Program to Generate Prime Numbers.
11. Program to Raise a Number to a Large Power.
12. Program for Removal of Duplicates.
13. Program to find the kth smallest Element.
14. Program to generate histogram.
15. Program for addition and multiplication of matrices.
16. Program that converts a number ranging from 1 to 50 to Roman equivalent
17. To delete n Characters from a given position in a given string.
18. Program to search for a Key value in a given list of integers using linear search method.
19. Program to sort the number in ascending and descending order.
20. Program for finding the factorial using recursive and non-recursive functions

COURSE OUTCOMES

On completion of the course, student will be able to
 CO1: Analyse and classify the given problem into various modules
 CO2: Analysing the complexity of problems, modularize the problems into small modules and then convert them into programs.
 CO3: Develop the codes containing looping and decision-making statements.
 CO4: Implement user defined functions.
 CO5: Apply recursion and call the function with appropriate p.

SBIA1301	MOLECULAR BIOLOGY AND GENOMICS	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

The students will be able to describe the marvel of molecular biology and as it is applied to molecular data science.

UNIT – I Introduction to nucleic acids

9hrs

Introduction- Nitrogenous bases - Purines and Pyrimidines - Nucleosides and Nucleotides -- Structure of nucleic acids – DNA - Watson & Crick model, RNA: m-RNA, t-RNA, r-RNA - Biological importance of nucleic acids. 16s rRNA and its significance.

UNIT – II Transcription and translation

9hrs

Concept of Central Dogma. DNA Replication - DNA Replication in prokaryotes. DNA Transcription. Genetic code and Translation. Gene regulation. Operon concept – Lac operon

UNIT – III Sequencing

9hrs

Genome sequencing methods: Maxim - Gilbert and Sanger's method, automated sequencing, pyrosequencing. Sequence assembly: Clone contig and shotgun approaches. DNA foot printing, Chromosome Walking, Chromosome Jumping.

UNIT – IV Mapping of genome

9hrs

Markers in genome mapping. Genetic maps: RFLP, SSLP and SNP. Physical maps: Restriction mapping, FISH and STS mapping. ESTs in genomic studies. DNA libraries – cDNA library and genome library

UNIT V – Post genomic era

9hrs

Genome Databases - Genome Annotation – Human genome project, Human Genetic Variation – Mutation - Single-Nucleotide Polymorphisms, DNA microarrays

Course outcomes

At the end of the course the students will

- CO1 – Describe the fundamentals of the gene and the genome.
- CO2 – Understand the basics of the macromolecules and their role in heredity.
- CO3 – Elucidate gene expression and its intricacies.
- CO4 – Establish the connection between the gene and the genome and its scaling.
- CO5 – To understand appreciate the projects and their relevance in the post genomic era.
- CO6 – To envision the applications in molecular data science as applied to physiology and medicine.

TEXT / REFERENCE BOOKS

1. Brown T.A, Gene Cloning an Introduction, VNR (UK) Co.Ltd. 1988.
2. Running, Mark. A Primer on Cell and Molecular Biology, Momentum Press, 2015.
3. Genomics : Volume I: Genomics, edited by Isidore Rigoutsos, and Gregory Stephanopoulos, Oxford University Press, Incorporated, 2006.
4. Mülhardt, Cornel. Molecular Biology and Genomics, Elsevier Science & Technology, 2007.
5. Genomics : Essential Methods, edited by Mike Starkey, and Ramnath Elaswarapu, John Wiley & Sons, Incorporated, 2010.
6. Brown, Stuart M., et al. Essentials of Medical Genomics, John Wiley & Sons, Incorporated, 2008.



SBIA1303	PROTEOMICS AND INTERACTOMICS	L	T	P	Credits	Total Marks
		3	1	0	4	100

COURSE OBJECTIVES

The student will be able to understand the methods to study about protein interactions with other biomolecules including label based and label free methods and suggest suitable approach for studying interactions.

Unit I Introduction

9hrs

Protein structure organization. Introduction to proteomics. Branches of proteomics. Applications and Scope of Proteomics.

Unit II Structural Proteomics

9hrs

Characterization of proteome, Protein isolation and purification – 2D gel electrophoresis, PAGE, Affinity Chromatography. Protein Sequencing - Edman Degradation. Protein Structure – Basics of X ray Crystallography, NMR and Mass spectrometry

Unit III Interaction Proteomics

9hrs

Interaction Proteomics: Protein - Protein Interactions – Characteristics and methods to study – Yeast two hybrid analysis, Phage display. Protein DNA interactions – DNA binding motifs

Unit IV Functional Proteomics and Interactomics

9hrs

Protein expression profiling - Protein Microarrays, Concepts of post transitional modifications. Introduction to interactomics. Label based and Label free methods to study interactions of proteins with other biomolecules

Unit V Developments in Proteomics

9hrs

Proteome databases. Proteome project. Comparative proteomics, Endogenous peptides and Peptidomics. Metaproteomics, Proteogenomics.

Course outcomes

At the end of the course the students will

- CO1 – describe about the field of proteomics and its applications
- CO2 – plan a method to study about a proteome from a species
- CO3 – illustrate the significance of interactions of proteins with other biomolecules
- CO4 – identify the right kind of technique to study interactions for multiple proteins
- CO5 – defend the utility of label based and label free methods of studying interactions
- CO6 – summarize on some of the latest concepts in proteomics

Texts / References

1. Richard Twyman Principles of Proteomics (Advanced Texts) Taylor & Francis 2004
2. Daniel C Liebler Introduction to Proteomics: Tools for the New Biology Humana books 2001
3. Dr. Richard Twyman Principles of Proteomics. Taylor & Francis 2013
4. C. David O'Connor, B. David Hames Proteomics: Methods Express Scion Publishing Ltd 2007



SBIA1304	PERL FOR BIOINFORMATICS	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

The student will be able to understand the basics of programming and use the Perl and its strengths for sequence manipulation and analysis.

UNIT I INTRODUCTION

9hrs

Biology and Computer Science - Getting Started with Perl - Perl's Benefits - Installing Perl - Running Perl Programs on various platforms - The Art of Programming - Programming Strategies - The Programming Process - Sequences and Strings - Representing Sequence Data - A Program to Store a DNA Sequence - Control Flow - Comments Revisited - Command Interpretation - Statements - Variables - Strings

UNIT II STARTING TO CODE AND GETTING HELP

9hrs

Concatenating DNA Fragments - Transcription: DNA to RNA - Using the Perl Documentation - Calculating the Reverse Complement in Perl - Proteins, Files, and Arrays - Reading Proteins in Files - Arrays - scalar and List Context - Motifs and Loops - Flow Control - Conditional Statements - open and unless

UNIT III WRITING MORE CODE

9hrs

Finding Motifs - Regular Expressions - Counting Nucleotides - Exploding Strings into Arrays - Operating on Strings - Writing to Files - Subroutines and Bugs - Subroutines - Scoping and Subroutines - Command-Line Arguments and Arrays - Passing Data to Subroutines - Modules and Libraries of Subroutines - Fixing Bugs in Your Code - The Perl Debugger

UNIT IV MANIPULATING MACROMOLECULES AND DATABASES

9hrs

Mutations and Randomization - A Program Using Randomization - A Program to Simulate DNA Mutation - Generating Random DNA - Analyzing DNA - The Genetic Code - Hashes - Data Structures and Algorithms for Biology - A Gene Expression Database - Gene Expression Data Using Unsorted Arrays - Gene Expression Data Using Hashes - Relational Databases - Translating Codons to Amino Acids - Translating DNA into Proteins - Reading DNA from Files in FASTA Format - Reading Frames - Translating Reading Frames

UNIT V RESTRICTION MAPS, REGULAR EXPRESSIONS AND MORE

9hrs

Regular Expressions - Logical Operators and the Range Operator - Finding the Restriction Sites - Perl Operations - GenBank - GenBank Files - Libraries - Separating Sequence and Annotation - Parsing Annotations - When to Use Regular Expressions - Protein Data Bank - Recursion - PDB Files - PDB File Format - Parsing PDB Files - BLAST - String Matching and Homology - BLAST Output Files - Parsing BLAST Output - The printf Function - Bioperl - Bioperl Tutorial Script

Course outcomes

- CO1 – Understand and write pseudo and code using Perl.
- CO2 – Write beginner level code in Perl for sequence manipulation.
- CO3 – Understand the deeper aspects of the Perl architecture and use them.
- CO4 – Identify the strengths of Perl and use them in sequence analysis.
- CO5 – Learn the fundamentals of object-oriented Perl.
- CO6 – Learn to automate Perl for bioinformatics applications.

TEXT / REFERENCE BOOKS

1. James Tisdall. Beginning Perl for Bioinformatics. O'Reilly Media, Inc. October 2E001. ISBN: 0596000804
2. James Tisdall. Mastering Perl for Bioinformatics. O'Reilly Media. March 2010
3. Michalski, Brent, et al. Beginning Perl, John Wiley & Sons, Incorporated, 2012
4. Dwyer, Rex A. Genomic Perl: From Bioinformatics Basics to Working Code, Cambridge University Press, 2002.
5. Dominus, Mark Jason. Higher-Order Perl: Transforming Programs with Programs, Elsevier Science & Technology, 2005.
6. Oualine, Steve. Wicked Cool Perl Scripts: Useful Perl Scripts That Solve Difficult Problems, No Starch Press, Incorporated, 2006.

Krumins, Peteris. Perl One-Liners: 130 Programs That Get Things Done, No Starch Press, Incorporated, 2013.



SBIA2301	PERL LAB	L	T	P	Credits	Total Marks
		0	0	4	2	100

SUGGESTED LIST OF EXERCISES

1. Simple Perl Scripts. Concatenation DNA fragments, Transcribing DNA into RNA
2. Scalar, Math, Strings, Operators to manipulate DNA/RNA/Protein sequence data.
3. Arrays. Extend the Perl script to do array manipulation. Parsing GenBank annotation using arrays
4. Uses of Hashes for the genetic code: translating codons into amino acids
5. Control. Extend the Perl script to show control structures. Uses of Do. Until Loops
6. Uses of subroutine to read FASTA files
7. Functions. Extend the Perl script to show examples of functions. Translate a DNA sequence in all six reading frames
8. System Calls. Extend the Perl script to call system functions:
9. Regular Expressions. Extend the Perl script to use regular expressions:
10. A Useful Script Extend the Perl script to do something useful. Extract sequence chains from PDB file.
11. Installing Bioperl, Uses of Bioperl modules for sequence manipulation, accessing local database

SBIA2302	MOLECULAR BIOLOGY LAB	L	T	P	Credits	Total Marks
		0	0	4	2	100

SUGGESTED LIST OF EXPERIMENTS

1. Preparation of various stock solutions required.
2. Isolation of genomic DNA from Bacteria
3. Visualization of DNA by Agarose Gel Electrophoresis.
4. Visualization of proteins by Polyacrylamide Gel Electrophoresis (SDS-PAGE).
5. Demonstration of Polymerase Chain Reaction (PCR) technique



SCSA1102	FUNDAMENTALS OF PYTHON PROGRAMMING	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

To understand why Python is a useful scripting language for developers and learn how to design and program Python applications.

UNIT I INTRODUCTION

9hrs

History of Python- Introduction to the IDLE interpreter (shell) -Expressions – Data Types - Built-in function - Conditional statements - Iterative statements- Input/output - **Compound Data Types - Nested compound statements** – Introduction to Object Oriented Concepts.

UNIT II FILES AND EXCEPTIONS HANDLING, MODULES, PACKAGES

9hrs

File Operations –Iterators - Exception handling - Regular Expressions- Creating Modules-Import Statement-Introduction to PIP- **Installing Packages via PIP-Using Python Packages**

UNIT III GUI PROGRAMMING

9hrs

GUI Programming in Python - Introduction to GUI library - Layout management - Events and bindings - **Fonts – Colours - Canvas - Widgets (frame, label, button, check box, entry, list box, message, radio button, text, spinbox).**

UNIT IV DATABASE AND NETWORK

9hrs

Database (using NoSQL): Connector Module –Cursor – Statements - Exceptions in database.
Network connectivity: Socket module - **Client – Server – Email – URL Access**

UNIT V CASE STUDY

9hrs

Web Programming using Python. **Image Processing – Facebook Analysis – Twitter Analysis.**

COURSE OUTCOMES

On completion of the course the student will be able to

- CO1**-Describe the Numbers, Math functions, Strings, List, Tuples and Dictionaries in Python
- CO2**-Do the decision Making and write functions in Python
- CO3**-Explain how to design GUI Applications in Python and evaluate different database operations
- CO4**-Design and develop Client Server network applications using Python
- CO5**-Ability to design real life situational problems and think creatively about solutions of them.
- CO6**-Apply the best features of mathematics, engineering and natural sciences to program real life problems.

Text Books/Reference Books

- Y. Daniel Liang, Introduction to Programming Using Python, Pearson, 2013.
- Paul Gries , Jennifer Campbell, Jason Montojo, Practical Programming:
- AnIntroduction to Computer Science Using Python 3, Pragmatic Bookshelf, 2/E 2014
- Magnus Lie Hetland, Beginning Python: From Novice to Professional, Apress



SCSA1301	DATABASE MANAGEMENT SYSTEMS	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

To understand terms related to database design and management and understand database security and performance issues

UNIT I Introduction to Databases

9Hrs

Databases and database users – Database system concepts and architecture – Data modelling using entity Relationship (ER) model – Enhanced ER model- Relational Model - The Relational Data Model and Relational Database Constraints - The Relational Algebra and Relational Calculus.

UNIT II Database Design

9Hrs

Overview of the Hierarchical Data Model - Overview of the Network Data Model – Relational database design: Mapping ER Model to Relational Model - Commercial query languages: QBE – Functional dependency – Normalization

UNIT III Query Processing

9Hrs

SQL Queries --Embedded SQL -My SQL: Basics, Queries in MySQL and Algorithms for Query Processing and Optimization - Introduction to Transaction Processing Concepts and Theory - Concurrency control techniques.

UNIT IV Recovery and Security

9Hrs

Database Recovery Techniques - Database Security - Distributed databases and Client- Server Architecture

UNIT V Object Database and Current Trends

9Hrs

Concepts for Object Database - Emerging Database Technologies and Application - Introduction to Data warehousing & Data mining –Application of Data mining

Course Outcomes

On completion of the course the student will be able to

- CO1-** Design ER-models to represent simple database application scenarios
- CO2-**Ability to understand and design data modelling using Entity-Relationship model
- CO3-**Implement SQL to a broad range of query and data update problems
- CO4-**Describe distributed databases and use database recovery mechanisms
- CO5-**Familiar with data warehousing and data mining applications
- CO6-**Apply Normalization techniques to normalize and improve the database design

TEXT / REFERENCE BOOKS:

1. Elmasri & Navathe, "Fundamentals of Database Systems", 6th Edition, Addison Wesley, 2011.
2. Abraham Silberschatz, Henry.F.Korth and S.Sudharshan, "Database System Concepts", 4th Edition, 2002.
3. Jiawei Han and Micheline Kamber, "Data Mining Concepts and Techniques", 3rd Edition, Elsevier, 2012
4. Jan L. Harrington, "Object oriented database design", Harcourt India private limited 2000



SBIA1401	STRUCTURAL BIOLOGY	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

The main objective of the course is to understand the basic relation between macromolecular structure and function that are derived by experimental techniques like XRD and NMR along with other chromatographic techniques.

Unit I Basics on Interactions

9 hrs

Stabilizing interactions: Covalent and non - covalent interactions, thermodynamic principles and chemical kinetics, criteria for aromaticity, reaction kinetics and energy profile diagram, pseudo-order reaction, isotope labeling, kinetically and thermodynamically controlled reaction, implications in peptide and protein structural stabilization and folding pathways.

Unit II Structure determination I

9 hrs

Structure determination of macromolecules: XRD - Bragg's law - crystal systems - Symmetry elements: - Crystal symmetry - rotation axis, mirror planes and center of inversion, proper/ improper axes of rotation, translational symmetry- screw axis and glide planes; Seven crystal systems - Application of XRD - Calculating electron density and Patterson maps (Fourier transform and Structure factors, convolutions), phases

Unit III Structure determination II

9 hrs

Nuclear Magnetic Resonance Criteria for NMR sensitivity - Shielding and de - shielding effects - NOIE effect - Chemical shifts - coupling constants - spin - spin splitting - Carbon-13 NMR - Instrumentation - Application of NMR to biology.

Unit IV Chromatography

9 hrs

Chromatography – Principles, Types – Thin Layer Chromatography, Column chromatography, Molecular sieving chromatography, Affinity chromatography, HPLC, GC - Radioactive Techniques-Liquid scintillation counting- Applications of chromatography.

UNIT V

9 hrs

Sequences Sequence alignments BLAST and other sequence search tools/ Sequence databases/ MSA / Phylogenetic trees / file types Annotation, names, ontology Structural alignment and Classification: Classification methods - necessity to compare proteins - similar sequence / function / structure.

Course Outcomes

At the end of the course the students will

- CO1** – Explain the relationship between protein sequence and protein structure.
- CO2** – Describe how structure translates into function within different biological fields such as catalysis, transport and regulation.
- CO3** – Estimate the validity of information in macromolecular structure databases
- CO4** – Use computer programs to visualise and analyse macromolecular structures from a functional perspective
- CO5** – Use bioinformatics tools for sequence alignment, sequence motif identification and prediction of secondary and tertiary structures.
- CO6** – Account for the purpose, theoretical background, ethical aspects and limitations of the above mentioned bioinformatics methods and use this knowledge to interpret relevant results.

Text / Reference Books

- Upadhyay., Upadhyay., Nath, "Biophysical Chemistry – Principles and Techniques", Himalaya Publishing House, 2002.
- Introduction to protein structure, C. Branden and J. Tooze
- X-Ray Structure Determination: A Practical Guide, 2nd Edition, by George H. Stout , Lyle H. Jensen.



SBIA1402	IMMUNOLOGY AND IMMUNOINFORMATICS	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

This course will enable the students to understand the immune system, its components and their functions

Unit I Introduction to immune system

9hrs

Innate and adaptive immune modulatory in vertebrates, Antigen processing and presentation Contemporary challenges to the immune system Infectious diseases Autoimmune diseases Categories of vaccines

Unit II Genetics of immunology

9hrs

Immunoglobulin classes and subclasses DR and LDR regions and sequence numbering, Immunogenetics & immunogenomics, Hybridoma technology: applications, Humanization of antibodies by design, Membrane receptors for antigen

Unit III Immune response

9hrs

The B-cell surface receptor for antigen (BCR), The T-cell surface receptor for antigen (TCR), The major histocompatibility complex (MHC) MHC polymorphism, Causes of MHC polymorphism, MHC supertypes, Epitopes, Affinity Maturation, Recognition of Antigen by B cells, Neutralizing Antibody

Unit IV Databases in immunology

9hrs

Databases in Immunology, Reverse vaccinology, rational vaccine design. Evolution and escape due to variations, Immunoinformatics databases & tools IMGT & IEDB BciPep, Epiteome, CED, AgAbDb

Unit V Computational immunology

9hrs

Reverse vaccinology and immunoinformatic, MHC- peptide, Specificity characterization, Ag-Ab: Sequential and conformational epitopes, HLA: nomenclature, HLA-peptide interactions & matrices, Epitope prediction algorithms: B-cell epitope predictions using sequence-based approaches, B-cell epitope predictions using structure-based approaches, T-cell epitope prediction methods

Course outcomes

At the end of the course the students will

- CO1 – Understand and explore sequence and structural databases relevant in the area of immunology
- CO2 – Study informatics-based approaches for prediction of epitopes, design of vaccines and immuno-diagnostic tools
- CO3 – Characterize and understand principles of antigen-antibody interactions
- CO4 – Understand algorithms and methods for prediction of epitopes
- CO5 – Explore and use approaches for vaccine design
- CO6 – Implement various computational tools and approaches to understand immunological issues

Text / Reference Books

1. Flower Darren R. Bioinformatics for Vaccinology Publisher: UK, John Wiley & Sons Inc. 2008. ISBN: 9780470027110
2. Flower Darren R. Immunoinformatics: Predicting Immunogenicity In Silico perl: New Jersey, Humana Press. 2007. ISBN: 9781603271189.
3. Kindt, Thomas J., Osborne Barbara A. Goldsby Richard A. Kuby Immunology 6th Edition. Publisher: New York, W. H. Freeman. 2007. ISBN: 9780716785903
4. Lund Ole , Nielsen Morten , Lundegaard Claus, Kesmir Can, Brunak SÅren. Immunological Bioinformatics. Publishers: London, MIT Press 2005 ISBN: 0262122804.
5. Foundation Novartis. Immunoinformatics: Bioinformatic Strategies for Better Understanding of Immune Function. Publisher: Chichester, John Wiley & Sons Inc. 2003. ISBN: 0470853565.



SBIA2101	CELL BIOLOGY AND BIOCHEMISTRY LAB	L
		0

1. Program to handle Input and Output statements
2. Program using Lists, Tuples, Dictionary and Sets
3. Program using conditional statements
4. Program using looping constructs
5. Program using functions and modules
6. Program using File handling
7. Program using Exception handling
8. GUI Programming using Python - Canvas Widget
9. GUI Programming using Python - Frame, label, button, check box, entry, listbox, message, radio button, text, spinbox
10. Database Programming using Python
11. Socket Programming using Python

SBIA2401	SEQUENCE ANALYSIS- II LAB	L	T	P	CREDITS	TOTAL MARKS
		0	0	4	2	100

1. DNA sequence analysis
 - a. Primer designing
 - b. ORF finder
 - c. Promoter analysis
 - d. Restriction site identification
 - e. Translation of DNA sequence
2. Protein sequence analysis
 - a. Protparam
 - b. Castp
 - c. Peptide cutter
 - d. Prosite



SBIA1501	MODELLING OF MOLECULES	L	T	P	Credits	Total Marks
		3	1	0	4	100

COURSE OBJECTIVES

The course intends to impart knowledge on the various concepts of molecular modelling and its application in drug designing with a focus on identification of lead molecules and the concept of pharmacophore.

Unit I Introduction to quantum mechanics

9hrs

Basic Concepts in Molecular Modelling: Introduction, Coordinate systems, Energy surfaces. Introduction to Quantum mechanics - Schrodinger wave equation, Time dependent and time independent forms.

Unit II Basics of molecular mechanics

9hrs

Features of molecular mechanics, force fields; Bond structure and bending angles –electrostatic, van der Waals and non –bonded interactions, hydrogen bonding; Energy minimization; Application of energy minimization.

Unit III Molecular mechanics

9hrs

Molecular Mechanics Introduction, Vander Waals and Dipole-Dipole Interaction, Types of Potential- Lennard- Jones, Truncated Lennard – Jones, Ionic and Polar potentials, Types of force field- AMBER, CHARMM, Merck Molecular Force Field, Consistent Force Field, MM2, MM3 and MM4 Force Field.

Unit IV Molecular dynamics

9hrs

Molecular Dynamics using simple models; Molecular Dynamics with continuous potentials and at constant temperature and pressure; Time-dependent properties; Solvent effects in Molecular Dynamics; Conformational changes from Molecular Dynamics simulation. Introduction, Newton's equation of motion, equilibrium point, radial distribution function, pair correlation functions, MD methodology, periodic box, algorithm for time dependence; leapfrog algorithm, Verlet algorithm, Boltzman velocity, time steps, duration of the MD run.

Unit V Simulation and Modelling

9hrs

Methods of computer aided drug design, ligand design methods, drug design approaches, Target identification and validation, lead optimization and validation, Structure and ligand-based drug design, modelling of target-small molecule interactions.

Course outcomes

At the end of the course the students will

- CO1 –Be able to apply and understand the modelling on the basis of quantum mechanical aspects.
- CO2 –To discuss the various types of interactions possible between the macromolecule and the ligand.
- CO3 –Identify the lead molecule by virtual screening.
- CO4 –Design the strategy towards drug design.
- CO5 –Design the strategy towards drug design based on the structural aspects of the macromolecule.
- CO6 –Adopt and implement the concepts of modelling to derive structures for molecules

Text / Reference books:

1. Andrew Leach - Molecular Modelling: Principles and Applications, 2nd Edition. Pearson Education EMA, 2001.
2. R.K. Prasad - Quantum Chemistry, 3rd Edition, 2006.



SBIA1502	DRUG DESIGN AND DEVELOPMENT	L	T	P	Credits	Total Marks
		3	1	0	4	100

COURSE OBJECTIVES

To understand the basics of drug design and to know the significance of pharmacophores in drug design.

UNIT I Introduction to Drugs

9hrs

Drugs- Classification – Chemical basis of classification, Legal classification of drugs. Soft drugs and Hard drugs, Prodrugs – classification, Application

Unit II Drug Targets and Activity

9hrs

Drug targets - Enzymes, Receptors, Proteins, Nucleic acids, Lipids. Drug solubility - Effect of pH, Effect of polarity and pKa. Drug Metabolism - Absorption, Distribution, Metabolism, Elimination. Pharmacokinetics and Pharmacokinetics of drugs

UNIT III Drug Discovery

9hrs

Introduction to Drug Discovery and Development, stages of drug discovery, Natural sources of drugs, Computers in drug discovery, Rational approaches to lead discovery based on traditional medicine, ADME/T analysis, Oral bioavailability, Drug likeliness, Lipinski's rule.

Unit IV Drug Design

9hrs

Types of drug design- Rational drug design, Structure based – Receptor based and ligand based. Virtual Screening, Hit to Lead identification, Lead optimisation. Pharmacophore analysis and mapping.

Unit V Docking and QSAR

9hrs

Molecular docking – steps, types of docking, Pose identification, applications, software used. QSAR - origin & development, Physicochemical parameters, partition coefficient, Hammett's substitution constant, Taft steric constant, Hansch analysis, Free Wilson analysis, 3D QSAR methods CoMFA & CoMSIA.

Course Outcomes

At the end of the course the students will

- CO1 –realize the concept of drugs and their classification
- CO2 –explore the various targets of drugs and identify the factors influencing the activity
- CO3 –Comprehend the process of drug discovery.
- CO4 –Explore the drug design concept and they basics
- CO5 –Appreciate the significance of docking and QSAR in drug design.
- CO6 –Create an approach to design drugs based on the structural knowledge available

Text / Reference Books

- 1) Shayne C. Gad, (2005), Drug Discovery Handbook, 2nd edition, John Wiley & Sons: New York.
- 2) Tudor I. Opera, (2005), Cheminformatics in Drug Discovery, 5th edition, John Wiley & Sons: New York.



SBIA1503	SYSTEMS BIOLOGY	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

To understand the basic concepts in Systems Biology and to apply them to real life problems in biology and medicine.

UNIT I Biological Basis

9hrs

Introduction – systems understanding – what is systems biology? - Reductionism or holism? - Mechanism or emergentist? - Homeostasis or robustness? - What type of systems disease is cancer? - What are the challenges facing systems and cancer biology? - is systems biology a scientific revolution?

UNIT II Integrative Analysis of Omics Data

9hrs

Integrative Analysis of Omics Data - Summary - Introduction - Omics Data and Their Measurement Platforms - Data Processing: Quality Assessment, Quantification, Normalization, and Statistical Analysis - Data Integration: From a List of Genes to Biological Meaning - **Outlook and Perspectives**

UNIT III Flux Analysis

9hrs

Flux Analysis in Biotechnology and Medicine - Introduction - Theoretical Foundations of Emerging Challenges - MFA - Metabolic Flux Analysis in Biotechnology - Metabolic Flux Analysis in Medicine - Metabolic Modelling for Design of Cell Factories - Building and Refining Genome-Scale Metabolic Models - **Strain Design Algorithms - Case Studies**

UNIT III Genome-Scale Metabolic Modelling

9hrs

Genome-Scale Metabolic Modelling and In silico Strain Design of Escherichia coli - The COBRA Approach - In silico Model-Based Strain Design of E. coli Cell Factories - Future Directions of Model-Guided Strain Design in E. coli References - Accelerating the Drug Development Pipeline with Genome-Scale Metabolic Network Reconstructions - Metabolic Reconstructions in the Drug Development Pipeline - **The Human Reconstruction - Community Models - Personalized Medicine**

UNIT IV Computational Modelling of Microbial Communities

9hrs

Computational Modelling of Microbial Communities - Ecological Models - Genome-Scale Metabolic Models - Drug Targeting of the Human Microbiome - The Human Microbiome - Association of the Human Microbiome with Human Diseases - Drug Targeting of the **Human Microbiome - Future Perspectives.**

UNIT V Modelling Biological Processes

9hrs

Signal Transduction - Toward Genome-Scale Models of Signal Transduction Networks - The Potential of Network Reconstruction - Information Transfer Networks - Approaches to Reconstruction of ITNs – Aging - Systems Biology of Aging - The Biology of Aging - The Mathematics of Aging - Future Challenges – Immune response - Modelling the Dynamics of the Immune Response - Background - **Dynamics of NF- κ B Signalling - JAK/STAT Signalling**

Course outcomes

At the end of the course the students will be able to

CO1 –To understand the basic concepts in Systems Biology

CO2 –To learn to apply systems biology to omics data.

CO3 –To appreciate the importance of flux balance analysis and its importance.

CO4 –To understand and perform genome scale modelling.

CO5 –To learn model microbial communities and their interactions

CO6 –To model specific biological processes.

Text / Reference books

1. Nielsen, Jens, and Stefan Hohmann. Systems Biology, John Wiley & Sons, Incorporated, 2017.
 2. Systems Biology, edited by Robert A. Meyers, John Wiley & Sons, Incorporated, 2012.
 3. Systems Biology: Philosophical Foundations, edited by Fred C. Boogerd, et al., Elsevier Science & Technology, 2007.
 4. Lodhi, Huma M., and Stephen H. Muggleton. Elements of Computational Systems Biology, John Wiley & Sons, Incorporated, 2010.
 5. Ingalls, Brian P. Mathematical Modelling in Systems Biology: An Introduction, MIT Press, 2013.
 6. Juan, Hsueh-Fen. Systems Biology: Applications in Cancer-Related Research, edited by Hsuan-Cheng Huang, World Scientific Publishing Co Pte Ltd, 2012.
- Young, Daniel L., and Seth Michelson. Systems Biology in Drug Discovery and Development, John Wiley & Sons, Incorporated, 2011.



SBIA1504	CHEMINFORMATICS	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

The course is designed to make the students realize the role of chemical molecules in biology with an emphasis on its physical parameters and conformations and their role in drug discovery.

UNIT I Introduction To Cheminformatics

9 Hrs.

Introduction to Cheminformatics - History and evolution of Cheminformatics - Chemical representation - Sequence, 2D, 3D structure, Types of chemical representation - linear notation, tabular storage, graphical representation

UNIT II Chemical Databases

9 Hrs.

Chemical Databases - Cambridge Structural Database CCDC CSD, Crystallographic Open Database COD Protein Data Bank PDB Ligand Explorer, ChempSpider.

UNIT III Combinatorial Chemistry

9 Hrs.

Combinatorial chemistry: Introduction, Liquid Phase synthesis, Solid phase synthesis, Designing Combinatorial Synthesis

UNIT IV Library Design

9 Hrs.

Combinatorial library design - Rational principles of compound selection for combinatorial library design and optimization approach, Modelling toxicity, Computer Assisted Synthesis design and structure based library design.

UNIT V Applications

9 Hrs.

Application of cheminformatics - QSPR Drug design - Target identification and Validation, lead finding and optimization - Pharmacophore-Based Drug Design - Structure-Based Drug design

Course Outcomes

At the end of the course the students will

CO1 –Describe the basics of cheminformatics and understand the representation of molecules

CO2 –Access the various chemical databases that form the basis for designing and studying new molecules.

CO3 –They will be able to appreciate the methods of synthesising new molecules.

CO4 –screen various compounds in the combinatorial library for the identification of lead molecules

CO5 –Analyse role of chemical molecules in drug design

CO6 –Comprehend the role of cheminformatics in drug design

Text / Reference Books

1. Andrew R Leach, V J Gillet An Introduction to Chemoinformatics Springer 2007
2. Jurgen Bajorath Chemoinformatics: Concepts, Methods, and Tools for Drug Discovery Humana Press 2010
3. Jurgen Bajorath Chemoinformatics for Drug Discovery Wiley 2013
4. Daria Bove Computational Chemistry: Theories, Methods and Applications Nova Science Publishers Inc, New York 2014



SBIA2502	CHEMINFORMATICS AND DRUG DESIGN LAB	L	T	P	Credits	Total Marks
		0	0	4	2	100

1. Visualization of Bio molecules – RASMOL, Pymol
2. Chemical databases
3. Drawing and Visualization of molecules – 2D and 3D - ChemSketch
4. Interconversion of file formats -OpenBabel
5. Molecular Modelling: Homology & Ab initio
6. SAVS - Validation of the Modelled Structure
7. ADMET properties
8. Energy Minimization
9. Interaction of protein with small molecules - Docking

SBIA2501	SYSTEMS BIOLOGY LAB	L	T	P	Credits	Total Marks
		0	0	4	2	100

SUGGESTED LIST OF EXPERIMENTS

1. Modelling and Visualization
2. E cell
3. Copasi
4. Construction and Analysis of gene expression networks
5. Mat lab I
6. Mat Lab II
7. Cell Designer
8. Cytoscape



SBIA1601	DATAMINING AND MACHINE LEARNING	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

The main objective is to understand various concepts of datamining and machine learning and implement various advanced mathematical, physical and statistical techniques that are being applied for analyzing information biological data.

Unit I Data warehousing

9hrs

Introduction – Operational data sources- data warehousing - Data warehousing design – Guidelines for data warehousing implementation - Data warehousing metadata An introduction to Genome databases – resources, file formats, functional annotations. Advance Database Systems and Applications, Data Mining Functionalities, Classification of Data Mining Systems, Major issues in Data Mining.

Unit II Data mining

9hrs

Data mining application – data mining techniques – data mining case studies- the future of data mining – data mining software - Association rules mining: Introduction- basics- task and a naïve algorithm- apriori algorithm – improve the efficient of the apriori algorithm – mining frequent pattern without candidate generation

Unit III Classification

9hrs

Classification : Introduction – decision tree – over fitting and pruning - DT rules- estimation predictive accuracy of classification methods - other evaluation criteria for classification method – classification software

Unit IV Cluster analysis

9hrs

cluster analysis – types of data – computing distances types of cluster analysis methods - partitioned methods – hierarchical methods – density based methods – dealing with large databases – quality and validity of cluster analysis methods - cluster analysis software.

Unit V Machine Learning

9hrs

Machine Learning Techniques - Introduction to Support Vector Machines and their applications, - Introduction to Neural Networks Classifying Samples from two populations using Multilayer Perceptron and back propagation - Using genetic algorithm and Perceptron for feature selection and supervised classification - Applications of ANN, SVM etc. Case studies on Identifying gene coding regions, Structure prediction, AI in healthcare

Course outcomes

At the end of the course the students will

- CO1 - Understand various data collection methods.
- CO2 - Comprehend the role of data mining in the field of biology
- CO3 - Approach data ware house and data mining as a process of data collection and processing.
- CO4 - Discover patterns from large amounts of data to analyze and extract patterns
- CO5 - Apply suitable methods of data analysis for a given problem.
- CO6 - Make predictions from the given data sets using classification and clustering algorithms

References

1. J. Han and M. Kamber, Jian Pei (2012) *Data mining: concepts and techniques*. 3rd edition, Elsevier
2. T. Hastie, R. Tibshirani, and J. Friedman (2001) *The Elements of Statistical Learning: data mining, inference and prediction*. Springer Verlag.
3. Data Mining Concepts and Techniques – Jiawei Han, Micheline Kamber, Academic Press Morgan Kaufman Publishers. 2006 Reference Books:
4. Data Mining: Practical machine learning tools Techniques with java implementation by Ian H.Witten, Eibe Frank, 2005.
5. Machine Learning and data mining in pattern recognition in third International conference MLDM, by Petra Perner and Azriel Rosenfeld, Springer.2003
6. Witten and Frank, Data Mining: Practical Machine Learning Tools and Techniques by, Elsevier.



SBIA1602	MICROARRAY DATA ANALYSIS	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

This course gives the technical knowledge on Microarray techniques and data analysis DNA Microarray and its statistical analysis of RNA data Statistical computing and Statistical Genetics

UNIT I DNA Microarrays

9 hrs

The Technical Foundations, Why are Microarray Important? What is a DNA Microarray?, Designing a Microarray Experiment-The Basic steps, Types..

UNIT II Microarray Databases

9 hrs

NCBI and Microarray Data Management, GEO (Gene Expression Omnibus), MAML, The benefits of GEO and MAML, The Promise of Microarray Technology in Treating Disease

UNIT III Microarray Data Normalization

9 hrs

Microarray Data Pre-processing, Data-Data normalization, Measuring Dissimilarity of Expression Pattern-Distance Motifs and Dissimilarity measures, Visualizing Microarray Data-Principal Component Analysis, Microarray Data.

UNIT IV Microarray Data Analysis

9 hrs

K Means Clustering, Hierarchical Clustering, Self-Organization Maps (SOM), Identifying Genes: Expressed usually in a sample-Expressed significantly in population-Expressed differently in two populations, Classifying Samples from two populations using Multilayer Perceptron, Support Vector Machines and their applications, Using genetic algorithm and Perceptron for feature selection and supervised classification.

UNIT V Microarray Applications

9 hrs

Gene Ontology and pathway analysis, Promoter analysis and gene regulatory network, Co expression analysis, CGH & Genotyping chips, Chromosome aberration and polymorphism via genome-wide scanning, Future direction of microarray approach, Pharmacogenomics, Toxicogenomics, Data mining.

Course outcomes

At the end of the course the students will

CO1 –To familiarize students with the DNA microarray technology and with the analysis of gene expression data

CO2 –To ensure familiarity with the different microarray techniques and their everyday use in a biological laboratory setting.

CO3 –To present some of the important low level analysis (Normalization, quality control, etc.).

CO4 –To provide a sound introduction to a good experimental design.

CO5 –To enable the students to access and analyse microarray data

CO6 – To understand the applications of microarray analysis in various fields of research

TEXT / REFERENCES BOOKS

1. Arun Jogota, "Microarray Data Analysis and Visualization", The Bay Press, 2001.
2. Ernst Wit and John McClure, "Statistics for Microarrays Design", Analysis and Inference, John Wiley & Sons, 2004.
3. Steen Knudsen, "Guide to analysis of DNA Microarray data", John Wiley & Sons, 2004.
4. Dov Stekel, "Microarray Bioinformatics", Cambridge University Press, 2003.
5. Uwe R. Müller, Dan V. Nicolau, "Microarray Technology and Its Applications", Springer, 2005.
6. Emanuele de Rinaldis, Armin Lahm, "DNA Microarrays: Current applications", Horizon Scientific Press, 2007



SCSA3016	DATA SCIENCE	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

- To understand the mathematical foundations required for data science.
- To describe a flow process for data science problems.
- To introduce basic data science algorithms and data visualization.
- To learn machine tools and techniques.
- To learn the ideas and tools for data visualization.

UNIT I Linear Algebra

9 Hrs.

Algebraic view – vectors 2D, 3D and nD, matrices, product of matrix & vector, rank, null space, solution of over determined set of equations and pseudo-inverse. Geometric view - vectors, distance, projections, eigenvalue decomposition, Equations of line, plane, hyperplane, circle, sphere, Hypersphere.

UNIT II Probability And Statistics

9 Hrs.

Introduction to probability and statistics, Population and sample, Normal and Gaussian distributions, Probability Density Function, Descriptive statistics, notion of probability, distributions, mean, variance, covariance, covariance matrix, understanding univariate and multivariate normal distributions, introduction to hypothesis testing, confidence interval for estimates.

UNIT III Exploratory Data Analysis And The Data Science Process

9 Hrs.

Exploratory Data Analysis and the Data Science Process - Basic tools (plots, graphs and summary statistics) of EDA - Philosophy of EDA - The Data Science Process - Data Visualization - Basic principles, ideas and tools for data visualization - Examples of exciting projects- Data Visualization using Tableau.

UNIT IV Machine Learning Tools, Techniques And Applications

9 Hrs.

Supervised Learning, Unsupervised Learning, Reinforcement Learning, Dimensionality Reduction, Principal Component Analysis, Classification and Regression models, Tree and Bayesian network models, Neural Networks, Testing, Evaluation and Validation of Models.

UNIT V Introduction To Python

9 Hrs.

Data structures-Functions-Numpy-Matplotlib-Pandas- problems based on computational complexity-Simple case studies based on python (Binary search, common elements in list), Hash tables, Dictionary.

Max. 45 Hrs.

COURSE OUTCOMES

On completion of the course, student will be able to

- CO1 - Explain the basic terms of Linear Algebra and Statistical Inference.
- CO2 - Describe the Data Science process and its components interaction.
- CO3 - Apply EDA and the Data Science process in a case study.
- CO4 - Classify Data Science problems.
- CO5 - Analyse and correlate the results to the solutions.
- CO6 - Simulate Data Visualization in exciting projects.

TEXT / REFERENCE BOOKS

1. Cathy O'Neil and Rachel Schutt. Doing Data Science, Straight Talk From The Frontline. O'Reilly. 2014.
2. Introduction To Linear Algebra - By Gilbert Strang, Wellesley-Cambridge Press, 5th Edition.2016.
3. Applied Statistics And Probability For Engineers – By Douglas Montgomery.2016.
4. Jure Leskovek, Anand Rajaraman and Jeffrey Ullman. Mining of Massive Datasets. v2.1, Cambridge University Press. 2014. (free online)
5. Avrim Blum, John Hopcroft and Ravindran Kannan. Foundations of Data Science.
6. Jiawei Han, Micheline Kamber and Jian Pei. Data Mining: Concepts and Techniques, 3rd Edition. ISBN 0123814790. 2011.
7. Trevor Hastie, Robert Tibshirani and Jerome Friedman. Elements of Statistical Learning, 2nd Edition. ISBN 0387952845. 2009. (free online)
8. LEARN PYTHON 3 THE HARD WAY, 3rd Edition, Zed A. Shaw
9. Python for Data Analysis Oreilly, Wes McKinney



SBIA3003	R PROGRAMMING FOR BIOLOGISTS	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

This course provides knowledge to handle biological problems in Bioinformatics and computational biology

UNIT I R Language

9hrs

Motivation of R, Attributes and some special values, Types of objects, Sequence generating and vector subsetting, Types of functions, Data Structures, Managing R session, Evaluations.

UNIT II Object Oriented Programming In R

9hrs

OOP in R, Basics of OOP, S3 OOP, S4 OOP, Documentation, Debugging, Managing S3 and S4 together.

UNIT III Input And Output In R

9hrs

Basic file handling, connections, File input and output, Tools for accessing files on the Internet.

UNIT IV Working With Character Data

9hrs

Built-in Capabilities, Regular Expression, Prefix, Suffix and Substrings, Biological sequences, Matching patterns.

UNIT V R Packages

9hrs

Package basics, Package Management and authoring, Initialization, Debugging and Profiling

Course outcomes

At the end of the course the students will

CO1 –Providing general properties and unique aspects of the R language

CO2 –To teach Object Oriented programming in R

CO3 –To teach methods for getting of data in and out

CO4 –To provide knowledge how to how to write own software packages

CO5 –Use R to apply statistics to analyze and interpret data.

CO6 –Choose appropriate analysis techniques for a variety of data types and formats.

TEXT / REFERENCES BOOKS

1. Robert G, "R programming in Bioinformatics", CRC press, Taylor and Francis Group, USA, 2008.
2. Own J, Robert. M, and Andrew R., "Introduction to Scientific programming and simulation using R", CRC Press, Taylor and Francis Group, USA, 2014.
3. Norman M. And Norman S. Matloff "The Art of R programming: A tour of statistical software Design", No Starch Press Inc, USA, 2011.



SITA3008	INTERNET OF THINGS	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

- To understand the concepts of Internet of Things
- To identify the various elements of an IoT System.
- To understand the various means of communication from Node / Gateway to Cloud Platforms.
- To identify types of data analytics and data visualization tools.
- To make students aware of security concerns and challenges while implementing IoT solutions.

UNIT I INTRODUCTION TO IOT

9 Hrs.

Introduction to IoT, Current technological trends and future prospects, - Evolution of IoT , Business Scope, Relation with embedded system, - Basic Architecture of an IoT, From M2M to IoT, M2M towards IoT, IoT Value Chains, An emerging industrial structure for IoT.

UNIT II ELEMENTS OF IOT

9 Hrs.

Application Sensors & Actuators - Edge Networking (WSN) – Gateways - IoT Communication Model – WPAN & LPWA, Overview of IoT supported Hardware platforms such as: Raspberry pi, ARM Cortex Processors, Arduino and Intel Galileo boards, Wearable Development Boards .

UNIT III COMMUNICATION AND CONNECTIVE TECHNOLOGIES

9 Hrs.

IoT Communication Model, Cloud computing in IoT, IoT in cloud architecture, Logging on to cloud, Selecting and Creating cloud service , cloud based IoT platforms - IBM Watson, Google cloud.

UNIT IV DATA ANALYTICS AND IOT PLATFORM

9 Hrs.

Big Data Analytics , Apache Hadoop, Using Hadoop MapReduce for Batch Data Analysis, Apache Storm, Data Visualization, Visualization tools for IoT.

UNIT V HANDS-ON PROJECTS

9 Hrs.

Industry 4.0 concepts. Sensors and sensor Node and interfacing using any Embedded target boards (Raspberry Pi / Intel Galileo/ARM Cortex/ Arduino), DIY Kits – Soil moisture monitoring, Weather monitoring, Air quality Monitoring, Movement Detection.

Max. 45 Hrs.

COURSE OUTCOMES

On Completion of the course, student will be able to

- CO1 - To learn and understand the technology and current trends in Internet of things.
- CO2 - To understand the various elements of IoT system and hardware devices.
- CO3 - To learn the programming languages and platforms for building IoT applications.
- CO4 - To understand the cloud computing and its relevance for developing IoT applications.
- CO5 - To design and implement IoT applications that manages big data with data analytics and visualization tools.
- CO6 - To implement hands-on projects using an appropriate software and hardware devices in various applications.

TEXT / REFERENCE BOOKS

1. The Internet of Things: Applications and Protocols, Wiley publications. Author(s): Oliver Hersent, David Boswarthick, Omar Elloumi.
2. Architecting the Internet of Things, Springer publications. Author(s): Dieter Uckelmann, Mark Harrison, Florian Michahelles.
3. Internet of Things with Arduino Cookbook, Packt Publications. Author(s): Marco Schwatz.
4. Internet of Things and Data Analytics, Wiley Publications



SBIA3004	NEXT GENERATION SEQUENCING	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

This course will provide an introduction to the technology, data analysis, tools and resources for next generation sequencing data. The content is intended to provide a broad overview of the subject areas, and to highlight key resources, approaches and methodologies.

Unit I Introduction to NGS

9hrs

NGS Platforms: Introduction to NGS, Roche/454 FLX, Illumina/Solexa Genome Analyzer, Applied Biosystems SOLiD system, Helicos Heliscope, Pacific Biosciences/single molecule real time (SMRT) sequencing.

Unit II Sequence data

9hrs

Genome assembly algorithms: Alignment of short-reads to reference genome using spaced seed (ELAND, SOAP), index-filtering algorithm (SeqMap), quality-score (RMAP), q-filter algorithm (SHRiMP), FM-index (Bowtie, BWA, SOAP2), suffix tree (MUMmer). Sequence Alignment formats: Sequence Alignment/Map (SAM) format, Binary Alignment/Map (BAM) format, Tools for conversion (SAMtools), Alignment viewers (IGV, MGViewer).

Unit III Sequence data assembly

9hrs

De-novo assembly: Overlap-layout-consensus (OLC) approach (Arachne, Phusion), de Bruijn and Euler path approach (Euler, SOAPdenovo), string graph assembler (SGA). Scaffolding: Supercontig, contig orientation, contig ordering, contig distancing and gap closing using SOAPdenovo, ABySS, OPERA and RACA.

Unit IV Softwares in NGS analysis

9hrs

Application of R in NGS analysis: Introduction to Bioconductor, Reading of RNA-seq data (ShortRead, Rsamtools, GenomicRanges), annotation (biomaRt, genomeIntervals), reads coverage and assign counts (IRanges, GenomicFeatures), differential expression (DESeq).

Unit V Applications

9hrs

Biological applications of NGS: Whole-genome sequencing, Exome sequencing, Transcriptome sequencing, Epigenome sequencing, Interactome sequencing, methylome sequencing.

Course outcomes

At the end of the course the students will

- CO1 –Describe the applications of the different NGS technologies, including the weakness and strengths of the approaches.
- CO2 –Explain the steps involved in a general NGS data analysis.
- CO3 –Explain key theoretical concepts of alignment and de novo assembly.
- CO4 –Synthesize and formulate a project and relevant question within the field.
- CO5 – Independently perform a basic NGS data analysis.
- CO6 –Apply analytical and reflective skills in analyzing results from individual steps and the final project.

Text / Reference Books:

- Next-generation DNA sequencing Informatics by Stuart M. Brown, Cold Spring Harbor Laboratory, 2013.
- RNA-seq Data Analysis: A Practical Approach by Eija Korpelainen, Jarno Tuimala, Panu Somervuo, Mikael Huss, Garry Wong. Chapman & Hall/CRC, 2014.
- Next generation sequencing: Translation to Clinical Diagnostics by Wong Lee-Jun C. (ed.), Springer, 2013.
- Next-generation genome sequencing: Towards Personalized Medicine by Michal Janitz, WileyVCH, 2008



SBIA3005	QUANTITATIVE MODELLING IN BIOLOGY	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

This course will educate students on various of pharmacogenomics and enable the student to implement the learning in a practical setting.

UNIT I INTRODUCTION

9hrs

Systems Biology and Synthetic Biology - Dynamic Mathematical Models - Basic Features of Dynamic Mathematical Models - State Variables and Model Parameters - Steady-State Behavior and Transient Behavior - Linearity and Nonlinearity - Global and Local Behavior - Deterministic Models and Stochastic Models - Dynamic Mathematical Models in Molecular Cell Biology - Drug Target Prediction Oscillatory Behavior in NFκB Signalling

UNIT II Modelling of Chemical Reaction Networks

9hrs

Chemical Reaction Networks - Closed and Open Networks - Dynamic Behavior of Reaction Networks - Simple Network Examples - Numerical Simulation of Differential Equations - Separation of Timescales and Model Reduction - Separation of Timescales: The Rapid Equilibrium Assumption - Separation of Timescales: The Quasi-Steady-State Assumption

UNIT III Analysis of Dynamic Mathematical Models

9hrs

Phase Plane Analysis - Direction Fields - Nullclines - Stability - Stable and Unstable Steady States - Linearized Stability Analysis - Limit-Cycle Oscillations - Bifurcation Analysis - Sensitivity Analysis - Local Sensitivity Analysis - Determination of Local Sensitivity Coefficients - Parameter Fitting

UNIT – IV Metabolic Networks

9hrs

Modelling of Metabolism - Example: A Pathway Model - Sensitivity Analysis of Metabolic Networks: Metabolic Control Analysis - Metabolic Pathways - Flux Control Pathways - Branched Pathways - Modelling of Metabolic Networks - Model Construction - Case Study: Modelling the Regulation of the Methionine Cycle - Stoichiometric Network Analysis - Metabolic Pathway Analysis - Constraint-Based Modelling: Metabolic Flux Analysis.

UNIT IV MODELLING BIOLOGICAL PROCESSES

9hrs

Signal Transduction Pathways - Signal Amplification - Bacterial Two-Component Signalling Pathways - G-Protein Signalling Pathways - Ultra sensitivity - Activation Cascades - Adaptation - Bacterial Chemotaxis - Memory and Irreversible Decision-Making - Apoptosis - Gene Regulatory Networks - Modelling of Gene Expression - Gene Regulatory Networks - Genetic Switches - The lac Operon - Oscillatory Gene Networks - Circadian Rhythms - Cell-to-Cell Communication - Computation by Gene Regulatory Networks - Promoters as Logic Gates - Stochastic Modelling of Biochemical and Genetic Networks - A Discrete Modelling Framework

Course outcomes

At the end of the course the students will

CO1 –Comprehend the need for quantitative models in biology.

CO2 –Understand the formulation of biochemical reaction networks.

CO3 –Identify and analyse dynamics mathematical models.

CO4 –Describe the modelling of metabolic networks.

CO5 –Discuss the modelling of specific biological processes.

CO6 –Discuss the futuristic aspects of quantitative modelling and its role in systems medicine.

Text / Reference Books

1. Ingalls, Brian P.. Mathematical Modelling in Systems Biology : An Introduction, MIT Press, 2013.
2. Systems Biology: A Textbook (Klipp et al., 2009)
3. System Modelling in Cellular Biology (Szallasi et al., 2006),
4. An Introduction to Systems Biology: Design Principles of Biological Circuits (Alon, 2007),
5. A First Course in Systems Biology (Voit, 2012).



SCSA3015	DEEP LEARNING	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

- To present the mathematical, statistical and computational basis of neural networks.
- To study the concepts of deep learning.
- To introduce dimensionality reduction techniques.
- To enable the students to know deep learning techniques to support real-time applications.
- To examine the case studies of deep learning techniques.

UNIT I Introduction

9 Hrs.

Introduction to machine learning- Linear models (SVMs and Perceptrons, logistic regression)- Intro to Neural Nets: What a shallow network computes- Training a network: loss functions, back propagation and stochastic gradient descent- Neural networks as universal function approximates.

UNIT II Deep Networks

9 Hrs.

History of Deep Learning- A Probabilistic Theory of Deep Learning- Back propagation and regularization, batch normalization- VC Dimension and Neural Nets-Deep Vs Shallow Networks- Convolutional Networks- Generative Adversarial Networks (GAN), Semi-supervised Learning.

UNIT III Dimensionality Reduction

9 Hrs.

Linear (PCA, LDA) and manifolds, metric learning - Auto encoders and dimensionality reduction in networks - Introduction to Convnet - Architectures – AlexNet, VGG, Inception, ResNet - Training a Convnet: weights initialization, batch normalization, hyper parameter optimization.

UNIT IV Optimization And Generalization

9 Hrs.

Optimization in deep learning– Non-convex optimization for deep networks- Stochastic Optimization- Generalization in neural networks- Spatial Transformer Networks- Recurrent networks, LSTM - Recurrent Neural Network Language Models Word-Level RNNs & Deep Reinforcement Learning - Computational & Artificial Neuroscience

UNIT V Case Study And Applications

9 Hrs.

Imagenet- Detection-Audio WaveNet-Natural Language Processing Word2Vec - Joint Detection- Bioinformatics- Face Recognition- Scene Understanding- Gathering Image Captions.

Max. 45 Hrs.

COURSE OUTCOMES

On completion of the course, student will be able to

- CO1 - Understand basics of deep learning.
- CO2 - Implement various deep learning models.
- CO3 - Realign high dimensional data using reduction techniques.
- CO4 - Analyze optimization and generalization in deep learning.
- CO5 - Explore the deep learning applications.
- CO6 - Design and creation of data models.

TEXT / REFERENCE BOOKS

1. Cosma Rohilla Shalizi, Advanced Data Analysis from an Elementary Point of View, 2015.
2. Deng & Yu, Deep Learning: Methods and Applications, Now Publishers, 2013.
3. Ian Goodfellow, Yoshua Bengio, Aaron Courville, Deep Learning, MIT Press, 2016.
4. Michael Nielsen, Neural Networks and Deep Learning, Determination Press, 2015.



SCSA1501	OPERATING SYSTEMS	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

- To have an overview of different types of operating systems.
- To understand the concept of process management.
- To understand the concept of storage management.
- To understand the concept of I/O and file systems.

UNIT I INTRODUCTION

9 Hrs.

Introduction - Operating system structures - System components - OS services - System calls - System structure -Resources Processes - Threads - Objects - Device management - Different approaches - Buffering device driver.

UNIT II PROCESS MANAGEMENT

9 Hrs.

Processes - Process concepts - Process scheduling - Operations on processes - Cooperating processes - CPU scheduling -Basic concepts - Scheduling criteria - Scheduling algorithms - Pre-emptive strategies - Non-pre-emptive strategies

UNIT III SYNCHRONIZATION AND DEADLOCKS

9 Hrs.

The critical section problem - Semaphores - Classic problems of synchronization - Critical regions - Monitors-Dead locks - Deadlock characterization - Prevention - Avoidance - Detection - Recovery.

UNIT IV MEMORY MANAGEMENT

9 Hrs.

Storage Management Strategies - Contiguous Vs. Non-Contiguous Storage Allocation - Fixed & Variable Partition Multiprogramming - Paging - Segmentation - Paging/Segmentation Systems - Page Replacement Strategies - Demand & Anticipatory Paging - File Concepts - Access Methods - Directory Structure - File Sharing - Protection - File – System Structure - Implementation.

UNIT V I/O SYSTEM, LINUX & SHELL PROGRAMMING

9 Hrs.

Mass Storage Structure - Disk Structure- Disk Scheduling - Disk Management - Swap Space Management – RAID Structure- Shell Operation Commands - File Management Operation - Internet Service - Telnet - FTP - Filters & RegularExpressions- Embedded Linux – Android Vs iOS - Unix Internals - Shell Programming - Variable, Arithmetic Operations, Control Structures, Handling Date, Time & System Information.

Max. 45 Hrs.

COURSE OUTCOMES

On completion of the course, student will be able to

CO1 - Understand the fundamental components of a computer operating system and how computing resources are managed by the operating system.

CO2 - Demonstrate and understand the different memory and I/O management techniques used in Operating Systems.

CO3 - Apply the concepts of CPU scheduling, synchronization and deadlocks in real computing problems.

CO4 - Have practical experience with software tools available in modern operating systems such as semaphores and monitors for process synchronization.

CO5 - Create design and construct the following OS components: System calls, Schedulers, Memory management systems, Virtual Memory and Paging systems in the modern operating system.

CO6 - Understand file system structure and implement a file system such as FAT.

TEXT / REFERENCE BOOKS

1. Abraham Silberschatz, Peter Galvin and Gagne, "Operating System Concepts", 6th Edition, Addison Wesley, 2002.
2. Harvey M.Deitel, "Operating System", 2nd Edition, Addison Wesley, 2000.
3. Gary Nutt, "Operating System, A modern perspective", 2nd Edition, Addison Wesley, 2000.
4. Richard Peterson, "Linux : The Complete Reference", 6th Edition, Tata McGraw Hills, 2007.
5. Andrew S. Tanenbaum, "Modern Operating Systems"

SBIA3007	BIGDATA IN HEALTHCARE	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

The course is designed to make the students realize the role and importance of big data in genomics and its implications for research in biology and medicine.

UNIT I Introduction

9hrs

Big Data - Big Data in Medicine? - The Three Vs of Big Data - Can Medicine Learn from Other Industries? - Big Data Framework in Medicine - Fields of Action in Medicine - It is Time for Big Data - The Example of Google Flu Trends - Big Data Adoption Model in Medicine - Big Data in the Context of Health - The Difference to Classic Clinical Analysis (Old Wine in New Bottles?) - Are Hospitals Prepared for Big Data? - **Personalized Medicine and Big Data** - **The (Genetic) Human in Numbers - Big Data**, "Personalized Medicine" and IT

UNIT II Data Management and Data Mining Approaches

9hrs

Data Management and Data Mining Approaches - Introduction Challenges of Managing and Mining Big Biomedical Data - Data Management Framework - Data Management Platforms - Data Mining in Translational Medicine - Data Mining Techniques and Methods - **Big Data and Data Mining - Applications of Data Mining in Translational Medicine**

UNIT III Information Management for Systems Medicine

9hrs

New Data Volumes - The Digitalization of the Knowledge Cycle - Consolidation of the Digital Medical Care and Research Systems - Integration and Differentiation - Good and Bad Data = Digital Expert Systems - **Human Beings are More Than the Sum of Their Data - Medicine of the Person is Made More Difficult**

UNIT IV Medical Big Data

9hrs

Medical Big Data - Data Protection - - Expectations - Legal Biological Principles - Data Sources - Health-Specific Use of Data - Risks - Data Protection Mechanisms - **Data Economy in Particular - Transparency**

UNIT V Big Data and Business

9hrs

Big Data in Healthcare from a Business perspective - Data as a Basis for Decisions - Strategy and Data - Big Data as a Part of Digitalization - Consulting as a Decision Support = The Path to an Informed Decision - The Transformation of Big Data to Smart Data - **New Technologies in the Area of Big Data – Benefits to the patient.**

Course Outcomes

CO1 –To understand the concept and nature of bigdata.

CO2 –To understand the basics of data management and mining approaches of genomic big data.

CO3 –To appreciate the role of bigdata in genomic medicine.

CO4 –Ethical use of medical bigdata.

CO5 –Use of genomic bigdata in entrepreneurship.

CO6 –Appreciate the enormous data available and the means to handle the data

Text /Reference books

1. Big Data in Medical Science and Healthcare Management : Diagnosis, Therapy, Side Effects, edited by Peter Langkafel, De Gruyter, Inc., 2015.
2. Translational Medicine : Tools and Techniques, edited by Aamir Shahzad, Elsevier Science & Technology, 2015.



SBIA3010	Health Informatics	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

- To Introduce the students the problems and challenges that confront the field of health informatics.
- To Introduce the students about the research avenues and salient features in the practice of health informatics.
- To Provide the students with the basic skills and knowledge in health informatics to apply in a career related to health care.
- To Lead students in the discussion around ethical and diversity issues in health informatics.
- To Provide additional direction to those interested towards the study in the field.

UNIT I Introduction & Overview: Health Information Technology & Health Informatics 9 Hrs.

Define information management, information technology and informatics - Various ways in which HIT has evolved to improve equality or enhance patient safety - Informatics drivers and trends - Standards development organizations: Health Level 7 (HL-7), Healthcare Information Technology Standards Panel (HITSP), and Office of the National Coordinator for Health Information Technology (ONC) Health IT Standards Committee-Health informatics – Needs, objectives and limitations - Medical Terminology for Health Informatics Professionals.

UNIT II Clinical Decision Support Systems (CDSS) And Workflow 9 Hrs.

Fundamental requirements of an effective Clinical Decision Support System-principles of health care data exchange and health care data standards relate to patient care, productivity and data analysis- common controlled vocabularies in use today: ICD, CPT, NDC, RxNorm, LOINC and SNOMED CT-advantages and disadvantages (unintended consequences) of Clinical Decision Support Systems (CDSS) in terms of cognitive support.

UNIT III Healthcare Database Management System 9 Hrs.

Health information systems – design, architecture and interoperability - Electronic health records (EHRs) - Electronic medical record (EMR) with electronic health record (EHR)- emergence of Personal Health Records and their implications for patients, health care providers and Decision Aids and health systems- EHR access agreement.

UNIT IV Computer Applications In Health Care And Biomedicine 9 Hrs.

The language of biomedical informatics: data, information & knowledge; standards and vocabularies - Evidence based practice (EVP) and clinical practice guidelines (CPG) - Evaluation methods for health informatics - Machine learning for extraction of medical knowledge - MATLAB programming for health informatics.

UNIT V Case Studies And Ethical Issues 9 Hrs.

Modelling, simulation and visualisation in health informatics - Ethical Issues in Health Informatics - Fundamentals of data science in medicine - Genomic and Systems Medicine - Advanced data analysis in medicine - Data driven decision making and evaluation in medicine - Legal Aspects of Health Information Systems - E-Healthcare & Ethics - E-Medicine Business Models - Case Studies in Health Information Management.

Max. 45 Hrs.

COURSE OUTCOMES

On completion of the course, student will be able to

- CO1 - Develop knowledge about problems and challenges that health addressed by informatics and apply the knowledge to the research and practice in health informatics.
- CO2 - Demonstrate basic skills and knowledge in health informatics for application in future health-related careers.
- CO3 - Demonstrate ability to ask, search for information and answer health-related questions.
- CO4 - To identify genomic variants associated with a disease phenotype and communicate this association.
- CO5 - Perform visualization and simple analysis of a data set to assess difficulty of predicting disease (cardiovascular and other) risk in a synthetic patient dataset.
- CO6 - Apply communication skills through an interview with an informatics professional and development of a written summary and analyse ethical and diversity issues in health informatics.

TEXT / REFERENCE BOOKS

1. Hoyt, RE and Yoshihashi, A, Eds. (2014). Health Informatics: Practical Guide for Healthcare and Information Technology Professionals, 6th Edition. Pensacola, FL, Lulu.com.
2. Wager KA, Lee FW, Glaser JP. Health Care Information Systems: A Practical Approach for Health Care Management. San Francisco, California: Jossey-Bass 2009.
3. Shortliffe EH, Cimino JJ. Biomedical Informatics: Computer Applications in Health Care and Biomedicine, 4th Edition.



New York, New York: Springer Science +Business Media, LLC 2014.





SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Accredited A++ Grade by NAAC | 12B Status by UGC | Approved by AICTE

www.sathyabama.ac.in

School of Bio & Chemical Engineering

Department of Biotechnology

1.1.3 Courses having focus on Employability / Entrepreneurship / Skill Development

CONTENTS

M.Sc. Bioinformatics and Data science

S. No.	Course Code	Course name	Page No
Employability			
1.	SBIA5101	Structural and functional genomics	3
2.	SBIA5102	Python for biologists	4
3.	SBIA5103	Biological databases and data analysis	5
4.	SMTA5102	Statistics and random process	7
5.	SBIA5203	Biomolecular Modelling	9
6.	SBIA5302	Computer aided drug design	11
7.	SBIA5304	Microarray data analysis	13
8.	SBIA5401	R programming	15
9.	SBIA7406	Data mining in bioinformatics	16
10.	SBIA7408	Next generation sequence data analysis	17
11.	SBIA7409	Health informatics	18
12.	SBIA7412	Cloud computing	19
Entrepreneurship			
13.	SBIA5102	Python for biologists	4



14.	SBIA5302	Computer aided drug design	11
15.	SBIA5304	Microarray data analysis	13
16.	SBIA5401	R programming	15
17.	SBIA7406	Data mining in bioinformatics	16
18.	SBIA7408	Next generation sequence data analysis	17
19.	SBIA7409	Health informatics	18
20.	SBIA7412	Cloud computing	19
Skill Development			
21.	SBIA5101	Structural and functional genomics	3
22.	SBIA5102	Python for biologists	4
23.	SBIA5103	Biological databases and data analysis	5
24.	SMTA5102	Statistics and random process	7
25.	SBIA6101	Python lab	8
26.	SBIA5203	Biomolecular Modelling	9
27.	SCSA6201	Oracle and SQL lab	10
28.	SBIA5302	Computer aided drug design	11
29.	SBIA5304	Microarray data analysis	13
30.	SBIA6301	Computer Aided Drug Design Lab	14
31.	SBIA5401	R programming	15
32.	SBIA7406	Data mining in bioinformatics	16
33.	SBIA7408	Next generation sequence data analysis	17
34.	SBIA7409	Health informatics	18
35.	SBIA7412	Cloud computing	19



SBIA5101	STRUCTURAL AND FUNCTIONAL GENOMICS	L	T	P	Credits	Total Marks
		3	*	0	3	100

COURSE OBJECTIVES

To have a knowledge on the conformational parameters and their influence on the structure of the macromolecule and to acquire a knowledge on the structure and activity relationship.

UNIT I INTRODUCTION

Structure & Organization of Prokaryotic & Eukaryotic genome - Nucleotide and protein sequencing methods –Chemical, Enzymatic, high through put method – Automated sequencing methods – shotgun – chromosome walking, Contig assembly - Levels of structures in Biological macromolecules.

UNIT II CONFORMATIONAL ANALYSIS

Biomolecules and their interactions - Forces that determine protein and nucleic acid structure, basic problems, polypeptide chains geometrics, potential energy calculations, observed values for rotation angles, hydrogen bonding, hydrophobic interactions and ionic interactions, disulphide bonds - **Prediction of proteins structure - Nucleic acids**, general characteristics of nucleic acid structure, geometrics, glycosidic bond rotational isomers - Ribose puckering - forces stabilizing ordered forms, base pairing, base stacking- tertiary structure of nucleic acids.

UNIT III STRUCTURAL ANALYSIS OF MACROMOLECULES

Size and shape of macromolecules - methods of direct visualization - **X-ray crystallography – X-ray diffraction, determination of molecular structures, electron microscopy, NMR**. Protein structure databases - Protein Data Bank - SCOP - CATH - structure superposition - RMSD - TM-score- structure alignment - Different structure alignment algorithms - DALI, CE, VAST, TM-align - protein folds in PDB.

UNIT IV STRUCTURE-FUNCTIONS RELATIONSHIP

DNA binding proteins, Prokaryotic and Eukaryotic transcription factors - DNA polymerases, Helix-turn-Helix motif in DNA binding, Trp repressor, Zn fingers, helix-turn helix motifs in homeodomain, Leucine zippers - DNA polymerases, Membrane proteins and receptors, Trans-membrane segments, bacterio-rhodopsin, photosynthetic centres, epidermal growth factor, insulin and PGDGF receptors and their interaction with effectors, protein phosphorylation, immunoglobulins, Serine proteases, ribonuclease, lysozyme.

UNIT V FUNCTIONAL GENOMICS

Introduction to Functional Genomics - cytological maps - Hap Map - SNPs and variation- Genotyping, microarray- analysis and applications. Microarray oligo sets and fabrication - integrative genomics and Meta genomics. **Quantitative Proteomics. Fluorescent imaging techniques. Quantitative RT-PCR analysis. High-throughput cloning and expression strategies. Next Generation Sequencing applications.**

COURSE OUTCOMES

On completion of the course the student will be able to

- CO1:** The student will acquire a basic knowledge on the basics of genomics.
- CO2:** They will be able to appreciate the influence of conformational parameters on the structure of the macromolecule.
- CO3:** Will be able to select the particular method of structural analysis methodology, given the protein.
- CO4:** Assess the structure and activity relationship of the macromolecule.
- CO5:** Acquire knowledge on the functional aspects of genomics.
- CO6:** Explore the strategies used in high throughput screening of biological data

TEXT / REFERENCE BOOKS

1. Cantor R., Schimmel P.R., Biophysical Chemistry, Vol. I, II, W.H. Freeman & Co., 1985.
2. Kensal E. van Holde, W. Curtis Johnson and P. Shing Ho, Principle of Physical Biochemistry, Prentice Hall, New York, 1998.
3. Pennington SR, Dunn MJ, Proteomics from Protein Sequence to Function, Viva Books Ltd, 2002.
4. G. Gibson and M. V. Muse. A primer of Genome Science. Sinauer Associates Inc; 2 edition, December 2004.

END SEMESTER EXAM QUESTION PAPER PATTERN

SBIA5102	PYTHON FOR BIOLOGISTS	L	T	P	Credits	Total Marks
		3	1	0	4	100

COURSE OBJECTIVES

To empower the students in the life sciences, who may have a pertinent scientific question that can be readily answered by computational techniques

UNIT I PROGRAMMING PRINCIPLES

Basic programming concepts - what is a program? - why python? - main features of python - comparing python with other languages - how is it used? - who uses python? - flavors of python - special python distributions - installing python - interactive mode - batch mode - other tools - additional resources - Basic Programming - Simple Values Booleans - Integers Floats Strings Expressions - Numeric Operators - Logical Operations - String Operations - Calls - Compound Expressions - Assigning Names - Defining Functions - Using Modules

UNIT II COLLECTIONS AND CONTROL STATEMENTS

Sets - Sequences - Strings, Bytes, and Byte arrays - Ranges - Tuples - Lists - Mappings - Dictionaries - Streams - Files - Generators - Collection- Related Expression Features - Comprehensions - Functional Parameters - Conditionals - Loops - Iterations - Exception Handlers - Extended Examples

UNIT III CLASSES, UTILITIES AND PATTERN MATCHING

Defining Classes - Class and Method Relationships - System Environment - Working with Text - Persistent Storage - Fundamental Syntax - The Actions of the re Module - Results of re Functions and Methods - Putting It All Together

UNIT IV STRUCTURED TEXT, WEB PROGRAMMING, RELATIONAL DATABASES AND GRAPHICS

HTML - XML - An XML File for a Complete Genome - Manipulating URLs: urllib.parse - Opening Web Pages: web browser - Web Clients - Web Servers - Tips, Traps, and Tracebacks - Representation in Relational Databases - Using Relational Data - SQL Basics - SQL Queries - Introduction to Graphics Programming

UNIT V BIOPYTHON AND PYTHON RECIPES

Project Organization - INSTALLING BIOPYTHON - Alphabet - Seq - Mutable Seq - Seq Record - Align - Align IO - ClustalW - SeqIO - AlignIO - BLAST - Biological Related Data - Entrez - PDB - PROSITE - Restriction - SeqUtils - Sequencing - SwissProt - Python recipes with commented source code - Collaborative Development: Version control with GitHub - PythonAnywhere - Scientific Python

COURSE OUTCOMES

At the end of the course the students will be able

- CO 1- To understand why Python is a useful scripting language for developers.
- CO 2- To learn how to design and program Python applications.
- CO 3- To learn how to use lists, tuples, and dictionaries in Python programs.
- CO 4- To learn how to identify Python object types.
- CO 5- To apply python in real time informatics applications.
- CO 6- To learn how to use biopython and online resources for long – term learning.

TEXT / REFERENCE BOOKS

1. Python Programming for Biology-Bioinformatics and Beyond, Tim J. Stevens, MRC Laboratory of Molecular Biology, Cambridge, Wayne Boucher, University of Cambridge
2. Managing Your Biological Data with Python by Allegra Via, Chapman and Hall/CRC; 1 edition (22 August 2017)
3. Practical Computing for Biologists by Steven H. D. Haddock, OUP USA; 2011 edition (5 November 2010)
4. Python for Biologists: A Complete Programming Course for Beginners, Dr Martin Jones, Createspace Independent Publishing Platform; 1 edition (7 September 2013)
5. Bioinformatics Programming Using Python by Mitchell L Model, Released December 2009, Publisher(s): O'Reilly Media, Inc. ISBN: 9780596154509
6. Python for Bioinformatics by Sebastian Bassi ISBN 9781138035263, Published July 30, 2017 by Chapman and Hall/CRC

SBIA5103	BIOLOGICAL DATABASES AND DATA ANALYSIS	L	T	P	Credits	Total Marks
		3	*	0	3	100

COURSE OBJECTIVES

To have a knowledge on the conformational parameters and their influence on the structure of the macromolecule and to acquire a knowledge on the structure and activity relationship.

UNIT I OVERVIEW AND CHALLENGES

Problem and Scope - Biological Data Integration - Developing a Biological Data Integration System - Specifications - Translating Specifications into a Technical Approach - Development Process - Evaluation of the System - The Life Science Discovery Process - An Information Integration Environment for Life Science Discovery - The Nature of Biological Data - Diversity - Variability - Data Sources in Life Science - Biological Databases Are Autonomous - Heterogeneous in Data Formats - Biological Data Sources Are Dynamic - Computational Analysis Tools Require Specific Input/Output Formats and Broad Domain Knowledge - Challenges in Information Integration - **Data Integration - Meta-Data Specification - Data Provenance and Data Accuracy - Ontology**

UNIT II DATA MANAGEMENT AND INTEGRATION

Data Management - Basics - **Two Popular Data Management Strategies and Their Limitations** - Traditional Database Management - Dimensions Describing the Space of Integration Solutions - - Browsing vs. Querying - Syntactic vs. Semantic Integration - Warehouse vs. Federation - Declarative vs. Procedural Access - Generic vs. Hard-Coded - Relational vs. Non- Relational Data Model - Use Cases of Integration Solutions - Browsing-Driven Solutions - Data Warehousing Solutions - Federated Database Systems Approach - Semantic Data Integration - Strengths and Weaknesses of the Various Approaches to Integration - Browsing and Querying: Strengths and Weaknesses - Warehousing and Federation: Strengths and Weaknesses - Generic and Hard-Coded Approaches: Strengths and Weaknesses - **Relational and Non-Relational Data Models**: Strengths and Weaknesses - A Hybrid Approach to Integration Is Ideal

UNIT III DESIGNING A BIOLOGICAL LEGACY

Biological Data - **Biological Tools and Workflows** - A Domain in Constant Evolution - Traditional Database Management and Changes - Data Fusion - Fully Structured vs. Semi-Structured - Scientific Object Identity - Concepts and Ontologies - Biological Queries - Searching and Mining - Browsing - Semantics of Queries - **Tool-Driven vs. Data-Driven Integration** - Query Processing - Biological Resources - Query Planning - Query Optimization - Visualization - Multimedia Data - Browsing Scientific Objects

UNIT IV MODEL DATABASES

SRS - An Integration and Analysis Tools Platform for Databanks - Integrating Flat File Databanks - Integration of XML Databases - Integrating Relational Databases - The SRS Query Language - Linking Databanks - The Object Loader - Scientific Analysis Tools - Interfaces to SRS - **Building biological databases with SQL** - Common database types - Relational databases -

UNIT V DATA MINING

Biomarker discovery - An introduction to biomarker discovery - Data preprocessing - Modeling - Validation - Case study - Discussion and future perspective

COURSE OUTCOMES

On completion of the course the student will be able to

- CO1-** Acquire the basic understanding of a biological database
- CO2-** Explore the various platforms for constructing the database.
- CO3-** Design the biological database and defining its attributes.
- CO4-** Assess the structure and modelling of complex databases.
- CO5-** Acquire knowledge on data mining and futuristic perspectives on biological data mining.
- CO6-** Explore the data mining role in biomarker discovery.

TEXTBOOKS / REFERENCE BOOKS

1. Zengyou, He. Data Mining for Bioinformatics Applications, Elsevier Science & Technology, 2015.
2. Bioinformatics : Managing Scientific Data, edited by Zoé Lacroix, and Terence Critchlow, Elsevier Science & Technology, 2003.
3. Bessant, Conrad, et al. Building Bioinformatics Solutions, Oxford University Press USA - OSO, 2014.

END SEMESTER EXAM QUESTION PAPER PATTERN

Max. Marks : 100

Exam Duration : 3 Hrs.

PART A : 6 Questions for 5 Marks each –without choice

30 Marks

PART B : 2 Questions from each unit with internal choice, each carrying 14 marks

70 Marks



SMTA5102	STATISTICS AND RANDOM PROCESS	L	T	P	CREDITS	TOTAL MARKS
		3	0	0	3	100

COURSE OBJECTIVES

To acquire skills in handling situations involving more than one random variable and functions of random variables. This gives exposure to statistical methods designed to contribute to the process of making scientific judgments in the face of uncertainty and variation.

UNIT 1 POINT ESTIMATION

Properties of estimators, Unbiasedness, Efficiency, Consistency, Sufficiency - Rao - Blackwell theorem, C.R. Rao Inequality - Methods of estimation: Method of Likelihood Estimation, Method of Minimum χ^2 Modified χ^2 - Method of moments.

UNIT II THEORY OF SAMPLING AND TEST OF HYPOTHESIS

Test of Hypothesis - large sample - test of significance - Single mean, difference of means, single proportion, difference of proportions -small sample - students 't' test - single mean - difference of means - Fisher's test or difference of variance - Chi square test - goodness of fit - independence of attributes.

UNIT III NON - PARAMETRIC TESTS

Sign test, Wilcoxon Signed Rank Test, Median test, Mann-Whitney U test, Kruskal - Wallis H test.

UNIT IV INTRODUCTION TO STOCHASTIC PROCESSES

Definition, Classification of Stochastic processes, Random walk, Markov chains, Transition Probability Matrix, Graphical representation of Markov chains, Classification of states and chains- Simple Problems- Chapman Kolmogorov Theorem.

UNIT V MARKOV PROCESS

Introduction to Markov Process- Poisson process-Postulates and properties-Birth and Death process-Hidden Markov Model - Introduction - Vitebi Algorithm, Forward and Backward Algorithm, Estimation algorithms.

TEXT / REFERENCE BOOKS

1. Veerarajan T, Probability, Statistics and Random Process, 4th edition, Tata McGraw Hill, 2008.
2. Medhi J, Stochastic Processes, Second Edition, New Age International Publishers. 2002.
3. Stephen Bernstein Schaum's Outline of Elements of Statistics I: Descriptive Statistics and Probability, McGraw-Hill, 1998.
4. Jerrold H zar, Bio statistical Analysis, 5th edition, Prentice Hall, 2010.
5. Irvin Miller John E Freund, Probability and Statistics for Engineers, Prentice Hall, 1977.

COURSE OUTCOMES

- CO1 Define characteristics of good estimators, List the methods of estimators, and Evaluate the MLE of estimators.
- CO2 Understand the basis and applications of sampling; Choose the appropriate tests of significance, Formulate the hypothesis for testing the significance of difference, Identify the size of the sample, level of significance, degrees of freedom of the samples.
- CO3 Advantages of non-parametric tests Choose the appropriate tests and Evaluate the various methods of Non-Parametric tests.
- CO4 List the classification of stochastic process, evaluate problems in tpm, Identify the classification of states and chains.
- CO5 Enable students to understand the tools used in bioinformatics and their corresponding algorithms used for evaluating the bio models.
- CO6 Explore the statistical methods designed to contribute to making scientific judgments in the face of uncertainty and to design suitable models.

SBIA6101	PYTHON LAB	L	T	P	Credits	Total Marks
		0	0	4	2	100

SUGGESTED LIST OF EXPERIMENT

1. Counting DNA Nucleotides
2. Transcribing DNA into RNA
3. Computing GC Content
4. Counting Point Mutations
5. Translating RNA into Protein
6. Finding a Motif in DNA and Finding a Protein Motif
7. Consensus and Profile
8. Inferring mRNA from Protein
9. Open Reading Frames
10. Calculating Protein Mass
11. Locating Restriction Sites
12. RNA Splicing
13. Longest Increasing Subsequence and Genome Assembly as Shortest Superstring
14. Finding a Spliced Motif
15. Transitions and Transversions
16. Biopython - Parsing GenBank records from the net and SwissProt sequences
17. Biopython - Converting between sequence file formats
18. Biopython - Multiple Alignments
19. Biopython - Determining the secondary structure
20. Biopython - Annotated Chromosomes
21. Biopython - Parsing KEGG records
22. Biopython - Simple quality filtering for FASTQ files
23. Biopython - Calculating a substitution matrix from a multiple sequence alignment
24. Biopython - Sequence parsing plus simple plots



SBIA5203	BIOMOLECULAR MODELLING	L	T	P	Credits	Total Marks
		3	1	0	4	100

COURSE OBJECTIVES

To enable the student to analyze and discuss the different constraints for a macromolecular modelling and debate the predicted structures of the models.

UNIT I BASIC CONCEPTS IN MOLECULAR MODELLING

Internal parameters - Z-matrix – Introduction to quantum chemistry –basic postulates – Schrodinger wave equation – Derivation – Hydrogen atom – Born-Oppenheimer approximation - Laws of thermodynamics - entropy – enthalpy - free energy calculations-chemical potential - Calculating thermodynamic properties using force field; Transferability of force field parameters, treatment of de-localized pi system; Force field for metals and inorganic systems – Application of energy minimization.

UNIT II MOLECULAR GEOMETRY

Conformational parameters - Potential energy Surface - Molecular mechanics: empirical forces fields - bond stretching - angle bending - torsional terms – non-bonded and electrostatic interaction.

UNIT III TYPES OF FORCE FIELDS AND ENERGY MINIMIZATION

Simplex – sequential univariate method - steepest descent - conjugate gradient method - Newton–Raphson method – Molecular Dynamics with continuous potentials and at constant temperature and pressure; Time dependent properties; Solvent effects in Molecular Dynamics - Conformational analysis - Molecular visualization - Molecular graphics – Rendering - Rasmol.

UNIT IV MOLECULAR DYNAMICS (MD) SIMULATION OF BIOPOLYMERS

Time steps - Setting up MD - energy conservation in MD Simulation -continuous potentials and constraint dynamics - MD at constant temperature and pressure - incorporating solvent effects - examples-random number generator - Monte Carlo simulation of biological macromolecules - MD softwares

UNIT V PREDICTION OF SECONDARY STRUCTURE

Membrane prediction – Comparative modeling - Sequence Alignment Homolog's - analogs - Homology modeling - steps in homology modeling – side chain modeling – loop modeling - fold recognition – ab initio prediction – Predicting protein structures by threading protein folding – active site/binding site prediction – tools – databases - CASP.

COURSE OUTCOMES

At the end of the course

- CO1- Students will learn the basics of molecular modeling aspects.
- CO2- They will be able to relate the conformational parameters to the conformation of the molecule.
- CO3- They will be able to predict the effect of the force field applied on a molecule.
- CO4- Students will be able to examine and analyze the significance of simulation of macromolecules.
- CO5- Students will be able to appreciate the predicted models.
- CO6- Assessment of novel algorithm in structure prediction

TEXTBOOKS / REFERENCE BOOKS

1. Leach A.R., Molecular Modelling - Principles and Applications, 2nd Edition, Prentice Hall, 2001.
2. Vinter J.G. and Mark Gardener, Molecular modeling and drug design, Mac Millan, 1994.
3. Prasad, R.K., Quantum Chemistry, Halsted Press, 1992.
4. Ramachandran, K.I, Deepa, G, Namboori, K, Computational Chemistry and Molecular Modeling: Principles & Application, Springer 2008.
5. 3. Young, D.C., Computational Chemistry: A Practical Guide for Applying Techniques to Real-World Problems, Wiley-Indescience 2001

SCSA6201	ORACLE AND SQL LAB	L	T	P	CREDITS	TOTAL MARKS
		0	0	4	2	100

LIST OF EXPERIMENTS

COURSE OBJECTIVES

1. Learn and practice data modelling using the entity-relationship and developing database designs.
2. Understand the use of Structured Query Language (SQL) and learn SQL syntax.
3. Apply normalization techniques to normalize the database
4. Understand the needs of database processing and learn techniques for controlling the consequences of concurrent data access.

COURSE OUTCOMES

CO 1: The learner will be able:

CO 2: To describe data models and schemas in DBMS.

CO 3: To understand the features of database management systems and Relational database

CO 4: To use SQL-the standard language of relational databases

CO 5: To understand the functional dependencies and design of the database

CO 6: To understand the concept of Transaction and Query processing.

1. Data Definition Language (DDL) commands in RDBMS
2. Data Manipulation Language (DML) and Data Control Language (DCL)
3. High level language extensions with cursors
4. High level language extension with Triggers
5. Procedures and Functions
6. Embedded SQL
7. Database design using E-R model and Normalization
8. Design and implementation of payroll processing system
9. Design and implementation of Banking system
10. Design and implementation of Library Information System
11. Design and implementation of Student Information System
12. Automatic Backup of Files and Recovery of Files



SBIA5302	COMPUTER AIDED DRUG DESIGN	L	T	P	CREDITS	TOTAL MARKS
		3	1	0	4	100

COURSE OBJECTIVES

To understand the basic concepts in drug design and to know the significance of computational methods and algorithm in drug design. Henceforth, to appreciate the logical reasoning and intellectual thinking in bridging the gap in drug discovery.

UNIT I INTRODUCTION TO DRUG DESIGN AND DEVELOPMENT

9hrs

Drug Discovery as a Process, Stages of drug discovery and development, Pathophysiology of a disease, Target Identification and Validation, 3-dimensional structure of the receptor-protein, nucleic acid, protein-nucleic acid complex, Structure Determination- X-ray crystallography, NMR spectroscopy, and mass spectrometry, Drug Target Validation.

UNIT II DRUG TARGETS

9hrs

Membrane proteins- receptors, ion channels and transporters, advantages and disadvantages of membrane proteins as drug targets, Nucleic acids-DNA, messenger RNA, and ribosomal RNA are important molecular targets for cancer, viral, and microbial chemotherapy, Rational design of DNA groove and recognition of RNA binding sites. Enzymes in cellular processes, enzymes as drug targets, development of agents that selectively block mRNA- antisense oligonucleotides, ribozymes, siRNA, gene silencing

UNIT III LEAD IDENTIFICATION AND MODIFICATION

9hrs

Biological assays- High-throughput screening, virtual screening-random , non-random screening, Combinatorial Library Design, pharmacophore modeling, functional group optimization, QSAR and QSPR methods- Partition coefficient, Hammett's substituent constant and Taft's steric constant. 3DQSAR- COMFA, COMSIA, bioisosteric replacement, fragmentation, rigid analogs, drug binding kinetics- competitive, non-competitive inhibition, binding parameters-binding affinity, dissociation constants, inhibitory constants

UNIT IV COMPUTER-AIDED DRUG DESIGN 9hrs

Molecular modeling- molecular mechanics and quantum mechanics Force field methods ,Minimum energy conformation-energy minimization, Lead based drug design-geometry optimisation, conformational searching. Pharmacophore mapping, molecular docking molecular dynamics

UNIT V DRUG DELIVERY

9hrs

Bioavailability, pharmacodynamics and pharmacokinetics. ADME/Toxicity predictions, Pre-Clinical and Clinical testing, Clinical Trials- Clinical Trial Design, Ethics of Human and Animal Experimentation, Intellectual Property, Commercial Considerations in Drug Development

Max 45 hrs

COURSE OUTCOMES

The student will be able to:

- CO 1-** Apply and understand the modelling on the basis of quantum mechanical aspects.
- CO 2-** Discuss the various types of interactions possible between the macromolecule and the ligand. **CO 3-** Analyse the improvement of current anticancer and antibacterial drugs through rational design. **CO 4-** Explore the interaction of drug target and drug
- CO 5-** Develop knowledge and techniques to identify drug targets
- CO 6-** Critically evaluate the strategy for lead identification and optimizing pharmacokinetic properties

TEXT / REFERENCE BOOKS

1. Shayne C. Gad, (2005), Drug Discovery Handbook, 2nd edition, John Wiley & Sons: New York.
2. Tudor I. Opera, (2005), Chemoinformatics in Drug Discovery, 5th edition, John Wiley & Sons: New York.



SBIA5304	MICROARRAY DATA ANALYSIS	L	T	P	Credits	Total Marks
		3	1	0	4	100

COURSE OBJECTIVES

Students are allowed to understand the key concepts of class comparison and class discovery of genes.

UNIT I GENE EXPRESSION

Basics of Gene expression- definition , gene expression studies, gene expression patterns- Applications of gene expression studies- Microarrays – definition , discovery, technique, making microarrays, spotted microarrays, In-situ synthesized oligonucleotide arrays, inkjet array synthesis, Affymetrix techniques, DNA CHIP technology, photolithography, spot quality, sample preparation and labelling, washing, image acquisition Sequencing by Hybridization Arrays- DNA Mass Array™ Technology- Printing DNA Microarrays-Types of microarrays - Designing a microarray experiment.

UNIT II IMAGE PROCESSING

Image processing, feature extraction, identifying positions of features- Normalization – data cleaning and transformation, within array normalization, between array normalization, measuring and quantifying microarray variability –variability between replicate features on an array-, variability between hybridizations to different arrays. Analysis of differentially expressed genes- significance analysis of microarrays.

UNIT III PREDICTION

Prediction of cross hybridization to related genes, thermodynamics of nucleic acid duplexes, prediction of T_m- probe secondary structure. Analysis of relationships between genes, tissues or treatments- similarity of gene or sample profiles – dimensionality reduction, principal component analysis, hierarchical clustering, machine learning methods for cluster analysis. Classification of tissues and samples – validation.

UNIT IV CARBOHYDRATE MICROARRAYS

Carbohydrate microarrays- Carbohydrate sources- Synthesis of oligosaccharides - Isolation of oligosaccharides from natural sources- Arrays of monosaccharides and disaccharides- Arrays of polysaccharides- Arrays of oligosaccharides-Perspectives- DNA micro array analysis-clustering-immunological applications.

UNIT V DATABASES AND TOOLS FOR MICROARRAYS

Bioinformatics in Arrays- Databases and tools for microarrays- Bioconductor, expression profiler, EST databases- Assessing levels of gene expression using EST's, TIGR gene indices, STACK, SAGE, CGAP, Xprofiler, ARRAY DB, cluster, tree view, Scanalyze, gene cluster, informatics aspects of microarray production- MGED and gene-ontology, description of MIAME (Minimum Information About a Microarray Experiment), Business Aspects of Biochip Technologies- Microarray Technology in Treating Disease.

COURSE OUTCOMES

At the end of the course the students will be able to

- CO 1- To learn the basics of DNA Microarray and its applications
- CO 2- To learn how to Analyze of the raw RNA sequence data
- CO 3- To understand why Statistical computing and Statistical Genetics is required
- CO 4- To apply data analysis to microarray related concepts
- CO 5- To apply the finding of the analytics in healthcare
- CO 6- To develop novel algorithms in the area

TEXT / REFERENCE BOOKS

1. Dov Stekel - Microarray bioinformatics, Cambridge,
2. Steen Knudsen - A Biologist's Guide to analysis of DNA microarray data, second edition, John Wiley & Sons, Inc.,
3. Andreas D Baxevanis and BF Francis Ouelette - Bioinformatics: A Practical Guide to the Analysis of Gene and Proteins , Wiley-Interscience,
4. David W. Mount - Bioinformatics: Sequences and genome analyses, Cold Spring Harbor Laboratory press,

SBIA6301	COMPUTER AIDED DRUG DESIGN LAB	L	T	P	Credits	Total Marks
		0	0	4	2	100

1. Molecular Visualization Softwares: Pymol and RasMol
2. Geometry Optimization
3. Binding Site Identification - Cast-P, POCASA, 3D ligand site, Metapocket,
4. Molecular Docking using discovery studio
5. Protein-ligand complex MD simulation using DS
6. Ligand Based drug design – QSAR



SBIA5401	R PROGRAMMING	L	T	P	Credits	Total Marks
		3	1	0	4	100

COURSE OBJECTIVES

This course provides knowledge to handle biological problems in Bioinformatics and computational biology

UNIT I – R LANGUAGE

Motivation of R, Attributes and some special values, Types of objects, Sequence generating and vector subsetting, Types of functions, Data Structures, Managing R session, Evaluations.

UNIT II - OBJECT ORIENTED PROGRAMMING IN R

OOP in R, Basics of OOP, S3 OOP, S4 OOP, Documentation, Debugging, Managing S3 and S4 together.

UNIT III - INPUT AND OUTPUT IN R

Basic file handling, connections, File input and output, Tools for accessing files on the Internet.

UNIT IV - WORKING WITH CHARACTER DATA

Built-in Capabilities, Regular Expression, Prefix, Suffix and Substrings, Biological sequences, Matching patterns.

UNIT V - R PACKAGES

Package basics, Package Management and authoring, Initialization, Debugging and Profiling

COURSE OUTCOMES

At the end of the course the students will be able to

- CO 1- Providing general properties and unique aspects of the R language
- CO 2- To teach Object Oriented programming in R
- CO 3- To teach methods for getting of data in and out
- CO 4- To provide knowledge how to how to write own software packages
- CO 5- To port this understanding to healthcare problems
- CO 6- To develop novel pipelines for genomic research

TEXT / REFERENCE BOOKS

- Robert G, "R programming in Bioinformatics", CRC press, Taylor and Francis Group, USA, 2008.
- Own J, Robert. M, and Andrew R., "Introduction to Scientific programming and simulation using R", CRC Press, Taylor and Francis Group, USA, 2014.
- Norman M. And Norman S. Matloff "The Art of R programming: A tour of statistical software Design", No Starch Press Inc, USA, 2011.

END SEMESTER EXAM QUESTION PAPER PATTERN

Max. Marks: 100

Exam Duration: 3 Hrs.

PART A: 6 Questions for 5 Marks each – without choice

30 Marks

PART B: 2 Questions from each unit with internal choice, each carrying 14 marks

70 Marks

SBIA7406	DATA MINING IN BIOINFORMATICS	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

This course will introduce the concepts of data ware house and data mining, which gives a complete description about the principles, used, architectures, applications, design and implementation of data mining and data ware housing concepts

UNIT I INTRODUCTION

Introduction, Importance of Data Mining, Relational Databases, Data Warehouses, Transactional Databases, Advance Database Systems and Applications, Data Mining Functionalities, Classification of Data Mining Systems, Major issues in Data Mining.

UNIT II PRIMITIVES AND SYSTEM ARCHITECTURES

Data Mining Primitives, Data Mining Query Language, Designing Graphical User, Interfaces Based on a Data Mining Query Language, Architectures of Data Mining Systems.

UNIT III CONCEPT DESCRIPTION AND ASSOCIATION RULES

Concept Description, Characterization and comparison, Data Generalization and Summarization-Based Characterization, Analytical Characterization, Mining Class Comparisons, Mining Association Rules in Large Databases, Association Rule Mining, Mining Single Dimensional Boolean Association Rules from Transactional Databases.

UNIT IV CLASSIFICATION AND PREDICTION

Classification and Prediction, Issues: Data preparation for classification and Prediction, Comparing classification Methods Classification by Decision Tree Induction: Decision Trees and Decision Tress induction

UNIT V CLUSTERING METHODS

Clustering Analysis, Types data in clustering analysis: Scaled variable, Binary variables, Variables of Mixed Types, Partitioning Methods: K-means and K-Medoids, Model-Based Methods, Data Mining Applications: Data mining for Biomedical and DNA Data Analysis

COURSE OUTCOMES

- CO1- Be familiar with mathematical foundations of data mining tools.
- CO2- Understand and implement classical models and algorithms in data warehouses and data mining
- CO3- Characterize the kinds of patterns that can be discovered by association rule mining, classification and clustering.
- CO4- Master data mining techniques in various applications like social, scientific and environmental context.
- CO5- Develop skill in selecting the appropriate data mining algorithm for solving practical problems.
- CO6-

TEXT/ REFERENCES BOOKS

1. Data Mining Concepts and Techniques – Jiawei Hen, Micheline Kambler, Academic Press Morgan Kaufman Publishers. 2006
2. Data Mining: Practical machine learning tools Techniques with java implementation by Ian H.Witten, Eibe Frank, 2005.
3. Machine Learning and data mining in pattern recognition in third International conference MLDM, by Petra Perner and Azriel Rosenfield, Springer.2003

END SEMESTER EXAM QUESTION PAPER PATTERN

Max. Marks : 100

Exam Duration : 3

Hrs. PART A : 6 Questions for 5 Marks each –without choice

30 Marks PART B : 2 Questions from each unit with internal choice, each carrying 14 marks

70 Marks

SBIA7408	NEXT GENERATION SEQUENCE DATA ANALYSIS	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

- NGS bioinformatics workflow steps following sequence generation
- NGS for transcriptomics; QC, mapping, visualisation tools
- NGS for genomics; assembly, alignment, QC and variant calling tools
- Exploring EMBL-EBI resources

UNIT I DNA SEQUENCING

DNA Sequencing, first generation DNA sequences, Drawbacks of the first generation sequencing method- Emergence of Next generation sequencing -Platform overview- Biological applications- Basic concepts- Recent scientific breakthroughs using NGS technology

UNIT II OVERVIEW OF NGS TECHNOLOGY

Pyro-sequencing, Illumina Genome Analyzer, Applied Biosystems Sequencing, Ion Torrent Sequencing, Polonator Technology, Nanopore Sequencing, Single Molecule Real Time DNA sequencing- Comparison of Next generation sequencing techniques, & applications-Drawbacks of NGS, NGS File formats- FASTA, FASTQ, FNA, CSFASTA, GFF, SAM and BAM

UNIT III NGS DATABASES AND ANALYSIS OF NGS DATA

NGS databases, Sequence Analysis: Pairwise and multiple sequence alignment methods and tools- Sample data processing with annotation

UNIT IV ALIGNMENT AND ANALYSIS

Genome alignment and analysis tools- BWA (Burrows-Wheeler Aligner), SAMtools, GATK (The Genome Analysis Toolkit), IGV (Integrative Genomics Viewer), HISAT, StringTie, Cuffcompare, Velvet, Oases, Trinity

UNIT V TRANSCRIPTOMICS ANALYSIS

Transcriptome (RNA) sequencing, Exome sequencing, Genome Annotation, Using NGS to detect sequence variants, ChIP-sequence, biological theories on ChIP-sequence analysis, Understanding the non – coding genome, Disease gene identification, DNA fragment evaluation, Peak identification, two condition comparison, Saturation analysis, Motif finding and related theories

COURSE OUTCOMES

After this course the students will be able to

- CO 1: Discuss a variety of applications and workflow approaches for NGS technologies
- CO 2: Use a range of bioinformatics software and tools to undertake basic analysis of NGS data
- CO 3: Understand the advantages and limitations of NGS analyses
- CO 4: Submit, browse and access a range of NGS data available in public repositories using EBI resources
- CO 5: Comprehend the different areas of NGS analysis utility
- CO 6: Apply the knowledge of NGS technology in transcriptome data analysis

TEXT BOOKS /REFERENCES

1. Ali Masoudi-Nejad, Zahra Narimani, Nazanin Hosseinkhan; "Next Generation Sequencing and Sequence Assembly", Methodologies and Algorithms, Springer; 2013.
2. Stuart M. Brown, "Next-Generation DNA Sequencing Informatics", Cold Spring Harbor Laboratory Press, 2013.
3. Mark I. Rees, "Challenges and Opportunities of Next-generation Sequencing for Biomedical Research", Academic Press, 2012.

END SEMESTER EXAM QUESTION PAPER PATTERN

Max. Marks : 100

Exam Duration : 3 Hrs.

PART A : 6 Questions for 5 Marks each –without choice

30 Mar

PART B : 2 Questions from each unit with internal choice, each carrying 14 marks

70 Marks

SBIA7409	HEALTH INFORMATICS	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

This course provides an insight into the fundamental aspects and applications of health informatics.

UNIT I INTRODUCTION

An Introduction to Structural Equation Modeling - Privacy and Data Protection in Healthcare - Biomedical Technology – Case - Allergic Severity Under Conjunctival Provocation - Texture Analysis of Heterogeneous MRI Data for Diagnostic Classification of Childhood Brain Tumors - Establishing a National Medical Device Registry

UNIT II CLINICAL INFORMATICS

Endoscopy Information System (EIS) - Remote Patient Monitoring System Implementation at a Cardiac Care Centre - Vagal Nerve Stimulation (VNS) Therapy System Implementation - Improving Emergency Room Performance - Automation of Fugl- Meyer Assessment Scale for Neurorehabilitation - Extensible Neuroimaging e - Repository for Clinical Trials of Tumours - Query Tool Enabling Clinicians and Researchers to Explore Patient Cohorts

UNIT III INTELLIGENT SYSTEMS

Clinical Decision Support Knowledge Management - Prediction of 5-Year Survival with Data Mining Algorithms – Abstractive Summarization of Drug Dosage Regimens - Exploitation of Medical Crisp Database for Fuzzy Systems - Spatial Epidemiology - E-Learning and Education - Role for Education - Proactive Support of Internet Browsing - Biomedical Informatics Education Programs and Ecosystems - e-Learning Approach to Hospital Information Systems

UNIT IV FORMALISATION AND INFORMATICS

Formalisation of Knowledge, Ontologies, Clinical Guidelines and Standards of Healthcare Clinical Pathways: Identifying Development, Implementation and Evaluation Challenges - “HoPRO” - Data Warehouse Design - Innovative Information Systems in Intensive Care Unit - A Framework to Enrich EHR System - e ENERCA) - Based Stroke Rehabilitation - Greek Public Hospitals' Websites

UNIT V ADVANCES

Health Information Management - Emergency Physician Productivity - Taming Big Data: An Information Extraction Strategy for Large Clinical Text Corpora - Institutional Repository of Digital Health Information Content - Healthcare IT - Neural Network in Healthcare - Healthcare Management - Public Health Informatics - Social Media and Health - Telemedicine

COURSE OUTCOMES

The students will be able to

- CO 1- Understand the basics concepts of health informatics.
- CO 2- To understand the intricacies of clinical informatics.
- CO 3- To elaborate on the understanding of decision support systems in a hospital.
- CO 4- Comprehend clinical documentation in the real world.
- CO 5- To ponder and on the various aspects of telemedicine, healthcare, and social networks in the practice of medicine

TEXT BOOKS / REFERENCES

1. Enabling Health Informatics Applications, edited by J. Mantas, et al., IOS Press, 2011.
2. Sullivan, Frank, and Jeremy Wyatt. ABC of Health Informatics, John Wiley & Sons, Incorporated, 2006.
3. Health Informatics : Improving Patient Care, edited by The Chartered Institute for IT BCS, BCS Learning & Development Limited, 2012.
4. International Perspectives in Health Informatics, edited by E.M. Borycki, et al., IOS Press, 2011.
5. Brender, Jytte. Handbook of Evaluation Methods for Health Informatics, Elsevier Science & Technology, 200

SBIA7412	CLOUD COMPUTING	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

- To gain knowledge in basics of Cloud computing.
- To understand Cloud Computing architecture and industry frameworks such as Map Reduce.
- To discuss practical applications of cloud computing.

UNIT I UNDERSTANDING CLOUD COMPUTING

History of Cloud computing -Cloud Computing Reference Model – NIST Architectural Framework -Types of Clouds- pros and cons of cloud computing –difference between web 2.0 and cloud - key challenges in cloud computing - Major Cloud players - Cloud Deployment Models - Virtualization in Cloud Computing- types of virtualization – web services – key concepts.

UNIT II CLOUD SERVICE MODELS

Software as a Service (SaaS) - Infrastructure as a Service (IaaS)- Platform as a Service (PaaS)- Service Oriented Architectur (SoA) - Elastic Computing - On Demand Computing - Parallelization in Cloud Computing – cloud resource management – dynamic resource allocation- Optimal allocation of cloud models.

UNIT III CLOUD DEPLOYMENT MODELS

Deployment models: Public cloud – Private Cloud –Hybrid cloud – Community cloud - Deployment of applications on the cloud - Hypervisor -Case studies- Xen, Hyper V, Virtual Box, Eucalyptus, Amazon Cloud Computing - Amazon S3 - Amazon EC2-, Windows Azure – creating deploying cloud services, Working with Google Appengine, Disaster Recovery.

UNIT IV VIRTUALIZATION

Cloud data centers – Energy efficiency in data centre – Mobile cloud computing service models– Need for virtualization – Types of Virtualization – Virtualization OS – VMNware, KVM – System VM – Process VM - Virtual Machine Monitor – Properties Interpretation and Binary Translation – HLL VM.

UNIT V CLOUD SECURITY

Cloud security – Security threats and solutions in clouds – Auditing protocols – dynamic auditing – storage security –Privacy preserving — Fully Homomorphic Encryption – Big data security- Cloud availability- DoS attacks – Fault tolerance management in cloud computing- Cloud computing in India.

COURSE OUTCOMES

At the end of the course the students will be able to

- CO1- Understand the basic concepts of cloud computing.
- CO2- Use cloud service models as infrastructure.
- CO3- Deploy cloud models in a real – time environment
- CO4- To efficiently use data centres.
- CO5- Grasp virtualization.
- CO6- To comprehend security issues involved.

TEXT / REFERENCE BOOKS

1. Cloud computing a practical approach - Anthony T.Velte, Toby J. Velte Robert Elsenpeter TATA McGraw- Hill , New Delhi, 2010.
2. Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online – Michael Miller – Que, 2008.
3. Cloud Computing Bible, Barrie Sosinsky, Wiley-India, 2010.
4. Cloud Computing: Principles and Paradigms, Editors: Rajkumar Buyya, James Broberg, Andrzej M. Goscinski, Wiley, 2011.
5. Cloud Computing: Principles, Systems and Applications, Editors: Nikos Antonopoulos, Lee Gillam, Springer, 2012.
6. Cloud Security: A Comprehensive Guide to Secure Cloud Computing, Ronald L. Krutz, Russell Dean Vines, Wiley-India, 2010.