Integrated M.Sc. Programme
in
Computer Science

DATA SCIENCE

Mahatma Gandhi University, Kottayam

Preamble

The syllabus and curriculum of Integrated M.Sc. Programme in Computer Science offered by the University is so designed to make a set of highly talented and skilled computer professionals enough to meet the requirements of software industry, research and development as well as in academic arena of Computer Science. The five year programme spread across 10 semesters inculcates the knowledge in Mathematics and Statistics essential for a computer professional in addition to computer programming skills.

The Integrated M.Sc. Programmes are designed as an alternative to the M.Sc. programmes after the B.Sc. degree . A candidate joining the programme have to take serious efforts during the period to acquire the essential skills. The integrated programmes enables the aspirant to complete the Post Graduate programme in a single stretch.

On successful completion of the programme, the candidate will be able to meet all expectations of the industry or can pursue research work leading to further laurels. The integrated programme syllabus consists of the fundamental papers of Computer Science as well as the latest additions on topics like Machine learning, Artificial Intelligence, Big Data and Internet of Things.

The syllabus of these Integrated Programmes give more stress on practical knowledge than conventional courses in Computer Science. The final semester is compulsorily meant for a project work / internship of six months duration, helping the candidate to be an IT professional. As part of Internship, the candidate can study any of the technology or package of industrial demand, which may not be covered in their syllabus.

Programme Objective

The Integrated M.Sc. Programme in Computer Science in Mahatma Gandhi University is meant to be introduced from Academic year 2020-21. The Primary objectives of the Programme are

a. To attract young aspirants to the field of Computer Science to join the newly introduced pattern of Integrated PG programmes , Complete UG with PG as a single package .
b. The programme can be a milestone for further research and development as well as a successful career in software industry.
c. Aspirants can acquire entrepreneurial skills in software development enabling the successful launching of start-ups or self-employment in IT related platforms.
d. To generate skilled professionals as per the demands of industry.

Eligibility Criteria for admission to Integrated M.Sc. in Computer Science Data Science

The eligibility for admission is recommended as Candidate shall be required to have passed the Plus Two in Kerala Syllabus or any other equivalent examination in Science stream with Physics, Chemistry and Mathematics/Computer Science as compulsory subjects.

OR

Candidate shall be required to have passed the Plus Two in CBSE / ICSE examination or any other equivalent examination in Science stream with Mathematics/Computer Science/Informatics Practices/ Information Technology as one of the optional subjects.

The Programme Overview

The integrated M.Sc. programme in Computer science includes English as the first language, a second language from the list of languages approved by the University, Mathematics and Statistics as complementary subjects in first four semesters in addition to core courses in Computer Science in various semesters.

Complementary papers “Graph Theory and Operations Research” and “Linear Algebra “ are common for both integrated courses in the first two semesters. For Integrated M.Sc. in Computer Science -Data Science Statistics as the complementary paper in the third and fourth semesters during these semesters.

Duration of the Programme

The integrated programme is designed to be of Five academic years containing 10 semesters. It is a single programme with no exit recommended after the first three years (graduation under normal circumstances.

COURSE STRUCTURE

AND

SCHEME

Integrated MSc Computer Science Data Science

Sl. No.	Course Code	Title	Hrs/ Week	Theory Hrs/ Weeks	Lab Hrs/ Weeks	Lab/ Theory	Type of Course	Credits	Total Hours
Semester I									
1	IEN1CC01	English 1	5	5		Theory	Common Course	4	90
2	ICSC1CR2	Programming in C	4	4		Theory	Core Course	3	72
3	ICSC1CR3	Introduction to Computer	3	3		Theory	Core Course	3	54
4	ICSC1CR4	Database Management Systems	3	3		Theory	Core Course	3	54
5	ICSC1CM5	Graph Theory and Operations Research	4	4		Theory	Complementary	4	72
6	ICSC1CP6	Software Lab I	6		6	Lab	Core Practical	4	108
		TOTAL	25	19	6			21	450
Semester II									
1	IML/IHN2CC01	Second Language	5	5		Theory	Common Course	4	90
2	ICSC2CR2	Object Oriented Programming Using C++	3	3		Theory	Core Course	3	54
3	ICSC2CR3	Data Structures using C++	3	3		Theory	Core Course	3	54
4	ICSC2CR4	Operating Systems	4	4		Theory	Core Course	4	72
5	ICSC2CM5	Linear Algebra	4	4		Theory	Complementary	4	72
6	ICSC2CP6	Software Lab II	6		6	Lab	Core Practical	4	108
		TOTAL	25	19	6			22	450
Semester III									
1	ICSD3CR1	Introduction to Data Science	4	4		Theory	Core Course	4	72
2	ICSC3CR2	Programming in Python	3	3		Theory	Core Course	3	54
3	ICSC3CR3	R Programming and Mathematics for Artificial Intelligence	4	4		Theory	Core Course	3	72
	ICSC3CR4	Computer Organization and Architecture	4	4		Theory	Core Course	4	72
5	ICSD3CM5	Probability and Statistics	4	4		Theory	Complementary	4	72
6	ICSD3CP6	Software Lab III: Python and R Programming	6		6	Lab	Core Practical	2	108
		TOTAL	25	19	6			20	450

Sl. No.	Course Code	Title	Hrs./ Week	Theory Hrs./ Weeks	Lab Hrs./ Weeks	Lab/ Theory	Type of Course	Credits	Total Hours
Semester IV									
1	IEN4CC01	English II	5	5		Theory	Common Course	4	90
2	ICSD4CM2	Probability Distributions and Statistical Inference	4	4		Theory	Complementary	3	72
3	ICSC4CR3	Data Mining	4	4		Theory	Core Course	4	72
4	ICSC4CR4	Software Engineering	4	4		Theory	Core Course	3	72
5	ICSC4CR5	Basics of Artificial Intelligence	4	4		Theory	Core Course	3	72
6	ICSD4CMP6	Complementary Lab: R Programming For Statistical Tools	2		2	Lab	Complementary Practical	2	36
7	ICSD4CP7	Software lab IV	2		2	Lab	Core Practical	1	36
		TOTAL	25	21	4			20	450
Semester V									
1	ICSC5CR1	Principles of Machine learning	3	3		Theory	Core Course	4	54
2	ICSC5CR2	Web Application Development Using PHP	4	4		Theory	Core Course	3	72
3	ICSC5CR3	Programming in Java	4	4		Theory	Core Course	3	72
4	ICSC5CR4	IT and Environment	3	3		Theory	Core Course	4	54
5	ICSD5PR5	Project Minor - Phase I	3		3		Core Project Minor		54
6	ICSD5CP6	Software Lab V: Java and PHP	8		8	Lab	Core Practical	3	144
		TOTAL	25	14	11			17	450
Semester VI									
1	ICSC6CR1	Linux and Shell Programming	3	3		Theory	Core Course	3	54
2	ICSC6CR2	Computer Networks	4	4		Theory	Core Course	4	72
3	ICSD6CR3	Mobile Application Development using Kotlin	4	4		Theory	Core Course	4	72
4	ICSC6EA1/2/3	Elective 1 [Bunch A]	3	3		Theory	Core Elective	3	54
5	ICSD6PR4	Project Minor Phase II	7		7	Lab	Core Project Minor	4	126
6	ICSD6CP5	Software Lab VI: Mobile Application Development using Kotlin	4		4	Lab	Core Practical	2	72
		TOTAL	25	14	11			20	450

Sl. No.	Code	Title	Hrs/ Week	Theory Hours/ Week	Lab Hours/ Week	Lab/ Theory Exam	Course Type	Credits	Total Hours
Semester VII									
1	ICSC7CR1	Computational Mathematics	4	4		Theory	Core Course	4	72
2	ICSD7CR2	Applied Statistics for Data Science	4	4		Theory	Core Course	4	72
3	ICSD7CR3	Advanced Python Programming for data Science	4	4		Theory	Core Course	4	72
4	ICSD7CR4	Data Engineering in Data Science	4	4		Theory	Core Course	3	72
5	ICSD7CP5	Software Lab VII : Python Programming Lab for data Science	5		5	Lab	Core Practical	3	90
6	ICSD7CP6	Software Lab VIII : Data Engineering Lab	4		4	Lab	Core Practical	2	72
		TOTAL	25	16	9			20	450
Semester VIII									
1	ICSD8EB1/2	Elective 2 [Bunch B]	4	4		Theory	Core Elective	4	72
2	ICSC8CR1	Advanced Deep Learning Techniques	4	4		Theory	Core Course	4	72
3	ICSD8CR2	Data Visualization	4	4		Theory	Core Course	4	72
4	ICSD8EC1/2	Elective 3 [Bunch C]	4	4		Theory	Core Elective	4	72
5	ICSD8CP3	Software Lab IX : Deep Learning Lab using R	4		4	Lab	Core Practical	2	72
6	ICSD8CP4	Software Lab X : Data Visualization Lab using Tableau	5		5	Lab	Core Practical	2	90
		TOTAL	25	16	9			20	450
Semester IX									
1	ICSD9ED1/2	Elective 4 [Bunch D]	4	4		Theory	Core Elective	4	72
2	ICSD9CR1	Text Analytics & Natural Language Processing	4	4		Theory	Core Course	3	72
3	ICSD9CR2	Web Analytics	4	4		Theory	Core Course	4	72
4	ICSC9EE1/2	Elective 5 [Bunch E]	4	4		Theory	Core Elective	4	72
5	ICSD9CP3	Software Lab XI : NLP using R/Python	4		4	Lab	Core Practical	2	72
6	ICSD9PR4	Case study and Minor project	5		5	Lab	Core Project Minor	3	90
		TOTAL	25	16	9			20	450
Semester X									
1	ICSDXPR1	Major Project	25		25	Lab	Core Project Major	16	450
2	ICSDXVV2	Comprehensive Viva Voce				Lab	Core Viva Voce	4	
		TOTAL	25		25			20	450

Electives: Integrated MSc Computer Science- Data Science

Bunch A	
ICSC6EA1	Cloud Computing
ICSC6EA2	Full stack programming Techniques
ICSC6EA3	Predictive Analytics
Bunch B	
ICSD8EB1	Advanced DBMS
ICSD8EB2	Business Intelligence & Analytics
Bunch C	
ICSD8EC1	Image and Video Analytics
ICSD8EC2	Geospatial Analysis
Bunch D	
ICSD9ED1	Healthcare data Analytics
ICSD9ED2	Social media Analytics
Bunch E	
ICSC9EE1	Block Chain Technology
ICSC9EE2	Big Data Analytics

SEMESTER - I

Si. No.	Course Code	Title	Hrs./ Week	Credits	Total hours
1	IEN1CC01	English - I (T)	5	4	90
2	ICSC1CR2	Programming in C (T)	4	3	72
3	ICSC1CR3	Introduction to Computer (T)	3	3	54
4	ICSC1CR4	Database Management Systems (T)	3	3	54
5	ICSC1CM5	Graph Theory and Operations Research (T)	4	4	72
6	ICSC1CP6	Software Lab I (L)	6	4	108
		Total	25	21	450

- Syllabus of English paper: IEN1CC01 English I: As approved by BoS of English (UG).
- Syllabus of complementary paper Mathematics: ICSC1CM5 Graph Theory and Operations Research: As approved by BOS of Mathematics (UG).

Course Title: Programming in C

Course Code: ICSC1CR2
Total Credits: 3

Course Objectives

On completion of the course:

- the student will be able to write a complete C program
- he/she will be able to use decision making statements and looping structures
- should have a clear concept on one-dimensional, two-dimensional arrays, modular programming using user defined functions
- clarity on concept of strings, structures and Unions
- should be able to use files for input and output
- basic ideas on dynamic storage allocation and command line arguments

Module 1:

Algorithm and flow chart (structure, desirable qualities, simple examples (sequential, branching and iterative)). Basic concepts in machine language program and program execution, assembly language program and assembler, High level language program and translators (compilers and interpreters). Procedural programming paradigm - examples. Steps in developing a program - (problem analysis, algorithm design, coding, debugging, testing, documentation). Approaches - top down and bottom-up approaches. C Language - structure of a C program - simple sequential program. Role of editor. Compilation, linking and execution under Windows and Linux. IDEs. Types of errors. (14 Hrs.)

Module 2:

Keywords, constants, variables, data types and variable names, assignment statement. Operators and expressions (including increment, decrement and sizeof()), precedence and order of execution, mixed mode expressions and type conversions. Elementary ideas in function (C program as a collection of functions), main function. Formatted input and output. Simple sequential programs. Decision making: The goto statement, if, if-else, nesting of if, else if ladder and switch statement, conditional expression. Example programs based on decision making. (12 Hrs.)

Module 3:

Control statements: The while loop, the do...while loop, the for loop, nesting of for loops, the break statement and continue statement. Example programs. Functions, basics, prototype, parameter passing, storage classes, recursion. Built-in functions. Example programs. (14 Hrs.)

Module: 4

Arrays, arrays and functions. Strings, string operations and algorithms, string functions in C. Example programs using arrays and strings (including simple search and sort, matrix operations). Pointers - basic concepts, pointer arithmetic, pointers and arrays, pointers and strings, pointers and functions. Dynamic memory allocation, Simple programs using pointers. (16 Hrs.)

Module 5:

Structures - basics, array of structures, pointers and structures, structure and function, self-referential structures, union. Programs using structures. The Pre-processor: File Inclusion, Macro Definition and Substitution, Macros with Arguments, Nesting of Macros, Conditional Compilation. (Simple illustrative examples) File Management: Defining and Opening a file, Closing Files, Input/Output Operations on Files, Predefined Streams, Error Handling during 1/0 Operations, Random Access to Files, Command Line Arguments. Simple examples of file creation. (16 Hrs.)

Book of Study

1. Programming in Ansi C: E Balagurusami 8th edition MC GRAW HILL INDIA publishers

References

1: Programming with C: Byron S Gottfried, schaumes outline series 4th edition
2: Programming in C: Ashok n Kamthane: Pearson education 3rd edition
3: Let us C: Yeshwant Kanetkar 16th edition BPB publishers

Course Title: Introduction to Computer

Course Code: ICSC1CR3

Total Credits: 3

Course Objectives

After successful completion of the course the students will be able to:

- understand basic functions of computer hardware, software components including memory & operating systems
- understand the concept of networking and internet
- understand IT and its impact on society.

Module: 1

Introduction: Functional units of a computer system, Different types of computers, Computer Software and Hardware, Types of Software (System software and Application software). Characteristics of Computers. Computer Languages (Machine, Assembly and Higher-Level languages- 36L,4GL,5GL). (8 Hrs.)

Module: 2

Interaction with Computers, Data Processing and Storage Information: Input devices, Output devices. Representation of Data, Processing of Data, the CPU, Memory, different types of RAM and ROM. Types of Storage devices (Magnetic storage devices, Optical storage devices, Solid state storage devices), SSD-types, performance, benefits; Graphics Processing Unit (GPU) . (12 Hrs.)

Module: 3

Introduction to Operating Systems, Networking: Definition of an Operating System, Different types of PC Operating Systems, File Management (file access methods, file operations, file naming). Computer Network: Basic elements of communication system, Data transmission modes. Data transmission speed, Data transmission media (twisted pair wire, coaxial cable, Microwave system, Communication satellite, Optical fibers), Modems, Categories of networks (PAN, CAN, LAN, WAN, MAN). (12 Hrs.)

Module: 4

Internet: Definition, Working of Internet, Major features of Internet, Major services (WWW, Electronic mail, FTP, Chat, Instant messaging, Telnet, Usenet News, Online services, Peer-to-peer services), TCP/IP, URL's, Web Browsers, Major elements of Internet Search engines, Popular Search Engines, Uses of the Internet, WWAN, Academic service (INFLIBNET, NOTEL, NICNET, BRNET).(12 Hrs.)

Module: 5

Introduction to Cyber World: Cyber space, introducing cyber laws, Scope of cyber laws (E-commerce ,online, contracts, IPRs(copyrights ,trademarks, and software patenting) ; cyber ethics, Cyber Addiction- types of internet addiction, causes of cyber addiction, effects of Internet addiction; Cyber Crimes-Introduction, categories of cybercrime, types of cybercrimes. (10 Hrs.)

Book of Study

1. Peter Norton's- Introduction to Computers, Sixth edition, Tata McGraw Hill
2. P.K Sinha & Priti Sinha - Computer Fundamentals, Fourth Edition, BPB Publications.
3. Barkha and U. Rama Mohan - Cyber Law Crimes, Asia Law House, New edition

References

1. V Rajaraman " Introduction to Information technology" Prentice-Hall of India.
2. Harley Hahn - The Internet Complete reference Tata Mc Graw-Hill edition
3. Dr. Farooq Ahamad- Cyber law in India (Law of Internet), New Era Law publication

Course Title: Database Management Systems

Course Code: ICSC1CR4
Total Credits: 3

Course Objectives

On completion of the course, the student should have:

- a clear concept on databases, data models, architecture and components od DBMS
- the concept of entity, attributes, associations and relationships concept of tables and its properties, table creation and manipulation of tables and databases using SQL.
- the concept of DDL and DML facilities.

Module: 1

Database Management System Concepts: Introduction, Significance of Databases, Advantages of database approach; Data Independence; Components of Database Systems, classification of Users, the Database Administrator (DBA) and his responsibilities; advantages and disadvantages of Database Management System. (9 Hrs.)

Module: 2

Entity attributes and Data Models for a Database; Entities and their Attributes, different types of Entities and Attributes, Association and relationships and their different types E-R Diagrams. Data Models, Hierarchical, Network and Relational data models. Benefits and Application of each Data models. (9 Hrs.)

Module: 3

DBMS Architecture and Schema, Data Dictionary. Three Level Architecture of DBMS, The External Level or Subschema, The Conceptual Level or Conceptual Schema, The Internal Level or Physical Schema, Data Definition Language, Data Manipulation Language: Database Management System Structure, Database Manager, Database Administrator, Data Dictionary: Brief introduction to Distributed databases and Client /Server Architecture. (10 Hrs.)

Module: 4

The Relational Approach to DBMS: The Concept of relations in Mathematics, Mathematical concept of sets, relations and functions, Relational OOOO approach to DBMS Attributes and Domains ; concept and properties of tables, cardinality and degree of relations, keys and different types of keys; strong entities and weak entities Entity integrity rule, the foreign key and rule of referential integrity. Representation of relational database schemas, integrity constraints and different types; Relational Algebra. Operators in Relational Algebra. (12 Hrs.)

Module: 5

The Structured Query Language (SQL). The need for SQL. Brief introduction to query languages and its evolution Basic structure of SQL queries, Data Definition Commands : Data types in SQL, CREATE , ALTER „DROP Commands Adding constraints in SQL, Basic operations in Data Manipulation Using SQL, INSERT, SELECT, DELETE, UPDATE, Substring comparison using LIKE operator, BETWEEN operator, SQL set operations UNION, EXCEPT, INTERSECT: order By and Group By clauses, complex queries in SQL, Nested queries, EXISTS and UNIQUE functions, Renaming of attributes and Joining of tables, Aggregate functions, Creating and Managing Views. (14 Hrs.)

Book of Study

1. Raghu Ramakrishnan & Johannes Gehrke, "Data Base Management Systems", Mc Graw Hill International Edition

References

1. Fundamentals of Database systems 7th edition: Elmasri & Navathe : Pearson Education
2. An introduction to database systems C.J Date 8th edition. Pearson education
3. Abraham Silberschatz, Hentry F.Korth and S.Sudharssan,"Database System Concepts", 4th Edition, Tata McGraw Hill.

Course Title: Graph Theory and Operations Research

Course Code: ICSC1CM5

Total Credits: 4

Course Objectives

On completion of the course, the student should have:

- To understand and apply the fundamental concepts in graph theory.
- To identify parts of a tree.
- To familiar with different tree traversal methods.
- To understand mathematical models used in Operations Research
- To apply linear programming problems, Transportation problems and assignment problems.

Module 1: Graphs

Graphs and Graph Models, Graph Terminology and Special types of Graphs, Representing Graphs and Graph Isomorphism, Connectivity, Euler and Hamilton Paths.

Text 1 Chapter 8 (Sections 8.1, 8.2, 8.3, 8.4 and 8.5 only) (18 Hrs.)

Module 2: Trees

Introduction to Trees, Application of Trees, Tree Traversal, and Spanning Trees.

Text 1 Chapter 9 (Sections 9.1, 9.2, 9.3 and 9.4 only) (18 Hrs.)

Module 3: Linear programming problems

Mathematical formulation of a L.P.P., General linear programming problems, solution of a L.P.P, graphical method for solving a L.P.P.

Simplex Method: Stack and surplus variables- reduction of any feasible solution to a basic feasible solution. Unbounded solution. Optimality conditions- artificial variable techniques- Big M method. (18 Hrs.)

Module 4: Transportation & Assignment Problems

Transportation model- solution by simplex method- north west corner rule, lowest cost entry method, Vogel method, MODI method, degeneracy, assignment problems. (18 Hrs.)

Book of Study

1. Kenneth H Rosen; Discrete Mathematics and Its Applications; 6th Edition; Tata Mc Graw-Hill Publishing Company Limited
2. Belly E Gillet – Introduction to Operations Research (A Computer Oriented Arithmetic Approach) (Tata Mc. GrawHill)

Course Title: Software Lab I

Course Code: ICSC1CP6

Total Credits: 4

The candidate needs to submit a Lab Record, duly signed by the teacher in charge and Head of the Department, (minimum of 20 Programs, (10+10, C and SQL), failing which he/she will not be allowed to attend the external software lab examination. The Lab record should be hard binded with name of college and the emblem of the college depicted on the first page and should be properly indexed.

1. Syllabus for C programs (54 Hrs.)

1. Simple Programs to familiarize printf() and scanf() functions.
2. Programs Based on Decision making statements, break, goto, continue, switch.
3. Programmes using Loop controls statements.
4. Programs Based on One dimensional and two-dimensional arrays (linear search, sort, matrix addition, multiplication, transpose etc.)
5. Programs on Strings and string handling functions.
6. Programs using the concept of Pointers, operations on pointers, Pointers to one dimensional array
7. Programs using the concept of functions, call by value, call by reference, Recursion.
8. Programs based on structure and union, array of structures, Pointer to structure, structure as argument to functions.
9. Simple programs using pointers and malloc ().

2. Syllabus for SQL Programs (54 Hrs.)

Problems involving the following topics to be included

1. Data definition commands - CREATE, ALTER, DROP, Adding Constraints Primary key, foreign key, unique key, check, not null.
2. Basic SQL queries INSERT, SELECT, DELETE, UPDATE, using multiple tables, ordering of rows using ORDER BY option, set operations using UNION, EXCEPT, INTERSECT, Substring Comparison using LIKE operator, BETWEEN operator.
3. Complex Queries Nested Queries, EXISTS and UNIQUE/DISTINCT functions, NULL values, renaming of attributes and joining of tables, Aggregate functions and grouping.
4. Managing views, Simple stored procedures.

SEMESTER - II

Si. No.	Course Code	Title	Hrs./ Week	Credits	Total hours
1	IML/IHN 2CC01	Second Language (T)	5	4	90
2	ICSC2CR2	Object Oriented Programming Using C++ (T)	3	3	54
3	ICSC2CR3	Data Structures using C++ (T)	3	3	54
4	ICSC2CR4	Operating Systems (T)	4	4	72
5	ICSC2CM5	Linear Algebra (T)	4	4	72
6	ICSC2CP6	Software Lab II (L)	6	4	108
		Total	25	22	450

- Syllabus of Second Language IML2CC01/IHN2CC01: As Approved by the respective BOS.
- Syllabus of complementary paper ICSC2CM5 - Linear Algebra As approved by BOS of Mathematics

Course Title: Object Oriented Programming Using C++

Course Code: ICSC2CR2
Total Credits: 3

Course Objectives

On completion of the course, the student will be able to understand:

- Object oriented programming concepts and introduction of C++ Programming language.
- Different control structures used in C++ and implementation of functions in C++.
- Importance of class and objects concept in programming
- Role of constructors and destructors and importance of Operator overloading
- Different types of inheritance and implementation of polymorphism.

Module 1:

Principles of Object-Oriented Programming, Beginning with C++ Object Oriented Technology, Disadvantages of conventional Programming, Programming Paradigms, Key concepts of Object-Oriented Programming, Advantages of OOP. Parts of C++ program, Types of tokens, Data types in C++, Type modifiers, Type casting, Constants, Constant pointers, Operators in CH, Referencing and dereferencing operators, Scope access operator, Memory management operators. (10 Hrs.)

Module 2:

Control structures and functions in C++ Decision making statements, Loops in CH, Function in C++ - The main function, Parts of a function, Passing arguments, Return by reference, Default arguments, Inline function, Function overloading, Principles of function overloading. (12 Hrs.)

Module 3:

Classes and Objects Structures in CH, Class in C++, Declaring Objects, public private protected Keywords, Defining member functions, Characteristics of member functions, Outside member function inline, Rule for inline functions, Data hiding, Memory allocation for objects, Static member variables and functions, Arrays of objects, Objects as function arguments, Friend functions, Friend classes, Local classes. (10 Hrs.)

Module 4:

Constructers, Destructors and Operator overloading

Constructors and destructors, Characteristics of constructors and destructors, Applications with constructors, Constructors with argument, Overloading Constructors, Constructors with Default arguments, Copy constructors, Destructors, Operator overloading, Over loading unary operators, Over loading binary operators, Overloading with friend function, Type conversion, Rules for overloading operators. (10 Hrs.)

Module 5:

Inheritance, Pointers, Binding, Polymorphism and virtual functions Inheritance, Access specifiers and simple inheritance, Types of inheritance, Virtual base classes, Constructors Destructors and inheritance, Advantages and disadvantages of inheritance. Pointer, Pointer declaration, void pointers, wild pointers, Pointers to objects, this pointer, Pointers to derived classes and base classes, Binding in C++, Pointer to derived class Objects, Virtual functions, Rules for virtual functions, Pure virtual functions, Abstract classes, Working of virtual functions, Object slicing. (12 Hrs.)

Book of Study

1. Ashok N. Kamthane, Object oriented Programming with ANSI & Turbo C++, First Edition, Pearson India

References

1. E. Balagurusamy - Object Oriented Programming with CH, Fifth edition, Tata McGraw Education Hill, 2011.
2. Ravichandran-Object Oriented Programming in C++, TMH, 3rd Edition

Course Title: Data Structures using C++

Course Code: ICSC2CR3

Total Credits: 3

Course Objectives

Upon successful completion of this course, students should be able to:

- Describe fundamental concepts of data structures.
- Illustrate the representation of arrays in memory and operations on it
- Compare and Contrast different searching and sorting techniques.
- Design operations on linear data structures such as stacks and queues.
- Implement operations on various types of linked lists.

Module 1:

Introduction to Data Structures, Definition, Classification of data structures, Primitive and Nonprimitive, Operations on data structures. Static and dynamic memory allocation. Dynamic memory allocation and pointers, Memory allocation operators in C++. User defined data types in C++. Recursion, Recursive functions in C++. (10 Hrs.)

Module 2:

Arrays, Linear array- Representation of array in memory, operations on linear array, Insertion, Deletion, Sorting and Searching. Two Dimensional Arrays - Representation of 2D array in memory, operations on 2D array. Multidimensional Arrays. (10 Hrs.)

Module 3:

Search and Sort: Search, Basic search techniques, Search algorithms. Searching techniques, Sequential search, Binary search. Sort, general Background, Definition, different types, Bubble Sort, Selection sort, Merge sort. (10 Hrs.)

Module 4:

Stack and Queue: Stack, Definition, Array representation of stack, Operations on stack. Infix, prefix and postfix notations, Conversion of an arithmetic expression from infix to postfix, Postfix evaluation, Applications of stack. Queue- Definition, Array representation of queue, Simple queue operations. Circular queues, Double ended queue, Priority queue. (12 Hrs.)

Module 5:

Linked List: Linked list-definition, Components of linked list, Representation of linked list, Advantages and disadvantages of linked lists, Types of linked list. Singly linked list. Operations on singly linked list, Creation, Insertion, Deletion, Search and Display. Doubly linked lists, Operations on doubly linked lists, Creation, Insertion, Deletion, Search and display. Circular linked list, Operations on circular linked list, Creation, Insertion, Deletion, Search and Display. (12 Hrs.)

Book of Study

1. G.S Baluja, Data Structures Through C++ (A Practical Approach), Danapat Rai & Co.

References

1. Ellis Horowitz and Sartaj Sajni, Fundamentals of Data Structures, Galgotia publications
2, Seymour Lipschutz, Theory and Problems of Data Structures, Schaums outline series

Course Title: Operating Systems

Course Code: ICSC2CR4

Total Credits: 4

Course Objectives

After completing the course, the student should be able to explain

- The fundamental concepts regarding an OS
- Concept of a process and management of processes

- Inter process synchronization methods and deadlock handling
- Various memory management techniques
- Concept of file and various file handling methods

Module 1:

Introduction OS Definition, Functions, Types of operating systems-Batch Operating System, Multi programming, Time sharing, Real time, distributed operating systems - Operating System Operations, Operating System Services, User Operating System Interface, System Calls, Types of System Calls. (14 Hrs.)

Module 2:

Processor Management: Job and process concept, Operating system view of process, process state, state transition diagram, PCB (Process control block), System state and process lists, process switch, threads, multi-threading operating system, operating system services for process management. Process Scheduling - Types of schedulers, scheduling and performance criteria, scheduling algorithms, multiple processor scheduling. (14 Hrs.)

Module 3:

Inter process synchronization and communication Concurrent Processes, need for inter process synchronization, critical section problem, mutual exclusion, mutual exclusion algorithms, semaphore definition, primitives, implementation of semaphores, monitors Deadlocks – Definition, Deadlock characterization, Resource allocation graph, methods for handling deadlocks, deadlock prevention, deadlock avoidance, safe state, resource allocation graph algorithm, Banker's algorithm, deadlock detection, recovery from deadlock. (16 Hrs.)

Module 4:

Memory Management Preliminaries, address binding, dynamic linking and loading, Overlays. logical versus physical address space, Swapping, Contiguous allocation - fragmentation, OO compaction, Paging-principles of page allocation, structure of page table, hardware support, multi-level paging, Segmentation-principles of operation, hardware, implementation of segment table, protection and sharing, fragmentation, segmentation with paging. Virtual Memory-Demand paging, Page replacement algorithms page allocation policies – Thrashing, hierarchical address translation tables, MMUS. (16 Hrs.)

Module 5:

File Management: File structure, File types, File access, File attributes, File operations. Directories - Flat directory systems, hierarchical directory systems. File system implementation Allocation methods, contiguous allocation, linked allocation, indexed allocation. (12 Hrs.)

Book of study

1. Silberschatz, Galvin, Gagne: Operating System Concepts, 7th Edition.

References

1. Andrew S. Tanenbaum, — Modern Operating System, Prentice Hall India.
2. Dhamdhere, system software and operating systems - Tata Mc Graw Hill.

3. H M Deitel, An Introduction to Operating System - Adison Wesley.
4. Tanenbaum, Modern Operating systems - Prentice Hall.
5. William Stallings, Operating Systems - Pearson Education.

Course Title: Linear Algebra

Course Code: ICSC2CM5
Total Credits: 4

Course Objectives

- Able to develop a better intuition for machine learning and deep learning algorithms
- Able to choose proper hyperparameters and develop a better model.

Module 1: Introduction To Vector Spaces

Vector Spaces: Rn and Cn, lists, Fnand digression on Fields, Definition of Vector spaces, Subspaces, sums of Subspaces, Direct Sums, Span and Linear Independence, bases, dimension. (15 Hrs.)

Module 2: Linear Maps

Definition of Linear Maps - Algebraic Operations on - Null spaces and Injectivity - Range and Surjectivity - Fundamental Theorems of Linear Maps - Representing a Linear Map by a Matrix - Invertible Linear Maps - Isomorphic Vector spaces - Linear Map as Matrix Multiplication - Operators - Products of Vector Spaces - Product of Direct Sum - Quotients of Vector spaces. (20 Hrs.)

Module 3: Eigenvalues, Eigenvectors and Eigenspaces

Eigenvalues and Eigenvectors - Eigenvectors and Upper Triangular matrices - Eigenspaces and Diagonal Matrices. (20 Hrs.)

Module 4: Inner Products and Norms

Inner Products, Norms, Orthonormal Bases, Self-Adjoint and Normal Operators, Spectal theorem, Polar Decomposition and Singular Value Decomposition. (proof of all theorems are excluded for module 4) (17 Hrs.)

Book of study

1. Sheldon Axler, Linear Algebra Done Right, Third Edition Springer, 2017.

Course Title: Software Lab II

Course Code: ICSC2CP6
Total Credits: 4

The candidate needs to submit a Lab Record, duly signed by the teacher in charge and Head of the Department, (minimum of 20 Programs, (10+10, CPP and DS) failing which he/she will not be allowed to attend the external software lab examination. The Lab record should be hard

binded with name of college and the emblem of the college depicted on the first page and should be properly indexed.

I. Syllabus for CPP programs (minimum of 10 questions) (54 Hrs.)

1. Programs based on default arguments, function overloading.
2. Programs based on array of objects, friend functions, passing objects as arguments to function.
3. Programs based on operator overloading (binary, unary) using member functions and friend functions.
4. Programs based on constructors, different types of constructors.
5. Programs based on inheritance, different types of inheritance, Polymorphism

II. Syllabus for Data structures using CPP (minimum of 10 questions) (54 Hrs.)

Student needs to code and implement CPP programs for the following:

Arrays - Insertion, Deletion, Polynomial addition using arrays Sort - Selection, Insertion, Quick Search - Linear search, Binary search Sparse matrix:- Sparse form representation, transpose and addition using the sparse form Stack: Implementation using arrays (linear stack), Infix to postfix conversion, Postfix evaluation Queue: - Implementation using arrays (linear queue), Implementation of circular queue Singly linked list - Implementation using dynamic memory allocation techniques, arrange the list based on the ascending or descending order of the information field, concatenate two linked lists, interchange any two nodes in a list, Implementation of circular list, Implementation of linked stacks and queues. Doubly linked list – Implementation of doubly linked list, Implementation of circular doubly linked list.

SEMESTER - III

Sl. No.	Course Code	Title	Hrs./ Week	Credits	Total Hours
1	ICSD3CR1	Introduction to Data Science (T)	4	4	72
2	ICSC3CR2	Programming in Python (T)	3	3	54
3	ICSC3CR3	R Programming and Mathematics for Artificial Intelligence (T)	4	3	72
4	ICSC3CR4	Computer Organization and Architecture (T)	4	4	72
5	ICSA3CM5	Probability and Statistics (T)	4	4	72
6	ICSD3CP6	Software Lab III: Python and R Programming (L)	6	2	108
		Total	25	20	450

Course Title: Introduction to Data Science

Course Code: ICSD3CR1
Total Credits: 4

Course Objectives

- To develop fundamental knowledge of concepts underlying data science and give a hands-on experience with real-world Problems.
- Know standard methods of data analysis and information retrieval.
- Able to formulate the problem of knowledge extraction as combinations of data filtration, analysis and exploration methods.

Module 1:

Introduction: Data Science - Evolution of Data Science - Big Data and Data Science hype – and getting past the hype - Why now? – Datafication - Current landscape of perspectives - Skill sets needed, Statistical Inference - Populations and samples - Statistical modeling, probability distributions, fitting a model. (14 Hrs.)

Module 2:

Data Analysis and Basic Tools: Exploratory Data Analysis and the Data Science Process -Basic tools of Exploratory Data Analysis - Philosophy of Exploratory Data Analysis - The Data Science Process - Case Study, Three Basic Machine Learning Algorithms - Linear Regression - k-Nearest Neighbors (k-NN) - k-means - Feature Generation and Feature Selection. (16 Hrs.)

Module 3:

Feature Extraction: Feature Generation (brainstorming, role of domain expertise, and place for imagination) - Feature Selection algorithms – Filters; Wrappers; Decision Trees; Random Forests. (14 Hrs.)

Module 4:

Recommendation Systems: Nearest Neighbor Algorithm - Algorithmic ingredients of a Recommendation Engine. Case Study. (14 Hrs.)

Module 5:

Dimensionality Reduction: Singular Value Decomposition - Principal Component Analysis. Data Visualization and Fraud Detection. (14 Hrs.)

References

1. Cathy O'Neil and Rachel Schutt. "Doing Data Science, Straight Talk from the Frontline". O'Reilly Edition, 2014.
2. V.Bhuvaneswari, T. Devi, "Big Data Analytics: A Practitioner's Approach", Sci-TechPublications, 2016.
3. Jure Leskovek, Anand Rajaraman and Jerey Ullman. Mining of Massive Datasets. v2.1, Cambridge University Press. 2014. (free online)

Course Title: Programming in Python

Course Code: ICSC3CR2
Total Credits: 3

Course objectives

After completing this course, the student will

- understand basic knowledge in Python programming.
- learn how to design and program Python applications.
- acquire object-oriented skills in Python.
- able to work with python standard library.

Module 1:

Programming Environment and Python Basics: Getting Started with Python Programming - Running code in the interactive shell, Editing, Saving, and Running a script. Using editors - IDLE, Jupyter. Basic coding skills – Writing simple programs. (10 Hrs.)

Module 2:

Building Python Programs: Data types, variables, operators. Control statements – branching controls, simple if, if - else, if - elif -else; looping, while, for. Functions - defining, calling, returning values, functions with default arguments, recursive functions, nested functions and lamda functions. Strings - operation, string functions. Work with dates and times. (10 Hrs.)

Module 3:

Containers: Lists - Basic list Operations and functions, List of lists, Slicing, Searching and sorting list, List comprehension. Work with tuples and Sets. Dictionaries - Dictionary functions, dictionary literals, adding and removing keys, accessing and replacing values, traversing dictionaries, reverse lookup. (10 Hrs.)

Module 4:

Object Oriented Programming: Design with classes, Inheritance – multi-level and multiple inheritance. Exceptions - Handle a single exception, handle multiple exceptions. Introduction to file I/O - Reading and writing text files, Manipulating binary files. More concepts: Decorators, generators and iterators. (12 Hrs.)

Module 5:

Scientific Python: NumPy - Basics, Creating arrays, Arithmetic, Slicing, Matrix Operations, Random Numbers. Plotting and visualization. Matplotlib - Basic plot, Ticks, Labels, and Legends. Pandas: operations on CSV files. Reading, Manipulating, and Processing Data. Python GUIs and event handling using tkinter. (12 Hrs.)

Book of Study

1. Kenneth A Lambert., Fundamentals of Python: First Programs, 2/e, Cengage Publishing,2016
2. Jeeva Jose, P Sojan Lal, Introduction to Computing and Problem solving with Python, Khanna Book Publishing, 2016
3. Wes McKinney, Python for Data Analysis, 2/e, Shroff / O'Reilly Publishers, 2017

Reference Books:

1. Allen B. Downey, Think Python: How to Think Like a Computer Scientist, 2/e, Schroff, 2016
2. Michael Urban and Joel Murach, Python Programming, Shroff/Murach, 2016
3. David M.Baezly, Python Essential Reference. Addison-Wesley Professional; 4/e, 2009.

Course Title: R Programming and Mathematics for Artificial Intelligence

Course Code: ICSC3CR3

Total Credits: 3

Course Outcome

- On completion first two units of the course, students will be able to use R language for programming purposes
- The remaining three units will enable the student to become confident in the mathematical portions needed in the field of artificial intelligence.
- The lab sessions for the paper is so designed to make the student an expert in R to solve problems in mathematics.

Module 1:

R Programming -Fundamentals, installation and use of software, data editing, use of R as a calculator, functions and assignments, arguments, scope, logic and statements in R, logical equivalence, Sets with R: Cardinality, Equality, Empty set, Subset, Union, Intersection, Complement, Cross product and Algebraic properties. (12 Hrs.)

Module 2:

R Programming - Exploring and cleaning data for analysis, Data organization, Arrays, and Matrices, Basics of Arrays in R, Matrix operations, Advanced Matrix operations, Additional Matrix facilities, Lists and Data frames. Mapping models to Machine Learning, Evaluating and Validating models, Probability distributions in R, Statistical models in R, Building, linear models, Generalized linear models, Nonlinear least squares and maximum likelihood models. (14 Hrs.)

Module 3:

Sets, Operations on sets, Venn Diagrams, Multi Sets, Binary Relations, Equivalence Relations, Ordering Relations, Operations on Relations, Partial Orders. Statements and Notation, Connectives, Quantified Propositions, Logical Inferences, Methods of Proof of an Implication, First Order Logic and other Methods of Proof, Rules of Inference for Quantified Propositions, Proof by Mathematical Induction. (15 Hrs.)

Module 4:

Linear Algebra – System of Linear equations, Solving System of Linear equations, Linear Independence, Vectors, Scalars, Addition, Scalar multiplication, dot product, vector projection, cosine similarity. Support Vector Machines, Implementation using Python, Classification using Support Vector Machines. (15 Hrs.)

Module 5:

Matrices, determinants, inverse of matrix. System of equations, Linear transformation - rank and nullity, Consistency, and Inconsistency of linear system of equations, rank nullity theorem, Echelon form of a matrix and Row reduced echelon form of matrix. Correlation coefficient, Eigen values and Eigen vectors. Principle Component analysis (PCA) – Concepts and properties. Dimensionality reduction with PCA. (16 Hrs.)

Book of Study

1. N Matloaff ,“The art of R Programming”, No Starch Press, Inc, 2011, Ist edition,ISBN-10: 1-59327-384-3, ISBN-13: 978-1-59327-384-2.
2. William B. Claster, “Mathematics and Programming for Machine Learning with R: From the Ground Up” CRC Press; 1st edition (27 October 2020), ISBN: 9780367507855.
3. For Maths
4. Kenneth H. Rosen, “Discrete Mathematics And Its Applications”, 7th Ed,McGrawHill, 2012.
5. Erwin Kreyszig, “Advanced Engineering Mathematics”, Wiley India, 9th Edition 2011.

References

1. Nina Zumel, John Mount, Jeremy Howard, Rachel Thomas, “Practical Data Science With R”, Manning Publications, Year: 2020, ISBN: 1617295876, 9781617295874.
2. “Mathematics for Data Science and Machine Learning using R” by Eduonix, September 2019, Packt Publishing, ISBN: 9781839210945.

3. Mark Gardener, "Beginning R: The Statistical Programming Language", ISBN: 978-1-118-16430-3 May 2012.

Web References

1. https://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf

Course Title: Computer Organization and Architecture

Course Code: ICSC3CR4
Total Credits: 4

Course objective

Upon successful completion of this course, students should be able to:

- Describe the fundamental organization of a computer system
- Explain addressing modes, instruction formats and program control statements
- Analyze the organization and performance of system memory hierarchy
- Describe basic concept of parallel computing.
- Describe fundamentals concepts of pipeline and vector processing

Module 1:

Introduction: Functional units of a computer, Basic operational concepts, Bus structure, Memory locations and addresses, Instructions and instruction sequencing, Instruction execution. Instruction Formats, Addressing Modes. (16 Hrs.)

Module 2:

Central Processing Unit: General Register Organization, Stack Organization, Instruction Formats, Instruction Classification, Addressing modes. (14 Hrs.)

Module 3:

Memory: Memory Hierarchy, RAM, ROM, Cache Memories, Virtual memory. (12 Hrs.)

Module 4:

Pipeline and Vector Processing: Parallel Processing, Architectural classification scheme-SISD, SIMD, MISD, MIMD, Pipelining, Arithmetic Pipeline, Instruction Pipeline, RISC Pipeline, Vector Processing, Array Processor. (16 Hrs.)

Module 5:

Multi Processors: Characteristics of Multiprocessors, Interconnection Structures, Interprocessor arbitration, Interprocessor communication and synchronization, Cache Coherence. (14 Hrs.)

Book of Study

1. Computer Organization, V. Carl Hamacher, Zvonko Vranesic, Safwat Zaky, 5th Edition, McGraw Hill Education.
2. Computer System Architecture, M. Morris Mano, Third Edition, Pearson/PHI.

References

1. Computer Organization and Architecture – William Stallings, Sixth Edition, Pearson/PHI.
2. Computer Architecture and Parallel Processing, Kai Hwang and F. A. Briggs, McGraw Hills
3. Computer Architecture & Organization– John P Hayes, Mc Graw Hill.
4. Structured Computer Organization – Andrew S. Tanenbaum, 4th Edition, Pearson/PHI.

Course Title: Probability and Statistics

Course Code: ICSD3CM5
Total Credits: 4

Course Objectives

- Acquire the mathematical foundations of probability and statistics for data science activities.
- Learn probability with underlying motivation being statistics for data science .
- Learn hands-on on generating random numbers and programming for statistics.

Module 1:

Introduction to Statistics, concepts of a statistical population and sample, Data types- qualitative and quantitative, discrete and continuous, primary and secondary. Different types of scale- nominal and ordinal, ratio and interval.

Data Collection and sampling techniques- SRS, systematic, stratified and cluster schedule and questionnaire. Data collection: direct, using third parties, sending questionnaire, by mail/telephone,

Classification and tabulation - One-way and two-way classified data, Preparation of frequency distribution, relative frequency and cumulative frequency distributions.

Graphs: Stem-and-leaf chart, Pie Char, Bar Chart, Histogram, Frequency polygon, Frequency curve and Ogives. (21 Hrs.)

Module 2:

Descriptive Measures: Averages- Arithmetic Mean, Median, Mode, Geometric Mean, Harmonic Mean and Weighted averages. Quantiles- quartiles, deciles, percentiles.

Measure of Dispersion: Absolute and relative measures dispersion - Range, Quartile Deviation, Mean Deviation and Standard Deviation, Co-efficient of variation, Box plot.

Moments: Raw moments, central moments and their inter relation, skewness- Pearson's, Bowley's and moment

Measures of skewness and Kurtosis- percentile and moment measure of kurtosis. (17 Hrs.)

Module 3:

Probability Random Experiments-Algebra of events- Mutually exclusive, equally likely and independent events. Classical, Frequency and Axiomatic approaches to probability. Monotone property, Addition theorem, Boole's inequality (finite case), and other simple properties.

Conditional probability. Multiplication theorem (up to 3 events). Independence of events. Total probability law. Bayes' theorem.
Random variables- discrete and continuous random variables. Probability mass and density functions, and distribution functions. Evaluation of conditional and unconditional probabilities. Change of variables- methods of Jacobian and distribution function (one variable case).
Bivariate Random Variables: bivariate probability mass and density functions. Marginal and conditional distributions. Independence of bivariate random variables.
Mathematical Expectation: Expectation of random variables and their functions. Definition of - Raw moments, central moments and their inter-relation, covariance. (17 Hrs.)

Module 4:

Correlation and Regression: Measure of association between two variables. Types of correlation, Scatter diagram, Karl Pearson's Coefficient of correlation and its mathematical properties, Spearman's Rank correlation and its interpretations, Regression Analysis, linear regression equations, properties of regression coefficients, Multiple regression model. (17 Hrs.)

Book of Study

1. S C Gupta and V K Kapoor; Fundamentals of Mathematical Statistics; Sultan Chand and Sons New Delhi

References

1.S P Gupta; Statistical Methods; Sultan Chand and Sons New Delhi
2. M R Spiegel; Theory and Problems of Statistics Schaum's Outline Series.

Course Title Software Lab III: Python and R Programming

Course Code: ICSD3CP6
Total Credits: 2

I. Programming in Python- Lab. (54 Hrs.)

1. Programs using conditional Branching and looping
2. Programs using functions and strings
3. Programs using Lists, Dictionaries, tuples and sets.
4. Program for traversing dictionaries
5. Programs using class, Inheritance and Exceptions.
6. Programs using files.
7. Arrays and Matrix using NumPy.
8. Plotting and Visualization using Matplotlib (Line, bar chart, pie chart etc..).
9. Data manipulations (data series and data frames) using Pandas
10. Simple programs using GUI

II. Programming using R language- Lab. (54 Hrs.)

1. R Program to add two vectors. The program covers:
    a. R Vector

b. R Operators

2. Find sum, mean and product of vector in r using built-in functions. The program covers:
   a. R Variables and Constants
   b. R Functions
3. R program to print the Fibonacci sequence by taking input from the user. The program covers:
   a. R if…else Statement
   b. R while loop
4. R program to find the factors of a number. The program covers:
   a. R Functions
   b. R for Loop
5. Calculator Application in R. The program covers:
   a. Using with and without R objects on console
   b. Using mathematical functions on console
6. Reading and writing different types of datasets
   a. Reading different types of data sets (.txt, .csv) from web and disk and writing in file in specific disk location.
   b. reading Excel data sheet in R.
   c. reading XML dataset in R
7. Solve this system using R:

Compute the inverse of the resultant matrix, Compute the determinant matrix, Compute the Eigenvalues/eigenvector. $x1 + x2 = 2$ $-x1 + x2 = 4$

8. Solve the system of linear equations using R.
   $5x + y = 15$, $10x + 3y = 9$
9. Write an R program to access the element at 3rd column and 2nd row, only the 3rd row and only the 4th column of a given matrix.
10. Descriptive statistics in R
    a. Write an R script to find basic descriptive statistics using summary, str, quartile function on mtcars & cars datasets.
    b. Write an R script to find subset of dataset by using subset (), aggregate () functions on iris dataset.

SEMESTER - IV

Si. No.	Course Code	Title	Hrs/ Week	Credits	Total Hours
1	IEN4CC01	English II (T)	5	4	90
2	ICSD4CM2	Probability Distributions and Statistical Inference (T)	4	3	72
3	ICSC4CR3	Data Mining (T)	4	4	72
4	ICSC4CR4	Software Engineering (T)	4	3	72
5	ICSC4CR5	Basics of Artificial Intelligence (T)	4	3	72
6	ICSD4CMP6	Complementary Lab: Statistical Analysis using R (L)	2	2	36
7	ICSD4CP7	Software lab IV (L)	2	1	36
		Total	25	20	450

- Syllabus of English II –As Approved by Board of Studies of English

Course Title: Probability Distributions and Statistical Inference

Course Code: ICSD4CM2
Total Credits: 4

Module 1:

Discrete Distributions: Degenerate, Uniform, Bernoulli, Binomial, Hyper geometric, Negative binomial, Geometric, Poisson

Mean, variance, m.g.f, their properties-fitting of Binomial and Poisson, memory less property of Geometric distribution, multinomial distributions and its applications. (17 Hrs.)

Module 2:

Continuous Distributions: Uniform, Exponential, Gamma, Cauchy, Pareto, and Laplace - mean, variance, m.g.f, characteristic function, their properties - memory less property of exponential distribution.

Normal and Lognormal Distributions: Properties, fitting of normal distribution, linear combination of normal variates, use of standard normal tables for various probability computation. Bivariate normal-marginal and conditional distributions.

Sampling Distributions: Concept of sampling distributions, Statistic(s) and standard error(s). Mean and variance of sample mean when sampling is from a finite population. Sampling distribution of mean and variance from normal distribution. Chi-square, t, F distributions and

statistics following these distributions. Relation among Normal, Chi- square, t and F distributions. (19 Hrs.)

Module 3:

Point Estimation: Concepts of Estimation, Estimators and Estimates. Point and interval estimation. Properties of good estimators- unbiasedness, efficiency, consistency and sufficiency.

Confidence Intervals: 100(1-α) % confidence intervals for mean, variance, proportion, difference of means and proportions and variances. (17 Hrs.)

Module 4:

Testing of Hypotheses: Statistical hypotheses, null and alternate hypotheses, simple and composite hypotheses, type-I and type-II errors. Critical Region. Size and power of a test, p-value, large sample tests - Z-tests for means, difference of means, proportion and difference of proportion, chi-square tests for independence, homogeneity and goodness of fit. Normal tests for mean, difference of means and proportion (when σ known), t-tests for mean and difference of means (when σ unknown), t-test for r = 0, paired t-test, test for proportion (binomial), chi-square test, F-test for ratio of variances. ANOVA

Non-Parametric Tests: Introduction to Non parametric tests, non-parametric equivalent of parametric tests. (19 Hrs.)

Book of Study

1. Rohatgi V.K. An Introduction Probability Theory and Mathematical Statistics, John Wiley and sons

References

1. Gupta S. C. and Kapoor V. K. (2002). Fundamentals of Mathematical Statistics, 11th edition, Sultan Chand and Sons.
2. George Casella, Roger L. Berger. Statistical Inference (2nd Ed).
3. Goon A. M., Gupta M. K., and Dasgupta B. (2005). Fundamentals of Statistics, Vol. I, 8th edition, World Press, Kolkatta.
4. Gibbons J.K (1971). Non-Parametric Statistical Inference, McGraw Hill.

Course Title: Data Mining

Course Code: ICSC4CR3

Total Credits: 4

Course Objectives

- To identify the scope and essentiality of Mining
- To analyze data, choose relevant models and algorithms for respective applications. To develop research interest towards advances in data mining.
- To introduce the basic concepts and techniques of Data mining

Module 1:

Introduction: What is Data mining? Data Mining Tasks, KDD process, 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. (12 Hrs.)

Module 2:

Data Pre-processing: 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 (14 Hrs.)

Module 3:

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, From Association Analysis to Correlation Analysis. (14 Hrs.)

Module 4:

Classification- Basic concepts, General approach to classification, Nearest neighbor models, Cross validation and re-sampling methods- kfold cross validation, Boot strapping, Measuring classifier performance- Precision, recall, ROC curves. Bayes Theorem, Bayesian classifier, Decision Trees- Entropy, Information Gain, Tree construction, ID3, Issues in Decision Tree learning- Avoiding Over-fitting, Reduced Error Pruning, The problem of Missing Attributes, Gain Ratio, Classification by Regression (CART) (16 Hrs.)

Module 5:

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. Grid Based- STING, CLIQUE, Outlier Analysis- what are outliers, Types of outliers, Outlier detection methods - Statistical Distribution-based Outlier Detection, Distance-Based Outlier Detection. Mining other kinds of data – Mining spatial data, mining multimedia data, mining text data, mining web data. (16 Hrs.)

Book of Study

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.

References

1. Arun K Pujari, Data Mining Techniques, , University Press

2. Sam Anahory & Dennis Murray, Data Warehousing in the Real World, Pearson Education, Asia.

3. Paulraj Ponnaiah, Data Warehousing Fundamentals,Wiley Student Edition

Course Title: Software Engineering

Course Code: ICSC4CR4

Total Credits: 4

Course Objectives

Upon the completion of the course, students should be able to

- Recognize the importance of basic processes in software Development life cycle.
- Understand the various activities associated with different models and their significance.
- To provide better understanding about the basic concepts of Software Engineering
- Familiarize the requirements in engineering and systematic approach in classical softwaredesign and development techniques.
- Familiarize with various software testing techniques and tools.
- Perceive the importance of Software Maintenance

Module 1:

Introduction: Evolution, Types of software development products; Software life cycle models: A few basic concepts, Waterfall model and its extension, Agile development models, Spiral model, Comparison of different life cycle models (14 Hrs.)

Module 2:

Software Project Management, Project Planning, Metrics for project size estimations, Project Estimation Techniques, Basic COCOMO model, Scheduling, Organization structure, Team structure, Staffing, Risk Management, Software Requirements Analysis and Specification: Requirements gathering and analysis, Software Requirements Specification (SRS) (14 Hrs.)

Module 3:

Software Design: overview of the design process, how to characterize a good software design, Cohesion and Coupling, Layered arrangements of modules, Approaches to software design, Function oriented design: Overview of SA/SD Methodology, Structured analysis, Developing the DFD model of a system, Structured Design, User Interface design: Characteristics of a good user interface, Basic concepts, Types of user interfaces (14 Hrs.)

Module 4:

Coding and Testing: Coding, Code review, Software documentation, Testing, Unit testing, Black box testing, white box testing: Basic concepts, Debugging Integration testing, system testing, Software Reliability and quality management: Software reliability, Software quality (14 Hrs.)

Module 5:

Software maintenance: Characteristics of software maintenance, Software reverse engineering, Software process models, Estimation of maintenance cost, Software Reuse: Basic issues in any Reuse Program, A Reuse approach, Reuse at Organization level , Emerging Trends:Client Server Software, Client Server architectures, CORBA, Service Oriented Architecture (SOA), Software as a Service (SaaS). (16 Hrs.)

Book of Study

1. Fundamentals of Software Engineering, Fifth Edition by Rajib Mall, PHI Learning Pvt. Ltd., February 2019.

References

1. Software Engineering 10th Edition by Ian Sommerville, PEARSON INDIA, October 2018.
2. Software Engineering – a Practitioner's approach Seventh Edition by Roger S Presman, 7th edition, McGraw Hill. 2017.

Course Title: Basics of Artificial Intelligence

Course Code: ICSC4CR5
Total Credits: 3

Course Objectives

Upon successful completion of this course students should be able to:

- Explain the basics of AI.
- Identify appropriate AI methods to solve a given problem.
- Will be able understand the concept of differentiation and its applications
- Should have sound knowledge in the field of probability and correlation and regression

Module 1:

Introduction to AI-Problem formulation, Problem Definition -Production systems, Control strategies, Search strategies. Problem characteristics, Production system characteristics - Specialized productions system- Problem solving methods – Problem graphs, Matching, Indexing and Heuristic functions -Hill Climbing-Depth first and Breath first, Constraints satisfaction – Related algorithms, Measure of performance and analysis of search algorithms. (14 Hrs.)

Module 2:

Knowledge representation, Knowledge representation using Predicate logic, Introduction to predicate calculus, Resolution, Use of predicate calculus, Knowledge representation using other logic-Structured representation of knowledge. Game playing. (14 Hrs.)

Module 3:

Differentiation, Limits and continuity rules of differentiation, Derivatives, Scalar derivatives, Partial derivatives, Differentiation of univariate functions, Partial differentiation and gradients, Gradient of vector valued function. Gradient of matrices. Optimization using gradient functions, constrained optimization, and Lagrange multipliers. Convex optimization. Back propagation in neural networks, implementation and application. (15 Hrs.)

Module 4:

Generating Functions of Sequences, Calculating Coefficients of Generating Functions, Recurrence Relations, Solving Recurrence Relations by Substitution and Generating Functions, The Method of Characteristic Roots, Solutions of Inhomogeneous Recurrence Relations, Complexity calculations of prominent algorithms. (15 Hrs.)

Module 5:

Probability, basics, Conditional Probability, Bayes Theorem, Distributions - Binomial, Poisson, normal distributions, and related problems. Descriptive Statistics, Regression, and correlation, Bayesian classification, implementation, applications. (14 Hrs.)

Books of Study

1. Artificial Intelligence, Deepak Khemani, Tata Mc Graw Hill Education.
2. Kenneth H. Rosen, "Discrete Mathematics and Its Applications", 7th Ed, McGraw-Hill, 2012.
3. Erwin Kreyszig, "Advanced Engineering Mathematics", Wiley India, 9th Edition 2011.
4. Walpole, R. E., Myers, R. H., Myers S L & Keying Ye, 'Probability and Statistics for Engineers and Scientists'. 8th ed, Pearson Education, 2007.

Course Title: Complementary Lab: Statistical Analysis Using R

Course Code: ICSD4CMP6
Total Credits: 2

Programs based on the following concepts: (36 Hrs.)

Doing Math and Simulation in R, Math Function, Cumulative Sums and Products-Minima and Maxima- Calculus, Functions for Statistical Distribution, Sorting, Linear Algebra Operation on Vectors and Matrices, Set Operation, Input /output, Accessing the Keyboard and Monitor, Reading and writer Files,

Probability Distributions, Normal Distribution- Binomial Distribution- Poisson Distributions Other Distribution, Basic Statistics, Correlation and Covariance, T-Tests, ANOVA.

Linear Models, Simple Linear Regression, -Multiple Regression Generalized Linear Models, Logistic Regression, - Poisson Regression.

Course Title: Software lab IV

Course Code: ICSD4CP7
Total Credits: 1

Software Lab IV- Data Mining Implementation using Python & R (36 Hrs.)

1. Implement Apriori algorithm for frequent item set generation using Python.
2. Implementation of Classification in Python using
    a. KNN
    b. Decision Tree

3. Implementation of Clustering in Python using
    a. K-means
    b. K-medoid
4. Regression Model using R
    a. Import a data from web page
    b. Name the dataset and now do Logistic Regression to find out relation between variables that are affecting the admission of a student in an institution based on his/her GRE score, GPA obtained and rank of the student.
    c. Also check whether the model if fit or not.
5. Classification model using R using above dataset
    a. Install relevant package for classification
    b. Choose classifier for classification problem
    c. Evaluate the performance of classifier
6. Implementation of Correlation and Covariance analysis using R
    a. Find the correlation matrix
    b. Plot the correlation plot on dataset and visualize an overview of relationships among data on iris dataset.
    c. Analysis of covariance: Variance (ANOVA), if data have categorical variables on iris data.

SEMESTER - V

Si. No.	Course Code	Title	Hrs./ Week	Credits	Total Hours
1	ICSC5CR1	Principles of Machine learning (T)	3	4	54
2	ICSC5CR2	Web application Development Using PHP (T)	4	3	72
3	ICSC5CR3	Programming in Java (T)	4	3	72
4	ICSC5CR4	IT and Environment (T)	3	4	54
5	ICSD5PR5	Project Minor - Phase I (L)	3	-	54
6	ICSD5CP6	Software Lab V: Java and PHP (L)	8	3	144
		Total	25	17	450

Course Title: Principles of Machine learning

Course Code: ICSC5CR1
Total Credits: 4

Course Objectives

On completion of the course, the student should have:

- To introduce students to the basic concepts and techniques of Machine Learning.
- To understand the supervised learning techniques such as Linear Regression, Logistic Regression, Support Vector Machine, Naïve Bayes Classifier.
- To understand the biological neural network and to model equivalent neuron models.

Module 1:

Introduction to Machine Learning – Machine learning basics, Types of machine learning, Applications of Machine Learning, Basic types of data in Machine learning, Data pre-processing, Predictive Vs Descriptive models, training a model, training versus testing, cross validation, overfitting & underfitting, Bias variance tradeoff, error measures, evaluating performance of a model. Introduction to feature engineering, ML tools in Python. (12 Hrs.)

Module 2:

Linear Regression – Problem formulation, Parameter Estimation, Bayesian Linear Regression, Multiple linear regression, fitting simple linear and multiple linear regression equations with examples, regularization techniques, case study and implementation. (10 Hrs.)

Module 3:

Logistic Regression - Interpreting Parameters in Logistic Regression, Inference for Logistic Regression, Logistic Models with Categorical Predictors, Multiple Logistic Regression, Fitting Logistic Regression Models and its implementation using real life examples. (10 Hrs.)

Module 4:

Support Vector Machine (SVM) Algorithm, Types of SVM, Hyperplane and Support Vectors, Working of SVM Applications of SVM; Naïve Bayes Classifier Algorithm, Bayes' Theorem, Types of Naïve Bayes Model, Working of Naïve Bayes' Classifier, Applications of Naïve Bayes Classifier (10 Hrs.)

Module 5:

Artificial Neural Network (ANN): Features, structure and working of Biological Neural Network (BNN), Comparison of BNN and ANN, History of neural network research, characteristics of neural networks, terminology, Applications of ANN, models of neuron McCulloch-Pitts model, Perceptron, Basic learning laws, Topology of neural network architecture. Backpropagation networks (BPN), Architecture of feed forward network, single layer ANN, multilayer perceptron, back propagation learning, input-hidden and output layer computation, backpropagation algorithm, selection of tuning, parameters in BPN, learning. (12 Hrs.)

Books of Study

1. Machine Learning, Saikat Dutt, Subramanian Chandramouli, Amit Kumar Das, Person, 2020.
2. Marc Peter Deisenroth, A. Aldo Faisal, Cheng Soon Ong, "Mathematics for Machine Learning", 2020.
3. Aurelien Geron, "Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow" Powered by Jupyter, published by O'Reilly Media.
4. Artificial neural network, B. Yegnanarayana - PHI Publication.
5. Neural networks, Fuzzy logic and Genetic Algorithms; S. Raj Sekaran and Vijayalakshmi Pari

Course Title: Web application Development Using PHP

Course Code: ICSC5CR2
Total Credits: 3

Course Objectives

Upon successful completion of this course, students should be able to:

- Develop web applications using Php and MySQL database.
- Use java scripts and jQuery in client side
- Use CSS concepts in Webpage designing

Module 1:

Introduction to PHP- Structure of PHP-Comments, Basic Syntax, Variables, Variable Assignment, Variable Typing, Constants, Predefined Constants, echo vs print Command, Functions, Variable Scope.

Expressions and Control Flow in PHP- Expressions, Conditionals- if Statement, else Statement, elseif Statement, switch Statement? Operator, Looping- while Loops, do...while Loops, for Loops, break, continue Statement.

PHP Functions and Objects- PHP Functions- Defining a Function, Returning a Value, Returning an Array, Passing Arguments by Reference, Returning Global Variables, Including and Requiring Files, PHP Objects- Declaring a Class, Creating an Object, Accessing Objects, Cloning Objects, Constructors, Destructors, Writing Methods, Declaring Properties, Declaring Constants, Property and Method Scope, Static Methods, Static Properties, Inheritance. (16 Hrs.)

Module 2:

PHP Arrays -Basic Access- Numerically Indexed Arrays, Associative Arrays, Assignment Using the array Keyword, foreach Loop, Multidimensional Arrays, Array Functions .Accessing MySQL Using PHP -Connecting to a MySQL Database, $_POST Array, create, insert, delete, update, select operations in MySQL database using PHP, Form Handling - Building Forms - Retrieving Submitted Data, Default Values, Input Types, HTML5 Enhancements- autocomplete Attribute, autofocus Attribute, placeholder Attribute, required Attribute, Override Attributes, width and height Attributes, min and max Attributes, step Attribute, form Attribute, list Attribute, color Input Type, number and range Input Types, Date and Time Pickers. (14 Hrs.)

Module 3:

Exploring JavaScript - JavaScript and HTML Text- Using Scripts Within a Document Head, Including JavaScript Files, Using Comments, Semicolons, Variables- String Variables, Numeric Variables, Arrays, Variable Typing, Functions, Global Variables, Local Variables, Document Object Model document.write, console.log, alert, Writing into Elements, with Statement, onerror, Using try...catch, Conditionals, Loops.

JavaScript Functions, Objects, and Arrays - JavaScript Functions- Defining a Function, Returning a Value, Returning an Array, JavaScript Objects- Declaring a Class, Creating an Object, Accessing Objects, prototype Keyword, JavaScript Arrays- Numeric Arrays, Associative Arrays, Multidimensional Arrays, Array Methods, Asynchronous Communication - XMLHttpRequest, Sending XML Requests. (14 Hrs.)

Module 4:

Introduction to CSS - Importing a Stylesheet, Importing CSS from Within, Embedded Style Settings, Using IDs, Using Classes, Using Semicolons, CSS Rules- Multiple Assignments, Using Comments, Style Types- Default Styles, User Styles, External Stylesheets, Internal Styles, Inline Styles, CSS Selectors- Type Selector, Descendant Selector, Child Selector, ID Selector, Class Selector, Attribute Selector, Universal Selector, Selecting by Group, CSS Cascade-Stylesheet Creators, Stylesheet Methods, Stylesheet Selectors, Measurements, Fonts and Typography- font-family, font-style, font-size, font-weight, Managing Text Styles- Decoration, Spacing, Alignment, Transformation, Indenting, CSS Colors- Short Color Strings, Gradients, Positioning Elements- Absolute Positioning, Relative Positioning, Fixed Positioning, Pseudoclasses, Box Model and Layout- Setting Margins, Applying Borders, Adjusting Padding, Object Contents. (14 Hrs.)

Module 5:

Introduction to jQuery- Including jQuery, jQuery Syntax, Avoiding Library Conflicts, Selectors, CSS Method, Element Selector, ID Selector, Class Selector, Combining Selectors, Handling Events, Event Functions and Properties- blur and focus Events, click and dblclick Events, keypress Event, mouse move Event, Other Mouse Events, Alternative Mouse Methods, submit Event, Special Effects, Hiding and Showing, toggle Method, Fading In and Out, Sliding Elements Up and Down, Manipulating the DOM- text vs html Methods, val and attr Methods, Adding and Removing Elements, Dynamically Applying Classes, Modifying Dimensions, DOM Traversal- Parent Elements, Child Elements, Sibling Elements, Selecting the Next and Previous Elements, Traversing jQuery Selections, is Method, Using jQuery Without Selectors- $.each Method, $.map Method, Asynchronous Communication. (14 Hrs.)

Book of Study

1. Learning PHP, MySQL & JavaScript, Robin Nixon 5th Edition, O'Reilly

References

1. Learn PHP 7: Object-Oriented Modular Programming using HTML5, CSS3, JavaScript, XML, JSON, and MySQL- Steve Prettyman, Apress
2. PHP, MySQL, JavaScript & HTML5 All-in-One for Dummies - Steve Suehring and Janet Valade, Wiley
3. Beginning JavaScript with DOM Scripting and Ajax from Novice to Professional, Christian Heilmann, Apress
4. Beginning jQuery: From the Basics of jQuery to Writing your Own Plug-ins- Jack Franklin Russ Ferguson, Second Edition, Apress
5. Sams Teach Yourself HTML, CSS & JavaScript Web Publishing in One Hour a Day, Seventh Edition

Course Title: Programming in Java

Course Code: ICSC5CR3
Total Credits: 3

Course Objectives

- To familiarize basic concepts of OO programming.
- To understand the concept of constructors, packages and multithreading.
- To inculcate concepts of GUI programming using swing.
- To be able to create applets and implement database connectivity.

Module 1:

Concepts of Object-oriented programming, Benefits of OOP, Features of java. Java environment, java tokens, Constant, variables, data types, operators, Control Statements-branching statements, looping statements, jump statements, labeled loops. (10 hrs.)

Module 2:

Defining a Class, Fields declaration, Method declaration, Creating object, Accessing class members, method overloading, Constructors, constructor overloading, super keyword, static Members, Inheritance, overriding methods, dynamic method dispatch, final (variables, methods and classes), abstract methods and classes, interfaces, visibility control. (10 hrs.)

Module 3:

Arrays- One dimensional arrays, declaration, creation, initialization of arrays, two dimensional arrays, String class. Packages: - java API packages overview (lang, util, io, awt, swing, applet), user defined packages-creating packages, using packages Exception Handling Techniques-try-catch-throw-throws-finally -Multithreading- creation of multithreaded program-Thread class Runnable interface, Thread life cycle. (12 hrs.)

Module 4:

Event Handling-Delegation Event Model-Event Classes-Sources of Events-Event Listeners-Event classes- Swing- architecture, components of swing- JLabel, JButton, JCheckBox, JRadioButton, JList, JComboBox, JTextField, JText Area, JPanel, JFrame, Layout Managers (Flow Layout, Grid Layout, Card Layout, Border Layout, Box Layout, Null Layout). (10 hrs.)

Module 5:

Applet Fundamentals -applet tag, applet life cycle, passing parameters to applets. Working with graphics -Line, Rectangle, Oval, Arc, color setting. JDBC architecture- JDBC connection, JDBCstatement object, JDBC drivers. (10 hrs.)

Book of Study

1. E. Balagurusamy- Programming with Java, Third Edition, McGraw Hill Companies.
2. K. Somasundaram - PROGRAMMING IN JAVA2, First Edition, Jaico Publishing House.

References

1. Patrick Naughton - Java2 The Complete Reference, Seventh Edition:
2. Cay S Horstmann & Gary Cornell - Core Java Volume 1- Fundamentals, Eighth edition.
3. Java 6 Programming Black Book 2007 Edition, Dreamtech press.

Course Title: IT and Environment

Course Code: ICSC5CR4

Total Credits: 4

Course Objectives

- To understand the importance of internet and use of IT in teaching and learning.
- To identify, formulate and solve environmental problems by utilizing the concept of environmental studies.
- To create awareness among people about protection of wild life & forests.
- To understand the impact of e-waste and use of green computing.
- Understanding of environmental policies and regulations.
- To aware with human rights.

Module 1: Introduction to Internet and Environment:
Internet- Internet as a knowledge repository, academic search techniques, creating cyber presence. Academic websites. Multidisciplinary nature of environmental studies -Definition, scope and importance, Need for public awareness. (10 Hrs.)
Module 2: Impact of IT in E-Learning:
Introduction to use of IT in teaching and learning, Learning Management System, Moodle, Edmodo, etc. Academic services– A note on INFLIBNET, NPTEL, NICNET. (10 Hrs.)
Module 3: IT and Society:
IT & Society- issues and concerns- digital divide, IT & development, the free software movement. IT industry: new opportunities and new threats, software piracy, cyber ethics, cybercrime, cyber threats, cyber security, privacy issues, cyber laws, cyber addictions, information overload, health issues guidelines for proper usage of computers, internet and mobile phones. Impact of IT on language & culture. (10 Hrs.)
Module 4: E-waste and Green Computing:
E-waste- Problems- Solutions-Impact of e-waste in living beings and environment- a study on e-waste management in India. Green computing, definition, meaning, scope. Green computing in India. (10 Hrs.)
Module 5: Human Rights:
An Introduction to Human Rights, Meaning, concept and development –History of Human Rights- Different Generations of Human Rights- Universality of Human Rights- Basic International Human Rights Documents - UDHR ,ICCPR,ICESCR.-Value dimensions of Human Rights
Human Rights and United Nations : Human Rights co-ordination within UN system- Role of UN secretariat The Economic and Social Council- The Commission Human Rights-The Security Council and Human rights The Committee on the Elimination of Racial Discrimination- The Committee on the Elimination of Discrimination Against Women- the Committee on Economic, Social and Cultural Rights- The Human Rights Committee- Critical Appraisal of UN Human Rights Regime.
Human Rights National Perspective : Human Rights in Indian Constitution – Fundamental Rights- The Constitutional Context of Human Rights-directive Principles of State Policy and Human Rights- Human Rights of Women- children –minorities- Prisoners- Science Technology and Human Rights- National Human Rights Commission- State Human Rights Commission- Human Rights Awareness in Education. (14 Hrs.)
Case Study:
The students need to view the film “Samaksham”, a film on environment produced by Mahatma Gandhi University Creations and submit a compulsory assignment reviewing film. The review is considered for internal mark assessment.
References
1. K.L. James, The Internet: A User's Guide 2nd Revised edition, PHI publication.

2. Bharucha Erach, Text Book of Environmental Studies for undergraduate Courses. University Press, IInd Edition 2013 (TB)
(https://www.ugc.ac.in/oldpdf/modelcurriculum/env.pdf).
3. Barkha and U Rama Mohan, Cyber Law & Crimes, 3rd Edition, Asia Law House.
4. Rakesh Johri, E-waste: Implications, regulations, and management in India and current global best practices, Teri publications.
5. Alan Evans, Kendall Martin, Mary Anne Poatsy, Technology in Action, Pearson.

Course Title Project Minor - Phase I

Course Code ICSD5PR5

General guidelines for Phase I and II

Each student needs to undertake a project work to implement various phases of Software Development. The project work is divided into two phases, Phase I in 5th and Phase II in the 6th Semester

The Phase I includes Problem identification and statement, system study and system design. For this he/she needs to identify a problem which is not yet automated. Analyze the Manual System existing there and suggest a framework of software befitting the problem. In the fifth semester (Phase I), system must be developed showing the data flow and needed DFD's and Database design. Evaluation of Project-Phase I will be internal for which the student needs to submit a spiral bound report of their work and appear before a team of faculty members formed by the head of the Department. The team should comprise of the Project Guide and two other Faculty members from the Department.

The Phase II of the project work in the Sixth semester is purely meant for development of software implementing the project identified and designed in Phase I. Student needs to develop a software using any of the Language or Package they have studied in their syllabus. Usage of any other language/package needs to get approved by the Committee/ Project Guide/HoD of the Department No student should be allowed to change the project work in Phase II and they need to develop the software for the work they identified in phase1. The internal mark for the Project work is to be awarded based on the student's performance in Phase I and II (50 percent Weightage for both the semesters). The Sixth semester project evaluation will be of External Nature along with its internal component. Students have to submit a hard bounded report for the evaluation.

Course Title: Software Lab V-Java and PHP

Course Code: ICSD5CP6

Total Credits: 3

I. Advanced Java Programming: Lab

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)

II. Web Application Development Using PHP: Lab

Develop programs for implementing the following concepts

1. Expressions and Control Flow in PHP.
2. PHP Functions -Returning a Value, Returning an Array, Passing Arguments by Reference, Returning Global Variables
3. Constructors, Destructors, Inheritance
4. PHP Arrays - Numerically Indexed Arrays, Associative Arrays, foreach Loop, Multidimensional Arrays.
5. MySQL Database - create, insert, delete, update, select operations using HTML form
6. HTML5 Enhancements- autocomplete Attribute, autofocus Attribute, placeholder Attribute, required Attribute, Override Attributes, width and height Attributes, min and max Attributes, step Attribute, form Attribute, list Attribute, colour Input Type, number and range Input Types, Date and Time Pickers
7. JavaScript - Conditionals, Loops, Functions, Objects, and Arrays
8. Asynchronous Communication - XMLHttpRequest, Sending XML Requests,
9. CSS Selectors- Type Selector, Descendant Selector, Child Selector, ID Selector, Class Selector, Attribute Selector, Universal Selector, Selecting by Group.
10. CSS- Fonts and Typography- font-family, font-style, font-size, font-weight, Managing Text Styles- Decoration, Spacing, Alignment, Transformation, Indenting.
11. CSS Colors- Short Color Strings, Gradients
12. Positioning Elements- Absolute Positioning, Relative Positioning, Fixed Positioning.
13. Pseudoclasses,
14. Setting Margins, Applying Borders, Adjusting Padding.
15. JQuery- Selectors, css Method, Element Selector, ID Selector, Class Selector, Combining Selectors
16. JQuery event handling
17. Special Effects, Hiding and Showing, toggle Method, Fading In and Out, Sliding Elements Up and Down.
18. Manipulating the DOM
19. Using jQuery Without Selectors- $.each Method, $.map Method
20. Asynchronous Communication

SEMESTER - VI

Si. No.	Course Code	Title	Hrs/ Week	Credits	Total Hours
1	ICSC6CR1	Linux and Shell Programming	3	3	54
2	ICSC6CR2	Computer Networks	4	4	72
3	ICSD6CR3	Mobile Application Development Using Kotlin	4	4	72
4	ICSC6EA1/2/3	Elective 1 [Bunch A]	3	3	54
5	ICSD6PR4	Project Minor Phase II	7	4	126
6	ICSD6CP5	Software Lab VI : Machine Learning Techniques	4	2	72
		Total	25	20	450

Course Title: Linux and Shell Programming

Course Code: ICSC6CR1
Total Credits: 3

Course objective

Upon completion of the course, students will be able to:

- gain working knowledge in Linux environment.
- get a clear view on Linux file system
- understand process scheduling in Linux.
- facilities for user creation and management and basics of shell programming.

Module 1:

Open Source Software : Free software foundation, Freedoms in free software, GNU project, Introduction and development of Linux, , advantages of Linux, Hardware requirement, Installing Linux, Linux File System overview, Linux Architecture, Boot Process, Kernel, shell - the user interface, GUI and CLI commands, Usage of input-output redirection (>, >> etc.), Basic Commands in Linux – creating directories, changing directories, listing directory contents, file related commands (create and edit text files, renaming, copying, deleting), Introduction to Shells, Different shells and their features, Online manuals in Linux, the Man command. (10 Hrs.)

Module 2:

The Linux File system: Partitioning the disk, Disk command, Important files and directories in Linux, the hierarchical file system. The root directory / and its important sub directories /bin, /etc, /dev /lib, /boot /home. /mnt, /tmp, /user. Navigating through the file system. Absolute and relative pathnames. The disk related commands, df, du, creating new partitions in Linux, deleting a partition, mounting and unmounting file systems , file types, file related commands Find,

touch, cat etc. searching for a pattern, concept of wild cards and regular expressions , grep, egrep commands. (10 Hrs.)

Module 3:

Process Management in Linux The Concept and properties of Processes, Creating processes, the Parent processes and child processes, PID's and its relevance, Killing processes and sending signals to a process (kill, killall, xkill), How to start and monitor processes, Identify CPU/memory intensive processes, adjust process priority, Processor scheduling in Linux. The Batch command, The at Command, nohup command, File processing commands, wc, cut, paste, sort , Mathematical Commands expr and factor. Different editors in Linux, Consol based editors and GUI based editors and comparison of basic features (Vi, ed, emacs, gEdit etc..). (12 Hrs.)

Module 4:

Users and Group Management, useradd, usermod, userdel, groupadd, groupmod, groupdel. Adding a New User Account, User Private Groups, Modifying / Deleting User Accounts, Group Administration, Password Aging Policies, Switching Accounts, passwd command, logging in as Super user, Networked Users and communication, Authentication, Configuration, Default File Permissions, Changing file ownership (chown), Changing file group ownership (chgrp), Permissions on files, Permissions on directories, How permissions are applied, Changing permissions (chmod), Access Control Lists (ACLs). (10 Hrs.)

Module 5:

Shell Programming: The role of shells in the Linux environment, : The bash shell, Shell commands, Other standard shells, Write a simple shell script to welcome users, Comments in a script ,Setting up permissions on a script, debug and Execute a script, Variables in shell, The export statement, Unset shell and environment variables, Getting User Input Via Keyboard , Bash variable existence check, Customize the bash shell environments: Recalling command history, Path name expansion, Create and use aliases, The tilde expansion, Startup scripts-(Using aliases, Changing bash prompt, Setting shell options , Setting system wide shell options), Commonly Used Commands and Utilities (ls,rm,cat etc),Developing shell scripts for adding a User, Changing Password of users etc. (10 Hrs.)

Book of Study

1. A Practical Guide to Linux Commands, Editors, and Shell Programming, 4th Edition, by Mark
G. Sobell, Matthew Helme, Prentice Hall, 2018. ISBN: 978-0-13-477460-2.
2. The redhat Linux Bible: Christopher Negus: Wiley Dreamtech India
3. Unix Shell Programming: Yeshwant kanetkar. BPB publications.

Course Title: Computer Networks

Course Code: ICSC6CR2

Total Credits: 4

Course Objectives

Upon completion of this course, the students will be able to:

- Understand the concepts of signals and OSI layer functions.
- Discuss the process of Multiplexing, switching and difference between guided and unguided media in networks.
- Describe, analyze various data link layer protocols.
- Describe and analyze various network, and transport layer protocols.
- Have a basic knowledge of the use of cryptography and network security.

Module 1:

Introduction to Networks, Data and signals-analog and digital, periodic analog signals, digital signals, bit rate, baud rate, bandwidth. Transmission impairments- attenuation, distortion and noise. Data communication protocols and standards, Network models - OSI model-layers and their functions. TCP/IP protocol suite. (14 Hrs.)

Module 2:

Bandwidth utilization Multiplexing: FDM, TDM, spread spectrum. Transmission Media- guided media and unguided media. Switching: message, Circuit and packet switched networks, datagram networks, virtual- circuit networks. (12 Hrs.)

Module 3:

Data link layer: Error Detection and Correction, Framing, flow and error control, Protocols - Noiseless channels (Simplest, Stop and Wait) and Noisy channels (Stop and Wait and Piggy Backing). Multiple Access Protocols. Random Access-ALOHA, CSMA. Wired LANs- IEEE standards, wireless LANs-Bluetooth, Cellular Telephony (16 Hrs.)

Module 4:

Network layer and Transport layer: Repeaters, Bridges, Gateways and routers. Logical addressing – IPV4 and IPV6 addressing, Internet protocol - IPV4 and IPV6. Connectionless and Connection Oriented Services: UDP and TCP. Congestion Control, Quality of Service. (16 Hrs.)

Module 5:

Application layer: HTTP, FTP, SMTP, DNS. Network security: Common Threats- Firewalls (advantages and disadvantages), Cryptography. (14 Hrs.)

Book of Study

1. B. A. Forouzan - Data communication and Networking, Fourth edition,TMH
2. Andrew S Tanenbaum - Computer Networks, Fourth Edition, Prentice Hall of India.

Course Title: Mobile Application Development Using Kotlin

Course Code: ICSD6CR3

Total Credits: 4

Course Objective

Upon successful completion of this course, students should be able to:

- Familiarize with Kotlin programming
- Develop Android applications using Kotlin

Module 1:

Kotlin Basics - Characteristics of Kotlin -Program Elements-Literals, Variables, Expressions and Statements, Keywords, Whitespace, Operators, Blocks, Comments, Basic Types-Numbers and Literal Constants, Characters, Booleans, Arrays, Strings and String Templates, Controlling Program Flow - ifs, when Statement, while Statement, for loops, Exception Handling, Handling Nulls. (10 Hrs.)

Module 1:

Functions- Declaring Functions, Single Expression Functions, Default Arguments, Named Parameters, Variable Number of Arguments, Extension Functions, Infix Functions, Operator Overloading.

Working with Types- Interfaces, Diamond Problem, Invoking Super Behaviour, Classes- Constructors, Inheritance, Properties, Data Classes, Visibility Modifiers, Access Modifiers, Object Declarations. (12 Hrs.)

Module 3:

Lambdas and Higher Order Functions- Higher Order Functions, Lambda and Anonymous Functions, Parameters in Lambda Expressions, Closures, with and apply.

Collections, Arrays - Arrays, Collections, Lists, Sets, Maps, Collections Traversal, Filter and Map.

Generics -Why Generics, Terminologies, Using Generics in Functions, Using Generics in Classes, Variance, Subclass vs Subtype, Reified Generics (12 Hrs.)

Module 4:

Activities and Layouts -Activity Class, Layout File, View and ViewGroup Objects, Containers, Event Handling-Introduction to Event Handling, Intents- Definition, Loose Coupling, Types of Intent- Explicit, Implicit Intents, Fragments.

Running in the Background- Basic Concepts, UI Thread, Threads and Runnables , Handler Class, Async Task, Anko’s doAsync . (10 Hrs.)

Module 5:

Data sharing and Storage -Shared Preferences, Sharing Data Between Activities, Internal Storage- Overview of File Storage, Internal and External Storage, Cache Directory.

Broadcast Receivers- Introduction to Broadcast Receivers, System Broadcast vs Custom Broadcast, Manifest Registration vs Context Registration, Basics of Broadcast Receivers, Implicit vs Explicit Broadcast Actions. (10 Hrs.)

Book of Study

1. Learn Android Studio 3 with Kotlin: Efficient Android App Development - Ted Hagos, Apress, Inc.

References

1. Android Development with Kotlin - Marcin Moskala, Igor Wojda, Packt.

2. Kotlin In-Depth [Vol-I] -A Comprehensive Guide to Modern Multi-Paradigm Language - Aleksei Sedunov, BPB Publications.
3. Learn Kotlin for Android Development- The Next Generation Language for Modern Android Apps Programming, Peter Spath, Apress.
4. Programming Kotlin Applications- Building Mobile and Server-Side Applications with Kotlin, By Brett McLaughlin, Wiley.

Course Title: Project Minor - Phase II

Course Code: ICSD6PR4
Total Credits: 4

General guidelines for Phase I and II

Each student needs to undertake a project work to implement various phases of Software Development. The project work is divided into two phases, Phase I in 5th and Phase II in the 6th Semester

The Phase I includes Problem identification and statement, system study and system design. For this he/she needs to identify a problem which is not yet automated. Analyze the Manual System existing there and suggest a framework of software befitting the problem. In the fifth semester (Phase I), system must be developed showing the data flow and needed DFD's and Database design. Evaluation of Project-Phase I will be internal for which the student needs to submit a spiral bound report of their work and appear before a team of faculty members formed by the head of the Department. The team should comprise of the Project Guide and two other Faculty members from the Department.

The Phase II of the project work in the Sixth semester is purely meant for development of software implementing the project identified and designed in Phase I. Student needs to develop a software using any of the Language or Package they have studied in their syllabus. Usage of any other language/package needs to get approved by the Committee/ Project Guide/HoD of the Department No student should be allowed to change the project work in Phase II and they need to develop the software for the work they identified in phase1. The internal mark for the Project work is to be awarded based on the student's performance in Phase I and II (50 percent Weightage for both the semesters). The Sixth semester project evaluation will be of External Nature along with its internal component. Students have to submit a hard bounded report for the evaluation.

Course Title: Software Lab VI - Mobile Application Development Using Kotlin

Course Code ICSD6CP5
Total Credits: 2

Develop GUI programs for android operating system and implement the following concepts.

1. Lambdas and Higher Order Functions
2. Collections, Arrays
3. Generics
4. Layouts, Activities, Intent, Fragments
5. UI Thread, Threads and Runnables, Handler Class
6. Async Task, Anko’s doAsync.
7. Shared Preferences
8. Broadcast Receivers (System Broadcast, Custom Broadcast).

SEMESTER – VII

Si. No.	Course Code	Title	Hrs/ Week	Credits	Total Hours
1	ICSC7CR1	Computational Mathematics	4	4	72
2	ICSD7CR2	Applied Statistics for Data Science	4	4	72
3	ICSD7CR3	Advanced Python Programming for data Science	4	4	72
4	ICSD7CR4	Data Engineering in Data Science	4	3	72
5	ICSD7CP5	Software Lab VII: Python Programming Lab for data Science	5	3	90
6	ICSD7CP6	Software Lab VIII: Data Engineering Lab	4	2	72
		Total	25	20	450

Course Title: Computational Mathematics

Course Code: ICSC7CR1
Total Credits: 4

Course objectives

Upon successful completion of this course, students should be able:

- To Understand set relations and functions, Use of Permutation and Combination for arranging objects.
- To do Predicate and Propositional Calculus for Precise reasoning.
- To Gain the methods of fuzzy logic, use the fuzzy set theory, and recognize fuzzy logic membership function, Understand Fuzzification and Defuzzification.
- To understand different concepts in automata theory and formal languages and determine solution to simple automata problems.
- To recognize real-world problems that are amenable to mathematical analysis, and formulate mathematical models of such problems

Module 1:

Sets, Relations and Functions: Set Operations, Representation and Properties of Relations, Equivalence Relations, Partially Ordering, Functions, Domain and Range, Types of Functions.
Counting and Mathematical Induction: Basics of Counting, Pigeonhole Principle, Permutations and Combinations, Inclusion- Exclusion Principle, Mathematical Induction.

Module 2:

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 3:

Fuzzy Logic:- Fuzzy Set Theory :- Fuzzy Versus Crisp – Crisp sets – Operations on Crisp Sets, Properties of Crisp Sets , Fuzzy Sets, Basic Fuzzy Set Operations, Properties of Fuzzy Sets – Crisp Relations, Operations on Crisp Relations - Fuzzy Relations –Operations on Fuzzy Relations , Properties, Membership Functions, Fuzzification, Defuzzification Methods.

Module 4:

Theory of Automata: Definition, Description of finite automata, Transition system and its properties, Acceptability of a string by a finite automaton, NFA, Equivalence of DFA and NFA, Minimization of finite automata, Construction of minimum automaton.
Regular sets and Regular grammars: Regular expressions, identities for regular expressions, finite automata and regular expressions, transition system containing ^-moves, conversion of nondeterministic systems to deterministic systems, Algebraic method using Arden‘s theorem, Construction of finite automata equivalent to a regular expression.

Module 5:

Languages and grammars: Basic definition and example, Definition of a grammar, derivation and the language generated by a grammar, Chomsky classification of languages, Context free languages and derivation trees

References

1. J.P. Tremblay & R Manohar- Discrete Mathematical Structures with Applications to Computer Science, McGraw Hill.
2. K.L.P Mishra & N. Chandrasekaran -Theory of Computer Science (Automata, Languages and Computation), Prentice Hall of India.
3. George J Klir& Bo Yuan- Fuzzy sets and Fuzzy logic Theory and applications, Prentice Hall of India.
4. Fuzzy Set Theory, Fuzzy Logic & Their Applications; Dr A K Bhargava- S Chand Publications.
5. Discrete Mathematics and Its Applications; Kenneth H. Rosen and Kamala Krithivasan- McGraw-Hill Publishers.

Course Title Applied Statistics for Data Science

Course Code ICSD7CR2
Total Credits: 4

Course Objectives

- A student would have in depth understanding of the key statistical concepts for applying strong knowledge base in Analytics domain.

- Students will learn to apply various statistical theories to solve real life situations by doing projects.
- Students will learn to apply R as a tool for statistical application.

Module 1:

Inferential Statistics: Unbiased Estimation, maximum likelihood estimators, Minimum variance unbiased estimators, Existence and construction of sufficient statistics, Complete family of distributions. Cramer Rao inequality, Rao-Blackwell theorem.

Module 2:

Time series Analysis: Time series as stochastic process, stationary time series- covariance stationarity, Modelling Time Series Data, Exponential Smoothing Methods - First-Order and Second order Exponential Smoothing, Forecasting, Exponential Smoothing for Seasonal Data, Exponential Smoothers, Autocorrelation function (ACF), partial auto correlation function (PACF), correlogram, AR, MA, ARMA, ARIMA Models, Yule- Walker equations, Box-Jenkins Model fitting and diagnostics. Forecasting future values,

Module 3:

Bayesian statistics: Bayesian parametric models, conjugate prior, Bayesian estimators – Hypothesis testing: testing framework, parametric testing, permutation test, multiple testing.

Module 4:

Design of Experiments: Basic principles of experimental design, uniformity trails, analysis of variance, one-way, two-way and three-way classification models, completely randomized design (CRD), randomized block design (RBD) Latin square design (LSD) and Graeco-Latin square designs, Analysis of covariance (ANCOVA), ANCOVA with one concomitant variable in CRD and RBD.

Module 5:

Probability and Statistics in R, Distributions, Hypothesis Tests in R, Simulation, Modeling, Estimate and visualize a regression model using R.

References

1.Peng R. D, Exploratory data analysis with R, Lulu.Com, 2012.

2. Peng R. D, R programming for data science, Leanpub, 2016.

3. Teetor P, R cookbook: Proven recipes for data analysis, statistics, and graphics, O' Reilly Media Inc., 2011.

4. Crawley M. J., The R book, John Wiley & Sons, 2012.

5. Montgomery D. C., Cheryl L. J., and Murat K. (2015) Introduction to Time Series Analysis and Forecasting. John Wiley & Sons.

6. Brockwell P.J and Davis R.A. (2002) Introduction to Time Series and Forecasting Second edition, Springer-Verlag.

7. Ruey S. Tsay (2005). Analysis of Financial Time Series, Second Ed. Wiley & Sons.

8. Abraham, B., &Ledolter, J. (2009). Statistical methods for forecasting (Vol. 234). John Wiley & Sons.
9. Chatfield, C. (2004). The Analysis of Time Series - An Introduction (Sixth edition), Chapman and Hall.
10.Berger, J. O. (2013). Statistical decision theory and Bayesian analysis. Springer Science & Business Media.
11. Raaiffe H. & Schlaiffer R. (2000) Applied Statistical Decision Theory, M.T.Press.
12. Das M.N. and Giri N.C. (1994) Design and analysis of experiments, Wiley Eastern Ltd.
13.Montgomery, C.D. (2012) Design and Analysis of Experiments, John Wiley, New York.
14. Dean, A. and Voss, D. (1999) Design and Analysis of Experiments, Springer Texts in Statistics.
15.S. C. Gupta and V.K.Kapoor (2017) Fundamentals of Mathematical Statstics, Sultan Chand & Sons.

Course Title: Advanced Python Programming for Data Science

Course Code: ICSD7CR3
Total Credits: 4

Course Objectives

On completion of course, students should be able:

- To learn about some Python functionality and techniques that are commonly used.
- To understand and use functionality of various Python libraries for different scientific and mathematical tasks.
- To gain basic insight of implementation of advanced concepts and use of various libraries for applying Machine Learning for problem solving.
- To acquire knowledge about the frameworks in Python.
- To analyze large data sets in Python fro data science.

Module 1:

Introduction: Review of Important Python Concepts, Overview of Advanced techniques in Python: Lambdas, Filter and map, is and id, Decorators, Iterators and Generators, Garbage Collector, environment, Exception handling, Interop module, Pickle, Marshal, Networking Concepts, Process and Threads, Sockets, Regular Expression, Heuristic search techniques.

Module 2:

Scientific and Numerical Computing with Python: Introduction to Scientific and Numerical computing, Introduction to various modules used for Scientific and Numerical programming: NumPy, SciPy, Scikit-Learn, MatplotLib and Keras & Pandas, Introduction of Internal Statistics, overview of common approaches to multivariate statistics.

Module 3:

Introduction to Processing of Data Sets: Overview of various Data sets, Data handling Techniques: using Structured and unstructured Files, Excel and SQL Files. Data Preprocessing and Data Analysis using Pandas and Seaborn, Data Visualization, Exploring duplicate data and missing data, Data fitting concepts, Introduction to collection modules, counter, data storage offline.

Module 4:

Introduction to Frameworks used with Python – TensorFlow: Concept of Computational Graph and Nodes Virtual Environment and Anaconda, Installing TensorFlow with GPU support on a Linux System, TF Datatypes, Placeholders, TF Variables, TF Session, Softmax, One Hot Encoding, Dropout, building hidden layers, Batching, Stochastic Gradient Descent, Building an Optimizer, Training and displaying outcome, Overview of various python frameworks

Module 5:

Implementation of Machine Learning concepts in Python: Introduction to Machine Learning Approaches, Overview of ML tasks: Supervised Learnings: Classifications, Regression. Unsupervised Learnings: Clustering, Semi-supervised Learning, Reinforcement Learning, Basics of implementation of Machine Learning modules using Python.

References

1. Rao N.R., "Core Python Programming", Dreamtech Publication India
2. Sarker M.O.F., "Python Network Programming Cookbook", Packt Publication
3. Sebastian Raschka, "Python Machine Learning", Packt Publication
4. Willi Richert, "Building Machine Learning Systems with Python", Packt publication
5. Fredrik Lundh, "Python Standard Library", O'Reilly Publications
6. Halterman R.,"Fundamentals of Python Programming", Southern Adventist University
7. Guttag J.V., "Introduction to Computation and Programming Using Python", Prentice Hall India
8. Chun W., "Core Python Programming", Prentice Hall India

Course Title: Data Engineering in Data Science

Course Code: ICSD7CR4

Total Credits: 3

Course Objectives

- General concepts in representing data, accessing it and analysing it.
- Provide application side of query processing, data Wrangling, Cleaning etc.
- Able to develop suitable data science ecosystem for the given application.
- Understand various data storage and retrieval techniques.
- Analyzing of data using Python and Data Wrangling.

Module 1:

Introduction: Data source, Big Data, Structured and unstructured data. Data Models and Storage: Relational databases, NoSQL database, normalized and denormalized data models, Data cleaning, Distributed Data Storage and Management, Hashing, Indexing.

Module 2:

Query processing: Querying big data using SQL and NoSQL, Elastic Search, Query optimization, speeding up, maintaining ACID property, Design Patterns, Data reliability, quality and provenance, Distributed query processing, Query optimization and Processing.

Data Warehousing: OLAP, OLTP.

Streaming Data analytics: In-memory Analytics, data pipelines and dashboards, Predictive Analytics (6 Lectures)

Module 3:

Data Wrangling - Importance of Data Wrangling - Tasks of Data Wrangling-Data Wrangling Tools-Introduction to Python-Python Basics-Data Meant to Be Read by Machines-CSV Data-JSON Data-XML Data.

Data Cleanup Basics-Identifying Values for Data Cleanup-Formatting Data-Finding Outliers and Bad Data-Finding Duplicates-Fuzzy Matching-RegEx Matching-Normalizing and Standardizing the Data-Saving the Data-Determining suitable Data Cleanup-Scripting the Cleanup Testing with New Data.

Module 4:

Exploring Data-Importing Data-Exploring Table Functions-Joining Numerous Datasets-Identifying Correlations-Identifying Outliers-Creating Groupings-Analyzing Data-Separating and Focusing the Data -Presenting Data-Visualizing the Data-Charts-Time-Related Data-Maps-Interactives-Words-Images, Video, and Illustrations-Presentation Tools-Publishing the Data-Open-Source Platforms.

Module 5:

Web Scrape -Analyzing a Web Page-Network/Timeline-Interacting with JavaScript-In-Depth Analysis of a Page-Getting Pages-Reading a Web Page-Reading a Web Page with LXML-XPath-Advanced Web Scraping-Browser-Based Parsing-Screen Reading with Selenium-Screen Reading with Ghost. PySpidering the Web-Building a Spider with Scrapy-Crawling Whole Websites with Scrapy.

References

1. M. KLEPPMANN (2017), Designing Data-Intensive Applications the Big Ideas Behind Reliable, Scalable, And Maintainable Systems, O'Reilly.
2. L. WEISE (2015), Advanced Data Management: For SQL, Nosql, Cloud and Distributed Databases, Walter De Gruyter Gmbh.
3. A. SILBERSCHATZ, H.F. KORTH, S. SUDARSHAN (2011), Database System Concepts, Mcgraw Hill Publications, 6th Edition.

4. Jacqueline Kazil & Katharine Jarmul," Data Wrangling with Python", O'Reilly Media, Inc,2016
5. H.G. MOLINA, J. ULLMAN, J. WIDOM (2014), Database Systems: The Complete Book, Pearson, 2nd Edition.
6. P. RAJ, A. RAMAN, D. NAGARAJ, S. DUGGIRALA (2015), Hig

Course Title: Software Lab VII - Python Programming Lab for data Science

Course Code: ICSD7CP5
Total Credits: 3

Course Objectives

Student will get acquainted with the various libraries and functions in python to help to understand the Data Science and machine Learning concept better.

Course Content

1. Usage of libraries Pyspark, Numpy, Scipy, Matplotlib, Pandsas, Python Script and Variable.
2. Reading different types of data sets (.txt, .csv) from Web and disk and writing in file in specific disk location.Reading Excel data sheet
3. Descriptive Statistical Analysis – Evaluation, Plotting and Interpretation Linear Regression: Read a data frame in csv/xls format containing the weather data such as pressure, min temp, max temp, humidity, and rainfall. Using the Pandas, MatPlotlib and SciPy plot the scatter plots and develop a linear interpolation between rainfall with all other parameters and evaluate the statistical significance of the model.
4. Creation of python programs on the modules Numpy Analyse the given series of data using pandas Python programs that uses the dictionaries, tuples and other data structures
5. Decision Tree Classification, attribute selection measures, and how to build and optimise Decision Tree Classifier using Python Scikit-learn
6. Apply regression Model techniques to predict the data
7. a) Install relevant classification packages.
   b) Choose a classifier for the classification problem.
   c) Evaluate the performance of the classifier.
8. Clustering algorithms for unsupervised classification. Plot the cluster data using python with Matplotlib visualizations.

Course Title: Software Lab VIII - Data Engineering Lab

Course Code: ICSD7CP6
Total Credits: 2

Course Objectives

- Perform Read and write operations on CSV, JSON and XML files

- Process the Excel file using Pandas
- Parse and Extract the Tables using Python library
- Apply the basis of Data cleanup operation on the given dataset
- Explore the web scraping in Python

Course Content

1. Write programs to parse text files, CSV, HTML, XML and JSON documents and extract relevant data. After retrieving data check any anomalies in the data, missing values etc.
Design a relational database for a small application and populate the database. Using SQL do the CRUD (create, read, update and delete) operations.
2. Import any CSV file to Pandas DataFrame and perform the following:
    (a) Handle missing data by detecting and dropping/ filling missing values.
    (b) Transform data using apply() and map() method.
    (c) Detect and filter outliers.
    (d) Perform Vectorized String operations on Pandas Series.
    (e) Visualize data using Line Plots, Bar Plots, Histograms, Density Plots and Scatter Plots.
3. Write a Python script to read each row from a given csv file and print a list of strings.
4. Write a Python program to read a given CSV file as a dictionary.
5. Write a Python program to convert Python dictionary object (sort by key) to JSON data. Print the object members with indent level 4
6. Write the python script to Read the XML file
7. Write a Pandas program to import excel data (child labour and child marriage data.xlsx) into a Pandas data frame and process the following a. Get the data types of the given excel data b. Display the last ten rows. c. Insert a column in the sixth position of the said excel sheet and fill it with NaN values
8. Develop the python script to parse the pdf files using pdfminer.
9. Extract the Table from the child labour and child marriage data.xlsx using pdfables library
10. Write a Python data wrangling scripts to insert the data into SQLite database
11. Develop the Python Shell Script to do the basic data cleanup on child labour and child marriage data.xlsx a. Check duplicates and missing data b. Eliminate Mismatches c. Cleans line breaks, spaces, and special characters
12. Import the data into `agate`then explores the table using agate methods and perform statistical correlations
13. Draw the chart between perceived corruption scores compared to the child labour percentages using matplotlib.
14. Write the python script to Map the Child Labour Worldwide using pygal.
15. Write a Python program to download and display the content of robot.txt for en.wikipedia.org

SEMESTER – VIII

Si. No.	Course Code	Title	Hrs/ Week	Credits	Total Hours
1	ICSD8EB1/2	Elective 2 [Bunch B]	4	4	72
2	ICSC8CR1	Advanced Deep Learning Techniques	4	4	72
3	ICSD8CR2	Data Visualization	4	4	72
4	ICSD8EC1/2	Elective 3 [Bunch C]	4	4	72
5	ICSD8CP3	Software Lab IX: Deep Learning Lab using R	4	2	72
6	ICSD8CP4	Software Lab X: Data Visualization Lab using Tableau	5	2	90
		Total	25	20	450

Course Title: Advanced Deep Learning Techniques

Course Code: ICSC8CR1
Total Credits: 4

Course Objectives

- To understand the theoretical foundations, algorithms and methodologies of Neural Network
- To design and develop an application using specific deep learning models
- To provide the practical knowledge in handling and analyzing real world applications

Module 1: Deep Learning Architectures
Machine Learning and Deep Learning, Representation Learning, Width and Depth of Neural Networks, Activation Functions: RELU, LRELU, ERELU, Unsupervised Training of Neural Networks, Restricted Boltzmann Machines, Auto Encoders, Deep Learning Applications. (14 Hrs.)
Module 2: Convolutional Neural Networks
Architectural Overview, Motivation, Layers, Filters, Parameter sharing, Regularization, Popular CNN Architectures: ResNet, AlexNet – Applications. (14 Hrs.)
Module 3: Transfer Learning
Transfer learning Techniques, Working, Approaches in Transfer Learning, Variants of CNN: DenseNet, PixelNet. (14 Hrs.)
Module 4: Sequence Modelling – Recurrent and Recursive Nets.
Recurrent Neural Networks, Bidirectional RNNs, Encoder-decoder sequence to sequence architectures - BPTT for training RNN, Long Short-Term Memory Networks. Auto Encoders:

Under complete Auto encoder, Regularized Auto encoder, stochastic Encoders and Decoders, Contractive Encoders. (16 Hrs.)

Module 5: LSTM, GRU

Image Segmentation, Image classification, Object Detection, Automatic Image Captioning, Image generation with Generative Adversarial Networks, LSTM as a classifier Model, Attention Models for Computer Vision. (14 Hrs.)

References

1. Ian Goodfellow, YoshuaBengio and Aaron Courville, "Deep Learning", MIT Press, 2017.
2. Josh Patterson, Adam Gibson "Deep Learning: A Practitioner's Approach", O'Reilly Media, 2017
3. Umberto Michelucci "Applied Deep Learning. A Case-based Approach to Understanding Deep Neural Networks" Apress, 2018.
4. Deep learning - Heaton, J. Ian goodfellow, yoshua bengio, and aaron courville, The MIT Press, First Edition

Course Title: Data Visualization

Course Code: ICSD8CR2

Total Credits: 4

Course Objectives

- To expose to visual representation methods and techniques that increase the understanding of complex data.
- To study good design practices for visualization.
- Practice the core principles using widely available tools like Tableau.
- Apply fundamental concepts of data visualization on projects.

Module 1: **Introduction to Data Visualisation**

Definition – Methodology – Seven Stages of Data Visualisation - Data Visualisation Tools. Visualising Data: Mapping Data onto Aesthetics – Visualising Amounts - Visualising Distributions: Histograms and Density Plots – Visualising Propositions: – Visualising Associations: Among Two or More Quantitative Variables – Visualising Time Series and Other Functions of an Independent Variable – Trends – Visualising Geospatial Data.

Module 2: **R: Interactive Data Visualisation**

Introduction to D3 - Fundamental Technology: The Web – HTML – DOM – CSS – JavaScript – SVG. D3 Setup – Generating Page Elements – Binding Data - Drawing with data – Scales: Domains and Ranges – Normalization – Creating a Scale – Scaling the Scatter Plot – Other Methods and Other Scales. Axes – Modernizing the Chart – Update the Data – Transition – Updates – Interactivity.

Module 3: **D3 Based Reusable Chart Library**

Setup and Deployment – Generate Chart – Customize Chart: Additional Axis – Show Axis Label – Change Chart Type – Format Values – Size – Color – Padding –Tooltip. Use APIs: Load and Unload – Show and Hide – Focus – Transform – Groups – Grid – Regions – Flow – Revert – Toggle –Legend – Sub chart – Zoom – Resize. Customize Style. Building Real time and Live Updating animated graphs with C3.

Module 4: **Tableau Introduction**

Environment Setup – Navigation – File & Data Types. TA SOURCE: Custom Data View – Extracting Data – Fields Operations – Editing Meta Data – Data Joining – Data Blending. Worksheets.

Module 5: **Basic and Advanced Charts in Tableau**

Bar Chart – Line Chart – Pie Chart – Scatter Plot – Bubble Chart –Gantt Chart – Histograms - Waterfall Charts. Dashboard – Formatting – Forecasting – Trend Lines.

References

1. Ben Fry, "Visualizing Data: Exploring and Explaining Data with the Processing Environment", O'Reilly, 1st Edition, 2008.
2. Scott Murray, "Interactive data visualization for the web: An Introduction to Designing with D3", O'Reilly, 2nd Edition, 2017.
3. Joshua N. Milligan, "Learning Tableau 2019: Tools for Business Intelligence, data prep, and visual analytics", Packt Publishing Limited, 2019.
4. Claus O. Wilke, "Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures", O.Reilly, 2019.

Course Title: Software Lab IX - Deep Learning Lab using R

Course Code: ICSD8CP3

Total Credits: 2

1. Implement a four layer deep neural network model for classification / regression. Compare the performance of the model by using various activation functions like RELU, LRELU, ERELU.
2. Implement an autoencoder model for feature extraction.
3. Implement an image classifier using CNN.
4. Design a pretrained model for image classification. Compare the performance using various pretrained models like ResNet, AlexNet and DenseNet
5. Implement a temperature predictor using RNN and LSTM. Include a comparison of performance when we used RNN and LSTM.
6. Design a model for Automatic Image Captioning using GRU.

Course Title: Software Lab X - Data Visualization Lab using Tableau

Course Code: ICSD8CP4

Total Credits: 2

1. Creating interactive data visualization (2D visualization) in the browser using D3.
2. Create D3-based reusable charts that enables deeper integration of charts into web applications.
3. Drawing histograms, density plots and scatter plots using Tableau.

SEMESTER – IX

Si. No.	Course Code	Title	Hrs./Week	Credits	Total Hours
1	ICSD9ED1/2	Elective 4 [Bunch D]	4	4	72
2	ICSD9CR1	Text Analytics & Natural Language Processing	4	3	72
3	ICSD9CR2	Web Analytics	4	4	72
4	ICSC9EE1/2	Elective 5 [Bunch E]	4	4	72
5	ICSD9CP3	Software Lab XI: NLP using R	4	2	72
6	ICSD9PR4	Case study and Minor project	5	3	90
		Total	25	20	450

Course Title: Text Analytics & Natural Language Processing

Course Code: ICSD9CR1

Total Credits: 3

Course Objectives

Text analytics concepts and applications

Fundamental of Information retrieval and natural language processing

Text analytics framework

Theoretical techniques and applications in text analytics (e.g. social media)

Module 1:

Introduction to Natural Language Processing-History of NLP, Text Analytics and NLP, Various Steps in NLP, Kick Starting an NLP project.

Types of Data-Structured, Semi Structured and Unstructured Data, Categorization of Data Based on Content

Module 2:

Basic Feature Extraction Methods-Cleaning Text Data-Tokenization, Feature Extraction from Texts, Feature Engineering.

Natural Language Processing and Python-Lists, Regular Expressions, Dictionaries, Writing Functions, Text Wrangling and Cleansing.

Module 3:

Building a Custom Corpus-Domain Specific Corpora, Corpus Data Management-Corpus Disk Structure.

Corpus Preprocessing and Wrangling- Breaking Down Documents, Corpus Transformation-Intermediate Preprocessing and Storage, Reading the Processed Corpus.

Module 4:

Classification for Text Analysis- Text Classification, Building a Classification Application- Cross Validation, Model Construction, Model Evaluation, Model Operationalization

Module 5:

Context-Aware Text Analysis- Grammar-Based Feature Extraction, n-Gram Feature Extraction, n-Gram Language Models.

Text Visualization-Visualizing Feature Space, Visualizing Clusters, Visualizing Classes.

References

1. Sohom Ghosh, Dwight Gunning, "Natural Language Processing Fundamentals", Packt Publishing.
2. Nitin Hardeniya, Jacob Perkins, Deepti Chopra, Nisheeth Joshi, Iti Mathur, "Natural Language Processing: Python and NLTK", Packt Publishing
3. Benjamin Bengfort, Rebecca Bilbro, and Tony Ojeda, "Applied Text Analysis with Python Enabling Language-Aware Data Products with Machine Learning", O'Reilly

Course Title: Web Analytics

Course Code: ICSD9CR2

Total Credits: 3

Course Objectives

- To provide common overview of web analytic activities
- Having problem solving ability- web analytics solutions.
- Ability to use web analytics Tools like Google Analytics, Yahoo Analytics

Module 1:

Introduction to Web Analytics: Definition, Types of Web Analytics, User Experience and Web Analytics Questions., A brief history of web analytics, Current Landscape and Challenges, Traditional Web analytics, Web analytics Activities, Measuring, Trinity.

Basic metrics of web analytics- Pageviews -Bounce rate −Pages per session − Demographic info – Devices. Source: Traffic sources − Organic traffic -Social traffic -Referral Direct traffic. Micro and Macro–Level Data Insights.

Data collection: Understanding the Data Landscape, Click stream Data, Outcomes Data, Research Data, Competitive Data.

Module 2:

Web Analytics Approach: Introduction, A model of Analysis, Showcasing the work, Context Matters, Contradicting the data. Working of Web Analytics: Introduction, Log File Analysis, Page Tagging, Metrics and Dimensions, Interacting with data in Google Analytics.

Module 3:

Introduction to Goals and Conversions-, Definition of Goals and Conversions, Conversion Rate, Goal Reports in Google Analytics, Finding the right things to measure as key, Performance

Indicators, Measure on a website that can constitute a goal. Learning about users: Introduction, Visitor Analysis.

Module 4:

Introduction: Search Query, Source and medium, Organic Search, Search Query -Navigational-Informational-Transactional, Search Query Analysis, Referral Traffic, Direct Traffic, Paid Search Keyword. Analyzing usage of content: introduction, Website content Reports

Module 5:

Click-Path Analysis: Introduction, Focus on Relationships between pages, Navigation Summary, Visitors Flow Report, analyzing how users move from one page type to another.

Segmentation: Introduction, Necessity, Procedure to segment, Ways to Segment, Useful ways to segmentUX questions.

Tools -Google Analytics – Piwik Web Analytics – Yahoo Web Analytics – Emerging Analytics: Social - Video - Mobile.

References

1. Avinash Kaushik - "Web Analytics 2.0: The Art of Online Accountability and Science of Customer Centricity" - 1st Edition - Sybex - 2009.
2. Avinash Kaushik - "Web Analytics: An Hour a Day" - 6th Edition - Sybex - PAP/ CDR Edition - 2007.
3. Brian Clifton - "Advanced Web Metrics with Google Analytics" - 3rd Edition- Sybex - 2012.

Course Title: Software Lab XI: NLP using R

Course Code: ICSD9CP3

Total Credits: 2

The candidate needs to submit a Lab Record, duly signed by the teacher in charge and Head of the Department, (minimum of 10 Programs), failing which he/she will not be allowed to attend the external software lab examination. The Lab record should be hard-binded with the name of college and the emblem of the college depicted on the first page and should be properly indexed.

1. Preprocessing of text (Tokenization, Filtration, Script Validation, Stop Word Removal, Stemming)
2. Morphological Analysis
3. N-gram model
4. POS tagging
5. Chunking
6. Named Entity Recognition
7. Virtual Lab on Word Generator and Word Analysis
8. Morphology
9. N-Grams
10. N-Grams Smoothing

11. Building POS Tagger
12. Building Chunker

Course Title: Case study and Minor Project

Course Code: ICSD8CP4
Total Credits: 3

Case study and Minor Project aims at giving students hands-on experience in applying the programming knowledge to develop a real application for data science, based on the case study conducted by the student on a real-world scenario. Students must take up individual projects. Evaluation of the project is external. The case study has to be presented in the project report, and submitted in hard bound format for evaluation.

The students will work on multiple case studies and projects from different domain specified by guide allotted to him. This course aims at discussing the key principles of knowledge discovery process through various case studies arising from different application areas. The students are expected to learn the main steps to traverse when they face new data analytics problems. With each case study, the tools for cleaning, processing and altering the data shall be visited. A particular attention shall be given to data inspection, feature reduction and model selection. Each case study will be completed by a thorough discussion and interpretation of the results.

SEMESTER – X

Si. No.	Course Code	Title	Hrs./Week	Credits	Total Hours
1	ICSDXPR1	Major Project	25	16	450
2	ICSDXVV2	Comprehensive Viva Voce		4	
		Total	25	20	450

Course Title: Major Project

Course Code: ICSDXPR1
Total Credits: 16

Industry or research internship should include partial/complete project implementation. Student should beallocated to the research guide in 7th semester itself and same guide should be continued for the**:** Industry Internship/ In house Research Project. Otherwise, the preferences/choices of the domain should be taken from the students. The guide needs to be allocated based on the preference/choices. The research project should be assigned to students. In case of Industry Internship, the assigned guide from college has to monitor and evaluate the progress of the student. The student has to exhibit the continuous progress through regular reporting and presentations and proper documentation. The continuous assessment of the progress needs to be documented unambiguously.

Course Title: Comprehensive Viva Voce

Course Code: ICSDXVV2
Total Credits: 4

The objective of comprehensive viva-voce is to assess the overall knowledge of the student in the relevant field of Computer Science acquired over 5 years of study in the integrated programme. The viva shall normally cover the subjects taught in all the semesters of programme. In doing so, the main objective of this course is to prepare the students to face interview both in the academic and the industrial sector.

Integrated MSc Computer Science- Data Science

ELECTIVES

Bunch A	
ICSC6EA1	Cloud Computing
ICSC6EA2	Full stack programming Techniques
ICSC6EA3	Predictive Analytics
Bunch B	
ICSD8EB1	Advanced DBMS
ICSD8EB2	Business Intelligence & Analytics
Bunch C	
ICSD8EC1	Image and Video Analytics
ICSD8EC2	Geospatial Analysis
Bunch D	
ICSD9ED1	Healthcare data Analytics
ICSD9ED2	Social media Analytics
Bunch E	
ICSC9EE1	Block Chain Technology
ICSC9EE2	Big Data Analytics

Elective: Bunch A

1. ICSC6EA1: Cloud Computing
2. ICSC6EA2: Full stack programming Techniques
3. ICSC6EA3: Predictive Analytics

Course Title: Cloud Computing

Course Code: ICSC6EA1
Total Credits: 3

Course Objectives

Cloud computing has evolved as a very important computing model, which enables information, software, and shared resources to be provisioned over the network as services in an on-demand manner.

On completion of this course provides an insight into what is cloud computing and the various services cloud is capable.

Module 1:

Introduction to Computing Paradigms, High-Performance Computing, Parallel Computing., Distributed Computing, Cluster Computing, Grid Computing, Cloud Computing., Biocomputing, Mobile Computing, Quantum Computing, Optical Computing. Nanocomputing.

Introduction: Benefits and Limitations-Cloud Architecture – Storage – Services –Service Providers - Types of Cloud Service Development – Services and Tools

Module 2:

Collaborating on Contact Management - Collaborating on Project Management- Collaborating on Word Processing, Spreadsheet, Presentations, Databases- Sharing Files and Photographs

Module 3:

Cloud Virtualization Technology – Virtualization Defined – Virtualization Benefits – Server Virtualization– Virtualization for x86 Architecture – Hypervisor Management Software – Logical Partitioning – VIO Server – Virtual Infrastructure Requirements

Module 4:

Deep Dive: Cloud Virtualization –Introduction - Storage Virtualization–Storage Area Networks– Network Attached Storage – Cloud Server Virtualization – Virtualized Data Center

Module 5:

Industrial platforms and new developments - Amazon web services: Compute services - Storage services- Communication services - Additional services - Google AppEngine: Architecture and core concepts - Application life cycle - Cost model Microsoft Azure: Azure core concepts - SQL Azure - Windows Azure platform appliance

References

1. Essentials of cloud Computing : K.Chandrasekhran , CRC press, 2014
2. Cloud Computing: Principles and Paradigms by Rajkumar Buyya, James Broberg and Andrzej M. Goscinski, Wiley, 2011.
3.Cloud Computing: Insights into New Era Infrastructure, Dr. Kumar Saurabh (2011). , Wiley India,
4. Mastering Cloud Computing Foundations and Applications Programming, Rajkumar Buyya, Christian Vecchiola, S. Thamarai Selvi (2013).
5. Cloud Computing: Fundamentals, Industry Approach and Trends, Rishabh Sharma (2014),Wiley India edition.

Course Title: Full-Stack Development

Course Code: ICSC6EA2
Total Credits: 3

Course Objectives

Learner can:

- Identify Structure and implementation of HTML/CSS.
- Apply intermediate and advanced web development practices.
- Implement basic JavaScript.
- Create visualizations in accordance with UI/UX theories.
- Develop a fully functioning website and deploy on a web server.

Module 1: HTML and CSS

Introduction to HTML, Browsers and HTML, Editor's Offline and Online, Tags, Attribute and Elements, Doctype Element, Comments, Headings, Paragraphs, and Formatting Text, Lists and Links, Images and Tables.

Introduction CSS, Applying CSS to HTML, Selectors, Properties and Values, CSS Colors and Backgrounds, CSS Box Model, CSS Margins, Padding, and Borders, CSS Text and Font Properties, CSS General Topics.

Module 2: JavaScript

Introduction to JavaScript, Applying JavaScript (internal and external), Understanding JS Syntax, Introduction to Document and Window Object, Variables and Operators, Data Types and Num Type Conversion, Math and String Manipulation, Objects and Arrays, Date and Time, Conditional Statements, Switch Case, Looping in JS, Functions.

Module 3: ReactJS

Introduction to ReactJS, Templating using JSX, Components, State and Props, Lifecycle of Components, Rendering List and Portals, Error Handling, Routers, Redux and Redux Saga, Immutable.js, Service Side Rendering, Unit Testing, Webpack.

Module 4: NodeJS

Node js Overview, Basics and Setup, Console, Command Utilities, Modules, Concepts, Events, Node js with Express js, Database Access.

Module 5: MongoDB

SQL and NoSql Concepts, Create and Manage MongoDB, Migration of Data into MongoDB, MongoDB with PHP, MongoDB with NodeJS, Services Offered by MongoDB, Connect MongoDB with Python.

References

1. The Full Stack Developer: Your Essential Guide to the Everyday Skills Expected of a Modern Full Stack Web Developer; Chris Northwood- Apress publications.
2. Full Stack Development with MongoDB; Manu Sharma-bpb publications.
3. Modern Full-Stack Development:Using TypeScript, React, Node.js, Webpack, and Docker; Frank W. Zammetti- Apress publications.
4. Mastering Html, Css & Javascript Web Publishing; Laura Lemay, Rafe Colburn, Jennifer Kyrnin- -bpb publications.

Course Title: Predictive Analytics

Course Code: ICSC6EA3
Total Credits: 3

Course Objectives

After completing this class, the student will develop the following competencies.

- Competency in Predictive Analytics Methods.
- Competency in Predictive Analytics Tools.
- Competency in the Predictive Analytics Cycle

Module 1:

Introduction: - Prediction Versus Interpretation, Key Ingredients of Predictive Models, Predictive Modeling Process. Data Pre-processing: - Data Transformations for Individual Predictors- Centering and Scaling, Transformations to Resolve Skewness, Data Transformations for Multiple Predictors- Transformations to Resolve Outliers, Data Reduction and Feature Extraction, Removing Predictors- Predictor Correlations, Adding Predictors, Binning Predictors.

Module 2:

Over-Fitting and Model Tuning-The Problem of Over-Fitting Model Tuning, Data Splitting, Resampling Techniques. Regression Models- Quantitative Measures of Performance, Linear Regression- Partial Least Squares, Penalized Models, Nonlinear Regression Models - Neural Networks, K-Nearest Neighbors.

Module 3:

Forecasting and time series analysis: Introduction to Decision Trees, Chi- Square Automatic Interaction Detectors (CHAID), Classification and Regression Tree (CART), Analysis of

Unstructured data, Naive Bayes algorithm. Classification Models: - Introduction of Classification Models - Discriminant Analysis and Other Linear Classification Models - Nonlinear Classification Model -Naïve Bayes - Support Vector Machines.

Module 4:

Introduction to Feature Selection -Consequences of Using Non-informative Predictors, Approaches for Reducing the Number of Predictors-Factors That Can Affect Model Performance- Measurement Error in the Outcome, Measurement Error in the Predictors.

Module 5:

Predicting Cognitive Impairment Predicting Caravan Policy Ownership, The Effect of Class Imbalance- Sampling Methods-Cost-Sensitive Training-Job Scheduling-Case Studies –Real world scenario where forecasting and Time series analysis.

Books of Study

1. Kuhn, Max, Kjell Johnson. “Applied predictive modeling”. Springer, 2018.
2. Montgomery “Applied statistics and probability for engineers” Third edition.

References

1. Siegel, Eric. “Predictive analytics”: The power to predict who will click, buy, lie, or die. John Wiley & Sons, 2013.
2. Abbott, Dean. “Applied predictive analytics”: Principles and techniques for the professional data analyst. John Wiley & Sons, 2014.
3. Miner, Gary, “Practical text mining and statistical analysis for non-structured text data applications”. Academic Press, 2012.

Elective: Bunch B

1. ICSD8EB1: Advanced DBMS
2. ICSD8EB2: Business Intelligence & Analytics

Course Title: Advanced DBMS

Course Code ICSD8EB1
Total Credits: 4

Course Objectives

- The design and implement Distributed Databases.
- To understand advanced DBMS techniques to construct tables and write effective queries, forms, and reports.
- Learn MongoDB to design queries.

Module 1:

Introduction to Database Systems and E-R Model : DBMS vs other systems: Overview, A Historical Perspective, Files System vs DBMS, Advantages of DBMS, Describing and storing data in a DBMS, Structure of a DBMS.

Module 2:

Structured Query Language and PL/SQL

Creating a table- Displaying table information - Altering an existing table – Dropping, renaming, Specifying Constraints in SQL, Basic Retrieval Queries in SQL, Additional Features of SQL, More SQL - Complex Queries, More Complex SQL Retrieval Queries, Views (Virtual Tables) in SQL: 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.

PL/SQL: Fundamentals - Block structure - comments - Data types – Other data types - Variable declaration - Assignment operation - Bind variables - Substitution variables - Printing, Control Structures and Embedded SQL: Control structures - Nested blocks - SQL in PL/SQL - Data manipulation - Transaction control statements.

Module 3:

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 4:

Introduction to NoSQL databases

NoSQL database concepts: Types of data -structured, unstructured, semi structured data, Bigdata, Types of NoSQL databases, NoSQL data modelling, Benefits of NoSQL, Comparison between SQL and NoSQL database systems.

NoSQL using MongoDB: Introduction to MongoDB shell, Running the MongoDB shell, MongoDB client, Basic operations with MongoDb shell, Basic datatypes, Arrays, Embedded documents. Querying with MongoDB: find() function, specifying which keys to return, query criteria, OR queries,

Module 5:

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.

References:

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.James R.Groff and Paul N. Weinberg,The complete reference SQL, 2nd edition,Tata McGraw Hill
4. T. M. Connolly and C. Begg -Database Systems: Practical approach to design, implementation, and management (sixth edition), Publisher: Pearson Education.
5. Vaswani Vikram, Complete Reference: MySQL, McGraw Hill Education.
6. Bayross Ivan, SQL, PL/SQL, The Programming language of ORACLE, BPB Publications, 3rd Rdition.
7. An Introduction to Database Systems, Date, C. J. Pearson Education, New Delhi, 2012.

Course Title: Business Intelligence & Analytics

Course Code ICSD8EB2

Total Credits: 4

Course Objectives

- To make them aware business analytics foundation.
- To make them ready to apply data analytics skills to the area of business intelligence
- Lean predictive analytics in business intelligence.

Module 1:

Business intelligence: Introduction, Concepts and Applications. Business Intelligence for better decisions, Business Intelligence types, tools, skills, applications. Decision making and Analytics - business intelligence, analytics and decision support. Foundation and technologies for decision making.

Module 2:

Business applications of Decision Trees, Regression, Artificial Neural Networks, Cluster Analysis, Association Rule Mining - Techniques, Algorithm, Exercise, Advantages and Disadvantages.

Module 3:

Big data and future directions for Business Analytics- Big Data Analytics, Business Analytics, Emerging Trends and Future Impacts. Business applications of Big Data, Technologies and Management Big data.

Module 4:

Predictive Analytics: Data mining in Business Intelligence- Text Mining,Web Mining - Business applications, practices and algorithms. Descriptive Analytics - Data warehousing, Business Reporting, Visual Analytics and Business Performance Management.

Module 5:

Understanding BI and Mobility, BI and Cloud Computing, Business Intelligence for ERP Systems, Social CRM and BI

References

1. Business Intelligence and Data Mining, Anil K. Maheshwari, PhD,Business Expert Press, LLC,2015

2. Business Intelligence And Analytics: System For Decision Support, Ramesh Sharda (Oklahoma State University), Dursun Delen(Oklahoma State University),Efraim Turban(University of Hawaii), Pearson Education,Inc., 2015
3. Fundamentals of Business Analytics, 2ed, R N Prasad, Seema Acharya

Elective: Bunch C

1. ICSD8EC1: Image and Video Analytics
2. ICSD8EC2: Geospatial Analysis

Course Title: Image and Video Analytics

Course Code: ICSD8EC1
Total Credits: 4

Course Objectives

Describe the fundamental principles of image and video analytics and have an idea of their application.

Able to apply state-of-the-art machine learning techniques (convolutional neural networks) to solving problems in image and video analysis

Module 1:

Introduction to Human visual perception, digital image representation, Image sampling and quantization, Mathematical tools for image processing: Vector and matric operations- Image Transforms (DFT, DCT, DWT, Hadamard).

Module 2:

Image and video segmentation and texture models; -Image and video demonizing- Image and Video enhancement- Image and Video compression. Image Pyramids for analysis and image compression;

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

Module 3:

Differential Motion Analysis methods, Change detection, Segmentation using motion, Image flow, segmentation using Moving camera, Optical flow, Analysis based on correspondence of interest points, detection of specific motion patterns, video tracking, motion models to aid tracking

Module 4:

Real-time video analytics and video mining, temporal and spatial event recognition, Vision-based activity recognition), Behaviour Analysis, Content-Based Analysis of Digital Video

Module 5:

Video Analytics: state of the art applications with reference to computer vision applications, Deep learning in video analytics, Human motion recognition and its applications, Video

Analytics for Business Intelligence, Virtual reality/Augmented reality applications, and Healthcare applications.

References

1. Sonka, Hlavac, Boyle, "Digital Image Processing and Computer Vision"- CENGAGE Learning, Indian Edition
2. Ramesh Jain, Kasturi, Schunck, "Machine Vision",McGraw-Hill
3. R.C. Gonzalez and R.E. Woods." Digital Image Processing", 3rd Edition. Addison Wesley, 2007.
4. White paper: Video Analytics: Technologies and use cases
https://wso2.com/whitepapers/innovating-with-video-analytics-technologies-and-use-cases/#07

Course Title: Geospatial Analysis

Course Code: ICSD8EC2
Total Credits: 4

Course Objectives:

By successfully completing the course, students should be able to

- command the theories and methods
- implement and practice typical geospatial methods
- analyze and visualize these implementations
- be familiar with open-source tools and data

Module 1

Introduction and terminology: Spatial analysis, GIS and software tools, Intended audience and scope, Software tools and Companion Materials, Terminology and Abbreviations, Common Measures and Notation. Conceptual Frameworks for Spatial Analysis: Basic Primitives, Spatial Relationships, Spatial Statistics, Spatial Data Infrastructure.

Module 2

Methodological Context: Analytical methodologies, Spatial analysis as a process, Spatial analysis and the PPDAC model, Geospatial analysis and model building, The changing context of GIScience. Building Blocks of Spatial Analysis: Spatial and Spatio-temporal Data Models and Methods, Geometric and Related Operations, Queries, Computations and Density, Distance Operations, Directional Operations, Grid Operations and Map Algebra.

Module 3

Data Exploration and Spatial Statistics, Statistical Methods and Spatial Data, Exploratory Spatial Data Analysis, Grid-based Statistics and Metrics, Point Sets and Distance Statistics, Spatial Autocorrelation, Spatial Regression.

Module 4

Surface and Field Analysis, Modeling Surfaces, Surface Geometry, Visibility, Watersheds and Drainage, Gridding, Interpolation and Contouring, Deterministic Interpolation Methods, Geostatistical Interpolation Methods. Network and Location Analysis, Introduction to Network and Location Analysis, Key Problems in Network and Location Analysis, Network Construction, Optimal Routes and Optimal Tours, Location and Service Area Problems, Arc Routing.

Module 5

Geocomputational methods and modeling: Introduction to Geocomputation, Geosimulation, Artificial Neural Networks (ANN), Genetic Algorithms and Evolutionary Computing, Big Data and Geospatial Analysis, Big Data and Research: Types of Big Data, Challenges of Big Data.

References:

1. Dr Michael J de Smith, Prof Michael F Goodchild, Prof Paul A Longley & Associates, Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools 6th edition, 2018.
2. Hassan Abdishakur, Geospatial Data Science Quick Start Guide, Packt Publishing Limited.
3. Joel Lawhead, Learning Geospatial Analysis with Python: Understand GIS fundamentals and perform remote sensing data analysis using Python 3.7, Packt Publishing Limited, 3rd Edition.

Elective: Bunch D

1. ICSD9ED1: Healthcare data Analytics
2. ICSD9ED2: Social media Analytics

Course Title: Health Care Data Analytics

Course Code: ICSD9ED1
Total Credits: 4

Course Objectives

Explore the various forms of electronic health care information for Analytics. To learn and implement various techniques for analyze health care data and to understand the predictive models for clinical data

Able to apply analytics for decision making in healthcare services. Able to apply data mining to integrate health data from multiple sources and develop efficient clinical decision support systems.

Module 1:

Introduction: Healthcare Data Analytics, Data Sources: Electronic Health Records– Biomedical Image -Sensor Data - Biomedical Signal - Genomic Data - Clinical Text- Mining of Biomedical Literature. EHR-Components of EHR, Benefits and Barrier in EHR, Phenotyping Algorithms.

Module 2:

Analysis: Biomedical Image Analysis- Bio Medical Image -Imaging Modalities-Object Detection-Image segmentation-Image registration – Feature Extraction , Mining of Sensor Data in Healthcare- Biomedical Signal Analysis- Genomic Data Analysis for Personalized Medicine.

Module 3:

Analytics: Natural Language Processing and Data Mining for Clinical Text: NLP Components-Mining Information fron Clinical Text-Challenges. Mining the Biomedical Literatures- Social Media Analytics for Healthcare.

Module 4:

Advanced Data Analytics: Advanced Data Analytics for Healthcare– Review of Clinical Prediction Models- Temporal Data Mining for Healthcare Data- Visual Analytics for Healthcare-Predictive Models for Integrating Clinical and Genomic Data-Information Retrieval for Healthcare-Privacy-Preserving Data Publishing Methods in Healthcare.

Module 5:

Models for Integrating Clinical and Genomic Data- Information Retrieval for Healthcare-Privacy- Preserving Data Publishing Methods in Healthcare.

Applications and Practical Systems for Healthcare: Data Analytics for Pervasive Health, Healthcare Fraud Detection , Data Analytics for Pharmaceutical Discoveries , Clinical Decision Support Systems, Computer-Aided Diagnosis, Mobile Imaging for Biomedical Applications

References

1. Chandan K. Reddy and Charu C Aggarwal, “Healthcare data analytics”, Taylor & Francis, 2015
2. Hui Yang and Eva K. Lee, “Healthcare Analytics: From Data to Knowledge to Healthcare Improvement, Wiley, 2016.

Course Title: Social Media Analytics

Course Code: ICSC9EE2

Total Credits: 4

Course Objectives

Upon successful completion of this course, students will be able to:

- Describe the different types of data commonly found on social platforms
- Use a social platform API to obtain data and understand the structure of those data
- Visualize that corpus along geographic and temporal axes
- Describe how and why different networks exist within the same data
- Compute a variety of networks measures from a social media dataset
- Describe the various types of text commonly found on social platforms
- Compute sentiment over social text

Module 1:

Introduction-New Challenges for Mining, Graph basics- Graph Representation, Types of Graphs, Connectivity in Graphs, Special Graphs, graph algorithms, Network measures centrality,

transitivity and reciprocity, balance and status, similarity, Network Models - Properties of Real-World Networks, Random Graphs, Small-World Model, Preferential Attachment Model

Module 2:

Data Mining Essentials- Data, Data Preprocessing, Data Mining Algorithms, Supervised Learning, Unsupervised Learning.

Module 3:

Communities and Interactions- Community Analysis, Community Evolution, Community Evaluation Information Diffusion in social media- Herd Behavior, Information Cascades, Diffusion of Epidemics.

Module 4:

Influence and Homophily- Measuring Assortativity, Influence, Homophily, Distinguishing Influence and Homophily

Recommendation in social media- Challenges, Classical Recommendation Algorithms, Recommendation Using Social, Evaluating Recommendations

Module 5:

Behavior Analytics- Individual Behavior, Individual Behavior Analysis, Individual Behavior Modelling, Individual Behavior Prediction, Collective Behavior

References

1. Reza Zafarani, Mohammad Ali Abbasi. Huan, "Social Media Mining- An Introduction", Cambridge University Press, 2014.
2. Jure Leskovec, AnandRajaraman, Jeffrey D. Ullman, "SMining of Massive Datasets".

Elective: Bunch E

1. ICSC9EE1: Block Chain Technology
2. ICSC9EE2: Big Data Analytics

Course Title: Blockchain Technology

Course Code: ICSC9EE1

Total Credits: 4

Course Objectives

Explain the working of Blockchain Technology. Integrating Blockchain technology to real world scenarios

Ability to understand what and why of Blockchain. Explore various components of Blockchain and its use . Can create their own Blockchain network application.

Module 1:

Introduction to Block chain Technology, the growth of blockchain technology, Block Chain as a distributed system, the history of blockchain and Bitcoin, Blockchain defined, Generic elements

of Blockchain, How Blockchain works, Benefits and limitations of blockchain, Tiers, Feature and Types. consensus mechanism, Type of consensus mechanism, Consensus in blockchain.

Module 2:

Introduction To Cryptocurrency: Bitcoin – Bitcoin Platform, Bitcoin Architectures - Digital Keys and Addresses – Transactions – Mining – Bitcoin Networks and Payments – Wallets – Alternative Coins – Theoretical Limitations – Bitcoin limitations. Ethereum concept and Ethereum classic.

Module 3:

Consensus Protocols and Security Issues Trust Essentials: Decentralized Systems, Consensus Protocols: Proof-of-Work (PoW), Proof-of-Stake (PoS), Delegated Proof-of-Stake (DPoS), Proof-of-Burn (PoB), Byzantine Fault Tolerance (BFT), Practical Byzantine Fault Tolerance (PBFT), Proof-of-Activity (PoA), Proof of Elapsed Time (PoET). Blockchain Security Threats, Challenges and Issues.

Module 4:

Enterprise Blockchain Platforms, Enterprise Blockchain Platform: Hyperledger, Hyperledger Architecture, Membership, Blockchain, Transaction, Chaincode, Hyperledger Fabric, Features of Hyperledger, Fabric Demo.

Module 5:

Blockchain Applications: Building on the Blockchain, Smart Contract and Ethereum Platform Introduction Ethereum, Architecture, Smart Contracts, Elements of Smart Contracts, Ethereum Operations, Incentive Model, Transactions in Ethereum, Introduction Solidity

References

1. A. Bahga, V. Madisetti (2017), Blockchain Applications: A Hands-On Approach, VPT.
2. M. Swan (2015), Blockchain: Blueprint for a New Economy, O'Reilly Media.
3. R. Wattenhofer (2016), The Science of the Blockchain, CreateSpace Independent Publishing Platform.
4. I. BASHIR (2017), Mastering blockchain, Packt Publishing Ltd.

Course Title: Big Data Analytics

Course Code: ICSC9EE2

Total Credits: 4

Course Objectives

- To explore the fundamental concepts of big data analytics.
- To learn to analyze the big data using intelligent techniques.
- To understand the various search methods and visualization techniques.
- To learn to use various techniques for mining data stream.
- To understand the applications using HDFS and Map Reduce Concepts
- To learn Hadoop ecosystem.

Module 1:

Understanding Big Data-Concepts and Terminologies. Big Data Characteristics-Volume, Velocity, Variety, Veracity, Value. Different types of data-structured, unstructured, semi structured, metadata. Business Motivations and Drivers for Big Data Adoption. Big Data Analytics Life Cycle.

Module 2:

Big Data Processing Concepts-Parallel Data Processing, Distributed Data Processing, Hadoop, Processing Workloads, Cluster, Processing in Batch Mode.

Hadoop Fundamentals-Introduction, Core Components, HDFS Daemons, Map Reduce Daemons, Resource Allocation with YARN, Workflow of MapReduce Job, HDFS High-Availability Daemons, Benefits and Challenges of HDFS.

File Sizes, Block Sizes, and Block Abstraction in HDFS, Data Replication, Data Locality, Network Topology, Network Bandwidth, and Rack Placement Policy

Module 3:

HDFS and MapReduce- Hadoop Distributed File System, MapReduce Framework, Hadoop Cluster Environment.

Map Reduce-Anatomy of a Map Reduce Job Run, Failures, Job Scheduling, Shuffle and Sort, Task Execution, Map Reduce Types and Formats, Map Reduce Features.

Module 4:

Hadoop Ecosystem

Pig : Introduction to PIG, Execution Modes of Pig, Comparison of Pig with Databases, Grunt, Pig Latin, User Defined Functions, Data Processing operators.

Hive : Hive Shell, Hive Services, Hive Metastore, Comparison with Traditional Databases, HiveQL, Tables, Querying Data and User Defined Functions.

Hbase : HBasics, Concepts, Clients, Example, Hbase Versus RDBMS.

Module 5:

Data Analytics Lifecycle. Review of Basic Data Analytic Methods Using R-Introduction to R, Exploratory Data Analysis, Statistical Methods for Evaluation

Advanced Analytics-Technology and Tools: In-Database Analytics-SQL essentials, In-Database Text Analysis, Advanced SQL

References

1. Thomas Erl,Wajid Khattak,Paul Buhler “Big Data Fundamentals: Concepts, Drivers & Techniques”, Prentice Hall.

2. Deepak Vohra, “Practical Hadoop Ecosystem”, Apress.

3. Tom White “Hadoop the Definitive Guide”, O’Reilly.

4. EMC Education Services. “Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data”, Wiley Publishing.