

**SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY**

**(Established under section 3 of UGC Act, 1956)
Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai - 600 119.**



SYLLABUS

SCHOOL OF BIO AND CHEMICAL ENGINEERING

DEPARTMENT OF BIOINFORMATICS

**BACHELOR OF SCIENCE IN BIOINFORMATICS AND DATA SCIENCE
[SIX SEMESTER DEGREE PROGRAMME]**

REGULATIONS 2019

CURRICULUM

Total marks for 1st Semester =700

Total marks for 2nd Semester =700

Semester 3										
S.no	Course type	Course code	Course Title	L	T	P	C	Marks		Page No
								CA E	ESE	
1	Theory	SBIA1301	Molecular Biology and Genomics	3	0	0	3	50	50	32
2	Theory	SBIA1302	Evolutionary biology	3	0	0	3	50	50	33
3	Theory	SBIA1303	Proteomics and Interactomics	3	1	0	4	50	50	35
4	Theory	SBIA1304	Perl For Bioinformatics	3	0	0	3	50	50	36
5	Theory		Elective	3	0	0	3	50	50	
6	Practical	SBIA2301	Perl Lab	0	0	4	2	50	50	38
7	Practical	SBIA2302	Molecular biology Lab	0	0	4	2	50	50	38
Total credits for 3 rd Semester = 20										
Total marks for 3 rd Semester =700										

Semester 4										
S.no	Course type	Course code	Course Title	L	T	P	C	Marks		Page No
								CA E	ESE	
1	Theory	SCSA1102	Fundamentals of Python Programing	3	0	0	3	50	50	40
2	Theory	SCSA1301	Database management system	3	0	0	3	50	50	41
3	Theory	SBIA1401	Structural biology	3	1	0	4	50	50	42
4	Theory	SBIA1402	Immunology and Immunoinformatics	3	0	0	3	50	50	43
5	Theory		Elective	3	0	0	3	50	50	
6	Practical	SCSA2102	Fundamentals of python programming lab	0	0	4	2	50	50	44
7	Practical	SBIA2401	Sequence analysis lab--II	0	0	4	2	50	50	44
8	PT	S738APT	Internship / Skill Development Training	0	0	2	1		100	
Total credits for 4 th Semester = 21										
Total marks for 4 th Semester =700										

Semester 5										
S.no	Course type	Course code	Course Title	L	T	P	C	Marks		Page No
								CAE	ESE	
1	Theory	SBIA1501	Modelling of molecules	3	0	0	3	50	50	46
2	Theory	SBIA1502	Drug design and development	3	0	0	3	50	50	47
3	Theory	SBIA1503	Systems Biology	3	0	0	3	50	50	48
4	Theory	SBIA1504	Cheminformatics	3	0	0	3	50	50	50
5	Theory		Elective	3	0	0	3	50	50	
6	Practical	SBIA2501	Systems biology lab	0	0	4	2	50	50	51
7	Practical	SBIA2502	Cheminformatics and Drug design Lab	0	0	4	2	50	50	51
8	Seminar	S738ASEM	Seminar	0	0	0	1		100	
Total credits for 5th Semester = 20										
Total marks for 5th Semester =800										

Semester 6										
S.no	Course type	Course code	Course Title	L	T	P	C	Marks		Page No
								CAE	ESE	
1	Theory	SBIA1601	Data mining and Machine Learning	3	0	0	3	50	50	53
2	Theory	SBIA1602	Microarray data analysis	3	0	0	3	50	50	55
3	Practical	S738APROJ	Project				15		100	
Total credits for 6th Semester = 21										
Total marks for 6th Semester =300										

Semester	Theory courses (including elective courses)			Practical Courses (including PT and project)		
	Total no.	Total Credits	Total Marks	Total no.	Total Credits	Total Marks
1	5	15	500	2	4	200
2	5	15	500	2	4	200
3	5	16	500	2	4	200
4	5	16	500	2	5	200
5	5	15	500	3	5	300
6	2	6	200	1	15	100
Overall Total	27	83	2700	10	37	1200

Overall total credits for B.Sc. programme	120
Overall total marks for B.Sc. programme	4000

PROFESSIONAL CORE ELECTIVES									
Sl. No.	Course code	Course title	L	T	P	C	Marks		Page No
							CAE	ESE	
1.	SBIA3001	Pharmacogenomics	3	0	0	3	50	50	58
2.	SBIA3002	Cancer Biology	3	0	0	3	50	50	59
3.	SCSA3016	Data Science	3	0	0	3	50	50	61
4.	SBIA3003	R Programing for Biologists	3	0	0	3	50	50	63
5.	SITA3008	Internet of Things	3	0	0	3	50	50	64
6.	SBIA3004	Next Generation Sequencing	3	0	0	3	50	50	65
7.	SBIA3005	Quantitative modelling in biology	3	0	0	3	50	50	67
8.	SBIA3006	Stem Cell Research	3	0	0	3	50	50	69
9.	SCSA3015	Deep Learning	3	0	0	3	50	50	70
10.	SCSA1501	Operating Systems	3	0	0	3	50	50	72
11.	SBIA3007	Bigdata in Health care	3	0	0	3	50	50	74
12.	SBIA3010	Health Informatics	3	0	0	3	50	50	75

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

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****PART A** : 10 questions of 2 marks each - No choice**PART B** : 2 questions from each unit of internal choice; each carrying 16 marks**Exam Duration : 3 Hrs.****20 Marks****80 Marks**

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. Aminoacids: 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

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A** : 10 questions of 2 marks each - No choice**20 Marks****PART B** : 2 questions from each unit of internal choice; each carrying 16 marks**80 Marks**

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

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

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****PART A** : 10 questions of 2 marks each - No choice**PART B** : 2 questions from each unit of internal choice; each carrying 16 marks**Exam Duration : 3 Hrs.****20 Marks****80 Marks**

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.

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A : 10 questions of 2 marks each - No choice****20 Marks****PART B : 2 questions from each unit of internal choice; each carrying 16 marks****80 Marks**

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.

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A :** 10 questions of 2 marks each - No choice**20 Marks****PART B :** 2 questions from each unit of internal choice; each carrying 16 marks**80 Marks**

SBIA1302	EVOLUTIONARY BIOLOGY	L	T	P	Credits	Total Marks
		3	1	0	4	100

COURSE OBJECTIVES

The students will be able to understand the basics of evolution and be able to perform evolutionary analyses using molecular data.

UNIT I INTRODUCTION**9hrs**

Introduction - What Is Evolution? - The History of Evolutionary Thought - The Evidence for Evolution - From DNA to Phenotypes - Phylogenetics and the History of Life - Interpretation of Phylogenetic Trees - Molecular Clock Dating - Historical Biogeography - Phylogeography - Concepts in Character Macroevolution: Adaptation, Homology, and Evolvability - Using Phylogenies to Study Phenotypic Evolution: Comparative Methods and Tests of Adaptation - Taxonomy in a Phylogenetic Framework - The Fossil Record - Human Evolution

UNIT II Natural Selection and Adaptation**9hrs**

Natural Selection and Adaptation - Natural Selection, Adaptation, and Fitness: Overview - Units and Levels of Selection - Theory of Selection in Populations Phenotypic Selection on Quantitative Traits - Responses to Selection: Experimental Populations - Responses to Selection: Natural Populations Evolutionary Limits and Constraints - Evolution of Modifier Genes and Biological Systems - Evolution of Form and Function - Biochemical and Physiological Adaptations

UNIT III Evolutionary Processes**9hrs**

Evolutionary Processes - Genetic Drift - Mutation - Geographic Variation, Population Structure, and Migration - Recombination and Sex - Genetic Load - Inbreeding - Selfish Genetic Elements and Genetic Conflict - Evolution of Mating Systems: Outcrossing versus Selfing.

UNIT IV Genes, Genomes, Phenotypes**9hrs**

Genes, Genomes, Phenotypes - Molecular Evolution - Genome Evolution - Comparative Genomics - Evolution of Sex Chromosomes - Gene Duplication - Evolution of New Genes - Evolution of Gene Expression - Epigenetics - Evolution of Molecular Networks - Evolution and Development: Organisms - Evolution and Development: Molecules - Genetics of Phenotypic Evolution - Dissection of Complex Trait Evolution - Searching for Adaptation in the Genome - Ancient DNA

UNIT V Speciation and Macroevolution**9hrs**

Speciation and Macroevolution - Species and Speciation - Speciation Patterns - Geography, Range Evolution, and Speciation - Speciation and Natural Selection - Speciation and Sexual Selection - Gene Flow, Hybridization, and Speciation - Coevolution and Speciation - Genetics of Speciation - Speciation and Genome Evolution - Adaptive Radiation - Macroevolutionary Rates - Macroevolutionary Trends - Causes and Consequences of Extinction - Species Selection - Key Evolutionary Innovations - Evolution of Communities

Course outcomes

At the end of the course the students will

- CO1** – describe the need for phylogenetic studies and understand its significance
- CO2** – infer the different methods to infer phylogeny used across history
- CO3** – Understand mechanisms of evolution.
- CO4** – Understand the fundamentals of macroevolution and coevolution.
- CO5** – Understand phylogeography, speciation and extinction.
- CO6** – designing a protocol to study phylogenetics with a hypothesis.

TEXT / REFERENCE BOOKS

1. The Princeton Guide to Evolution, edited by David A. Baum, and Douglas J. Futuyma, Princeton University Press, 2013.
2. Davidson, Eric H., and Isabelle Peter. Genomic Control Process: Development and Evolution, Elsevier Science & Technology, 2015.
3. Why does Evolution Matter? The Importance of Understanding Evolution, edited by Gabriel Trueba, Cambridge Scholars Publishing, 2014.
4. Martin, Lynn B., et al. Integrative Organismal Biology, John Wiley & Sons, Incorporated, 2014.
5. Wagner, Günter P. Homology, Genes, and Evolutionary Innovation, Princeton University Press, 2014.

END SEMESTER EXAMINATION QUESTION PAPER PATTERN

Max. Marks : 100

Exam Duration : 3 Hrs.

PART A : 10 questions of 2 marks each - No choice

20 Marks

PART B : 2 questions from each unit of internal choice; each carrying 16 marks

80 Marks

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

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A** : 10 questions of 2 marks each - No choice**20 Marks****PART B** : 2 questions from each unit of internal choice; each carrying 16 marks**80 Marks**

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. Oualline, Steve. Wicked Cool Perl Scripts: Useful Perl Scripts That Solve Difficult Problems, No Starch Press, Incorporated, 2006.
7. Krumins, Peteris. Perl One-Liners: 130 Programs That Get Things Done, No Starch Press, Incorporated, 2013.

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****PART A : 10 questions of 2 marks each - No choice****PART B : 2 questions from each unit of internal choice; each carrying 16 marks****Exam Duration : 3 Hrs.****20 Marks****80 Marks**

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, listbox, message, radiobutton, 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

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****PART A** : 10 questions of 2 marks each - No choice**PART B** : 2 questions from each unit of internal choice; each carrying 16 marks**Exam Duration : 3 Hrs.****20 Marks****80 Marks**

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

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****PART A :** 10 questions of 2 marks each - No choice**PART B :** 2 questions from each unit of internal choice; each carrying 16 marks**Exam Duration : 3 Hrs.****20 Marks****80 Marks**

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 - NOE 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

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

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****PART A** : 10 questions of 2 marks each - No choice**PART B** : 2 questions from each unit of internal choice; each carrying 16 marks**Exam Duration : 3 Hrs.****20 Marks****80 Marks**

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 immunoinformatics, 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 PRes 2005 ISBN: 0262122804.
5. Foundation Novartis. Immunoinformatics: Bioinformatic Strategies for Better Understanding of Immune Function. Publisher: Chichester, John Wiley & Sons Inc. 2003. ISBN: 0470853565.

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A** : 10 questions of 2 marks each - No choice**20 Marks****PART B** : 2 questions from each unit of internal choice; each carrying 16 marks**80 Marks**

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 dependant 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.

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A** : 10 questions of 2 marks each - No choice**20 Marks****PART B** : 2 questions from each unit of internal choice; each carrying 16 marks**80 Marks**

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.

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****PART A** : 10 questions of 2 marks each - No choice**PART B** : 2 questions from each unit of internal choice; each carrying 16 marks**Exam Duration : 3 Hrs.****20 Marks****80 Marks**

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 emergentism? - 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

- Nielsen, Jens, and Stefan Hohmann. Systems Biology, John Wiley & Sons, Incorporated, 2017.
- Systems Biology, edited by Robert A. Meyers, John Wiley & Sons, Incorporated, 2012.
- Systems Biology: Philosophical Foundations, edited by Fred C. Boogerd, et al., Elsevier Science & Technology, 2007.
- Lodhi, Huma M., and Stephen H. Muggleton. Elements of Computational Systems Biology, John Wiley & Sons, Incorporated, 2010.
- 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.

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A :** 10 questions of 2 marks each - No choice**20 Marks****PART B :** 2 questions from each unit of internal choice; each carrying 16 marks**80 Marks**

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, Chempider.

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 Cheminformatics Springer 2007
2. Jurgen Bajorath Cheminformatics: Concepts, Methods, and Tools for Drug Discovery Humana Press 2010
3. Jurgen Bajorath Cheminformatics for Drug Discovery Wiley 2013
4. Daria Bove Computational Chemistry: Theories, Methods and Applications Nova Science Publishers Inc, New York 2014

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****PART A** : 10 questions of 2 marks each - No choice**PART B** : 2 questions from each unit of internal choice; each carrying 16 marks**Exam Duration : 3 Hrs.****20 Marks****80 Marks**

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 Pernert and Azriel Rosenfeld, Springer.2003
6. Witten and Frank, Data Mining: Practical Machine Learning Tools and Techniques by, Elsevier.

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A : 10 questions of 2 marks each - No choice****20 Marks****PART B : 2 questions from each unit of internal choice; each carrying 16 marks****80 Marks**

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

END SEMESTER EXAMINATION QUESTION PAPER PATTERN

Max. Marks : 100

Exam Duration : 3 Hrs.

PART A : 10 questions of 2 marks each - No choice

20 Marks

PART B : 2 questions from each unit of internal choice; each carrying 16 marks

80 Marks

SBIA3001	PHARMACOGENOMICS	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

This course introduces students to the emerging fields of pharmacogenomics and personalized 'precision' medicine that are now transforming the face of modern health care.

UNIT I Introduction**9hrs**

Pharmacogenomics: Past, Present, and Future - The Genetic Basis of Pharmacogenomics - Methodologies in Pharmacogenomics- The Pharmacogenetics of Drug Metabolism - Pharmacogenomics and Drug Transport/Efflux - Pharmacodynamics and Pharmacogenomics.

UNIT II Pharmacogenomics: Past, Present, and Future**9hrs**

The Genetic Basis of Pharmacogenomics - Methodologies in Pharmacogenomics - The Pharmacogenetics of Drug Metabolism - Pharmacogenomics and Drug Transport/Efflux. - Pharmacodynamics and Pharmacogenomics

UNIT III Important Issues in Pharmacogenomics**9hrs**

The Role of Pharmacists in Pharmacogenomics - Pharmacogenomics and Clinical Testing = Pharmacogenomics in Pharmacy Education - Ethical Considerations in Pharmacogenomics - Future Promise of Pharmacogenomics in Clinical Practice

UNIT IV Advances in pharmacogenomics**9hrs**

Traditional Therapeutic Drug Monitoring and Pharmacogenomics: Are They Complementary? - Pharmacogenomics Aspect of Warfarin Therapy - Pharmacogenetics and Cancer Chemotherapy - Pharmacogenomic Considerations in Anesthesia and Pain Management - Pharmacogenomics of Immunosuppressants

UNIT V Practice and future of pharmacogenomics**9hrs**

Pharmacogenomics of Cardiovascular Drugs - Pharmacogenomic Aspects of Antiretroviral Therapy - Pharmacogenetics of Psychoactive Drugs - Pharmacogenomic Aspects of Adverse Drug- Drug Interactions - Microarray Technology and Other Methods in Pharmacogenomics Testing - Pharmacogenetic Testing in the Clinical Laboratory Environment

Course outcomes

At the end of the course the students will

- CO1** –Comprehend the historical basis of pharmacogenomics along with the basics of pharmacodynamics.
- CO2** –Understand drug metabolism and its genetic basis.
- CO3** –Identify the issues in pharmacogenomics
- CO4** –Describe the pharmacogenomics of cancer and its implementation.
- CO5** –Analyse the latest advancements in pharmacogenomics
- CO6** –Discuss the futuristic aspects of pharmacogenomics and its role in systems and personalized medicine.

TEXT / REFERENCE BOOKS

1. Pharmacogenomics in Clinical Therapeutics, edited by Loralie J. Langman, and Amitava Dasgupta, John Wiley & Sons, Incorporated, 2012.
2. Zdanowicz, Martin M.. Concepts in Pharmacogenomics, ASHP, 2010.

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A :** 10 questions of 2 marks each - No choice**20 Marks****PART B :** 2 questions from each unit of internal choice; each carrying 16 marks**80 Marks**

SBIA3002	CANCER BIOLOGY	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

This course will educate students on various genetic and molecular changes that normal cells undergo to turn into malignant cancer cells and also describe factors that contribute to cancer development. It presents an overview on cancer prevention and currently available therapeutic treatments.

Unit I Fundamentals Of Cancer Biology

9hrs

Definition, Hall marks of cancer, Types of Cancer, Common Symptoms. Cell signalling in carcinogenesis- Growth factors and their receptors. Genes frequently mutated in cancer - Cellular oncogenes, Tumour suppressor genes. Causes of cancer- Mutations and cancer, Viral causes of cancer, the importance of DNA repair systems

Unit II Carcinogenesis

9hrs

The main stages of carcinogenesis – an overview, Early steps characterised in cancer, Cellular principles of invasion and metastasis, Stromal microenvironment and carcinogenesis

Unit III Principles Of Carcinogenesis

9hrs

Chemical Carcinogenesis -Metabolism of Carcinogenesis, Natural History of Carcinogenesis, Targets of Chemical Carcinogenesis. Physical Carcinogenesis- X - Ray radiation - mechanism of radiation Carcinogenesis.

Unit IV Cancer Metastasis

9hrs

Clinical significances of invasion, heterogeneity of metastatic phenotype, Metastatic cascade, Basement Membrane disruption, Three-step theory of Invasion, Proteinases and tumour cell invasion.

Unit V Detection Of Cancer And Cancer Therapy

9hrs

Detection of Cancers, Prediction of aggressiveness of Cancer, Advances in Cancer detection. Different forms of therapy, Chemotherapy, radiation therapy, and Immuno therapy: advantages and limitations. Challenges in cancer treatment.

Course outcomes

At the end of the course the students will

- CO1** –Identify the six hall marks of cancer
- CO2** –Analyse the mechanism of carcinogenesis and evaluate the effect of chemical carcinogens and assess the role of physical carcinogens
- CO3** –Identify the ways in which the oncogenes are activated and outline the current state of knowledge about the biochemical and molecular basis of cancer
- CO4** –Describe the multistage nature of cancer and be able to discuss factors which influence this process
- CO5** –List out the various cancer treatment strategies and their suitability
- CO6** –Discuss contemporary approaches to the development of therapies, including specific molecular alterations in cancer as therapeutic targets and prognostic factors to guide treatment.

Texts / Reference Books

- Robert A. Weinberg The Biology of Cancer, Garland Science 2013
- Lewis J. Kleinsmith Principles of Cancer Biology: Pearson New International Edition 2013
- Robin Hesketh Introduction to Cancer Biology Cambridge University Press 2012
- Lauren Pecorino Molecular Biology of Cancer: Mechanisms, Targets, and Therapeutics Oxford University Press 2005

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A : 10 questions of 2 marks each - No choice****20 Marks****PART B : 2 questions from each unit of internal choice; each carrying 16 marks****80 Marks**

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

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A :** 10 questions of 2 marks each - No choice**20 Marks****PART B :** 2 questions from each unit of internal choice; each carrying 16 marks**80 Marks**

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.

END SEMESTER EXAMINATION QUESTION PAPER PATTERN

Max. Marks : 100

Exam Duration : 3 Hrs.

PART A : 10 questions of 2 marks each - No choice

20 Marks

PART B : 2 questions from each unit of internal choice; each carrying 16 marks

80 Marks

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

END SEMESTER EXAMINATION QUESTION PAPER PATTERN

Max. Marks : 100

Exam Duration : 3 Hrs.

PART A : 10 questions of 2 marks each - No choice

20 Marks

PART B : 2 questions from each unit of internal choice; each carrying 16 marks

80 Marks

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

END SEMESTER EXAMINATION QUESTION PAPER PATTERN

Max. Marks : 100

Exam Duration : 3 Hrs.

PART A : 10 questions of 2 marks each - No choice

20 Marks

PART B : 2 questions from each unit of internal choice; each carrying 16 marks

80 Marks

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 NF κ B 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

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

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A : 10 questions of 2 marks each - No choice****20 Marks****PART B : 2 questions from each unit of internal choice; each carrying 16 marks****80 Marks**

SBIA3006	STEM CELL RESEARCH	L	T	P	Credits	Total Marks
		3	0	0	3	100

COURSE OBJECTIVES

The objective is to make the students aware of the basics of stem cells, experimental techniques to isolate them and understand their applications.

Unit I Fundamentals of Stem Cell research**9hrs**

Introduction - Properties of stem cells - Types of stem cell - Similarities and differences between embryonic and adult stem cells - pluripotency - totipotency

Unit II Isolation of stem cells**9hrs**

Isolation of stem cells from embryo cultures - Invitro fertilization - blastocyst - inner cell mass - Growing ES cells in lab - laboratory tests to identify ES cells - properties of ES cells and Embryonic stem cell markers - Governance of hESC research, Ethical issues at stake, regulation in European and non - European countries.

Unit III Differentiation**9hrs**

Somatic stem cells, Test for identification of adult stem cells, Adult stem cell differentiation - Trans differentiation -Plasticity - Different types of adult stem cells.

Unit IV Stem Cells and Tissue Engineering**9hrs**

Tissue engineering triad, ECM components and their role in tissue development, Tissue engineering application - Production of complete organ - Kidney - Eyes - Heart - Brain.

Unit V Applications**9hrs**

Gene therapy - genetically manipulated stem cells - stem cells and Animal cloning - transgenic animals and stem cells - Therapeutic applications – Parkinson's disease - Neurological disorder - limb amputation - heart disease - spinal cord injuries - diabetes -burns - HLA typing- Alzheimer's disease.

Course Outcomes

At the end of the course the students will

- CO1 –appreciate the significance of stem cells
- CO2 –the strategy associated in the governance of stem cell research in various countries.
- CO3 –Identify the adult stem cells.
- CO4 –Appreciate the role of stem cells in tissue engineering.
- CO5 – Appreciate the role of stem cells in therapy.
- CO6 – Describe the key role played by stem cells in the living system

Text / Reference Books

1. Kursad and Turksen, Embryonic Stem cells, Humana Press, 2002
2. Stem cell and future of regenerative medicine. By committee on the Biological and Biomedical applications of Stem cell Research, 12 National Academic press, 2002

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A :** 10 questions of 2 marks each - No choice**20 Marks****PART B :** 2 questions from each unit of internal choice; each carrying 16 marks**80 Marks**

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- Generating 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.

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A : 10 questions of 2 marks each - No choice****20 Marks****PART B : 2 questions from each unit of internal choice; each carrying 16 marks****80 Marks**

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 drivers.

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 VsiOS - 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"

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A : 10 questions of 2 marks each - No choice****20 Marks****PART B : 2 questions from each unit of internal choice; each carrying 16 marks****80 Marks**

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

- 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 EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A** : 10 questions of 2 marks each - No choice**20 Marks****PART B** : 2 questions from each unit of internal choice; each carrying 16 marks**80 Marks**

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.

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A :** 10 questions of 2 marks each - No choice**20 Marks****PART B :** 2 questions from each unit of internal choice; each carrying 16 marks**80 Marks**

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 quality 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.

END SEMESTER EXAMINATION QUESTION PAPER PATTERN**Max. Marks : 100****Exam Duration : 3 Hrs.****PART A :** 10 questions of 2 marks each - No choice**20 Marks****PART B :** 2 questions from each unit of internal choice; each carrying 16 marks**80 Marks**