**MASTER OF SCIENCE IN COMPUTER SCIENCE PROGRAMME STRUCTURE AND SYLLABUS From 2019-2020 Admission Onwards (UNDER MAHATMA GANDHI UNIVERSITY PGCSS REGULATIONS 2019)**

Semester I

CA500101 - Computational Mathematics

CA010101 - Advanced web Technology

CA010102 - Operating Systems

|  |  |
| --- | --- |
| CA500102 | - Advanced Java Programming |
| CA010103  Semester II | - Lab I [ Java& PHP] |
| CA500201 | - Advanced Data Structures |
| CA010201 | - Computer Networks |
| CA010202 | - Research Methodology and Technical Writing |
| CA500202 | - Database Management system and SQL |
| CA010203  Semester III | - Lab II [ DS using Java, SQL] |
| CA010301 | * Digital Image Processing * Elective I |
| CA010302 | - Python Programming |
| CA500301 | - Software Engineering |
| CA010303 | - Lab III [ DIP using Python] |
| CA010304  Semester IV | - Mini Project using IOT |
| CA010401 | - Data Mining |
|  | - Elective II |
|  | - Elective III |
| CA010402 | - Main Project |
| CA010403 | - Course Viva |

Elective Group A

CA800301 - Introduction to Cyber Security

CA800402 - Applied Cryptography

CA800403 - Ethical Hacking

Elective Group B

CA810301 - Statistical Computing for Data Analytics

CA810402 - Big Data Management Using R

CA810403 - Data Analytics

Elective Group C

CA820301 - Soft Computing

CA820402 - Advanced Python Programming

CA820403 - Pattern Recognition

### **First Semester Courses**

|  |  |
| --- | --- |
| CA500101 | Computational Mathematics |
| CA010101 | Advanced web Technology |
| CA010102 | Operating System |
| CA500102 | Advanced Java Programming |
| CA010103 | Lab I [ Java & PHP] |

**CA500101--Computational Mathematics**

## Module I

Mathematical Logic: Propositional Calculus: Statements and notations, Connectives: negation, conjunction, disjunction, statement formulas and truth tables, conditional and biconditional, Well-formed formulas, tautologies, equivalence of formulas, tautological implication. Normal forms: Disjunctive and conjunctive normal forms.

Predicate calculus: Predicates, statement functions, variables and quantifiers, predicate formulas, free & bound variables, universe of discourse.

## Module II

Basic Statistics: Measure of central value: Introduction, types of average- arithmetic mean: calculation of arithmetic mean-discrete series, continuous series. Median:calculation of median**-**discrete series, continuous series. Mode: calculation of mode- discrete series, continuous series.

Measures of dispersion: Absolute and relative measures of dispersion, Range, Mean deviation:calculation of mean deviation-individual observations, discrete series, and continuous series.Standard deviation: calculation of standard deviation- individual observations, discrete series, continuous series, coefficient of variation.

## Module III

Correlation& Regression analysis**:** introduction, correlation and causation ,types of correlation, Karl Pearson’s coefficient of correlation-direct method of finding out correlation coefficient, calculation of correlation coefficient when change of scale and origin is made. Regression: introduction, regression equation of y on x, regression equation of x on y. **Module IV**

Theory of Automata: Definition, Description of finite automaton, Transition systems and its properties, Acceptability of a string by a finite automata, Non deterministic finite state machines, Equivalence of DFA and NDFA, Minimization of finite automata -construction of minimum automaton. Regular sets and regular grammar: Regular expressions, Transition system containing null moves, construction of finite automata equivalent to a regular expression. **Module V**

Fuzzy logic: Introduction, Crisp set an overview, Fuzzy sets basic types, Basic concepts , Characteristics and significance of paradigm shift.

## Reference Text

1. J.P. Tremblay & R Manohar- Discrete Mathematical Structures with Applications to Computer Science ,McGraw Hill.
2. S. P. Gupta- “ Statistical Methods”, Sultan Chand & Sons.
3. K.L.P Mishra & N. Chandrasekaran -Theory of Computer Science(Automata ,Languages and Computation) ,Prentice hall of India.
4. George J Klir& Bo Yuan- Fuzzy sets and Fuzzy logic Theory and applications, Prentice hall of India.

**CA010101--Advanced Web Technology**

## Module I

Internet introduction, WWW, understanding client/server role, web browsers, web servers, HTML 5 core elements and attributes, text formatting and presentational tags, links, adding images, image maps, lists, tables, HTML 5 form controls- text input, check box, radio button, select box, file select box, buttons, number, date, time, calendar, and range, <nav>,<section>, <article>, <header>, <footer>.

CSS Introduction, <link> and <style>, CSS properties, text pseudo classes.

## Module II

Javascript, document object model, variables, operators, popup boxes , functions, conditional statements, looping, events, built-in –objects, form validation.

Introduction to PHP, server side scripting, php comments, variables, echo and print, PHP operators, data types, branching statements, loops.

## Module III

Arrays, PHP functions, working with forms, $\_GET, $\_POST, $\_REQUEST, String functions, include and require, session and cookie, error handling in PHP.

## Module IV

Object Oriented Programming using PHP- classes, objects, constructor, destructor, inheritance, polymorphism, function overriding.

Introduction to MySQL, Database & table creation, database operations-select, insert, update, delete, drop, database connections, functions for managing database connections.

## Module V

CodeIgniter (PHP MVC Framework) – MVC Overview, Explaining Models, Views,

Controllers, Installation, Setup Dreamweaver/NetBeans IDE, Folder Structure

Configuration -Libraries & Helpers, Active Record Class-- Selecting Data, Inserting Data, Updating Data, Deleting Data , Working with Simple Database Program

## Reference Text

1. Steven Holzner, The complete reference PHP 5.2, 5thEdition , Tata McGraw-Hill
2. Steven Holzner, The complete reference HTML5 & CSS, 5th Edition, Tata McGraw-Hill Edition.
3. Steve Suehring, Tim converse and Joyce Park, *PHP6 and MySQL* , Wiley publication
4. IVanBayRoss , HTML, DHTML, JavaScript, Perl CGI, 4th Edition, BPB Publications

**CA010102-- Operating Systems**

## Module I

Computer system architecture – single processor systems , multiprocessor systems , clustered systems.Operating system operations- dual mode and multimode operation. Process management, Memory management, Storage management. Computing Environments- Traditional computing, Mobile computing, Distributed systems, Client Server computing, Peer-to-Peer computing, Virtualization, Cloud computing, Real-time embedded systems.

System structures - Operating system services , System calls , Types of system calls ,Operating system structure-Simple structure, Layered approach, Microkernals ,Modules, Hybrid systems.

## Module II

Process management - Process concept - Process state, PCB, Process Scheduling -Scheduling queues, Schedulers, Context switch, Operations on processes - creation, termination, Interprocess Communication- Shared memory systems , MessagePssing systems.

Multithreaded Programming - Overview , Multithreading Models.

Process Scheduling – Basic Concepts, Scheduling criteria , Scheduling algorithms- FCFS, SJF, Priority scheduling, RR scheduling, Multilevel queue scheduling, Multilevel Feedback queue scheduling,

## Module III

Process Synchronization - The critical section problem- Peterson’sSolution, Synchronization hardware , MutexLocks , Semaphores, Monitors,Monitor usage

Deadlocks – System model, Deadlock characterisation, Methods for handling deadlocks, Deadlock prevention, Deadlock avoidance, Deadlock detection, Recovery from deadlock.

## Module IV

Memory management- Memory management strategies - Basic hardware , Address binding ,

Logical Vs Physical address space, Dynamic loading , Dynamic linking and shared libraries , Swapping ,Contiguous memory allocation ,segmentation , Paging - Basic method , Hardware support, Protection,Shared pages.

Virtual memory management :- Demand paging - Basic concepts ,Performance of demand paging, Page Replacement, Page Replacement algorithms - FIFO, Optimal page replacement, LRU page replacement .

## Module V

Case study -The Linux System - Features , Advantages,Linux history , Design Principles ,Kernel Modules,Process Management, Scheduling - Process Scheduling, Real-time Scheduling , Virtual Memory , File Systems , Interprocess Communication , Security .

Various types of shells available in Linux - Comparison between various shells - Linux Commands for files and directories - cd, ls, cp ,rm, mkdir, rmdir, pwd, file , more, less . Creating and viewing files using cat .

Reference Text

1. Abraham Silberschatz, Galvin, Gange, Operating SsystemConcepts, 9th Edition ,Wiley Publishers .
2. Milan kovic, Operating Systems, Second Edition .
3. Official Red hat Linux Users Guide- Red hat, Wiley Dreamtech India.
4. Christopher Negus, Red Hat Linux Bible -2005 Edition,Wiley Dreamtech India.
5. YeswantKanethkar, Unix Shell Programming,First Edition, BPB .

**CA500102—Advanced Java Programming**

**ModuleI** : Object Oriented Programming Concepts and Basics of Java.

Java Programming Environment – JDK, Java Virtual Machine, Bytecode, Features of Java Flow Control Statements – Conditional Statements, Iteration Statements, Jump Statements

Arrays –One Dimensional Array, Multi-dimensional Array , Object Oriented Programming Concepts- ( Objects and Classes, Encapsulation, Inheritance, Polymorphism) , Type of

Inheritance , Method Overloading, Method Overriding, Dynamic Method Despatch

**Module II** :Input/Output Handling

Constructors- Constructor Overloading , this, super, final, abstract and static Keywords,

Interfaces- Defining an Interface, Implementing Interface, Extending Interfaces. String -

String Handling Fundamentals, Comparison of String and StringBuffer Class, Special String Operations- Character Extraction, String Comparison, Searching String, Modifying a String,

String Copy ,Input and Output Streams – Byte Stream , Character Stream

**ModuleIII**:Packages; Exception Handling and Thread

Packages – Defining Packages, Built in Packages(java.lang, java.util, java.io, java.net, javax.swing), Importing Packages, Implementation of User Defined Packages, Access Protection in Java, Exception Handling - try, catch, throw, throws and finally Statements, Java’s Built-in Exceptions, Creating User Defined Exceptions. Threads- Thread Lifecycle, Thread Priorities, The Thread Class, Runnable Interface, Creating a Thread – Implementing Runnable, Extending Thread ,Inter Thread Communication, Suspending Resuming and Stopping Threads.

**Module IV**: GUI Programming

Basic Event Handling – Delegation Event Model, Important Event Classes And Listener Interfaces, Handling Mouse and Keyboard Events,Adapter Classes, Swing -Window Fundamentals – Class Hierarchy, Frame, Creating a Simple Window Based Application, ImageIcon, JLabel, JTextField, JTextArea, JButton, JCheckBox, JRadioButton, JList,

JComboBox, JTable, JTabbedPane, JScrollPane, Layout Management – The FlowLayout, BorderLayout, GridLayout, CardLayout

**ModuleV** : File, Database and Networking

File Management - Reading and Writing Files (FileInputStream and FileOutputStream Classes), Networking Basics- Networking Classes and Interfaces, InetAddress, TCP/IP Client Sockets, URL Connection, TCP/IP ServerSockets, JDBC - The Design of JDBC, JDBC Configuration, Executing SQL Statements- Scrollable and Updatable ResultSets, RowSets, Transactions.

## Reference Text

1. Herbert Schildt,Java 2 The Complete Reference, Tata McGraw Hill (5thEdn.)
2. James. P. Cohoon,Programming java5.0, , Jack. W. Davison (Tata McGraw Hill)
3. C Thomas Wu, An introduction to Object Oriented Programming with Java, , Tata McGraw Hill, (2006)
4. Wigglesworth and McMillan ,Java Programming: Advanced Topics, , Cengage Learning India, 3rdEdn.
5. Bernard Van Haecke, JDBC:Java Database Connectivity, , IDG Books India (2000)

## CA010103 – Lab I Advanced Java Programming& PHP

1. Basic Concepts and File Handling
   1. Inheritance, Polymorphism
   2. Constructors
   3. Interface
   4. Package
   5. One Dimensional and Two Dimensional Array Manipulation
   6. String Handling (Character Extraction, String Comparison, Searching String, Modifying a String, String Copy)
   7. Exception (Built-in and User Defined)
   8. Thread (Using Runnable Interface and Thread Class)
   9. File management (File reading, Writing, Appending and Content Replacing)
2. GUI, Database and Networking
   1. Event Handling (Keyboard and Mouse Events)
   2. Working with Swing ( ImageIcon, JTextField, JTextArea, JButton, JCheckBox, JRadioButton, JComboBox, JList, JTable)
   3. Layout Management (The FlowLayout, BorderLayout, GridLayout, CardLayout)
   4. Simple Programs of Database Connectivity

## PHP

1. Create a calendar of the month of January 2019 using HTML.
2. Design a page for the inauguration of your department association using HTML & CSS.
3. Create and validate a bio data form using JavaScript and HTML.
4. Create a JavaScript program to display todays date and current time.. 5. Program to check whether the string is palindrome or not using PHP.
5. Create a Login page by using PHP and session.
6. Create a simple online quiz page using PHP and MySQL.
7. Create a simple user registration page in PHP.
8. Create an E- mail registration form using Code igniter.
9. Create an online shopping cart using Code igniter.

**. Second Semester Courses**

|  |  |
| --- | --- |
| CA500201 | Advanced Data Structure |
| CA010201 | Computer Networks |
| CA010202 | Research Methodology and Technical Writing |
| CA500202 | Database Management system and SQL |
| CA010203 | Lab II [ DS using Java, SQL] |

**CA500201-- Advanced Data Structures**

## Module I

Concept of data structures, types of data structures, examples.

Introduction to algorithms, Performance analysis-Space complexity, Time complexity, Amortised complexity, Asymptotic notations, Performance measurement; various algorithm designing techniques-Divide and conquer, Greedy method, Dynamic programming, Backtracking, Branch and bound, *Np*-hard and *Np*-completeness problems.

## Module II

Arrays: Organization, Representation and implementation of arrays, examples. Implementation of Stacks and Queues, Circular Queues, Priority Queues, Double ended queues, Applications of stacks and queues.

Sorting and Searching techniques: Linear and Binary search, Selection sort, Merge sort, Simple insertion sort, Quick sort, Shell sort, Radix sort.

## Module III

Lists: Representation and implementation of singly linked list, Circular linked lists, doubly linked list, Linked list representation of stacks and queues, examples.

Dynamic storage management. Boundary tag system. Garbage collection and compaction.

## Module IV

Trees: Representation and Implementation, Binary trees, insertion and deletion of nodes in binary tree, binary tree traversals, Binary search trees, Threaded Binary trees, Balanced trees ( AVL trees), B- trees- Insertion and Deletion of nodes, Tree search

## Module V

Graphs: Directed Graphs, Shortest Path Problem, Undirected Graph, Spanning Trees, Techniques for graphs –Breadth First Search (BFS) and traversal, Depth First Search (DFS) and traversal

Hashing: Static hashing, hash tables, hash functions, overflow handling. **Reference Text**

1. Robert Lafore, Data structures and Algorithms in Java, Pearson Publications
2. Clifford A Shaffer , Data Structures and Algorithm analysis in Java
3. Ellis Horowitz, SartajSahni, SanguthevarRajasekaran, Computer Algorithms/C++, University press publications
4. G S Baluja, Data structures Through C++,
5. Ellis Horowitz and SartajSahni, Fundamentals of Data structure

**CA010201--Computer Networks**

## Module I

The OSI Model- Layered architecture, Peer-to-Peer process, Encapsulation. Layers in the OSI Model- Physical layer, Data link layer, Network layer, Transport layer, Session layer, Presentation layer, Application layer. TCP/IP protocol suite. Addressing- Physical addresses, Logical addresses, Port addresses, Specific addresses. Physical Layer:- Transmission media- Guided media- Twisted pair, Coaxial, Fiber-optic Cables. Unguided Media- Radio waves, microwaves, infrared waves. Switching-Packet switched networks, Datagram Networks, Virtual circuit networks.

**Lab-** Different types of LAN cables, connectors, Hub, Switch, Router, and Configuration of star LAN (Ethernet LAN)

## Module II

Data Link Layer:- Framing, Flow and Error Control, Protocols, Noiseless channels- simplest, stop-and-wait protocols. Noisy channels- Stop and wait ARQ, Go-Back NARQ, Selective Repeat ARQ. Piggybacking. Random access protocols- Aloha (Pure& slotted), CSMA, CSMA/CD, CSMA/CA. Standard Ethernet- MAC sub layer, Frame format. Fast Ethernet, Gigabit Ethernet. Wireless LAN (IEEE 802.11)- Architecture, MAC sub layer, Frame format, Addressing mechanism, Physical layer. Bluetooth- Architecture, Bluetooth layers, Frame format. Connecting Devices- Hubs, Switches, Routers, Gateway.

**Lab**-connecting two LAN using a switch

## Module III

Network Layer: IPv4 Addresses- Address space, Notations, Classful addressing, Classless addressing, NAT. IPv6 Addresses- Structure, Address space. Internet Protocol (IP)- IPv4 Datagram format, IPv6- Advantages, Packet format. Transition from IPv4 to IPv6- Dual stack, Tunneling, Header translation. Address mapping protocols: ARP, RARP, BOOTP, DHCP. Error Reporting protocol: ICMP-Types of Messages, Message format, Error Reporting, Query. Multicasting Protocol: IGMP- Group management, IGMP messages, Message format, IGMP operations. Forwarding- Forwarding Techniques, Forwarding process, Routing table. Unicast routing protocols- Distance vector routing (RIP), Link state routing (OSPF), Path vector routing (BGP).

**Lab:** Configuring Wireless LAN (WiFi)

**Module IV** Transport layer: User Datagram Protocol (UDP)- Well-known ports for UDP, Datagram format, UDP operation, Use of UDP. TCP- TCP services, TCP features, TCP segment format, TCP connection- connection establishment, connection termination. SCTP-SCTP services, SCTP features, SCTP packet format, SCTP association- association establishment, data transfer, association termination. Congestion control- open loop congestion control, closed loop congestion control.

## Module V

Application layer: Domain Name System- Name space, Domain name space, Distribution of Name space, DNS in the Internet- Generic domains, Country domains, Inverse domains. Resolution- Resolver, Recursive resolution, iterative resolution. DNS message. Types of records. DDNS (Dynamic Domain Name System). TELNET- Logging, Network virtual terminal, options, mode of operation. E-mail- Architecture, User Agent, Message Transfer Agent (SMTP), Message Access Agent: POP, IMAP. Web-based mail. FTP: Basic model of FTP, Control connection, Data connection, Anonymous FTP. HTTP protocol- HTTP transaction, Message formats, Persistent and Non persistent connection, Proxy server. **Reference Text**

1. Behrouz Forouzan- Data Communication & Networking – Fourth Edition - The McGraw-Hill Companies, Inc.
2. James F. Kurose & Keith W. Ross - Computer Network – Top down approach – 7th Edition - Pearson Education, Limited
3. Behrouz A. Forouzan - TCP/IP Protocol Suite – 4th Edition - [McGraw-Hill](https://vaibhav2501.files.wordpress.com/2012/02/tcp_ip-protocol-suite-4th-ed-b-forouzan-mcgraw-hill-2010-bbs.pdf)

**CA010202-Research Methodology and Technical Writing**

## Module -I

Research Methodology**:** Meaning of Research, Objectives of Research, Motivation in Research, Types of Research, Research Approaches, Significance of Research, Research Methods versus Methodology, Research and Scientific Method, Research Process, Criteria of Good Research.

**Module -II** Research Design**:**

Reading and Reviewing-Research literature, Finding Research Papers, Critical Reading, Developing a literature Review, Guidelines for Research Skills and Awareness, Validity of Research, Reliability in Research.Meaning of Research Design, Need for Research Design, Features of good design, Different Research Designs.

## Module -III

Data Collection and Analysis**:** Introduction, Need for Data Collection**,** Methods ofData

Collection, Principles for Accessing Research Data, Data Processing, Data Analysis,

Presentation of Data, Error Analysis, Scientific Models. Scientific Methodology **-** Introduction Rules and Principles of Scientific Method, Hypothesis,Testing of Hypothesis, Basic concepts, Procedure,Important parametric tests: z-test ,t-test, *χ*2-square test, F test.

## Module -IV

Reporting and thesis writing

Presentation of algorithms, Environment of Algorithms, Asymptotic Cost. Graphs. Technical Reports- Structuring General format, Report-Bibliography referencing and footnotes.Research in Practice**-** Literature Review, Journals, Conference Proceedings, journal Impact Factor, citation Index, h Index .Application of Computer in Research **--**MS office and its application in Research, Use of Internet in Research – Websites, search Engines, E-journal and E-Library.

## Module -V (10 Hours)

Ethics in Research –Research Ethics, Importance of Ethics in Research, Ethics values and Principles, Some Ethical issues ,Plagiarism, Misuse of Privileged Information, Misuse of

Data, Authorship and other publication issues, meaning of Copy Right, Copy Right and Information Technology **Reference Text**

1. Kothari, C.R., 1990. Research Methodology: Methods and Techniques. New Age International. Publishers( Second revised edition)
2. Justin Zobel,Writing For Computer Science, Springer (Third Edition)
3. K Prathapan,Research Methodology for Scientific Writing ,I.K International Publishing House Pvt.Ltd
4. Garg, B.L., Karadia, R., Agarwal, F. and Agarwal, U.K., 2002. An introduction to Research Methodology, RBSA Publishers.
5. S.P Satarkar, S.V., 2000. Intellectual Property Rights and Copy right. Ess Publications.

## Additional reading

1. Carlos, C.M., 2000. Intellectual property rights, the WTO and developing countries: the TRIPS agreement and policy options. Zed Books, New York.
2. Day, R.A., 1992.How to Write and Publish a Scientific Paper, Cambridge University Press.
3. Fink, A., 2009. Conducting Research Literature Reviews: From the Internet to Paper. Sage Publications
4. Leedy, P.D. and Ormrod, J.E., 2004 Practical Research: Planning and Design, Prentice Hall.
5. Sinha, S.C. and Dhiman, A.K., 2002. Research Methodology, Ess Publications. 2 volumes.
6. Trochim, W.M.K., 2005. Research Methods: the concise knowledge base, Atomic Dog Publishing. 270p.

**CA500202--Database Management System and SQL**

## Module I

Database, need for DBMS, users, DBMS architecture, data models, views of data, data independence, database languages, Relational Model-Basic concepts, keys, integrity constraints, ER model-basic concepts, ER diagram, weak entity set, ER to Relational, relationships, generalization, aggregation, specialization

## Module II

Codd‘s rules, Relational model concepts , Relational algebra- Select, Project, Join, Relational calculus-tuple relational calculus and domain relational calculus, Specifying constraints management systems, Anomalies in a database, Functional dependencies, NormalizationFirst, Second, Third, Boyce Codd normal forms, multi-valued dependency and Fourth normal form, Join dependency and Fifth normal form.

Relational database query languages-Basics of SQL, Data definition in SQL- Data types, Creation, Insertion, Viewing, Updation, Deletion of tables, Modifying the structure of the tables, Renaming, Dropping of tables, Data constraints-I/O constraints, ALTER TABLE command.

## Module III

Database manipulation in SQL- Computations done on the table- Select command, Logical operators, Range searching, Pattern matching, Grouping data from tables in SQL, GROUP BY, HAVING clauses, Joins-Joining multiple tables, Joining tables to itself, DELETE, UPDATE, Views-Creation, Renaming the column of a view, Destroys view- Program with SQL, Security-locks, Types of locks, Levels of locks, Cursors - working with cursors, error handling, Developing stored procedures,-Creation, Statement blocks, Conditional execution, Repeated execution, Cursor-based repetition, Handling Error conditions, Implementing triggers, Creating triggers, Multiple trigger interaction.

## Module IV

Concept of transaction, ACID properties, serializability, states of transaction, Concurrency control, Locking techniques, Time stamp based protocols, Granularity of data items,

Deadlock, Failure classifications, storage structure, Recovery & atomicity, Log base recovery, Recovery with concurrent transactions, Database backup & recovery, Remote Backup System, Database security issues

## Module V

Object Oriented Database Management Systems (OODBMS) - concepts, need for OODBMS, composite objects, issues in OODBMSs, advantages and disadvantages of OODBMS. Distributed databases - motivation - distributed database concepts, types of distribution, architecture of distributed databases, the design of distributed databases, distributed transactions, commit protocols for distributed databases

## Reference Text

1. Elmasri and Navathe, Fundamentals of Database Systems, 5th Edition, Pearson
2. Abraham Silbersehatz, Henry F. Korth and S.Sudarshan, Database System Concepts, 6 th Edition, Tata McGraw-Hill.

3.JamesR.Groff and Paul N. Weinberg The complte reference SQL Second edition,Tata McGraw Hill

## CA010203-Lab-II-DS & SQL Advanced Data Structures (Using Java)

1. Array implementation – Insertion of new element into a specified position, Deletion of an element from the specified position within the array
2. Stack implementation – PUSH, POP and Traverse
3. Queue implementation –Insertion, deletion and Traverse
4. Circular Queue implementation –Insertion, deletion and Traverse
5. Deque (Double ended queue) implementation –Insertion, deletion and Traverse
6. INFIX to POSTFIX Conversion
7. INFIX to PREFIX conversion
8. POSTFIX evaluation
9. Searching - Linear and Binary search using arrays
10. Sorting – Selection sort, Merge sort, Simple insertion sort, Quick sort, Shell sort, Radix sort
11. Lists implementation - Singly linked list, Circular linked list, Doubly linked list
12. Dynamic array implementation- Linked list representation and implementation of stack and queue operations
13. Creation of binary tree, counting no. of nodes and display the nodes in a tree
14. Searching a node in a binary tree
15. Insertion and deletion of nodes in a B-Tree
16. Graphs – Implementation of BFS and DFS

## SQL

1. Creating database tables and using data types (create table, modify table, drop table). 2. Data Manipulation (adding data with INSERT, modify data with UPDATE, deleting records with DELETE).

1. Implementing the Constraints (NULL and NOT NULL, primary key and foreign key Constraint, unique, check and default constraint).
2. Retrieving Data Using SELECT (simple SELECT, WHERE, IN, BETWEEN, ORDERED BY, DISTINCT and GROUP BY).
3. Aggregate Functions (AVG, COUNT, MAX, MIN, SUM).
4. String functions.
5. Date and Time Functions.
6. Use of union, intersection, set difference.
7. Implement Nested Queries & JOIN operation.
8. Performing different operations on a view.
9. Stored Procedure Programming – Simple Procedures – decision making – Loops – Error handlers – Cursors – Functions - Triggers – Calling Stored Procedure from Triggers.

### **Third Semester Courses**

|  |  |
| --- | --- |
| CA010301 | Digital Image Processing |
|  | Elective I |
| CA010302 | Python Programming |
| CA500301 | Software Engineering |
| CA010303 | Lab III [ DIP using Python] |
| CA010304 | Mini Project using IOT |

**A010301--Digital Image Processing**

## Module-I

Fundamentals of Image Processing –Definition of Image, Digital Image and Digital Image Processing, Examples of fields that use Digital Image Processing, Fundamental steps in image processing, Components of Image Processing system, Elements of Visual perception, Image sensing and acquisition, Image sampling and quantization, Relationships between pixels– Color image fundamentals – Color Models-RGB, CMY, HSI

## Module-II

Image Enhancement in spatial domain – Basic Intensity transformation functions – Image

Negatives, Log Transformations, Power Law Transformations, Piecewise Linear

Transformations, Histogram processing, Enhancement using arithmetic, logic operationsImage Subtraction and Image averaging – Fundamentals of spatial filtering ,Smoothing spatial Filters.

## Module-III

Image Enhancement in Frequency domain – Introduction to Fourier transform: 1- D, 2 –D DFT and its Inverse Transform, Properties of 2-D DFT, Image Smoothing and Sharpening using Frequency Domain Filters- Ideal, Butterworth and Gaussian filters and Homomorphic filtering.

## Module-IV

Image restoration and Compression: A Model of Image degradation and restoration process – Noise models-Gaussian Noise, Rayleigh Noise, Gamma Noise, Exponential Noise, Impulse

Noise, Restoration using Mean Filters, Order Statistics filters, Adaptive filters. CompressionNeed for data compression, Huffman, Run Length Encoding, Shift codes, Arithmetic coding, JPEG standard, MPEG.

## Module-V

Image Segmentation –Fundamentals, Edge detection-Gradient operator, Marr-Hildreth edge detector, canny edge detector, Thresholding- Global Thresholding using otsu’s method, Variable Thresholding, Region based segmentation – Region growing, Region splitting and merging, Segmentation using morphological watersheds.

## Reference Text

1. Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing, Pearson, Third Edition, 2014.
2. Kenneth R. Castleman, Digital Image Processing Pearson, 2006.
3. Anil K Jain, Fundamentals of Digital Image Processing, Prentice Hall, Fourth Edition,1989.
4. William K. Pratt, Digital Image Processing, John Wiley, Fourth Edition, New York, 2002.
5. Milan Sonka et al, Image processing, analysis and machine vision Brookes/Cole, Vikas Publishing House, Fourth edition, 2007.

**CA010302 --Python Programming**

## Module I

Introduction-Features of Python, Installation, Basic Syntax, Variables and Datatypes, Operators- Arithmetic operators, Assignment operators, Comparison operators, Logical operators, Identity operators, Membership operators, Bitwise operators, Casting. Conditional Statements-if, if-else, Nested if-else. Looping Statements-for, while, Nested loops. Control Statements-break, continue, Pass.

## Module II

Strings and text files; manipulating files and directories, os and sys modules; text files: reading/writing text and numbers from/to a file; creating and reading a formatted file (csv or tab-separated).

String Manipulations-Accessing Strings, Basic Operations-extract, replace, len, lower, upper, split, substrings, String slices,strings and number system: converting strings to numbers and vice versa, String Methods.

## Module III

Lists, tuples, and dictionaries; basic list operators, replacing, inserting, removing an element; searching and sorting lists; dictionary literals, adding and removing keys, accessing and replacing values; traversing dictionaries.

Functions: hiding redundancy, complexity; arguments and return values; formal vs actual arguments, named arguments, Recursive functions.

## Module IV

Modules & Packages: Creating Modules, import Statement, Locating Modules, Namespaces and Scope, Packages, Date and Time Modules.

Simple Graphics and Image Processing: “turtle” module; simple 2d drawing - colors, shapes; digital images, image file formats, image processing Simple image manipulations with 'image' module.

## Module V

File Operations-Reading config files in python ,Writing log files in python ,Understanding read functions, read(), readline() and readlines(), Understanding write functions, write() and writelines() ,Manipulating file pointer using seek Programming using file operations

## Reference Text

1. Kenneth A. Lambert,The Fundamentals of Python: First Programs, 2011, Cengage Learning, ISBN: 978-1111822705.
2. Mark Summerfield,Programming in python 3:second edition.
3. Charles Dierbach, “Introduction to Computer Science using Python”, Wiley, 2015
4. R Nageswara Rao, Python Programming

**CA500301--Software Engineering**

## Module I

Introduction-Software engineering, Software process, SE practices, Process models-Generic process models, Prescriptive process model, Specialised process model, The unified process model.

## Module II

Agile Development-Agility, Agility and cost of change, Agile process, Extreme programming, Adaptive software development, Scrum, Dynamic system development method ,Feature driven development, Agile Modeling, Agile Unified Process.

Introduction to UML: Class Diagram, Deployment Diagram, Use-Case Diagram, Sequence Diagram, Communication Diagram, Activity Diagram, State Diagram.

## Module III

Understanding Requirements-Requirement engineering ,Building the Requirement modelRequirement modeling approaches-Scenario based modelling, UML Model that supplement the Use Case, Data modelling concepts Class Based modeling -Class responsibility collaborator modelling, Flow oriented modelling, Creating a behavioural model.

## Module IV

Software Design-Design concepts-The Design Model, Architectural Design- Architectural styles and design, Architectural mapping using data flow, Component level design-Design guidelines, Conducting component level design, Component based development, User Interface Design-Golden rules, Interface design steps.

Testing- Software testing strategy-A Strategic Approach to software testing.Testing conventional applications-White box testing, Black box testing, Testing object-oriented applications-Object-oriented testing methods.

## Module V

Software project management-Software measurement, Metrics for software quality, Software project estimation-Decomposition technique ,Empirical estimation model-The COCOMO11 Model. Project scheduling-basic principles-Defining a task set, Defining a task network, scheduling .Risk management-Software Risks, Risk identification, Risk projection, Risk refinement, THE RMMM PLAN.

## Reference Text

1. Roger S. Pressman, “Software Engineering: A Practitioner’s Approach” .McGrawHill International Edition. 7th Edition.
2. Carlo Ghezzi, Mehdi Jazayeri. Dino Mandrioli : “Fundamentals of software Engineering” , Second Edition.
3. Richard Fairley : Software Engineering Concepts”, Tata McGraw Hill Edition 1997.
4. Martin I. Shooman : “Software Engineering – Design Reliability and Management”,McGraw Hill International Edition.

## CA010303—Lab III DIP &Python Python

1. Programs using elementary data items, lists, dictionaries and tuples
2. Programs using conditional branches, loops.
3. Programs using functions
4. Programs using exception handling
5. Programs using classes and objects
6. Programs using inheritance
7. Programs using polymorphism 8. Programs to implement file operations.

9. Programs using modules.

## Digital Image Processing

Program to input gray scale image and color image, convert image to array of numbers and perform rotations on the image.

1. Program for conversion between colour spaces RGB, CMY, HSI.

Program to find histogram value and display histogram of a grayscale and color image.

1. Program to apply basic intensity transformations.
2. Program to Implement 2-D DFT and Transform domain Filtering.
3. Program to read a grayscale image, corrupt the image using any noise models and apply mean filters or adaptive median filters to remove the noise.

Program for edge detection using gradient operators.

1. Program to Segment the image using Thresholding.

## CA010304—Mini Project using IOT

**CA010305--Internet of Things** **Module I:**

Advanced Programming with Python: Basic operators and variables, Decision Making and Loops, Modules, Exception Handling, Classes, Function, Multithreading, GUI Programming, email using SMTP.

Introduction to IoT**:**

Evolution of internet, Components and architecture of IoT, Types and requirements of IoT network, Protocols, Standards and Communication Technologies., IoT application areas.

**Module II:**

Arduino IDE**:** Arduino Software Development, Interaction of Arduino board With Computers, GPIO Programming with Arduino, ADCs, Custom Library for Arduino IDE.

Sensor Interfacing with Arduino IDE**:** Temperature sensor LM 35, Humidity sensor DHT-11., PIR sensor, Distance Measurement using HC SR 04, Gas Sensor interfacing using MQxx Series, LCD interfacing, SMS using GSM Module, Weight measurement using load cell.

**Module III:**

Raspberry Pi**:**  Linux basics, Linux commands, RPi models, RPi programming languages and Operating Systems, , GPIO Interfacing, 1 wire driver, SPI and I2C protocol, MCP3008 ADC.

Sensor Interfacing with RPi**:** Remote desktop of RPi Using SSH and VNC, LED and switch Interfacing, DS18B20 temperature sensor, Heart beat sensor, RPi FM Radio, Stepper motor, Servo motor, Gas sensor interfacing, Relay interfacing, Arduino and RPi.

**Module IV:**

Image Processing: Introduction to OpenCV, GUI Features, Image Processing in open CV: Image thresholding, Smoothing images, Image Edge Detection, Template Matching, Foreground Extraction using GrabCut Algorithm. Rpi camera interfacing: Camera Calibration, Face Detection using Haar Cascades, Feature Matching.

Cloud: Deployment models of cloud, Cloud configuration using Amazon cloud/thingspeak.

**Module V:**

**Project:**

1. Real time patient monitoring system using IoT and Cloud
2. Real time remote user authentication using face recognition
3. Implementation of a smart vehicle using IoT
4. IoT base waste management/ smart city application
5. Barcode/QR code based library access system
6. Implementation of a smart home using sensors and open CV model
7. Weather station Using IoT and OpenCV
8. Traffic monitoring using IoT and OpenCV
9. Real time video streaming to thingspeak / youtube using RPi
10. Cloud based attendance monitoring using face recognition

## Reference Text

1. Internet of Things: architecture and design principles, TRaj Kamal, McGraw Hill Company
2. Internet of Things: Architectures, Protocols and Standards, By Simone Cirani, Gianluigi Ferrari, Marco Picone, Luca Veltri, Wiley and sons
3. The Internet of Things: How Smart TVs, Smart Cars, Smart Homes, and Smart cities are changing the world By Michael Miller
4. K M Abubeker et al , “IoT based real time patient monitoring and analysis using Raspberry Pi 3, ieeexplore digital library, ICECDS-2017, DOI: 10.1109/ICECDS.2017.8389932

### **Fourth Semester Courses**

|  |  |
| --- | --- |
| CA010401 | Data Mining |
|  | Elective II |
|  | Elective III |
| CA010402 | Project |
| CA010403 | Viva-voce |

**A010401-- Data Mining**

## Module I

Introduction**:** What is Data mining? Data Mining Tasks, KDD process, Data Mining Functionalities, Mining Frequent Patterns, Associations and Correlations, Classification and Prediction, Cluster Analysis, Classification of Data Mining systems, Major issues in Data Mining, Data objects and Attribute types- Nominal, Binary, Ordinal and Numeric attributes, Measuring the central tendency- Mean, Median and Mode. Data Warehouse, Multidimensional Data Model-Data Cubes, Schemas for multidimensional models-Stars, Snowflakes and Fact Constellations.

## Module II

Data Preprocessing: Needs of Pre-processing the Data, Data Cleaning- Missing Values, Noisy Data, Data Cleaning as a Process. Data Integration- Redundancy and correlation analysis, Data Reduction- Attribute Subset Selection, Dimensionality Reduction, Numerosity

Reduction, PCA. Data Transformation strategies, Data transformation by Normalization, Discretization by Binning, Histogram Analysis

## Module III

Association Analysis- Frequent patterns, Basic terminology in association analysis- Binary representation, Itemset and support count, Association Rule, Support and Confidence, Frequent Item set generation- The Apriori Algorithm, Generating Association Rules from Frequent Itemsets, FP Growth algorithm, Pattern evaluation Methods- How strong association rules can be uninteresting and misleading, From Association Analysis to Correlation Analysis, Constraint-Based Frequent pattern Mining, Metarule-Guided Mining of Association Rules.

## Module IV

Classification :- Basic concepts, General approach to classification, Decision Tree Induction,

Basic Decision Tree algorithm, Attribute Selection Measures- Information Gain, Gain Ratio,

Gini Index, Tree Pruning. Bayes Classification methods- Bayes‘ Theorem, Naïve Bayesian Classification, Rule-based Classification - Using IF-THEN Rules for Classification, Rule Extraction from a Decision Tree, Rule Induction Using a Sequential Covering Algorithm. Metrics for evaluating classifier performance, Cross validation. Classification by Back propagation- A Multilayer Feed-Forward Neural Network, Defining a Network Topology, Backpropagation, Inside the Black Box: Backpropagation and Interpretability.

## Module V

Cluster Analysis: Introduction, Basic Clustering methods- Partitioning methods- k-Means and k-Medoid. Hierarchical Methods - Agglomerative and Divisive Hierarchical Clustering. Density Based Methods - DBSCAN, OPTICS, DENCLUE. Grid Based- STING, CLIQUE, Outlier Analysis- what are outliers, Types of outliers, Outlier detection methods - Statistical Distribution-Based Outlier Detection, Distance-Based Outlier Detection.

## Reference Text

1. Jiawei Han & Micheline Kamber,Data Mining, Concepts and Techniques, , 3rd Edition.
2. Pang Ning Tan, Michael Steinbach and Vipin Kumar,Introduction to Data Mining, Pearson India Education Services
3. Arun K Pujari, Data Mining Techniques, , University Press
4. Sam Anahory& Dennis Murray,Data Warehousing in the Real World, Pearson Education, Asia.
5. PaulrajPonnaiah, Data Warehousing Fundamentals,Wiley Student Edition

**2. Electives**

**Group A**

**CA800301**-**Introduction to Cyber Security**

## Module 1

What Is Computer Security?, Values of Assets, Threats, Confidentiality, Integrity, Availability,, Types of Threats, Types of Attackers, Harm, Risk and Common Sense, Method–Opportunity–Motive, Controls, Authentication, Identification Versus Authentication, Authentication Based on Phrases and Facts: Something You Know, Authentication Based on Biometrics: Something You Are, Authentication Based on Tokens: Something You Have.

## Module 2

Access Control, Access Policies, Implementing Access Control, Procedure-Oriented Access Control, Role-Based Access Control.Cryptography, Problems Addressed by, Encryption,

Terminology.Malicious Code—Malware, Malware—Viruses, Trojan Horses, and Worms, Technical Details: Malicious Code.Countermeasures for Users.Email Attacks, Fake Email, Fake Email Messages as Spam, Fake (Inaccurate) Email Header Data Phishing, Protecting Against Email Attacks.

## Module 3

Security in Operating Systems, Security Features of Ordinary Operating Systems, Protected Objects, Operating System Tools to Implement Security, Functions, Security in the Design of Operating Systems ,Simplicity of Design, Layered Design, Kernelized Design Reference Monitor, Correctness and Completeness, Secure Design Principles Trusted Systems, Trusted System Functions.

## Module 4

Threats to Network Communications Interception: Eavesdropping and Wiretapping, Modification, Fabrication: Data Corruption, Interruption: Loss of Service, Port Scanning.

Denial of Service, How Service Is Denied, Flooding Attacks in Detail, Distributed Denial-ofService, Scripted Denial-of-Service Attacks ,Bots, Botnets, Firewalls, What Is a Firewall?, Design of Firewalls Types of Firewalls, Intrusion Detection Systems, Types of IDSs. **Module 5**

Security Requirements of Databases, Integrity of the Database, Element Integrity, Auditability, Access Control, User Authentication,Availability,

Integrity/Confidentiality/Availability.

Information Technology Act 2000, Cyber Crimes- Computer Crime, Nature of Crimes, Penalty for damage to computer, Computer system, Tampering with Computer source documents, Hacking, Computer related offences.

## Reference Text

1.Charles P. Pfleeger ,Shari Lawrence Pfleeger ,Jonathan Margulies - Security in Computing, Fifth Edition

2**.**Barkhs and U. Rama Mohan, “Cyber Law Crimes”, Asia Law House, New Edition

3. Sood,“Cyber Laws Simplified”, Mc Graw Hill

**CA800402** – **Cryptography**

## Module 1

Classical Encryption Techniques-Symmetric Cipher Model, Substitution Techniques, Transposition Techniques, Steganography. Traditional Block Cipher Structure, The Data Encryption Standard, The Strength of DES, Block Cipher Design Principles.

## Module 2

Advanced Encryption Standard-AES Structure, AES Transformation Functions, AES Key Expansion, AES Implementation. Multiple Encryption and Triple DES. Principles of Pseudorandom Number Generation, Pseudorandom Number Generators, Stream Ciphers, RC4, True Random Number Generators.

## Module 3

Principles of Public-Key Cryptosystems, The RSA Algorithm, Diffie-Hellman Key Exchange, Cryptographic Hash Functions, Applications of Cryptographic Hash Functions, Secure Hash Algorithm (SHA).

## Module 4

Message Authentication Codes, Message Authentication Requirements, Message Authentication Functions, Requirements for Message Authentication Codes, MACs Based on Hash Functions: HMAC.

## Module5

Key Management and Distribution, Symmetric Key Distribution Using Symmetric Encryption, Symmetric Key Distribution Using Asymmetric Encryption, Distribution of Public Keys, X.509 Certificates, Public-Key Infrastructure, Digital Signatures- Digital Signatures.

Reference Text

1.William Stallings, Cryptography and Network Security-Principles and Practice, Sixth Edition 2014.

2. BruiceSchneier, Applied Cryptography, Second edition, Wiley publications, 2006

**CA800403**- **Ethical Hacking**

## Module 1

Ethics of Ethical Hacking, Ethical Hacking and the Legal System, Proper and Ethical Disclosure, Social Engineering Attacks, Physical Penetration Attacks, Insider Attacks.

## Module 2

Vulnerability Analysis- Passive Analysis, Advanced Static Analysis with IDA Pro, Advanced Reverse Engineering, Client-Side Browser Exploits, From Vulnerability to Exploit.

## Module 3

Hacking windows – Network hacking – Web hacking – Password hacking. A study on various attacks – Input validation attacks – SQL injection attacks – Buffer overflow attacks - Privacy attacks.

## Module 4

TCP / IP – Checksums – IP Spoofing port scanning, DNS Spoofing. Dos attacks – SYN attacks, Smurf attacks, UDP flooding, DDOS – Models. Firewalls – Packet filter firewalls, Packet Inspection firewalls – Application Proxy Firewalls. Batch File Programming.

## Module 5

Basic Linux Exploits, Advanced Linux Exploits, Collecting Malware and Initial Analysis, Hacking Malware.

Reference Text

1. Ankit Fadia“ Ethical Hacking” 2nd Edition Macmillan India Ltd
2. Allen Harper, Shon Harris, Jonathan Ness, Chris Eagle, Gideon Lenkey, and Terron Williams Gray Hat Hacking The Ethical Hacker’s Handbook ,, Third Edition.

## Group B

**CA810301** – **Statistical Computing for Data Analytics**

**Module – I** Data Analytics Life Cycle

Introduction to Big data Business Analytics - State of the practice in analytics role of data scientists - Key roles for successful analytic project - Main phases of life cycle - Developing core deliverables for stakeholders.

**Module – II** Statistics

Sampling Techniques - Data classification, Tabulation, Frequency and Graphic representation

-

Measures of central value - Arithmetic mean, Geometric mean, Harmonic mean, Mode, Median, Quartiles, Deciles, Percentile - Measures of variation – Range, IQR, Quartile deviation, Mean deviation, standard deviation, coefficient variance, skewness, Moments & Kurtosis.

**Module – III** Probability And Hypothesis Testing

Random variable, distributions, two dimensional R.V, joint probability function, marginal density function. Random vectors - Some special probability distribution - Binomial, Poison, Geometric, uniform, exponential, normal, gamma and Erlang. Multivariate normal distribution - Sampling distribution – Estimation - point, confidence - Test of significance, 1&

2 tailed test, uses of t-distribution, F-distribution, χ2 distribution.

**Module – IV** Predictive Analytics

Predictive modeling and Analysis - Regression Analysis, Multicollinearity , Correlation analysis, Rank correlation coefficient, Multiple correlation, Least square, Curve fitting and good ness of fit.

**Module – V** Time Series Forecasting And Design Of Experiments

Forecasting Models for Time series : MA, SES, TS with trend, season - Design of Experiments, one way classification, two way classification, ANOVA, Latin square, Factorial Design.

Reference Text

1. Chris Eaton, Dirk Deroos, Tom Deutsch et al., “Understanding Big Data”, McGrawHIll, 2012.
2. Alberto Cordoba, “Understanding the Predictive Analytics Lifecycle”, Wiley, 2014.
3. Eric Siegel, Thomas H. Davenport, “Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die”, Wiley, 2013.
4. James R Evans, “Business Analytics- Methods, Models and Decisions”, Pearson 2013 5. R N Prasad, Seema Acharya, Fundamentals of Business Analytics, Wiley 2015

CA810402 – **Big data Management using R**

## Module I

Introduction to Big DataAnalytics Big Data Overview - Data Structures-Analyst Perspective on Data Repositories. State of the Practice in Analytics – BI versus Data Science-Current analytical architecture. Drivers of big data Emerging big data Ecosystem and a new approach to Analytics. Key Roles for the New Big Data Ecosystem, Examples of Big Data Analytics.

## Module II

Data Analytics Lifecycle **:**Data Analytics Lifecycle Overview – Key roles for a successful

Analytics project. Background and overview of data analytics life cycle. Phase 1: Discovery -

Learning the business domain-resources-framing the problem-identifying key stakeholdersInterviewing the analytics sponsor-developing initial hypotheses-Identifying Potential Data sources. Phase 2: Data Preparation - Preparing the Analytic Sandbox-performing ETLTLearning about the data. Data conditioning-Survey and visualize-Common tools for the data preparation phase. Phase 3: Model Planning: Data exploration and variable selection. Model selection – common tools for the model planning phase- Phase 4: Model Building – common tools Phase 5: Communicate Results - Phase 6: Operationalize.

## Module III

Text Analysis: Text Analysis Steps - A Text Analysis Example - Collecting Raw Text - Representing Text - Term Frequency - Inverse Document Frequency (TFIDF) Categorizing Documents by Topics. **Advanced analytics**: Analytics for unstructured data. Use cases –Map reduce-Apache Hadoop. The Hadoop echosystem: Pig-Hive-HBase-Mahout-NoSQL.

## Module IV

Communicating and Operationalizing an analytics project. Creating the final deliverables.

Developing core material for multiple Audiences-Project goals- Main findings – approach-

Model description- Key points supported with data. Model details- RecommendationsAdditional tips on final presentation-Providing technical specification and code. Data visualization basics: Key Points Supported with data. Evaluation of a graph-Common representation methods-How to clean up a graphic - Additional considerations.

## Module V

Introduction to R **–** Basics - Download & Install R, RStudio - R Data Types: Arithmetic & Logical Operators - R Matrix: Create, Print, add Column, Slice - Data Frame: Create, Append, Select, Subset, Factor in R: Categorical & Continuous Variables - **Data Preparation -** R Data Frame: Create, Append, Select, Subset **-** List in R: Create, Select Elements - R Sort a Data Frame using Order() - R Dplyr: Data Manipulation(Join) & Cleaning(Spread) - Merge Data Frames in R: Full and Partial Match - Functions in R Programming, IF, ELSE, ELSE IF, For Loop in R using List and Matrix - While Loop in R - apply(), lapply(), sapply(), tapply() Function in R - Import Data into R: Read CSV, Excel - Replace Missing Values(NA) in R - R Exporting Data to Excel, CSV, Text File - Correlation in R: Pearson & Spearman with Matrix Example - R Aggregate Function: Summarise &Group\_by() - R Select(), Filter(), Arrange(), Pipeline - **Data Analysis -** Scatter Plot in R using ggplot2 - Boxplot in R - Bar Chart & Histogram in R.

## Reference Text

1. EMC Education Services, “Data Science and Big Data Analytics”, WILEY
2. Bart Baesens – “Analytics in a Big Data World “, WILEY
3. Mark Hornick, Tom Plunkett - “Using R to Unlock the Value of Big Data “
4. R programming for Data Science – Roger D Peng

CA810403 – **Data Analytics**

**Module I:** Introduction to Big Data

Introduction to Data Analytics Platforms–Traits of Big data -Challengesof Conventional Systems - Web Data – Evolution Of Analytic Scalability - Analytic Processes and Tools - Analysis vs Reporting - Modern Data Analytic Tools - Statistical Concepts: Sampling Distributions - Re- Sampling - Statistical Inference - Prediction Error, Ethics in Big Data Analytics.

**Module II:** Data Analysis

Regression Modeling - Multivariate Analysis - Bayesian Modeling - Inference and Bayesian Networks - Support Vector and Kernel Methods - Analysis of Time Series: Linear Systems Analysis -

Nonlinear Dynamics - Rule Induction - Neural Networks: Learning And Generalization - Competitive Learning - Principal Component Analysis and Neural Networks - Fuzzy Logic: Extracting Fuzzy Models from Data - Fuzzy Decision Trees - Stochastic Search Methods.

**Module III:** Mining Data Streams

Introduction To Streams Concepts–Stream Data Model andArchitecture - Stream Computing - Sampling Data in a Stream – Filtering Streams – Counting Distinct Elements in a Stream – Estimating Moments – Counting Oneness in a Window – Decaying Window - Real time Analytics Platform (RTAP) Applications.

**Module IV:** Frequent Itemsets

Mining Frequent Itemsets - Market Based Model–Apriori Algorithm–Handling Large Data Sets in Main Memory – Limited Pass Algorithm – Counting Frequent Itemsets in a Stream **Module V: Clustering (16 Hours)** Clustering Techniques–Hierarchical–K-Means–Clustering High DimensionalData – CLIQUE And PROCLUS – Frequent Pattern based Clustering Methods – Clustering in Non- Euclidean Space

– Clustering for Streams and Parallelism.

## Reference Text

1. Michael Berthold, David J. Hand, “Intelligent Data Analysis”, Springer, 2007.
2. AnandRajaraman and Jeffrey David Ullman, “Mining of Massive Datasets”, Cambridge University Press, 2012.
3. Bill Franks, “Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics”, John Wiley & sons, 2012.
4. Glenn J. Myatt, “Making Sense of Data”, John Wiley & Sons, 2007
5. Pete Warden, “Big Data Glossary”,

O’Reilly, 2011.

1. Jiawei Han, MichelineKamber “Data Mining Concepts and Techniques”, Second Edition, Elsevier, Reprinted 2008.

**Group C**

**CA820301** – **SoftComputing.**

## Module I

Introduction :- Neural networks, Artificial network, Advantages of networks, Application scope of neural networks, Fuzzy logic, genetic algorithm, Hybrid systems – classification soft computing. Artificial neural network, Fundamental concept, Artificial neural network, biological neural network, Brains Vs computer comparison between biological neuron and artificial neuron

## Module II

,Evolution of neural networks, Basic models of ANN, important terminologies of ANNS, McCulloch-Pitts neuron, Hebb network. Supervised learning network – perception networks, Adaline, multiple Adaptive linear neurons (Madaline),Functional link networks ,tree neural networks,wavelet neural networks.

## Module III

Back propagation networks, radial basis function network,time delay neural network,Associative memory networks. Auto associative memory network, Hetero associative memory network BAM, Hop filed network, unsupervised Learning networks,Fixed weight competitive nets, Kohonenself organising maps, Learning vector quantization.

## Module IV

Introduction to Fuzzy logic ,classical relations and fuzzy relations,tolerance and equivalence relations,non interactive fuzzy sets,membership functions, features of membership functions ,fuzzification,methods of membership value assignments,defuzzification, lambda- cuts , defuzzificationmethods,fuzzy arithmetic and fuzzy measures,fuzzyintegrals,fuzzyrulebaseand approximate reasoning,truthvalues and tables in fuzzylogic,fuzzy proposition,formation of rules, fuzzy reasoning,fuzzy inference systems,**overview** of fuzzyexpert system.

## Module V

Genetic algorithm, Applications of GA, Biological background ,Genetic algorithms Vs traditional algorithms, Basic terminologies in genetic algorithm, simple GA ,General GA, operators in GA ,Encoding, selection, crossover, mutation.

Reference Text

1. S.N.SivanandamS.N.Deepa ,”Principles of soft computing “ second edition, Wiley India Pvt. Ltd .
2. J.S.R Jang, C.T. son ,E. MIZUTANI “Neuro Fuzzy and soft computing” first edition, ,pearson education.

3 .**S**Rajasekaran, G. A.Vijayalakshmi **“**Neural networks, Fuzzy logic and genetic Algorithms :

synthesis and application”, second edition ,prentice Hall of India.

**CA820402**– **Advanced Python Programming**

## Module I

Python Object Oriented Programming-Concept of class, object and instances, Constructor, class attributes and destructors, Real time use of class in live projects, Inheritance , overlapping and overloading operators, Adding and retrieving dynamic attributes of classes ,Programming using Oops support.

## Module II

Python Regular Expression-Powerful pattern matching and searching Power of pattern searching using regex in python Real time parsing of networking or system data using regex Password, email, url validation using regular expression Pattern finding programs using regular expression

## Module III

Python Exception Handling-Avoiding code break using exception handling, Safe guarding file operation using exception handling, Handling and helping developer with error code, Programming using Exception handling.

## Module IV

Python Database Interaction -SQL Database connection using python, Creating and searching tables, Reading and storing config information on database, Programming using database connections.

Graphical user interfaces; event-driven programming paradigm; tkinter module, creating simple GUI; buttons, labels, entry fields, dialogs; widget attributes - sizes, fonts, colors layouts, nested frames.

## Module V

Python Multithreading-Understanding threads, Forking threads, Synchronizing the threads, Programming using multithreading.

Contacting User through Emails Using Python- installing smtp python module, Sending email, Reading from file and sending emails to all users addressing them directly for marketing

Python CGI Introduction -Writing python program for CGI applications, Creating menus and accessing files, Server client program

## Reference Text

1. Mark Summerfield, Programming in python 3, second edition.
2. Clinton W Brownley, Foundations for Analytics with Python.

**CA820403**– **Pattern Recognition**

## Module I

Introduction – Pattern recognition systems – The design cycle –Learning and Adaptation -

Bayesian Decision theory - Introduction - Continuous features - two-category classification - Minimum error rate classification - Classifiers, Discriminant functions and Decision Surfaces – The normal density - Discriminant Functions for the Normal Density- Error probabilities and Integrals

## Module II

Parameter estimation and supervised learning - Maximum likelihood estimation - Bayesian estimation – Bayesian Parameter Estimation Gaussian case and general theory - Nonparametric techniques – Density estimation - Parzen Windows - kn-Nearest Neighbour Estimation - Nearest-Neighbour Rule – k-Nearest Neighbour Rule.

## Module III

Linear Discriminant Functions - Linear discriminant functions and decision surfaces – Generalized linear discriminant functions – Two-category linearly separable case - Nonseparable behavior - Linear programming algorithms - Support vector machines - Multilayer neural networks – Feedforward operation and classification - Backpropagation algorithm - Error surfaces - Backpropagation as feature mapping.

## Module IV

Stochastic methods – Stochastic search- Boltzmann learning – Nonmetric methods - Decision trees – CART – Other tree methods - Grammatical methods - Grammatical inference.

## Module V

Unsupervised learning and clustering – Mixture densities and identifiability – Maximumlikelihood estimates - Applications to normal mixtures - Unsupervised Bayesian learning - Data description and clustering.

Reference Text

1. R.O.Duda, P.E.Hart and D.G.Stork, Pattern Classification, Second edition, John Wiley, 2006
2. Gonzalez R.C. & Thomson M.G., Syntactic Pattern Recognition - An Introduction, Addison Wesley.
3. Fu K.S., Syntactic Pattern Recognition And Applications, Prentice Hall
4. RajanShinghal, Pattern Recognition: Techniques and Applications, Oxford University Press, 2008.