PERIYAR UNIVERSITY

SALEM 636 011

M.Sc. DEGREE

COMPUTER SCIENCE-DATA ANALYTICS

REGULATIONS AND SYLLABUS

(Effective from the Academic year 2019-2020 and thereafter)

Regulations

Effective from the Academic year 2019 - 20

1. OBJECTIVE OF THE PROGRAMME

To Develop the Post Graduates in **COMPUTER SCIENCE-DATA ANALYTICS** with strong knowledge of theoretical COMPUTER SCIENCE-DATA ANALYTICS subjects who can be employed in research and development units of industries and academic institutions.

2. ELEIGIBILITY FOR ADMISSION

A candidate who has passed B.Sc Computer Science/B.C.A/B.Sc Computer Technology/B.Sc Information Science/Technology degree of this University or any of the degree of any other University accepted by the syndicate as equivalent thereto subject to such conditions as may be prescribed therefore shall be permitted to appear and qualify for the M. Sc COMPUTER SCIENCE-DATA ANALYTICS degree examination of this University after a course of study of two academic years.

3. DURATION OF THE PROGRAMME

The programme for the degree of **Master of Science in COMPUTER SCIENCE-DATA ANALYTICS** shall consist of **two Academic years** divided into four semesters. Each semester consist of 90 working days.

Periyar University, Salem

CURRICULUM AND SCHEME OF EXAMINATIONS

Two year M. Sc-COMPUTER SCIENCE-DATA ANALYTICS Programme

Course	Number of Credit s	Hours Per Week	Examination Duration (hrs)	Marks		
				I.A	ESE	Total
Semester-I						
Core Course-19UPCSC3C01 Design and Analysis of Algorithm	4	4	3	25	75	100
Core Course-19UPCSC3C02 Advanced Web Technology	4	4	3	25	75	100
Core Course-19UPCSC3C03 Advanced Data Base Management Systems	4	4	3	25	75	100
Core Course-19UPCSC3C04 Compiler Design	4	4	3	25	75	100
Elective-01	4	4	3			
Core Course-19UPCSC3C05 Algorithm – Lab	2	4	3	40	60	100

Core Course-19UPCSC3C06 Advanced Web Technology – Lab	2	4	3	40	60	100
SWAYAM / MOOC /SOFT SKILL-01 (Optional)	ADD ON COURSE					
Semester-II						
Core Course-19UPCSC3C07 Distributed Operating System	4	4	3	25	75	100
Core Course-19UPCSC3C08 Advanced Java Programming	4	4	3	25	75	100
Core Course-19UPCSC3C09 Cryptography and Network Security	4	4	3	25	75	100
Elective-02	4	4	3			
Elective-03	4	4	3			
Core Course-19UPCSC3C10 * Advanced Java-Lab	2	4	3	40	60	100
Core Course-19UPCSC3C11* Data Analytics using R – Lab	2	4	3	40	60	100
Non-major Elective /	4	4	3	25	75	100

Supportive Course						
SWAYAM / MOOC / SOFT SKILL-02 (Optional)	ADD ON COURSE					
Semester-III						
Core Course-19UPCSC3C12 Digital Image Processing	4	4	3	25	75	100
Core Course-19UPCSC3C13 Internet of Things	4	4	3	25	75	100
Core Course-19UPCSC3C14 Machine Learning	4	4	3	25	75	100
Elective-04	4	4	3			
Elective-05	4	4	3			
Core Course-19UPCSC3C15* Image Processing – Lab	2	4	3	40	60	100
Core Course-19UPCSC3C16 * Machine Learning-Lab/ Mini Project	2	4	3	40	60	100
SWAYAM / MOOC / SOFT SKILL-03 (Optional)	ADD ON COURSE					
Grand Total	90					

Semester-IV						
Option-I						
Core Course-19UPCSC3C17 **Dissertation and Viva Voice (Industry/Research)**	16	-	-	50	150	200
Option-II						
Elective-06	4	4	3			
Elective-07	4	4	3			
Core Course-19UPCSC3C18 **Dissertation and Viva-Voce** **(Industry/Research)**	8	12	-	50	150	200

Elective Courses (EC) **

Course code	Name of the Course	No. of Hours			Credits
		L	T	P	
19UPCSC3E01	Discrete Mathematics for Computing	3	1	-	4
19UPCSC3E02	Statistical Computing	3	1	-	4
19UPCSC3E03	Fundamentals of Data Science	3	1	-	4
19UPCSC3E04	Data Visualization Techniques	3	1	-	4
19UPCSC3E05	Predictive Analytics	3	-		3
19UPCSC3E06	Predictive Analytics Lab	-			1
19UPCSC3E07	Text Analytics	-	-		3
19UPCSC3E08	Text Analytics Lab	2			1
19UPCSC3E09	Web Analytics	3	-		4
19UPCSC3E10	Market Analytics	-			4
19UPCSC3E11	Operations Research	-	-		3
19UPCSC3E12	Operations Research Lab	2			1
19UPCSC3E13	Optimization Techniques	3	1		4
19UPCSC3E14	Probability and Stochastic Process	-			4
19UPCSC3E15	Linear Programming	3	1		4
19UPCSC3E16	Big Data Analytics	-			3
19UPCSC3E17	Big Data Analytics Lab	3	-	-	1
19UPCSC3E18	Information Retrieval	-	-	2	4
19UPCSC3E19	Natural Language Computing	3	1	-	3
19UPCSC3E20	Natural Language Computing Lab	3	1	-	1
19UPCSC3E21	Image and Video Analytics	3	1	-	4
19UPCSC3E22	Social Networking and Mining	3	-	-	4
19UPCSC3E23	Web Intelligence	-	-	2	4
19UPCSC3E24	Deep Learning Techniques	3	1	-	4
19UPCSC3E25	Bioinformatics	3	-	-	4

**** The Elective Courses may be updated a per the current trends in Computer Science by their respective boards**

M. Sc- COMPUTER SCIENCE-DATA ANALYTICS

SYLLABUS

SEMESTER-I

COURSE- 19UPCSC3C01 **Credits: 4**

DESIGN AND ANALYSIS OF ALGORITHMS

Course Objective:

- To learn effective problem solving methodologies in Computing applications
- To analyze the algorithmic procedure efficiently and to determine the computational complexity of algorithms
- To learn stepwise procedure to solve problems.
- Students can solve problems by a model of computations

Unit I

Introduction: Algorithm Definition – Algorithm Specification – Performance Analysis-Asymptotic Notations. Elementary Data Structures: Stacks and Queues – Trees – Dictionaries – Priority Queues – Sets and Disjoint Set Union – Graphs

Unit II

Divide and Conquer: The General Method – Defective Chessboard – Binary Search – Finding the Maximum and Minimum – Merge Sort – Quick Sort – Selection - Strassen‟s Matrix Multiplication.

Unit III

The Greedy Method: General Method - Container Loading - Knapsack Problem - Tree Vertex Splitting – Job Sequencing With Deadlines - Minimum Cost Spanning Trees - Optimal Storage On Tapes – Optimal Merge Patterns - Single Source Shortest Paths.

Unit IV

Dynamic Programming: The General Method – Multistage Graphs – All-Pairs Shortest Paths – Single-Source Shortest Paths - Optimal Binary Search Trees - String Editing - 0/1 Knapsack - Reliability Design - The Traveling Salesperson Problem - Flow Shop Scheduling. Basic Traversal and Search Techniques: Techniques for Binary Trees – Techniques for Graphs – Connected Components and Spanning Trees – Biconnected Components and DFS.

Unit V

Backtracking: The General Method – The 8-Queens Problem – Sum of Subsets – Graph Coloring – Hamiltonian Cycles – Knapsack Problem Branch and Bound: Least Cost searchhod - 0/1 Knapsack Problem.

Text Book:

1. Ellis Horowitz, Satraj Sahni and Sanguthevar Rajasekaran, Fundamentals of Computer Algorithms, Universities Press, Second Edition, Reprint 2009.

References:

1. Data Structures Using C - Langsam, Augenstien, Tenenbaum, PHI
2. Data structures and Algorithms, V.Aho, Hopcropft, Ullman , LPE
3. Introduction to design and Analysis of Algorithms - S.E. Goodman, ST. Hedetniem- TMH.
4. Carlos A.Coello Coello, Gary B.Lamont, David A.Van Veldhuizen, “Evolutionary Algorithms for Solving Multi-Objective Problems”, Springer 2nd Edition, 2007.

COURSE- 19UPCSC3C02 **Credits: 4**

ADVANCED WEB TECHNOLOGY

Course Objective:

- To explore the backbone of web page creation by developing .NET skill.
- To enrich knowledge about HTML control, web control classes and ADO.NET
- To understand the need of usability, evaluation methods for web services
- Apply the knowledge of ASP.NET object, ADO.NET data access and SQL to develop a client server model.
- Recognize the difference between Data list and Data grid controls in accessing data.

Unit - I

Overview of ASP.NET - The .NET framework – Learning the .NET languages : Data types – Declaring variables- Scope and Accessibility- Variable operations- Object Based manipulation- Conditional Structures- Loop Structures- Functions and Subroutines. Types, Objects and Namespaces : The Basics about Classes- Value types and Reference types- Advanced class programming- Understanding name spaces and assemblies. Setting Up ASP.NET and IIS

Unit – II

Developing ASP.NET Applications - ASP.NET Applications: ASP.NET applications– Code behind- The Global.asax application file- Understanding ASP.NET Classes- ASP.NET Configuration. Web Form fundamentals: A simple page applet- Improving the currency converter- HTML control classes- The page class- Accessing HTML server controls. Web controls: Web Control Classes – Auto PostBack and Web Control events- Accessing web controls. Using Visual Studio.NET: Starting a Visual Studio.NET Project- Web form Designer- Writing code- Visual studio.NET debugging. Validation and Rich Controls: Validation- A simple Validation example- Understanding regular expressions- A validated customer form. State management - Tracing, Logging, and Error Handling.

Unit – III

Working with Data - Overview of ADO.NET - ADO.NET and data management- Characteristics of ADO.NET-ADO.NET object model. ADO.NET data access : SQL basics– Select , Update, Insert, Delete statements- Accessing data- Creating a connection- Using a command with a DataReader - Accessing Disconnected data - Selecting multiple tables – Updating Disconnected data. Data binding: Single value Data Binding- Repeated value data binding- Data binding with data bases. Data list – Data grid – Repeater – Files, Streams and Email – Using XML

Unit - IV

Web Services - Web services Architecture : Internet programming then and now- WSDL–SOAP- Communicating with a web service-Web service discovery and UDDI. Creating Web services : Web service basics- The StockQuote web service – Documenting the web service- Testing the web service- Web service Data types- ASP.NET intrinsic objects. Using web services: Consuming a web service- Using the proxy class- An example with TerraService.

Unit – V

Advanced ASP.NET - Component Based Programming: Creating a simple component – Properties and state- Database components- Using COM components. Custom controls: User Controls- Deriving Custom controls. Caching and Performance Tuning: Designing and scalability– Profiling- Catching- Output catching- Data catching. Implementing security: Determining security requirements- The ASP.NET security model- Forms authentication- Windows authentication.

Text Book

1. Mathew Mac Donald, “ASP.NET Complete Reference”, TMH 2005.

Reference Books

1. Crouch Matt J, “ASP.NET and VB.NET Web Programming”, Addison Wesley 2002.
2. J.Liberty, D.Hurwitz, “Programming ASP.NET”, Third Edition, O‟REILLY, 2006.

COURSE- 19UPCSC3C03 **Credits: 4**

ADVANCED DATABASE MANAGEMENT SYSTEMS

Course Objective:

- To acquire Knowledge of Database Models, Applications of Database Models and Emerging Trends
- To learn Various Data models and Database Architecture
- To learn Knowledge patterns, Object Oriented Databases

Unit-I

Relational and parallel Database Design: Basics - Entity Types - Relationship Types - ER Model - ER-to-Relational Mapping algorithm. Normalization: Functional Dependency - 1NF - 2NF - 3NF - BCNF -4NF and 5NF. Architecture -I/O Parallelism - Interquery Parallelism - Intraquery Parallelism - Intraoperation Parallelism - Interoperation Parallelism.

Unit-II

Distributed and Object based Databases: Architecture - Distributed data storage - Distributed transactions - Commit protocols - Concurrency control - Query Processing. Complex Data Types - Structured Types and Inheritance - Table Inheritance - array and Multiset - Object Identity and Reference Types - Object Oriented versus Object Relational.

Unit-III

Spatial Database: Spatial Database Characteristics - Spatial Data Model - Spatial Database Queries - Techniques of Spatial Database Query - Logic based Databases: Introduction - Overview - Propositional Calculus - Predicate Calculus - Deductive Database Systems - Recursive Query Processing.

Unit-IV

XML Databases: XML Hierarchical data model - XML Documents - DTD - XML Schema - XML Querying - XHTML - Illustrative Experiments.

Unit-V

Temporal Databases: Introduction - Intervals - Packing and Unpacking Relations - Generalizing the relational Operators - Database Design - Integrity Con/straints - Multimedia Databases: Multimedia Sources - Multimedia Database Queries - Multimedia Database Applications.

Text Books

1. Abraham Silberschatz, Henry F Korth , S Sudarshan, "Database System Comcepts", 6th edition , McGraw-Hill International Edition , 2011
2. C.J.Date, A.Kannan, S.Swamynathan, "An Introduction to Database Systems", 8th Edition, Pearson Education Reprint 2016.

Reference Books

1. Ramez Elmasri, Shamkant B Navathe, "Fundamental of Database Systems", Pearson, 7th edition 2016.
2. Thomas Connolly, Carolyn Begg., "Database Systems a practical approach to Design , Inmplementation and Management ", Pearson Education, 2014.

COURSE- 19UPCSC3C04 **Credits: 4**

COMPILER DESIGN

Course Objective:

- To learn principles, algorithms and techniques used to construct various phases of compiler.
- To acquire knowledge about finite automata, regular expressions, context-free grammars, compiler parsing techniques, Syntax Directed definitions and translation scheme
- To understand intermediate machine representations and actual code generation

Unit – I

Lexical analysis - Language Processors, The Structure of a Compiler, Parameter passing mechanism – Symbol table - The role of the lexical analyzer - Input buffering - Specification of tokens - Recognition of tokens – Finite automata - Regular expression to automata.

Unit – II

Syntax Analysis - The role of the parser - Context-free grammars - Writing a grammar - Top down Parsing - Bottom-up Parsing - LR parsers- LALR parsers.

Unit – III

Semantic Analysis - Inherited and Synthesized attributes – Dependency graphs – Ordering the evaluation of attributes – S-attributed definitions – L-attributed definitions – Applications of Syntax Directed translation – Syntax Directed translations schemes - Storage organization – Stack allocation of space.

Unit – IV

Intermediate Code Generation - Variants of Syntax trees – Three Address code – Types and Declarations - Translation of Expressions – Type checking - Control flow - Back patching - Switch Statements - Procedure calls.

Unit - V

Code Generation and Code Optimization - Issues in the design of a code generator - The target language – Address in the Target Code – Basic Block and Flow graphs – Optimization of Basic Blocks - A simple code generator – Peephole Optimization.

Text Book

1. Alfred V. Aho, Monica S.Lam, Ravi Sethi and Jeffrey D. Ullman, “Compilers- Principles, Techniques and Tools”, Second Edition, Pearson Education Asia, 2009.

Reference Books

1. A.V. Aho, Ravi Sethi, J.D. Ullman, Compilers - Principles, Techniques and Tools, Addison- Wesley, 2003.
2. Fischer Leblanc, Crafting Compiler, Benjamin Cummings, Menlo Park, 1988.
3. Kennath C.Louden, Compiler Construction Principles and Practice, Vikas publishing House, 2004.
4. Allen I. Holub, Compiler Design in C, Prentice Hall of India, 2001.
5. S.Godfrey Winster, S.Aruna Devi, R.Sujatha, “Compiler Design”, yesdee Publishers, Third Reprint 2019.

COURSE- 19UPCSC3C05 **Credits: 4**

ALGORITHM - LAB

1. Apply the divide and conquer technique to arrange a set of numbers using merge sort.
2. Apply the divide and conquer technique to implement Strassen‟s matrix multiplication Algorithm
3. Computer the transitive closure of a given directed graph using Warshall‟s Algorithm.
4. Implement 0/1 knapsack problem using Dynamic programming
5. Find minimum cost spanning tree of a given undirected graph using Kruskal‟s Algorithm.
6. Find minimum cost spanning Tree of a given undirected graph using Prim‟s Algorithm.
7. Implement 8 Queen‟s problem using backtracking
8. Implement Knapsack problem using backtracking
9. Solve Dljkstra‟s Algorithm using greedy technique
10. Solve subset sum problem using backtracking

COURSE- 06 **Credits: 2**

ADVANCED WEB TECHNOLOGY - LAB

Course Objective:

- Learn how to create websites using Asp.Net
- Implement the advanced web concepts using Asp.Net and Ado.Net
- Learn to use Asp.Net web controls
- Design web applications using Asp.Net and Ado.Net

List of Programs:

Programs using ASP.NET Server controls

1. Create a website for a bank and include types of navigation.
2. Write a program to perform Asp.Net State.
3. Design Image Mapping using Asp.Net
4. Create the following using web controls
   a) Money conversion
   b) Temperature conversion
5. Write a program to create an advertisement using Ad rotator.
6. Create a user control that contains a list of colors. Add a button to the Web Form which when clicked changes the color of the Form to the color selected from the list.
7. Create a user control that displays the current date and time. Include it in a Web Form and refresh it each time a button is clicked.
8. Create a user control that receives the user name and password from the user and validates them. If the user name is “Radiant” and the password is “asp.net” then the user is authorized, otherwise not.

Programs using ADO.NET and ASP.NET

1. Create a web application to insert 3 records inside the SQL database table having following fields(DeptId, DeptName, EmpName, Salary). Update the salary for any one employee and increment it to 15% of the present salary. Perform delete operation on 1 row of the database table.
2. Create a Web App to display all the Empname and Deptid of the employee from the database using SQL source control and bind it to GridView. Database fields are(DeptId, DeptName, EmpName, Salary)

Course Outcomes:

- Apply .NET concepts to design and developweb applications
- Create a basic website using Asp.Net concepts
- Design web page and connect to the backend databases
- Applying different functionalities in Asp.Net and Ado.Net

COURSE- 19UPCSC3C07 **Credits: 4**

DISTRIBUTED OPERATING SYSTEM

Course Objective:

- To study core concepts of distributed operating system.
- To understand hardware, software and communication in distributed OS
- To learn the distributed resource management components.
- Practices to learn concepts of OS and Program the principles of Operating Systems

Unit I

Introduction – Operating System Definition – Functions of Operating System – Types of Advanced Operating System – Design Approaches – Synchronization Mechanisms – concepts of a Process – Critical Section Problem – Process Deadlock – Models of Deadlock – Conditions for Deadlock – System with single-unit requests, Consumable Resources , Reusable Resources.

Unit II

Distributed Operating Systems: Introduction- Issues – Communication Primitives – Inherent Limitations –Lamport‟s Logical Clock , Vector Clock, Global State , Cuts – Termination Detection – Distributed Mutual Exclusion – Non Token Based Algorithms – Lamport‟s Algorithm - Token Based Algorithms –Distributed Deadlock Detection – Distributed Deadlock Detection Algorithms – Agreement Protocols

Unit III

Distributed Resource Management – Distributed File Systems – Architecture – Mechanisms – Design Issues – Distributed shared Memory – Architecture – Algorithm – Protocols – Design Issues – Distributed Scheduling – Issues – Components – Algorithms.

Unit IV

Failure Recovery and Fault Tolerance – Concepts – Failure Classifications – Approaches to Recovery – Recovery in Concurrent Systems – Synchronous and Asynchronous Check pointing and Recovery –Check pointing in Distributed Database Systems – Fault Tolerance Issues – Two-Phase and Nonblocking Commit Protocols – Voting Protocols – Dynamic Voting Protocols.

Unit V

Multiprocessor and Database Operating Systems –Structures – Design Issues – Threads – Process Synchronization – Processor Scheduling – Memory management – Reliability/Fault Tolerance – Database Operating Systems – concepts – Features of Android OS, Ubuntu, Google Chrome OS and Linux operating systems.

Text Books

1. MukeshSinghalN.G.Shivaratri, "Advanced Concepts in Operating Systems", McGraw Hill 2000.

2. Distributed Operating System – Andrew S. Tanenbaum, PHI.

Reference Books

1. Abraham Silberschatz, Peter B.Galvin, G.Gagne, "Operating Concepts", 6th Edition Addison Wesley publications 2003.

2. Andrew S.Tanenbaum, "Modern Operating Systems", 2nd Edition Addison Wesley 2001

COURSE- 19UPCSC3C08 **Credits: 4**

ADVANCED JAVA PROGRAMMING

Course Objective

- To deepen student‟s programming skills by analyzing the real world problem and implement the concepts in real time projects
- To learn the ethical, historical, environmental and technological aspects of Advanced Java Programming and it impacts the social and economic development of society

Unit-I

Design Patterns: Introduction to Design patterns - Catalogue for Design Pattern - Factory Method Pattern, Prototype Pattern, Singleton Pattern- Adapter Pattern- Proxy Pattern-Decorator Pattern- Command Pattern- Template Pattern- Mediator Pattern-Collection Framework – Array List class – Linked List class – Array List vs. Linked List - List Iterator interface - Hash Set class- Linked Hash Set class-Tree Set class Priority Queue class - Map interface-Hash Map class- Linked Hash Map class –Tree Map class - Comparable interface - Comparator interface-Comparable vs. Comparator

Unit-II

Applet Fundamentals- Applet Class - Applet lifecycle- Steps for Developing Applet Programs- Passing Values through Parameters- Graphics in Applets- GUI Application - Dialog Boxes - Creating Windows - Layout Managers – AWT Component classes – Swing component classes- Borders – Event handling with AWT components - AWT Graphics classes - File Choosers - Color Choosers – Tree – Table –Tabbed panels–Progressive bar - Sliders.

Unit-III

JDBC -Introduction - JDBC Architecture - JDBC Classes and Interfaces – Database Access with MySQL -Steps in Developing JDBC application - Creating a New Database and Table with JDBC - Working with Database Metadata; Java Networking Basics of Networking - Networking in Java- Socket Program using TCP/IP - Socket Program using UDP- URL and Inet address classes.

Unit-IV

Servlet: Advantages over Applets - Servlet Alternatives - Servlet Strengths - Servlet Architecture - Servlet Life Cycle – Generic Servlet, Http Servlet - First Servlet - Invoking Servlet - Passing Parameters to Servlets - Retrieving Parameters - Server-Side Include – Cookies- JSP Engines - Working with JSP - JSP and Servlet - Anatomy of a JSP Page- Database Connectivity using Servlets and JSP.

Unit-V

Lambda Expressions- Method Reference- Functional Interface- Streams API, Filters- Optional Class- Nashorn- Base 64 Encode Decode- JShell(RPEL)- Collection Factory Methods- Private Interface Methods- Inner Class Diamond Operator- Multiresolution Image API.

Text Books

1. Bert Bates, Karthy Sierra , Eric Freeman, Elisabeth Robson, “Head First Design Patterns”, O‟REILLY Media Publishers.(1st-Unit).
2. Herbert Schildt, “Java: A Beginner Guide”, Oracle Pres-Seventh Edition. (2nd and 3rd Unit).
3. Murach‟s, “Java Servlets and JSP”, 2nd Edition, Mike Murach & Associates Publishers; 3rd Edition. (4th Unit).
4. Warburton Richard, “Java 8 Lambdas”, Shroff Publishers & Distributors Pvt Ltd. (5th Unit).

Reference Books

1. Paul Deitel and Harvey Deitel, “Java: How to Program”, Prentice Hall Publishers; 9th Edition.
2. Jan Graba, “An Introduction to Network Programming with Java-Java 7 Compatible”, 3rd Edition, Springer.

COURSE- 19UPCSC3C09 **Credits: 4**

CRYPTOGRAPHY AND NETWORK SECURITY

Course Objective:

- To understand Cryptography Theories, Algorithms and Systems.
- To understand necessary Approaches and Techniques to build protection mechanisms in order to secure computer networks.
- To know about the malicious software & firewalls.

Unit I

Introduction - Security trends – Legal - Ethical and Professional Aspects of Security - Need for Security at Multiple levels - Security Policies – Model of network security – Security attacks - services and mechanisms – OSI security architecture – Classical encryption techniques: substitution techniques - transposition techniques - steganography- Foundations of modern cryptography: perfect security – information theory – product cryptosystem – cryptanalysis.

Unit II

Symmetric Encryption and Message Confidentiality - Symmetric Encryption Principles - Symmetric Block Encryption Algorithms - Stream Ciphers and RC4 - Chipher Block Modes of Operation - Location of Encryption Devices - Key Distribution. Public-key Cryptography and Message Authentication: Approaches to Message Authentication - Secure Hash Functions and HMAC - Public-Key Cryptography Principles - Public-Key Cryptography Algorithms - Digital Signatures - Key Management.

Unit III

Authentication Applications - Kerberos - x.509 Authentication Service - Public-Key Infrastructure. Electronic Mail Security: Pretty Good Privacy (PGP) - S/MIME.

Unit IV

IP Security - IP Security Over view - IP Security Architecture - Authentication Header - Encapsulating Security Payload - Combining Security Associations. Web Security: Web Security Considerations - Secure Socket Layer(SSL) and Transport Layer Security(TLS) - Secure Electronic Transaction(SET).Network Management Security: Basic Concepts of SNMP - SNMPv1 Community Facility - SNMPv3.

Unit V

Intruders - Intruders - Intrusion Detection - Password Management. Malicious Software: Virus and Related Threats - Virus Countermeasures - Distributed Denial of Service Attacks. Firewalls: Firewall Design Principles - Trusted Systems - Common Criteria for Information Technology Security Evaluation.

Text Books

1. Behrouz A. Ferouzan - “Cryptography & Network Security” - Tata Mc Graw Hill - 2007 - Reprint 2015.
2. Stallings William - “Cryptography and Network Security - Principles and Practice 2017.
3. William Stallings - “Network Security Essentials Applications and Standards ”Third Edition - Pearson Education - 2008.

Reference Books

1. Man Young Rhee - “Internet Security: Cryptographic Principles” - “Algorithms And Protocols” - Wiley Publications - 2003.
2. Charles Pfleeger - “Security In Computing” - 4th Edition - Prentice Hall Of India - 2006.
3. Ulysess Black - “Internet Security Protocols” - Pearson Education Asia - 2000.
4. Charlie Kaufman And Radia Perlman - Mike Speciner - “Network Security - Second Edition - Private Communication In Public World” - PHI 2002.

5. Bruce Schneier And Neils Ferguson - “Practical Cryptography” - First Edition - Wiley Dreamtech India Pvt Ltd - 2003.
6. Douglas R Simson “Cryptography – Theory And Practice” - First Edition - CRC Press - 1995.
7. Http://Nptel.Ac.In/.

COURSE- 19UPCSC3C10 **Credits: 4**

ADVANCED JAVA LAB

Course Objective:

- To enable the students to implement different java packages.
- To develop the students with the skills to implement different java tools.

Use JAVA Programming Language to implement the following:

1. To create applets incorporating the following Features:
   a. Create a color palette with matrix of buttons
   b. Set background and foreground of the control text area by selecting a color from color palette.
   c. In order to select Foreground or background use check box control as radio buttons
   d. To set background images
2. Use GridLayout to design a calculator and simulate the functions of simple calculator.
3. To create Input output and Random files
4. To develop chat application with datagram sockets and datagram packets.
5. To invoke servlet from HTML forms.
6. To invoke servlet from Applets.
7. To invoke servlet from JSP.
8. Simple client/server application.
9. JDBC to interact with database.
10. To create multiple chat applications using TCP packets.

DATA ANALYTICS USING R - LAB

Course: 19UPCSC3E01 **Credits: 2**

Objectives:

- To implement mathematical aggregation operators in "R-script".
- To understand the Statistical operations in "R".

1. To get the input from user and perform numerical operations (MAX, MIN, AVG, SUM, SQRT, ROUND) using in R.

2. To perform data import/export (.CSV, .XLS, .TXT) operations using data frames in R.

3. To get the input matrix from user and perform Matrix addition, subtraction, multiplication, inverse transpose and division operations using vector concept in R.

4. To perform statistical operations (Mean, Median, Mode and Standard deviation) using R.

5. To perform data pre-processing operations i) Handling Missing data ii) Min-Max normalization

6. To perform dimensionality reduction operation using PCA for Houses Data Set

7. To perform Simple Linear Regression with R.

8. To perform K-Means clustering operation and visualize for iris data set

9. Write R script to diagnose any disease using KNN classification and plot the results.

10. To perform market basket analysis using Association Rules (Apriori).

ELECTIVE - 01 LIST

COURSE: 19UPCSC3E01 **Credits: 4**

DISCRETE MATHEMATICS FOR COMPUTING

(Theorems and Proofs are not expected)

Course objective

- To understand the applications of functions and relations
- To understand the basic concepts of mathematical logic and predicate calculus
- To understand the concept of method of induction
- To develop the skills in solving recurrence relations.

Unit-I

Wellformed formulas – truth table of well formed formula – tautology, contradiction and contingency –equivalence of formulas. Algebra of propositions – Functionality complete sets – Normal forms of well formed formulas- Rules of Inference for propositional calculus – well formed formulas of predicate calculus – Rules of Inference for predicate calculus – predicate formulas involving two or more quantifiers.

Unit-II

Set theory – relations- functions – set identities – Binary relations – properties of binary relations in a set – Equivalence relations and partial orderings – Representation of a relation by a matrix - presentation of a relation by a digraph - Basics of Counting – Integers and Induction.

Unit-III

Formulation as Recurrence Relations-solving recurrence Relation by Iteration-solving Recurrence Relations- Solving Linear Homogeneous Recurrence Relations of Order Two-Solving Linear Non homogeneous Recurrence Relations. Permutations-Combinations-Permutations with repetitions-Combinations with repetition-permutations of sets with indistinguishable objects.

Unit-IV

Definition and examples-properties of lattices –lattices as algebraic systems-Sub lattices and lattice Isomorphism-special classes of lattice –distributive lattices and Boolean algebras.

Unit-V

Connected Graphs-Euler Graphs-Hamiltonian circuits and paths – planar graphs – matrix representation of graphs.

Course outcome

On successful completion of the course, students will

- solve problems in mathematical logic
- identify and apply basic concepts of set theory, arithmetic, logic, proof techniques, binary relations, graphs and trees
- solve recurrence relations
- construct lattice applications
- understand the applications of Graph Theory in Computer Science.

Text Book

1. N.Chandrasekaran and M.Umaparvathi, Discrete mathematics, PHI Learning Private Limited, New Delhi, 2010.

Unit 1:(2.1-2.11) **Unit 2:**(1.3-1.7, 4.1-4.2, 5.1-5.5)**Unit 3:**(6.1-6.5,3.1-3.6)

Unit 4:(8.1-8.6) **Unit 5:**(10.1-10.5 and 10.8)

Reference Books

1. J.P.Trembley and R.Manohar, Discrete Mathematical Structures with Applications to Computer Science, Tata McGraw Hill, New Delhi, 1997.
2. T. Sengadir, Discrete Mathematics and Combinatorics, Pearson New Delhi 2009.
3. RakeshDube,AdeshPandeyRitu Gupta, Discrete Structures and Automata Theory, Narosa publishing House New Delhi 2007.

STATISTICAL COMPUTING

COURSE 19UPCSC3E02 **Credits: 4**

Course Objective:

- To understand the applications of various correlation methods
- To study and model the sampling concepts
- To acquire knowledge on Hypotheses test

Unit-I

Correlation - Definition of Correlation- Scatter Diagram- Kari Pearson‟s Coefficient of Linear Correlation- Coefficient of Correlation and Probable Error of r- Coefficient of Determination - Merits and Limitations of Coefficient of Correlation- Spearman‟s Rank Correlation.

Unit-II

Regression Analysis - Regression and Correlation(Intro)- Difference between Correlation and Regression Analysis- Linear Regression Equations -Least Square Method- Regression Lines- Properties of Regression Coefficients- Standard Error of Estimate.

Unit-III

Probability Distribution and mathematical Expectation- Random Variable- Defined - Probability Distribution a Random Variable- Expectation of Random Variable- Properties of Expected Value and Variance.

Unit-IV

Sampling and Sampling Distributions - Data Collection- Sampling and Non-Sampling Errors – Principles of Sampling-- Merits and Limitations of Sampling- Methods of Sampling- Parameter and Statistic- Sampling Distribution of a Statistic- Examples of Sampling Distributions- Standard Normal - Student‟s t - Chi-Square (x2) and Snedecor‟s F- Distributions.

Unit-V

Statistical Inference- Estimation and Testing of Hypothesis - Statistical Inference- Estimation- Point and interval- Confidence interval using normal - t and x2Distributions- Testing of Hypothesis- Significance of a mean - Using t Distribution.

Text Book

1. K.L. Sehgal - “Quantitative Techniques and Statistics” - First Edition - Himalaya Publishing House - 2011.

Reference Books

1. N. P. Bali - P. N. Gupta - C. P. Gandhi - “A Textbook of Quantitative Techniques” - First Edition - Laxmi Publications - 2008.
2. U. K. Srivastava - G. V. Shenoy - S. C. Sharma - “Quantitative Techniques for Managerial Decisions” - Second Edition - New Age International Publishers - 2005.
3. David Makinson - “Sets - Logic and Maths for Computing” - Springer - 2011.
4. Christopher Chatfield -”Statistics for Technology- A Course in Applied Statistics - Third Edition” - CRC Press - 2015.

COURSE - 19UPCSC3E03 **Credits: 4**

FUNDAMENTALS OF DATA SCIENCE

Course Objective:

- To learn about the fundamental science of data and concepts of big data
- To learn the lifecycle of the data analytics and data analytics using R
- To learn advanced data analytics methods and big data analytics

Unit - I

Introduction to Big Data Science- Definition of Big Data - Big data characteristics & considerations - Data repositories- analyst perspective - Business drivers for analytics - Typical analytical architecture - Business Intelligence Vs Data science - Drivers of Big data analytics - Role of data scientist in Big data ecosystem - Applications of Big data analytics.

Unit - II

Data Analytics Lifecycle -Need of Data analytic lifecycle - Key roles for successful analytic projects - various phases of Data analytic lifecycle: Discovery - Data Preparation - Model Planning - Model Building - Communicating Results - Operationalization.

Unit - III

Basic Data Analytics methods using R: Introduction to R: GUI of R - Getting data into & out of R - Data types in R - Basic operations - Basic statistics - Generic functions - Data visualization using R - Data exploration & presentation - Statistics for model building & evaluation.

Unit - IV

Advanced Analytics- Theory & Methods: Clustering - Association Rules - Apriori algorithm - Linear Regression - Logistics Regression - Naïve Bayesian classifiers - Decision Trees

Unit - V

Advanced Analytics –Big Data: Time series analysis - Text analysis -Technology and Tools: MapReduce and Hadoop- Communicating and Operationalizing an Analytics Project - Creating the Final Deliverables

Text Books

1. David Dietrich - Barry Hiller - “Data Science & Big Data Analytics” - EMC education services - Wiley publications - 2012
2. Trevor Hastie - Robert Tibshirani - Jerome Friedman - "The Elements of Statistical Learning" - Springer - Second Edition - 2011.

Reference Books

1. Mark gardner - “Beginning R: The Statistical Programming Language” - Wrox Publication
2. Adam Fowler - “NoSQL For Dummies” - John Wiley & Sons - ISBN-1118905628

COURSE - 19UPCSC3E04 **Credits: 4**

Data Visualization Techniques

Course Objective:

- To develop skills to both design and critique visualization
- To understand why visualization is an important part of data analysis
- To understand the components involved in visualization design
- To understand the type of data impacts the type of visualization

Unit-I

Data Preparation :Importing Data - Text files -Excel spreadsheets -Statistical packages - Databases - Cleaning Data : Selecting variables - Selecting observations - Creating/Recoding variables - Summarizing data - Using pipes - Reshaping data - Missing data - Introduction to ggplot2 -ggplot- geoms - grouping scales - facets -labels- themes - Placing the data and mapping options- Graphs as objects

Unit-II

Univariate Graphs - Categorical : Bar Chart -Pie Chart - Tree Map- Quantitative - Histogram - Kernel Density plot - Dot Chart - Bivariate Graphs - Categorical vs. Categorical : Stacked bar chart - Grouped bar chart - Segmented bar chart - Improving the color and labeling - Other plots - Quantitative vs. Quantitative :Scatterplot - Line plot- Categorical vs. Quantitative: Bar chart (on summary statistics) - Grouped kernel density plots - Box plots -Violin plots -Ridgeline plots - Mean/SEM plots - Strip plots - Beeswarm Plots -Cleveland Dot Charts - Multivariate Graphs - Grouping - Faceting

Unit- III

Maps: Dot density maps-Choropleth maps:Data by country-Data by US state-Data by US county -Time-dependent graphs: Time series- Dummbbell charts - Slope graphs - Area Charts - Statistical Models : Correlation plots - Linear Regression - Logistic regression - Survival plots - Mosaic plots

Unit- IV

3-D Scatterplot : Biplots - Bubble charts - Flow diagrams -Sankey diagrams- Alluvial diagrams - Heatmaps - Radar charts - Scatterplot matrix - Waterfall charts- Word clouds -Customizing Graphs - Axes: Quantitative axes - Categorical axes - Date axes- Colors: Specifying colors manually-Color palettes: Points & Lines: Points - Lines - Fonts - Legends: Legend location- Legend title - Labels- Annotations: Adding text - Adding lines - Highlighting a single group - Themes- Altering theme elements - Pre-packaged themes

Unit- V

Saving Graphs : Via menus - Via code - File formats -External editing - Interactive Graphs - leaflet - plotly -rbokeh - rCharts - highcharter- Best Practices: Labeling - Signal to noise ratio - Color choice- y-Axis scaling - Attribution

Text Book

1. Rob Kabacoff , Data Visualization with R, Bookdown, 2018.

   Chapters: 1-13 https://rkabacoff.github.io/datavis/

Reference Book

1. Kirthi Raman - Mastering Python Data Visualization -Packt Publishing -

   2015

ELECTIVE-02 LIST

COURSE - 19UPCSC3E05 **Credits: 4**

PREDICTIVE ANALYTICS

Course Objective:

- To learn fundamentals of data, predictive analytics methods and techniques.
- To understand given problem and analyze the results efficiently

Unit -I

Introduction- Entering the Arena: Exploring Predictive Analytics-Adding Business Value-Starting a Predictive Analytics Project-Surveying the Marketplace: Predictive Analytics in the Wild-Online Marketing and Retail-Implementing a Recommender System-Exploring Your Data Types and Associated Techniques: Recognizing Your Data Types-Identifying Data Categories-Generating Predictive Analytics-Connecting to Related Disciplines.

Unit- II

Complexities of Data: Finding Value in Your Data-Constantly Changing Data-Complexities in Searching Your Data-Differentiating Business Intelligence from Big-Data Analytics-Visualization of Raw Data-Incorporating Algorithms in Your Models: Applying Models-Modeling Data-Healthcare Analytics Case Studies - Social and Marketing Analytics Case Studies.

Unit- III

Identifying Similarities in Data: Explaining Data Clustering - Converting Raw Data into a Matrix-Identifying K-Groups in Your Data-Finding Associations Among Data Items-Applying Biologically Inspired Clustering Techniques-Predicting the Future Using Data Classification: Introducing Data Classification to Your Business-Exploring the Data-Classification Process-Using Data Classification to Predict the Future: Decision trees, Support vector machine, Naïve Bayes classification algorithm, Neural networks, The Markov Model, Linear regression.

Unit- IV

Developing a Roadmap: Convincing Your Management to Adopt Predictive Analytics-Preparing Data: Listing the Business Objectives -Processing Your Data-Structuring Your Data- Building a Predictive Model: Getting Started-Developing and Testing the Model-Going Live with the Model- Visualization of Analytical Results: Visualization As a Predictive Tool-Evaluating Your Visualization-Visualizing Your Model‟s Analytical Results.

Unit-V

The Part of Tens-Ten Reasons to Implement Predictive Analytics: Outlining Business Goals-Knowing Your Data-Organizing Your Data-Satisfying Your Customers-Reducing Operational Costs-Increasing Returns on Investments (ROI)-Increasing Confidence-Making Informed Decisions-Gaining Competitive Edge-Improving the Business-Ten Steps to Build a Predictive Analytic Model-Building a Predictive Analytics Team-Setting the Business Objectives-Preparing Your Data-Sampling Your Data-Avoiding “Garbage In, Garbage Out”-Creating Quick Victories-Fostering Change in Your Organization-Building Deployable Models-Evaluating Your Model-Updating Your Model.

Text book

1. Anasse Bari, Mohammad Chaouchi, Tommy Jung, Predictive Analytics For Dummies, 2nd Edition, 2017

COURSE - 19UPCSC3E06 **Credits: 4**

TEXT ANALYTICS

Course Objective:

- To interpret the basics of text analysis.
- To infer about text mining - text analytics and web analytics.
- To illustrate the domains that makes up text analytics and web analytics.

Unit - I

History of Text Mining 6: Roots of text mining - Information extraction and text mining - Development of enabling technology in text mining - Sentiment analysis and opinion mining.

Unit - II

Basics of Text Analytics 6: Definition - Business challenges addressed: information organization and access - Discovery of patterns – Discovery.

Unit - III

Seven Practice Areas of Text Analytics 6: Seven practice areas of text analytics - Finding the appropriate solution to a problem - Overall relationship - Visualizing the domains of text analytics.

Unit - IV

Web Analytics and Web Mining 9: Value of web analytics - Components of web mining - Concepts and terminology in web analytics - Web analytics and web mining - Optimal paths to successful web analytics evolution in a company.

Unit - V

Future of Text and Web Analytics 9: Text analytics and text mining - Future of web analytics - Future of text mining - Integration of web analytics with standard business intelligence tools – New areas that may use text analytics.

Text Books

1. Gary Miner John Elder IV - Robert Nisbet - Dursun Delen - Thomas Hill - Andrew Fast - "Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications" -1st Edition - Academic Press - 2012.
2. Brian Clifton Sybex - "Advanced Web Metrics with Google Analytics" - 3rd Edition - Pearson education - 2012.

COURSE - 19UPCSC3E07 **Credits: 4**

WEB ANALYTICS

Course Objective:

- To learn effective problem solving methodologies in Computing applications
- To provide overview and establish the need for web analytics.
- To understand and apply metrics to analyze the web data.
- To provide exposure to usage of web analytic tools.

Unit - I

Introduction to Web Analytics 7: A Brief history of Web Analytics –Web Analytics Terminology – Traditional Web Analytics – Web Analytics 2.0 – Capturing Data- Tools Selection – Quality Aspects –Implementing Best Practices.

Unit - II

Web Data Collection 9: Web Traffic Data – Web Transactional Data – Web Server Data – Page Weights– Usability Studies – User Submitted Information – Integrating Form based data – Web Data Sources – Server Log Files – Page Tags – Clickstream Data –Outcomes Data – Research Data –Competitive Data.

Unit - III

Web Analytics Strategy 7: Component of Web Analytics Strategy – Customer Centric Focus – Business Problem Solving Focus – Reporting vs Analysis – IT and Business Strength – Clickstream vs Web 2.0 – Vendor Specific Options and Issues.

Unit - IV

Metrics and KPIs 7: Measuring Reach – Measuring Acquisition – Measuring Conversion – Measuring Retention – Focus on „Critical Few"- Key Performance Indicators – Case Studies.

Unit - V

Data Analysis: Customer centricity – Lab Usability Studies – Usability Alternatives – Surveys – Heuristic Evaluations - Web enabled user research options – Competitive Intelligence Analysis Content organization tool – Process measurement tools- Visitor Segmentation Tools- Campaign Analysis – Commerce Measurement Tools -Google Analytics – Piwik Web Analytics – Yahoo Web Analytics – Emerging Analytics: Social - Video - Mobile.

Reference Books

1. Avinash Kaushik - “Web Analytics 2.0: The Art of Online Accountability and Science of Customer Centricity” - 1st Edition - Sybex - 2009.
2. Brian Clifton - “Advanced Web Metrics with Google Analytics” - 3rd Edition - Sybex - 2012.
3. Eric Peterson - “Web Analytics Demystified:A Marketer's Guide to Understanding How Your Web Site Affects Your Business” -1st Edition - Celilo Group Media - 2004.
4. Avinash Kaushik - “Web Analytics: An Hour a Day” - 6th Edition - Sybex - PAP/ CDR Edition - 2007.
5. Justin Cutroni - “Google Analytics” - 2nd Edition - O'Reilly Media - 2010.

COURSE- 19UPCSC3E08 **Credits: 4**

MARKET ANALYTICS

Course Objective:

- To learn effective problem solving methodologies in Computing applications
- To introduce the principles and strategic concepts of marketing analytics.
- To understand cost concepts (TOTAL HOURS - fixed - variable) - profit margins - and lifetime value of the customer.
- To get an overview of the benefits and objectives of quantitative marketing.

Unit - I

Introduction: Introduction to Marketing Analytics – Market Insight – Market sizing and trend analysis.

Unit - II

Market Segmentation: Market segmentation – Segment identification - analysis - and strategy - Competitive analysis- Competitor identification - analysis and strategy.

Unit - III

Business Strategy and Operations: Business Strategy - Analytics-based strategy selection - Business Operations - Forecasting - predictive analytics - and data mining.

Unit - IV

Product - Service and Price Analytics: Product and Service Analytics - Conjoint analysis and product/service metrics - Price Analytics - Pricing techniques and assessment.

Unit - V

Distribution and Promotion Analytics: Distribution Analytics - Analytics-based channel evaluation and selection - Promotion Analytics - Promotion budget estimation and allocation. Sales Analytics and Analytics in Action: Sales Analytics - Metrics for sales - profitability - and support- Analytics in Action - Pivot tables and data-driven presentations.

Text Books

1. Stephan Sorger - "Marketing Analytics: Strategic Models and Metrics" - 1st Edition - Create Space Independent Publishing Platform - 31-Jan-2013.
2. Stephan Sorger - "Marketing Planning: Where Strategy Meets Action" - 1st Edition - Prentice Hall PTR - 03-Sep-2011.
3. Cesar A.Brea - "Pragmalytics : Practical approaches to the Marketing analytics in the Digital Age" -1st Edition - iUniverse - 2012.

ELECTIVE – 03 LIST

COURSE – 19UPCSC3E09 **Credits: 4**

OPERATIONS RESEARCH

Course Objective

- To understand the concept of optimization
- To develop mathematical model of real life cases
- To study Optimization algorithms

Unit – I

Linear Programming Problem (LPP): Formulations and graphical solution of (2 variables) canonical and standard terms of linear programming problem.

Unit – II

Algebraic Solution: Simplex algorithm - Simplex methods – solving problems with slack variable

Unit – III

Transportation Model: North West corner Method - Least cost method - and vogel‟s approximation method. Assignment Model : Hungarian assignment model – Travelling sales man problem.

Unit – IV

Replacement Problem: Replacement policy for equipment that deteriorate gradually - Replacement of item that fail suddenly-Individual and group replacement - Problems in mortality and staffing.

Unit – V

Project Scheduling PERT/CPMNetworks – Fulkerson‟s Rule – Measure Of Activity – PERT Computation – CPMComputation – Resource Scheduling.

Text Books

1. KantiSwarup - P.K. Gupta &Manmohan – Operation Research 1996.
2. S.Kalavathy: Operations Research – Second Edition – Vikas Publishing House Pvt.Ltd. - 2

ELECTIVE- 19UPCSC3E10 **Credits: 4**

OPTIMIZATION TECHNIQUES

Course Objective:

- To learn effective problem solving methodologies in Computing applications

Unit - I

Introduction: Fundamentals of optimization - Principles of Evolutionary Processes and Genetics - History of Evolutionary Computation.

Unit – II

Common Methods: Genetic algorithms - Evolution strategies - Evolution programming - Genetic programming - Classifier systems - Swarm Intelligence.

Unit – III

Basic Data Structures and Operators: Genetic Representation - Search operators - Selection schemes - Selection pressure and its impact on evolutionary search. Advanced Features: Constraint handling - Co-evolution and speciation - Parameter Adaptation and Control.

Unit – IV

Evolutionary Computation Approaches: Evolutionary multi-objective optimization - Parallel evolutionary algorithms - Dynamic optimization - Hybrid evolutionary algorithms.

Unit - V

Applications: Knapsack problem - Minimum spanning tree - Travelling Salesman Problem - Scheduling - Robotics and automation - Inventory Control.

Reference Books

1. Xinjie Yu and Mitsuo Gen - ”Introduction to Evolutionary Algorithms” - Springer - New York - 2010.
2. Jun Sun - Choi-Hong Lai and Xiao-Jun Wu - “Particle Swarm Optimization: Classical and Quantum Perspectives” - Taylor and Francis - USA - 2012.
3. Carlos A Coello Coello - Gary B Lamont and David A Van Veldhuizen - “Evolutionary Algorithms for Solving Multi-Objective Problems” - Springer - New York - 2007.
4. Back T - D. B. Fogel and Z. Michalewicz - “Evolutionary Computation 1: Basic Algorithms and Operators” - Taylor and Francis Group - New York - 2000.
5. Back T - D. B. Fogel and Z. Michalewicz - “Evolutionary Computation 2: Advanced Algorithms and Operators” - Taylor and Francis Group - New York - 2000.
6. Thomas Back - “Evolution Strategies - Evolutionary Programming Genetic Algorithms” - Oxford University Press - New York - 1996.

ELECTIVE- 19UPCSC3E11 **Credits: 4**

Probability and Stochastic Process

Course Objective:

- To learn the basic methodology of "probabilistic thinking" and to apply it to probabilistic problems in signal processing and Communication Engineering;
- To understand basic concepts of probability theory - random variables - Conditional probability and conditional expectation - joint distribution and independence - mean square estimation.
- To understand the difference between time averages and statistical averages

Unit – I

Probability: Probability introduced through Sets and Relative Frequency - Experiments and Sample Spaces - Discrete and Continuous Sample Spaces - Events - Probability Definitions and Axioms - Mathematical Model of Experiments - Probability as a Relative Frequency - Joint Probability - Conditional Probability - Total Probability - Baye‟s Theorem - Independent Events. Random Variable: Definition of a Random Variable - Conditions for a Function to be a Random Variable - Discrete - Continuous - and Mixed Random Variables

Unit – II

Distribution & Density Functions: Distribution and Density functions and their Properties – Binomial - Poisson - Uniform - Gaussian - Exponential - Rayleigh and Conditional Distribution - Methods of defining Conditional Event - Conditional Density - and Properties. Operation on One Random Variable – Expectations: Introduction - Expected Value of a Random Variable - Function of a Random Variable - Moments about the Origin - Central Moments - Variance and Skew - Chebychev‟s Inequality - Characteristic Function - Moment Generating Function - Transformations of a Random Variable: Monotonic Transformations for a Continuous Random Variable - Non-

monotonic Transformations of Continuous Random Variable - Transformation of a Discrete Random Variable.

Unit – III

Multiple Random Variables: Vector Random Variables - Joint Distribution Function - Properties of Joint Distribution - Marginal Distribution Functions - Conditional Distribution and Density – Point Conditioning - Conditional Distribution and Density – Interval conditioning - Statistical Independence - Sum of Two Random Variables - Sum of Several Random Variables - Central Limit Theorem (Proof not expected) - Unequal Distribution - Equal Distributions. Operations on Multiple Random Variables: Expected Value of a Function of Random Variables: Joint Moments about the Origin - Joint Central Moments - Joint Characteristic Functions - Jointly Gaussian Random Variables: Two Random Variables case - N Random Variable case - Properties - Transformations of Multiple Random Variables - Linear Transformations of Gaussian Random Variables.

Unit – IV

Stochastic Processes – Temporal Characteristics: The Stochastic Process Concept - Classification of Processes - Deterministic and Nondeterministic Processes - Distribution and Density Functions - Concept of Stationarity and Statistical Independence - First-Order Stationary Processes - Second-Order and Wide-Sense Stationarity - Nth Order and Strict-Sense Stationarity - Time Averages and Ergodicity - Mean Ergodic Processes - Correlation-Ergodic Processes - Autocorrelation Function and its Properties - Cross-Correlation Function and its Properties - Covariance and its Properties - Linear System Response of Mean and Mean-squared Value - Autocorrelation Function - Cross Correlation Functions - Gaussian Random Processes - Poisson Random Process.

Unit – V

Stochastic Processes – Spectral Characteristics: Power Spectrum: Properties - Relationship between Power Spectrum and Autocorrelation Function - Cross-Power Density Spectrum - Properties - Relationship between Cross-Power Spectrum and Cross-Correlation Function - Spectral Characteristics of System Response: Power Density Spectrum of Response - Cross-Power Spectral Density of Input and Output of a Linear System.

Text books

1. Probability - Random Variables & Random Signal Principles – Peyton Z. Peebles - 4Ed. - 2001 - TMH.

2. Probability and Random Processes – Scott Miller - Donald Childers - 2 Ed - Elsevier - 2012.

Reference Books

1. Probability - Random Variables and Stochastic Processes – Athanasios Papoulis and S. Unnikrishna Pillai - 4 Ed. - TMH.

2. Theory of Probability and Stochastic Processes- Pradip Kumar Gosh - UniversityPress

3. Probability and Random Processes with Application to Signal Processing – Henry Stark and John W. Woods - 3 Ed. - PE

4. Probability Methods of Signal and System Analysis – George R. Cooper - Clave D. MC Gillem - 3 Ed. - 1999 - Oxford.

5. Statistical Theory of Communication – S.P. Eugene Xavier - 1997 - New Age Publications.

COURSE - 19UPCSC3E12 **CREDITS- 4**

LINEAR PROGRAMMING

Course Objective:

- To learn effective problem solving methodologies in Computing applications
- To introduce methods of optimization to engineering students.
- To know numerous application in civil - environmental - electrical (control) engineering - and industrial engineering.
- To maintain a balance between theory - numerical computation - problem setup for solution by optimization software - and applications to engineering systems.

Unit - I

Linear Programming: Standard form of a linear programming problem – Geometry of linear programming problems – Definitions and theorems – Solution of a system of linear simultaneous equations – Pivotal reduction of a general system of equations – Motivation to the simplex method – Simplex algorithm.

Unit - II

Non-Linear Programming: Unconstrained Non-Linear Programming- 1-dimensional minimization methods- Classification - Fibonacci method and Quadratic interpolation method -Constrained Non-Linear Programming : Characteristics of a constrained problem – Classification - Basic approach of Penalty Function method- Basic approaches of Interior and Exterior penalty function methods- Introduction to convex Programming Problem.

Unit - III

Transportation Problem: Finding initial basic feasible solution by north – West corner rule - least cost method and Vogel‟s approximation method – Testing for optimality of balanced transportation problems.

Unit - IV

Unconstrained Optimization Techniques: Analytical method -Newton‟s method- Golden-section search method - Univariate method - Powell‟s method - Steepest descent method.

Unit - V

Dynamic Programming: Dynamic programming multistage decision processes – Types – Concept of sub optimization and the principle of optimality – Computational procedure in dynamic programming – Examples illustrating the calculus method of solution - Examples illustrating the tabular method of solution.

Text Books

1. S.Rao - “Engineering optimization: Theory and practice” - 4th Edition – New Age International - 2009.
2. H.S. Kasene & K.D. Kumar - “Introductory Operations Research” -3rd Edition -Springer India Pvt .LTd. - 2002.
3. H.A. Taha - “Operations Research: An Introduction” - 6th Edition - PHI Pvt. Ltd - 2004.

COURSE - 19UPCSC3E13 **CREDITS- 4**

Big Data Analytics

Course Objective:

- To understand the basic concepts of big data
- To learn Hadoop, MapReduce, Hive, HBase and Pig

Unit - I Fundamentals of Big Data

Understanding Big Data: Concepts and Terminology – Big Data Characteristics – Types of Data – Case Study Background – Drivers for Big Data Adoption: Information and Communication Technology – Big Data Analytics Lifecycle

Unit - II Fundamentals of Hadoop

Core components of Hadoop- Apache Hadoop – HDFS Daemons – MapReduce Daemons – HDFS High Availability Daemons – Benefits and Challenges of HDFS – File Sizes, Block Sizes and Block Abstraction in HDFS – Data Replication – How does HDFS Store, Read, and Write Files? – Data Serialization Options – File System Shell Commands for HDFS

Unit - III HDFS and MapReduce

Choosing Key and Value Types for MapReduce Jobs – The Relationship of Input Keys to Output Keys – Sorting Keys and Values – Sort and Shuffle Process – MapReduce Job Configuration and Submission

Hadoop Distributed File System – MapReduce Framework – Setting the Environment – Hadoop Cluster Modes – Running a MapReduce Job with the MR1Framework - Running a MapReduce Job with the Yarn Framework – Running Hadoop Streaming

Unit - IV Hive and HBase

Apache Hive: Setting the Environment – Configuring Hadoop, Hive – Starting HDFS, Hive Server, CLI – Creating and Using a Database– Creating a Managed Table – Loading data into a Table – Creating a Table using LIKE – Adding Data into a Table from Queries – Adding Data using INSERT INTO TABLE - Adding Data using INSERT OVERWRITE – Creating a table using CREATE TABLE AS SELECT – Altering, Truncating and Dropping a Table– Creating an External

Table – Apache HBase: Setting the Environment - Configuring Hadoop, Hive and HBase – Starting the HBase and HBase Shell – Creating HBase Table – Adding Data to a Table – Listing all Tables – Getting a Row of Data – Scanning a Table – Counting the Number of Rows in a Table – Altering a Table – Deleting a Table Row, Column – Disabling and Enabling a Table – Truncating and Dropping a Table – Determining If Table Exists – Creating a Hive External Table stored by HBase

Unit - V Pig

Introduction – Installing and Running Pig – Grunt – Pig‟s Data Model – Introduction to Pig Latin – Advanced Pig Latin – Developing and Testing Pig Latin Scripts – Making Pig Fly – Writing Evaluation and Filter Functions – Writing and Loading Store Function

Text Books

1. Alan Gates, “Programming Pig”, Oreilly Publication, 2011.
2. Deepak Vohra, “Practical Hadoop Ecosystem: A Definitive Guide to Hadoop-Related Frameworks and Tools”, Apress, 2016.
3. Thomas Erl, Wajid Khattak, Paul Buhler, “Big Data Fundamentals Concepts, Drivers & Techniques”, Service Tech Press, 2015.

Reference Books

1. Noreen Burlingame , “The little book on Big Data”, New Street publishers, 2012.
2. Anil Maheshwari, “ Data Analytics”, McGraw Hill Education, 2017.

COURSE - 19UPCSC3E14 **CREDITS- 4**

INFORMATION RETRIEVAL

Course Objective:

- To understand the basics of information retrieval process.
- To evaluate the performance of IR system.
- To explore information sharing on semantic web.
- To understand the various applications of Information Retrieval giving emphasis to multimedia and distributed IR - web Search.

Unit - I

Introduction: Basic Concepts of IR - Data Retrieval & Information Retrieval - IR system block diagram. Automatic Text Analysis: Luhn's ideas - Conflation Algorithm - Indexing and Index Term Weighing - Probabilistic Indexing - Automatic Classification. Measures of Association - Different Matching Coefficient - Classification Methods - Cluster Hypothesis - Clustering Algorithms - Single Pass Algorithm - Single Link Algorithm - Rocchio's Algorithm.

Unit - II

Storage and Searching Techniques: Storage: Inverted file - Suffix trees & suffix arrays - Signature Files - Scatter storage or hash addressing - Clustered files.IR Models: Basic concepts - Boolean Model - Vector Model Searching strategies: Boolean Search - Serial search - cluster based retrieval - Query languages - Types of queries - Patterns matching - structural queries.

Unit - III

Retrieval Performance Evaluation and Ontology: Performance evaluation: Precision and recall - alternative measures Ontology: Ontology based information sharing - Ontology languages for semantic web - Ontology creation.

Unit - IV

Distributed and Multimedia IR: Distributed IR: Introduction - Collection Partitioning - Source Selection - Query Processing - web issues. MULTIMEDIA IR: Introduction - Data Modeling - Query languages - Generic multimedia indexing approach - One dimensional time series - two dimensional color images - Automatic feature extraction.

Unit - V

Web Searching and Web Recommendation: Searching the Web: Challenges - Characterizing the Web - Search Engines - Browsing - Meta-searchers - Web crawlers - Meta-crawler - Web data mining - Finding needle in the Haystack - Searching using Hyperlinks - Page ranking algorithms. Collaborative Filtering and Content Based Recommendation of Documents and Products - Information Extraction and Integration: Extracting Data from Text. Semantic Web - Collecting and Integrating Specialized Information on the web.

Text Books

1. Yates & Neto - "Modern Information Retrieval" - Pearson Education - ISBN 81-297-0274-6.
2. C.J. Rijsbergen - "Information Retrieval" - (www.dcs.gla.ac.uk).
3. Heiner Stuckenschmidt - Frank van Harmelen - “Information Sharing on the Semantic Web” - Springer International Edition-ISBN 3-540-20594-2.

Reference Books

1. Christopher D. Manning - PrabhakarRaghavan and HinrichSchutze “Introduction to Information Retrieval” - Cambridge University Press - ISBN 978-0-521-86571-5
2. Mark leven - “Introduction to search engines and web navigation” - John Wiley and sons Inc. - ISBN 9780-170-52684-2.
3. V. S. Subrahamanian - Satish K. Tripathi “Multimedia information System” - Kulwer Academic Publisher.

4. ChabaneDjeraba -"Multimedia mining A highway to intelligent multimedia documents" - Kulwer Academic Publisher - ISBN 1-4020-7247-3.
5. Ricci - F - Rokach - L. Shapira - B.Kantor - "Recommender Systems Handbook" - First Edition - 2011.
6. Stefan Buttcher - Charles L. A. Clarke - Gordon V. Cormack - Information Retrieval Implementing and Evaluating Search Engines - The MIT Press - Cambridge - Massachusetts London - England - 2010.

COURSE - 19UPCSC3E15 **CREDITS- 4**

NATURAL LANGUAGE COMPUTING

Course objective

- To get introduced to language processing technologies for processing the text data.
- To understand the role of Information Retrieval and Information Extraction in Text Analytics.
- To acquire knowledge on text data analytics using language models.

Unit-I

Natural Language Processing – Linguistic Background -– Mathematical Foundations - Morphological Analysis-Tokenization- Stemming-Lemmatization – Boundary Determination.

Unit-II

Reading unstructured data - Representing text data - Part of speech tagging – Syntactic representation - Text similarity - WordNet based similarity- Shallow parsing –Semantic representation.

Unit-III

Information retrieval and Information extraction - Named Entity Recognition – Relation Identification-Template filling.

Unit-IV

Language model - Probabilistic Models - n-gram language models- Hidden Markov Model- Topic Modelling - Graph Models -Feature Selection and classifiers -Rule-based Classifiers - Maximum entropy classifier – Clustering-Word and Phrase-based Clustering.

Unit-V

Tools – Natural Language Tool kit, Apache OpenNLP. Applications of Text Analytics – Applications in Social media - Life science - Legal Text–Visualization -Case studies.

Text Books

1. Christopher D. Manning and Hinrich Schutze, “Foundations of Statistical Natural Language Processing”, MIT Press, 1999.
2. Steven Struhl, “Practical Text Analytics: Interpreting Text and Unstructured Data for Business Intelligence”, Kogan Page, 2015.
3. Matthew A. Russell, “Mining the Social Web”, O'Reilly Media, 2013.
4. Steven Bird, Ewan Klein and Edward Loper, “Natural Language Processing with Python”, 1st Edition, O'Reilly Media, 2009.

Outcomes

- Students will be able to:
- Process the text data at syntactic and semantic level.
- Extract the ¬key information from Text data.
- Analyze the text content to provide predictions related to a specific domain using language models

COURSE - 19UPCSC3E16 **CREDITS- 4**

Image and Video Analytics

Course objective

- To teach the fundamentals of digital image processing, image and video
- analysis.
- To understand the real time use of image and video analytics.
- To demonstrate real time image and video analytics applications and others.

Unit-I

Digital image representation- Visual Perception- Sampling and Quantization- Basic Relations between Pixels- Mathematical Tools Used in Digital Image Processing: Fundamental Operations –Vector and Matric Operations- Image Transforms (DFT, DCT, DWT, Hadamard).

Unit-II

Fundamentals of spatial filtering: spatial correlation and convolution- smoothingblurring- sharpening- edge detection - Basics of filtering in the frequency domain: smoothing-blurring- sharpening--Histograms and basic statistical models of image.

Unit-III

Colour models and Transformations – Image and Video segmentation-Image and video demonising- Image and Video enhancement- Image and Video compression.

Unit-IV

Object detection and recognition in image and video-Texture models Image and Video 25 classification models- Object tracking in Video.

Unit-V

Applications and Case studies- Industrial- Retail- Transportation & Travel- Remote sensing-Video Analytics in WSN: IoT Video Analytics Architectures.

Reference Books

1. R.C. Gonzalez and R.E. Woods." Digital Image Processing". 3rd Edition. Addison Wesley, 2007.
2. W. Härdle, M. Müller, S. Sperlich, A. Werwatz, "Nonparametric and Semi parametric Models", Springer, 2004.
3. Rick Szelisk, "Computer Vision: Algorithms and Applications", Springer 2011.
4. Jean-Yves Dufour, "Intelligent Video Surveillance Systems", Wiley, 2013.
5. Caifeng Shan, Fatih Porikli, Tao Xiang, Shaogang Gong, "Video Analytics for Business Intelligence", Springer, 2012.
6. AsierPerallos, Unai Hernandez-Jayo, Enrique Onieva, Ignacio Julio García Zuazola, "Intelligent Transport Systems: Technologies and Applications", Wiley, 2015.
7. Basudeb Bhatta, "Analysis of Urban Growth and Sprawl from Remote Sensing Data", Springer, 2010

Outcomes

Students will be able to:

- Describe the fundamental principles of image and video analysis and have an
- idea of their application.
- Apply image and video analysis in real world problems.

COURSE - 19UPCSC3E17 **CREDITS- 4**

SOCIAL NETWORKING AND MINING

Course objective

- To understand the components of the social network.
- To model and visualize the social network.
- To mine the users in the social network.
- To understand the evolution of the social network.
- To mine the interest of the user.

Unit-I

Introduction- Introduction to Web - Limitations of current Web – Development of Semantic Web – Emergence of the Social Web – Statistical Properties of Social Networks -Network analysis - Development of Social Network Analysis - Key concepts and measures in network analysis - Discussion networks - Blogs and online communities - Web-based networks.

Unit-II

Modeling And Visualization- Visualizing Online Social Network Visualizations - Graph Representation - Centrality- Clustering - Node-Edge Diagrams - Visualizing Social Networks with Matrix-Based Representations- Node-Link Diagrams - Hybrid Representations - Modelling and aggregating social network data – Random Walks and their Applications –Use of Hadoop and Map Reduce – Ontological representation of social individuals and relationships.

Unit-III

Mining Communities- Aggregating and reasoning with social network data- Advanced Representations - Extracting evolution of Web Community from a Series of Web Archive - Detecting Communities in Social Networks - Evaluating Communities – Core Methods for Community Detection & Mining - Applications of Community Mining Algorithms - Node Classification in Social Networks.

Unit-IV

Text and Opinion Mining- Text Mining in Social Networks -Opinion extraction – Sentiment classification and clustering - Temporal sentiment analysis - Irony detection in opinion mining - Wish analysis - Product review mining – Review Classification – Tracking sentiments towards topics over time.

Unit-V

Tools for Social Network Analysis- UCINET – PAJEK – ETDRAW – StOCNET – Splus – R – NodeXL – SIENA and RSIENA – Real world Social Networks (Facebook- Twitteretc.)

Text Books

1. Charu C. Aggarwal, "Social Network Data Analytics", Springer; 2011.
2. Peter Mika, "Social Networks and the Semantic Web", 1st edition, Springer, 2007.
3. BorkoFurht, "Handbook of Social Network Technologies and Applications", 1st edition, Springer, 2010.
4. GuandongXu, Yanchun Zhang and Lin Li, "Web Mining and Social Networking –Techniques and applications", 1st edition, Springer, 2011.
5. Giles, Mark Smith, John Yen, "Advances in Social Network Mining and Analysis", Springer, 2010.
6. Ajith Abraham, Aboul Ella Hassanien, VáclavSnáel, "Computational Social
7. Network Analysis: Trends, Tools and Research Advances", Springer, 2009.
8. Toby Segaran, "Programming Collective Intelligence", O"Reilly, 2012.
9. Sule Gündüz-Ogüdücü, A. Şima Etaner-Uyar, "Social Networks: Analysis and Case Studies", Springer, 2014.

Outcomes

- Students will be able to:
- Work on the internal components of the social network.
- Model and visualize the social network.
- Mine the behavior of the users in the social network.
- Predict the possible next outcome of the social network.
- Mine the opinion of the user.

COURSE - 19UPCSC3E18 **CREDITS- 4**

WEB INTELLIGENCE

Course objective

- To know the importance of qualitative data, get insights and techniques.
- To develop customer-centric approach in dealing with data.
- To know the principles, tools and methods of web intelligence.
- To apply analytics for business situations.

Unit-I

Web Analytics – Basics – Traditional Ways – Expectations – Data Collection – Clickstream Data – Weblogs – Beacons – JavaScript Tags – Packet Sniffing – Outcomes data – Competitive data – Search Engine Data.

Unit-II

Qualitative Analysis – Customer Centricity – Site Visits – Surveys – Questionnaires – Website Surveys – Post visits – Creating and Running-Benefits of surveys – Critical components of successful strategy.

Unit-III

Web Analytic concepts – URLS – Cookies – Time on site – Page views – Understand standard reports – Website content quality – Navigation reports (top pages, top destinations, site overlay). – Search Analytics – Internal search, SEO and PPC – Measuring Email and Multichannel Marketing - Competitive intelligence and Web 2.0 Analytics – Segmentation – Connectable reports.

Unit-IV

Google Analytics: Analytics - Cookies - Accounts vs Property - Tracking Code - Tracking Unique Visitors - Demographics - Page Views & Bounce Rate Acquisitions - Custom Reporting.

Unit-V

Goals & Funnels – Filters - Ecommerce Tracking - Real Time Reports - Customer Data Alert - Adwords Linking - Adsense Linking -Attribution Modeling - Segmentation - Campaign Tracking - Multi-Channel Attribution.

References:

1. Avinash Kaushik, “Web Analytics 2.0: The Art of Online Accountability and
2. Science Of Customer Centricity “, 1st edition, Sybex, 2009.
3. Michael Beasley, “Practical Web Analytics for User Experience: How Analytics
4. can help you Understand your Users”, Morgan Kaufmann, 2013.
5. Magy Seif El-Nasr, Anders Drachen, Alessandro Canossa, eds., “Game Analytics: Maximizing the Value of Player Data”, Springer, 2013.
6. Bing Liu, “Web Data Mining: Exploring Hyperlinks, Content, and Usage Data”, 2nd Edition, Springer, 2011.
7. Justin Cutroni, “Google Analytics”, O‟Reilly, 2010.
8. Eric Fettman, Shiraz Asif, Feras Alhlou , “Google Analytics Breakthrough”, John Wiley & sons, 2016.

Outcomes:

Students will be able to:

- Know the concepts and terminologies related to web analytics.
- Explore various parameters used for web analytics and their impact.
- Explore the use of tools and techniques of web analytics.
- Get experience on websites, web data insights and conversions

COURSE - 19UPCSC3E19 **CREDITS- 4**

DEEP LEARNING

Unit -I

Deep Learning: AI and deep learning - The history and rise of deep learning - Why Deep Learning? - The motivation of deep architecture – Applications - Future potential and challenges.

Getting Yourself Ready for Deep Learning: Basics of linear algebra - Deep learning with GPU - Deep learning software frameworks - Setting up deep learning on AWS

Unit -II

Getting Started with Neural Networks: Multilayer perceptrons - How a network learns - Deep learning models - Practical examples

Deep Learning in Computer Vision: Origins of CNNs- Convolutional Neural Networks -Fine-tuning CNNs - Popular CNN architectures

Unit -III

NLP - Vector Representation: Traditional NLP - Deep learning NLP – Applications.

Advanced Natural Language Processing: Deep learning for text - Recurrent neural networks - Long short-term memory network – Applications

Unit -IV

Multimodality: What is multimodality learning? - Challenges of multimodality learning- Image captioning - Visual question answering - Multi-source based self-driving

Deep Reinforcement Learning: What is reinforcement learning (RL)? - Deep reinforcement learning - Implementing reinforcement learning

Unit - V

Deep Learning Hacks: Massaging your data - Tricks in training - Fine-tuning - Model compression

Deep Learning Trends: Recent models for deep learning - Novel applications

TEXT BOOK:

Anurag Bhardwaj, Wei Di, Jianing Wei,"**Deep Learning Essentials**", Packt Publishing, 2018. Chapters:1-10

References

Goodfellow, I., Bengio,Y., and Courville, A., Deep Learning, MIT Press, 2016..

Bishop, C. ,M., Pattern Recognition and Machine Learning, Springer, 2006.

COURSE - 19UPCSC3E20 **CREDITS- 4**

BIOINFORMATICS

Course objective

- To understand Bio informatics from computing perspective.
- To comprehend bio informatics databases, file formats and its applications.
- To understand the applications of Bio informatics

Unit -I

History of bioinformatics-History of Bioinformatics-role of Bioinformatics in biological sciences- scope of bioinformatics -introduction to internet-WWW-network basics- LAN & WAN standards-network topologies and protocols- FTP-HTTP - division of Bioinformatics- Bioinformatics and internet-challenges in Bioinformatics.

Unit -II

Databases in bioinformatics-Databases in Bioinformatics- Genbank- NCBI-EMBL- DDBJ -UniGene- SGD- EMI Genomes- -protein databases-PIR-SWISSPROT-TrEMBL-PrositePRINTS -structural databases-PDB- SCOP- CATH-PDB_SELECT- PDBSUM- DSSP- FSSPDALI- PRODOM- protein families & pattern databases-Pfam- KEGG - sequence storage sequence accuracy-EST-STS- sequence retrieval systems- Entrez-SRS- sequence query refinement using Boolean operators- limits- preview- history and index.

Unit -III

Sequence submission-Sequence submission tools-BANKIT-SEQUIN-WEBIN-SAKURAliterature databases-PubMed and medline. Data mining and its techniques- data warehousing- Sequence annotation- principles of genome annotation- annotation tools & resources.

Unit -IV

Applications of bioinformatics-Applications of Bioinformatics-phylogenetic analysissteps in phylogenetic analysis-microarrays-DNA and protein microarraysBioinformatics in pharmaceutical industry- informatics & drug-discovery – pharma informatics resources drug discovery and designing-SNP.

Unit -V

File formats-File formats-raw/plain format-NCBI-Genbank flat file format-ASN.1- GCGFASTA- EMBL- NBRF- PIR-swissprot sequence formats- PDB format-Introduction to structure prediction methods.

References:

1. Attwood T.K, Parry-Smith, "Introduction to Bioinformatics", Addison Wesley Longman, 1999.
2. David W Mount, "Bioinformatics: Sequence and Genome Analysis", 2nd edition, CBS Publishers, 2004.
3. Arun Jagota, "Data Analysis and Classification for Bioinformatics", Pine Press, 2001.
4. Des Higgins and Willie Taylor, "Bioinformatics Sequence, Structures and Databanks", Oxford University Press, 2000.
5. Jason T.L.Wang, Mohammed J. Zaki, Hannu T.T. Toivonene and Dennis Shasha, "Data Mining in Bioinformatics", Springer International Edition, 2005.
6. K. Erciyes, "Distributed and Sequential Algorithms for Bioinformatics", Springer, 2015.