Tumkur University

Department of Studies and Research in Computer Science

Syllabus for

Master of Science in Data Science – M. Sc. (DS) Programme

Choice Based Credit System (CBCS Scheme)

2023 - 24 onwards

Eligibility for M.Sc. Data Science:

Eligibility for Admission to the Programme Candidates for admission to the first-year programme leading to the Degree of Master of Science in Data Science (M.Sc. – DS) will be required to possess: A pass in B.Sc. Computer Science or its equivalents / B.Sc. Data Science / B.Sc. Data Analytics / B.Sc. Mathematics or its equivalents / B.Sc. Statistics or its equivalents / B.Sc. Physics / B.Sc. Electronics / B.E. / B.Tech. / CSE / IT / ECE / EEE and E & I or its equivalents/ Bachelor of Vocational course in computer science or related

.

SEMESTER – I

Sl. No	Paper	Title of the paper	Instruction Hrs. per week	No. of Credits	Duration of the Exam.	Marks		
						Internal Assessment	Semester End Exam.	Total
1	CPT 1.1	Mathematical Foundations for Data Science	4	4	3 Hrs.	30	70	100
2	CPT 1.2	Principals of Data Science	4	4	3 Hrs.	30	70	100
3	CPT 1.3	Python Programming	4	4	3 Hrs.	30	70	100
4	SPT 1.4 A	Data Structure using C	4	4	3 Hrs.	30	70	100
	SPT1.4 B	Advanced Database Management systems	4	4	3 Hrs.	30	70	100
5	CPP 1.5	Social Networking and Soft Skill Enhancement Course	4	2	3Hrs.	15	35	50
6	CPP1.6	Fundamental Tools for Data Science Lab	4	2	3Hrs.	15	35	50
7	CPP 1.7	Python Programming Lab	4	2	3Hrs.	15	35	50
8	SPP1.8 A	Data Structure using C Lab	4	2	3Hrs.	15	35	50
	SPP1.8 B	Advanced Database Management Systems Lab	4	2	3Hrs.	15	35	50
TOTAL			32	24				600

NOTE: Theory: 1 Hr. =1Credit; Practical: 2 Hrs. =1Credit

CPT: Core Paper Theory   SPT: Special Paper Theory
CPP: Core Paper Practical   SPP: Special Paper Practical

SEMESTER – II

Sl.No	Paper	Title of the paper	Instruction Hrs. per week	No. of Credits	Duration of the Exam.	Marks		
						Internal Assessment	Semester End Exam.	Total
1	CPT 2.1	Machine Learning	4	4	3 Hrs.	30	70	100
2	CPT 2.2	Statistics with R Programming	4	4	3 Hrs.	30	70	100
3	SPT 2.3 A	Multivariate Data Analysis	4	4	3 Hrs.	30	70	100
	SPT 2.3 B	Probability Distribution	4	4	3 Hrs.	30	70	100
4	OET 2.4	Information and Communication Technology	4	4	3 Hrs.	30	70	100
5	CPP 2.5	Machine Learning Lab	4	2	3 Hrs.	10	40	50
6	CPP 2.6	Statistics with R Programming Lab	4	2	3 Hrs.	15	35	50
7	SPP 2.7 A	Multivariate Data Analysis Lab	4	2	3 Hrs.	15	35	50
	SPP 2.7 B	Data Analysis Using Excel Lab	4	2	3 Hrs.	15	35	50
8	OEP 2.8	Information and Communication Technology Lab	4	2	3 Hrs.	15	35	50
	TOTAL		32	24				600

NOTE: Theory: 1 Hr. =1Credit; Practical: 2 Hrs. =1Credit

CPT: Core Paper Theory   SPP: Special Paper Practical

CPP: Core Paper Practical   OET: Open Elective Paper

SPT: Special Paper Theory   OEP: Open Elective practical

SEMESTER – III

Sl.No	Paper	Title of the paper	Instruction Hrs. per week	No. of Credits	Duration of the Exam.	Marks		
						Internal Assessment	Semester End Exam.	Total
1	CPT 3.1	Advanced Software Engineering	4	4	3 Hrs.	30	70	100
2	CPT 3.2	Artificial Intelligence	4	4	3 Hrs.	30	70	100
3	SPT 3.3 A	Deep Learning	4	4	3 Hrs.	30	70	100
	SPT 3.3 B	Cloud Computing	4	4	3 Hrs.	30	70	100
4	OET 3.4	Web Programming	4	4	3 Hrs.	30	70	100
5	CPP 3.5	Minor Project	4	2	3Hrs.	15	35	50
6	CPP 3.6	Artificial Intelligence Lab	4	2	3Hrs.	15	35	50
7	SPP 3.7 A	Deep Learning Lab	4	2	3Hrs.	15	35	50
	SPP 3.7 B	Cloud Computing Lab	4	2	3Hrs.	15	35	50
8	OEP 3.8	Web Programming Lab	4	2	3Hrs.	15	35	50
	TOTAL		32	24				600

NOTE: Theory: 1 Hr. =1Credit; Practical: 2Hrs=1Credit

CPT: Core Paper Theory   SPP: Special Paper Practical

CPP: Core Paper Practical   OET: Open Elective Paper

SPT: Special Paper Theory   OEP: Open Elective practical

SEMESTER – IV

Sl.No	Paper	Title of the paper	Instruction Hrs. per Week	No. of Credits	Duration of the Exam.	Marks		
						Internal Assessment	Semester End Exam.	Total
1	CPT 4.1	Research Methodology	4	4	3 Hrs.	30	70	100
2	CPT 4.2	Data Visualization	4	4	3 Hrs.	30	70	100
3	SPT 4.3 A	Natural Language Processing	4	4	3 Hrs.	30	70	100
	SPT 4.3 B	Big Data Analytics	4	4	3 Hrs.	30	70	100
4	CPP 4.4	Mobile App Development Lab	4	2	3 Hrs.	10	40	50
5	CPP 4.5	Data Visualization Lab	4	2	3 Hrs.	15	35	50
6	SPP 4.6 A	Natural Language Processing Lab	4	2	3 Hrs.	15	35	50
	SPP 4.6 B	Big Data Analytics Lab	4	2	3 Hrs.	15	35	50
7	PROJECT 4.7	Major Project	8(4+4)	6	3 Hrs.	15+30	Viva –Voce 35 Dissertation 70	150
	TOTAL		32	24				600

NOTE: Theory: 1 Hr. =1Credit; Practical: 2Hrs=1Credit

CPT: Core Paper Theory   SPT: Special Paper Theory

CPP: Core Paper Practical   SPP: Special Paper Practical

Internal Assessment Marks Allotment for Theory:

Internal Test:	10 Marks
Seminar:	5 Marks
Assignment:	5 Marks
Attendance:	5 Marks
Class room exercise:	5 Marks
TOTAL:	**30 Marks**

Internal Assessment Marks Allotment for Practical:

Internal Test:	10 Marks
Records and Attendance:	5 Marks
TOTAL:	**15 Marks**

First Semester			
Title of the Paper: Mathematical Foundations for Data Science			
Paper	**CPT 1.1**	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course outcome:

- Summarize the vector space properties and analyze different forms of the Linear Transformations.
- Derive the Norms and Inner Product Spaces.
- Recognize and compute the single and multi-sample tests for m descriptive and inferential statistics in many different fields statistics in many different fields.
- Understand the concepts of probability and its applications.

UNIT I   **12 Hours**

Vector space and Transformations:

Vector Spaces, Subspaces, Linear Combinations and Systems of Linear Equations, Linear Dependence and Linear Independence, Bases and Dimension, Maximal Linearly Independent Subsets; Linear Transformations, Null Spaces, and Ranges, The Matrix Representation of a Linear Transformation, Composition of Linear Transformations, and Matrix Multiplication, Invertibility and Isomorphisms, The Change of Coordinate Matrix, The Dual Space.

UNIT II   **12 Hours**

Inner Products and Eigen Values & Eigen Vectors:

Prerequisites: Inner Products and Norms (No theorem proof). The Adjoint of a Linear Operator, Normal and Self-Adjoint Operators, Unitary and Orthogonal Operators and Their Matrices, Orthogonal Projections, The Gram-Schmidt Orthogonalization Process and Orthogonal Complements. Eigenvalues and Cayley-Hamilton Theorem (Statement and properties with relevant problems only), Eigenvectors, and properties of Eigen values and Eigen vectors.

UNIT III   **7 Hours**

Descriptive Statistics:

Meaning of Statistics and its definition-Functions-Characteristics-limitations. Collection of data Classification of data, preparation of frequency distribution and tabulation of data.

Graphical representation of median and mode by - histograms, Cumulative frequency curves (Ogives). Measure of Central Tendency - Arithmetic Mean (Average), Partition values – Median, quartiles, and Mode and its applications.

UNIT IV   **7 Hours**

Methods of Dispersion Range, Quartile deviation, Standard deviations and Coefficient of Variation, Measure of Skewness and Kurtosis.

UNIT V   **10 Hours**

Probability:

Random experiments, trial, sample space, events. Approaches to probability- classical, empirical, subjective and axiomatic. Addition rules of probability. Conditional probability, independence of events and multiplication rule of probability. Bayes theorem (no proof any theorem) and its applications.

Text Books:

1. S. Friedberg, A. Insel, and L. Spence - Linear Algebra, Fourth Edition, PHI, 2009.

2. Kenneth Hoffman, Ray Kunze- Linear Algebra, Second Edition, Eastern Economy Edition, 2011.

3. Gupta. S.C and Kapoor V.K. Fundamentals of Mathematical Statistics, Sultan Chand and sons, (2001)

4. Freund J.E., Mathematical Statistics, Prentice hall, (2001)

Reference Books:

1. Ron Larson and David C. Falvo, Elementary Linear Algebra, Sixth Edition, Houghton Mifflin Harcourt Publishing Company, 2009.

2. Jimmie Gilbert and Linda Gilbert – Linear Algebra and Matrix Theory, Academic Press, An imprint of Elsevier.

3. S. Kumeresan – Linear Algebra, A Geometric approach, Prentice Hall India, 2000.

4. S.P.Gupta, "Statistical methods"- Sultan Chand & Sons, New Delhi, 2012 Edition

5. P K Srimani and M Vinayaka Murthy, "Probability and Statistics", Subhas Stores, 2000

6. Ross Sheldon, A First Course in Probability, Macmillan (6th edition).W.N. Venables, D.M Smith, "An introduction to R"

7. D.C. Montogomery and G.C.Runger, "Applied Statistics and Probability for engineers", New Jersey, John Wiley and Sons, 3rd edition.

First Semester			
Title of the Paper: Principles of Data Science			
Paper	CPT 1.2	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcomes:

- To get the data science and its application area.
- To understand the core concepts and emerging technologies in data science.
- To develop applied experience with data science software, programming, applications and processes.
- To develop practical skills needed in modern analytics.
- To give a hands-on experience with real-world data analysis.

UNIT I   **8 Hours**

INTRODUCTION TO DATA SCIENCE:

Definition – Basic Terminology- Data science Venn diagram- Types of Data- Structured versus Unstructured data- Quantitative versus Qualitative data- The Four Levels of Data- Five steps of Data Science- Data Science Process Overview –Data science classification-Data Science Algorithms- Business Intelligence and Data Science- Components of Data Science.

UNIT II   **12 Hours**

DATA PROCESS AND EXPLORATION:

Introduction-Prior Knowledge-Data Preparation-Modeling-Applications-Objectives of Data Exploration-Datasets- Descriptive statistics- Data Visualization: Introduction- Types of Data visualization- Technologies for visualization - Various visualization techniques - The Five Cs of Data Visualization.

UNIT III   **10 Hours**

DATA MODELLING AND ANALYTICS:

Data Science Methodology- Analytics for Data Science- Data Analytics Examples- Data Analytics Life Cycle- Data Discovery- Data preparation- Model Planning- Model Building Operationalization.

UNIT IV   **10 Hours**

FEATURE SELECTION AND FORECASTING:

Introduction-Feature Selection: Classifying feature selection methods- Anomaly Detection: Introduction- Distance and Density based outlier detection-Local Outlier Factor-Time series Forecasting- Decomposition-Smoothing based methods-Regression based methods-Machine Learning methods.

UNIT V   **8 Hours**

DATA SCIENCE TOOLS AND APPLICATIONS:

Introduction to Data Science Tools- SAS- APACHE FLINK -BigML- Excel- Tableau-Matplotlib Tensor Flow- Weka- Applications: Hands-on with Solving Data Problems-Introduction-Collecting and Analyzing Twitter Data- Collecting and Analyzing YouTube Data.

Text Books:

1.Sanjeev J. Wagh, Manisha S. Bhende, Anuradha D. Thakare, Fundamentals of Data Science, 1 st Edition, 2022

2. Daimi, Kevin, Ed. Hamid R. Arabnia, Principles of Data Science, Springer, 2020.

3. Vijay Kotu, Bala Deshpande, Data Science: Concepts and Practices, Morgan Kaufmann Publishers, Second edition, 2019

4.D J Patil, Hilary Mason, Mike Loukides, Ethics and Data Science, O' Reilly, 1st edition, 2018 5.Sinan Ozdemir, Principles of Data Science, Packt Publishing, December 2016

Reference Books:

1.Jure Leskovek, Anand Rajaraman and Jeffrey Ullman, Mining of Massive Datasets. v2.1, Cambridge University Press, 2014.

2. Cielen, Davy, Arno DB Meysman, Mohamed Ali, Introducing Data Science: Big Data, Machine Learning, and more, using Python Tools, Manning Publications Co., 2016

First Semester			
Title of the Paper: Python Programming			
Paper	CPT 1.3	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcome:

- Analyze the significance of python program development environment and apply it to solve real world applications.
- Demonstrate the usage of lists arrays and loops concepts. Learn the syntax and semantics of the Python programming language.
- Illustrate the process of structuring the data using lists, tuples
- Demonstrate the use of built-in functions to navigate the file system.
- Implement the Object Oriented Programming concepts in Python

UNIT I   **8 Hours**

Installing Python, Simple program using Python, Expressions and Values, Variables and Computer Memory, error detection, multiple line statements, designing and using functions, functions provided by Python, Tracing function calls in memory model, omitting return statement. Working with Text: Creating Strings of Characters, Using Special Characters in Strings, Creating a Multiline String, Printing Information, Getting Information from the Keyboard.

UNIT II   **10 Hours**

A Boolean Type , Choosing Statements to Execute, Nested If Statements , Remembering the Results of a Boolean Expression Evaluation , A Modular Approach to Program Organization, Importing UNITs , Defining Your Own UNITs, Testing Code Semi automatically Grouping Functions Using Methods: UNITs, Classes, and Methods , Calling Methods the Object-Oriented Way, Exploring String Methods, Underscores.

UNIT III   **10 Hours**

Storing Collections of Data Using Lists: Storing and Accessing Data in Lists, modifying Lists, Operations on Lists, Slicing Lists, Aliasing, List Methods, Working with a List of Lists. Repeating Code Using Loops: Processing Items in a List, Processing Characters in Strings, Looping Over a Range of Numbers, Processing Lists Using Indices, Nesting Loops in Loops,

Looping Until a Condition Is Reached, Repetition Based on User Input, Controlling Loops Using Break and Continue Reading and Writing.

UNIT IV   10 Hours

Files: Kinds of files, Opening a File, Techniques for Reading Files, Files over the Internet, Writing Files, and Writing Algorithms that Use the File-Reading Techniques, Multiline Records. Storing Data Using Other Collection Types: Storing Data Using Sets, Storing Data Using Tuples, Storing Data Using Dictionaries, Inverting a Dictionary, Using the In Operator on Tuples, Sets, and Dictionaries, Comparing Collections.

UNIT V   10 Hours

Collection of New Information Object-Oriented Programming: Understanding a Problem Domain, Function "Is instance," Class Object, and Class Book , Writing a Method in Class Book, Plugging into Python Syntax: More Special Methods ,Creating Graphical User interface: Building a Basic GUI, Models, Views, and Controllers, Customizing the Visual Style Introducing few more Widgets, Object-Oriented GUIs, Keeping the Concepts from Being a GUI Mess.

Text Books:

1. Practical Programming: An introduction to Computer Science Using Python, second edition, Paul Gries, Jennifer Campbell, Jason Montojo, The Pragmatic Bookshelf.

2. Learning with Python: How to Think Like a Computer Scientist Paperback – Allen Downey , Jeffrey Elkner, 2015

Reference Books:

1. Introduction to Python for Computational Science and Engineering (A beginner's guide), Hans Fangohr. Exploring Python, Timothy A. Budd, Mc Graw Hill Education Python for Informatics: Exploring Information, Charles Severance. Learning Python, Fourth Edition, Mark Lutz, O'Reilly publication.

First Semester			
Title of the Paper: Data Structure Using C			
Paper	SPT 1.4 A	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcome:

- Differentiate primitive and non-primitive structures
- Design and apply appropriate data structures for solving computing problems.
- Apply sorting and searching algorithms to the small and large data sets.

UNIT I   8 Hours

Introduction and Overview: Definition, Elementary data organization, Data Structures, data structures operations, Abstract data types, algorithms complexity, time-space trade off. Preliminaries: Mathematical notations and functions, Algorithmic notations, control structures, Complexity of algorithms, asymptotic notations for complexity of algorithms. String Processing: Definition, Storing, Strings, String as ADT, String operations, word/text processing, Pattern Matching algorithms.

UNIT II   10 Hours

Arrays: Definition, Linear arrays, arrays as ADT, Representation of Linear Arrays in Memory, Traversing Linear arrays, Inserting and deleting, Multi-dimensional arrays, Matrices and Sparse matrices Linked list: Searching in a Singly linked list, Memory allocation, Garbage collection, Insertion into a singly linked list, Deletion from a singly linked list, Doubly linked list, Header linked list, Circular linked list.

UNIT III   10 Hours

Stacks: Definition, Array representation of stacks, Linked representation of stacks, Stacks as ADT, Arithmetic Expressions: Polish Notation, Conversion of infix expression to postfix expression, Evaluation of post fix expression, Application of stack, Queues: Definition, Array representation of queue, Linked list representation of queues. Types of queue: Simple queue, Circular queue, Double-ended queue, Priority queue, Operations on Queue, Applications of queue.

UNIT IV   **10 Hours**

Binary Tree: Definitions, Tree Search, Traversal of Binary Tree, Tree Sort, Building a Binary Search Tree, Height Balance: AVL Trees, Contiguous Representation of Binary Trees: Heaps, Lexicographic Search Tree; Tries, External Searching: m-array tree, B- Trees, B+ Tree, B*Tree, Applications of Trees, Graphs: Mathematical Back ground, Computer Representation, Graph Traversal, Topological Sorting, Greedy Algorithm, Graphs as Data Structure.

UNIT-V   **10 Hours**

Searching: Introduction and Notation, Sequential Search, Binary Search, Comparison of Methods. Sorting: Introduction and Notation, Insertion Sort, Selection Sort, Shell Sort, Divide and Conquer, Merge sort for Linked List, Quick sort for Contiguous List. Hashing: Sparse Tables, Choosing a Hash function, Collision Resolution with Open Addressing, Collision Resolution by Chaining.

Text Books:

1. Seymour Lipschutz,"Data Structures with C", Schaum's out Lines, Tata Mc Graw Hill, 2011.

2. Robert Kruse, C.L.Tondo, Bruce Leung, Shashi Mogalla, "Fta Structures and Program Design using C", Person Education, 2009.

Reference Books:

1. Marks Allen Weiss, "Data Structures and Algorithm Analysis in C", Second Edition, Pearson Education, 2013.

First Semester			
Title of the Paper: Advanced Database Management Systems			
Paper	SPT 1.4 B	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	4

Course Outcomes:

- To get the basic database concepts, applications, data models, schemas and instances.
- To familiarize entity relationship model for a database.
- To understand the use of constraints and relational algebra operations.
- To understand the basic concepts of transaction processing and concurrency control.
- To emphasize the importance of normalization in databases.

UNIT I   10 Hours

Introduction to Database Management Systems, Purpose of Database Systems, View of Data, Database Languages, Database System Structure, Data Models, Database Design and ER Model: Entity, Attributes, Relationships, Constraints, Keys.

UNIT II   10 Hours

Structured Query Language Basics SQL Overview: Data Types and Literals, DDL, DML, DCL, TCL. Data Definitions, Basic Structure Operations, Additional Operations, Set Operations, Null Values, Aggregate Functions, Nested Sub-Queries. Modifications of Database: Deletion, Insertion and Updates.

UNIT III   8 Hours

Intermediate Structured Query Language: Joins Expressions, Views, Transactions. Relational Integrity: Domain, Referential Integrities, Enterprise Constraints. Data Types and Schemas, Authorizations, Functions and Procedures, Triggers.

UNIT IV   10 Hours

Relational Query Languages & Database Design: Relational Algebra, Tuple relational Calculus, Design Process, Entity Relationship Model, ER Diagram, Design Issues, Extended E-R Features, converting E-R & EER diagram into tables.

UNIT V   **10 Hours**

Relational Model and Normalization:

Relational Database Design Relational Model: Basic concepts, Attributes and Domains, CODD's Rules, Database Design: Features of Good Relational Designs, Normalization, Atomic Domains and First Normal Form, Decomposition using Functional Dependencies, Algorithms for Decomposition, 2NF, 3NF, BCNF.

Text Books:

1. Abraham Silberchatz, Henry K.Forth, Sudharshan, Database System Concepts, 7th edition, McGraw Hill, 2020

Reference Books:

1. R. Elmasri, S.B. Navathe, Fundamentals of Database Systems, Seventh Edition, Pearson Education, 2016.

2 .Bipin C Desai, "An introduction to Database Systems", Galgotia Publications, 2015.

First Semester			
Title of the Paper: Social Networking and Soft Skill Enhancement Course			
Paper	CPP 1.5	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	48	Credits	02

I. SOCIAL NETWORK ETIQUETTES

Introduction to social network – Social Networking Etiquettes - Pros and Cons - Usage of Facebook, Instagram, WhatsApp, Telegram, Youtube, Evolution of Android and IOS, Introduction to Linkedin & Benefits.

List of Practical Programs:

- Official Mail id Creation
- Linkedin Id Creation,
- Linkedin Profile Bulilding ,
- Facebook Id and Creation
- Modifying the existing FB ID
- Create video and upload in the YouTube
- Create an account in Telegram and WhatsApp

II. BROWSING CULTURE

Introduction to browsing – Search Engines-Google - Bing -Yahoo!-AOL -MSN –DuckDuckGo ,browsers, phishing – Cookies - URL – https:// extensions , browsing history, Incognito mode-VPN – Pros and Cons – Book mark.

List of Practical Programs:

- Work with different search engines
- Check the browsing history

III. NETWORKING

Basics of networking - LAN, MAN, WAN, Introduction to network topologies, Protocols , IP Commands (Command line prompt), Define online complier and editor

- Find Your System IP
- Ping Command,
- Firewall Fortinet, Basic DOS Commands

IV. PROFESSIONALISM

Dress Code, Body Language, Appropriate, Attire, Communication Skills, Interview preparation – Introducing yourself - How to greet Superiors, Importance of Eye Contact During conversation.

Resume Skills- Structure of standard resume, types and customizing resume.

List of Practical Programs:

- Micro presentation
- Group Discussion

V. DIGITAL IDENTIFICATION

Introduction to NAD - Importance of Aadhar, PAN Card, Passport, Bank Account, Bar Code, QR scan, Payment Gateway (Gpay, Phone Pe, UPI, BHIM, Paytm), Mobile Banking

List of Practical Programs:

- NAD registration Step by Step,
- Linking bank account with net banking ,
- Register for payment gateway).
- TOTAL : 30 PERIODS
-

WEB REFERENCES :

I: Social Network Etiquettes:

1. https://sproutsocial.com/glossary/social-media-etiquette/

2.https://www.shrm.org/resourcesandtools/tools-and-samples/hrqa/pages/socialnetworkingsitespolicy.aspx

3. https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02711/full

4.https://medium.com/@sirajea/11-reasons-why-you-should-usetelegram-instead-of-whatsapp-ab0f80fbfa79

5. https://buffer.com/library/how-to-use-instagram/

6. https://www.webwise.ie/parents/what-is-youtube/

7. https://www.androidauthority.com/history-android-os-name-789433/

8. https://www.mindtools.com/pages/article/linkedin.htm

II: Browsing Culture:

1.https://sites.google.com/site/bethanycollegeofteacheredn/unit--ictconnecting-with-world/national-policy-on-information-andcommunication-technology-ict/accessing-the-web-introduction-to-thebrowser-browsing-web

2. https://www.wordstream.com/articles/internet-search-engineshistory

3. https://www.malwarebytes.com/phishing/

4. https://www.adpushup.com/blog/types-of-cookies/ 24 Syllabus IOT

5. https://www.eff.org/https-everywhere

6. https://www.sciencedirect.com/topics/computer-science/browsinghistory\

7. https://www.vpnmentor.com/blog/pros-cons-vpn/

8.   https://www.tech-wonders.com/2016/10/use-hush-privatebookmarking-extension-chrome.html

III: Networking

1. https://www.guru99.com/types-of-computer-network.html

2. https://www.studytonight.com/computer-networks/networktopology-types

3. https://www.cloudflare.com/learning/network-layer/what-is-aprotocol/

4. https://www.howtogeek.com/168896/10-useful-windows-commandsyou-should-know/

5. https://paiza.io/en

IV: Professionalism

1. https://career.vt.edu/develop/professionalism.html

2. https://englishlabs.in/importance-dress-code/

3. https://www.proschoolonline.com/blog/importance-of-body-languagein-day-to-day-life

4. https://www.thespruce.com/etiquette-of-proper-attire-1216800

5. https://shirleytaylor.com/why-are-communication-skills-important/

6. https://www.triad-eng.com/interview-tips-for-engineers/

7.   https://www.indeed.co.in/career-advice/interviewing/interviewquestion-tell-me-about-yourself

8. https://toggl.com/track/business-etiquette-rules/

V: Digital Identification

1. https://nad.ndml.in/nad-presentation.html

2. https://www.turtlemint.com/aadhaar-card-benefits/

3. https://www.bankbazaar.com/pan-card/uses-of-pan-card.html

4. https://www.passportindex.org/passport.php

5.https://consumer.westchestergov.com/financial-education/moneymanagement/benefits-of-a-bank-account

6. https://en.wikipedia.org/wiki/QR_code

First Semester			
Title of the Paper: Fundamental Tools for Data Science Lab			
Paper	CPP 1.6	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	48	Credits	02

List of Practical Programs:

1. Download, install and explore the features of NumPy, SciPy, Jupyter, Statsmodels and Pandas packages.
2. Working with Numpy arrays
3. Working with Pandas data frames
4. Reading data from text files, Excel and the web and exploring various commands for doing descriptive analytics on the Iris data set.

5. Use the diabetes data set from UCI and Pima Indians Diabetes data set for performing the following:
   a. Univariate analysis: Frequency, Mean, Median, Mode, Variance, Standard Deviation, Skewness and Kurtosis.
   b. Bivariate analysis: Linear and logistic regression modelling
   c. Multiple Regression analysis
   d. Also compare the results of the above analysis for the two data sets.

6. Apply and explore various plotting functions on UCI data sets.
   a. Normal curves
   b. Density and contour plots
   c. Correlation and scatter plots
   d. Histograms
   e. Three-dimensional plotting
7. Visualizing Geographic Data with Base map

First Semester			
Title of the Paper: Python Programming Lab			
Paper	**CPP 1.7**	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	48	Credits	02

List of Practical Programs:

1.   a   Demonstrate about Basics of Python Programming.

     b   Demonstrate about fundamental Data types in Python Programming. (i.e., int, float, complex, bool and string types)

     c   Demonstrate the working of following functions in Python.

       i) id()   ii) type( )   iii) range( )

     d   Write a Python program to demonstrate various base conversion functions.

     e   Write a Python program to demonstrate various type conversion functions.

2.   a   Demonstrate the following Operators in Python with suitable examples.

       i) Arithmetic Operators   ii) Relational Operators   iii) Assignment Operator

       iv) Logical Operators   v) Bit wise Operators   vi) Ternary Operator

       vii) Membership Operator   viii) Identity Operators

3.   a   Write Python programs to demonstrate the following:

       i) input ()   ii) print( )   iii)'sep' attribute

       iv) 'end 'attribute   v) replacement Operator ({})

     b   Demonstrate the following Conditional statements in Python with suitable examples. i) if statement   ii) if else statement   iii) if – elif – else statement

     c   Demonstrate the following Iterative statements in Python with suitable examples.

       i) while loop   ii) for loop

d   Demonstrate the following control transfer statements in Python with suitable examples.

i) break   ii) continue   iii) pass

4.   Write Python programs to print the following Patterns:

```
A                E E E E E E E E          *        *
A B              D D D D D D             * *      * *
A B C             C C C C C             * * *    * * *
A B C D            B B B               * * * *  * * * *
A B C D E            A                * * * * * * * * * *
```

5.   a   Write a Python program to demonstrate various ways of accessing the string.

i) By using Indexing (Both Positive and Negative)

ii) By using Slice Operator

b   Demonstrate the following functions/methods which operates on strings in Python with suitable examples:

i) len( )	ii) strip( )	iii) rstrip( )	iv) lstrip( )
v) find( )	vi) rfind( )	vii) index( )	viii) rindex()
ix) count( )	x) replace( )	xi) split( )	xii) join( )
xiii) upper( )	xiv) lower( )	xv) swapcase( )	xvi) title( )

xvii) capitalize( ) xviii) startswith()   xix) endswith()

6.   a   Python program to perform read and write operations on a file.

b   Python program to copy the contents of a file to another file.

c   Python program to count frequency of characters in a given file.

7.   a   Demonstrate the different ways of creating list objects with suitable example programs.

b   Demonstrate the following functions/methods which operates on lists in Python with suitable examples:

i) list( )	ii) len( )	iii) count( )	iv) index ( )
v) append( )	vi) insert( )	vii) extend()	viii) remove( )
ix) pop( )	x) reverse( )	xi) sort( )	xii) copy( )
xiii) clear( )			

c   Demonstrate the following with suitable example programs: i) List slicing ii) List Comprehensions

8.   a   Demonstrate the different ways of creating tuple objects with suitable example programs.

    b   Demonstrate the following functions/methods which operates on tuples in Python with suitable examples:

i) len( )   ii) count( )   iii) index( )   iv) sorted( )

v) min ( )   vi)max( )   vii) cmp( )   viii) reversed( )

9.   a   Demonstrate the different ways of creating set objects with suitable example programs.

    b   Demonstrate the following functions/methods which operates on sets in Python with suitable examples:

i) add( )   ii) update( )   iii) copy( )   iv) pop( )

v) remove( )   vi)discard( )   vii) clear( )   viii) union()

ix) intersection( )   x) difference( )

10.   a   Demonstrate the different ways of creating dictionary objects with suitable example programs.

    b   Demonstrate the following functions/methods which operates on dictionary in Python with suitable examples:

i) dict( )   ii) len( )   iii) clear( )   iv)   get( )

v) pop( )   vi)popitem( )   vii) keys( )   viii) values( )

ix) items( )   x) copy( )   xi) update( )

11.   a   Demonstrate the following kinds of Parameters used while writing functions in Python.

i) Positional Parameters   ii) Default Parameters

iii) Keyword Parameters   iv) Variable length Parameters

b Write a Python program to return multiple values at a time using a return statement.

c Write a Python program to demonstrate Local and Global variables.

d Demonstrate lambda functions in Python with suitable example programs

12. Implement the following Searching and Sorting techniques in Python by using functions

 i) Linear Search ii) Binary Search iii) Bubble Sort iv) Merge Sort

13. a Demonstrate the following in-built functions to use Regular Expressions very easily in our applications.

 i) compile( ) ii) finditer( ) iii) match( ) iv) fullmatch( )

 v) search( ) vi) findall( ) vii) sub( ) viii) subn( )

 ix) split( )

b Write a Regular Expression to represent all RGM language (Your own language) identifiers.

Rules:

1. The allowed characters are a-z,A-Z,0-9,#.

2. The first character should be a lower case alphabet symbol from a to k.

3. The second character should be a digit divisible by 3.

4. The length of identifier should be at least. Write a python program to check whether the given string is RGM language identifier or not?

c) Write a Regular Expression to represent all 10 digit mobile numbers.

Rules:

1. Every number should contain exactly 10 digits.

2. The first digit should be 7 or 8 or 9 Write a Python Program to check whether the given number is valid

First Semester			
Title of the Paper: Data Structure Using C Lab			
Paper	**SPP 1.8 A**	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	48	Credits	2

List of Practical Programs:

1. Write a program to convert a prefix notation to postfix notation.

2. Write a program to Evaluate a given postfix expression and its values for the variables

3. Write a program to simulate the working of circular queue providing the following operations–Insert, Delete and Display.

4. Write a program to Demonstrate recursion, Generate Fibonacci sequence, Solve Towers of Hanoi Problem.

5. Write a program to simulate the working of a linked list providing the following operations: Insert at the beginning/end; Insert at the position; Display list

6. Write a program to simulate the working of a circular linked list providing the following operations Delete from the beginning/end; Delete a given element; Display list

7. Write a program to simulate the working of a Dequeue

8. Write a program to simulate the working of a double linked list to implement stack.

9. Write a program to create a binary tree and implement the tree traversal techniques of in order, preorder and post order.

10. Write a program to implement quick sort

11. Write a program to implement the search techniques of Linear Search; Binary Search

12. Write a program to create a class called STACK using an array of integers. Implement the following operations by overloading the operators '+' and '--': s1 = s1 + element; where s1 is an object of the class STACK and element is an integer to be pushed on the top of the stack s1 = --s1; where s1 is an object of the class STACK. '- 22 'operator pops the element.

13. Handle the STACK empty and full conditions. Also display the contents of the stack after each operation, by overloading the << operator.

14. Write a program to create a class called QUEUE with member functions to add an element and to delete an element from the queue. Using the member functions, implement a queue of integers. Demonstrate the operations by displaying the contents of the queue after every operation.

First Semester			
Title of the Paper: Advanced Database Management Systems Lab			
Paper	SPP 1.8 B	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	48	Credits	02

List of Practical Programs:

1. Creation of database for the following details:
   a. Employee details   b. Student details   c. Hospital management
   d. Railway reservation   e. Hostel management

2. BOOK(Book_id, Title, Publisher_Name, Pub_Year) BOOK_AUTHORS(Book_id, Author_Name) PUBLISHER(Name, Address, Phone) BOOK_COPIES(Book_id, Branch_id, No-of_Copies) BOOK_LENDING(Book_id, Branch_id, Card_No, Date_Out, Due_Date) LIBRARY_BRANCH(Branch_id, Branch_Name, Address)

   Write the queries to

   a. Retrieve details of all books in the library – id, title, name of publisher, authors, number of copies in each branch, etc.
   b. Get the particulars of borrowers who have borrowed more than 3 books, but from Jan 2017 to Jun 2017.
   c. Delete a book in BOOK table. Update the contents of other tables to reflect this data manipulation operation.
   d. Partition the BOOK table based on year of publication. Demonstrate its working with a simple query.
   e. Create a view of all books and its number of copies that are currently available in the Library.

3. Consider the schema for Movie Database: ACTOR(Act_id, Act_Name, Act_Gender) DIRECTOR(Dir_id, Dir_Name, Dir_Phone) MOVIES(Mov_id, Mov_Title, Mov_Year, Mov_Lang, Dir_id) MOVIE_CAST(Act_id, Mov_id, Role) RATING(Mov_id, Rev_Stars)

   Write the queries to

   a. List the titles of all movies directed by 'Hitchcock'.
   b. Find the movie names where one or more actors acted in two or more movies.
   c. List all actors who acted in a movie before 2000 and also in a movie after 2015 (use JOIN operation).
   d. Find the title of movies and number of stars for each movie that has at least one rating and find the highest number of stars that movie received. Sort the result by movie title.

e. Update rating of all movies directed by 'Steven Spielberg' to 5.

4. Consider the schema for Company Database: EMPLOYEE(SSN, Name, Address, Sex, Salary, SuperSSN, DNo) DEPARTMENT(DNo, DName, MgrSSN, MgrStartDate) DLOCATION(DNo,DLoc) PROJECT(PNo, PName, PLocation, DNo) WORKS_ON(SSN, PNo, Hours)

   Write the queries to

   a. Make a list of all project numbers for projects that involve an employee whose last name is 'Scott', either as a worker or as a manager of the department that controls the project.
   b. Show the resulting salaries if every employee working on the 'IoT' project is given a 10 percent raise. C.
   c. Find the sum of the salaries of all employees of the 'Accounts' department, as well as the maximum salary, the minimum salary, and the average salary in this department
   d. Retrieve the name of each employee who works on all the projects controlledby department number 5 (use NOT EXISTS operator).
   e. For each department that has more than five employees, retrieve the department number and the number of its employees who are making more than Rs. 6,00,000.

5. Perform the following:
   Viewing all databases, viewing all Tables in a Database, Creating Tables (With and Without Constraints), Inserting/Updating/Deleting Records in a Table, Saving (Commit) and Undoing (rollback).

6. Perform the following:
   Altering a Table, Dropping/Truncating/Renaming Tables, backing up / Restoring a Database.

7. For a given set of relation schemes, create tables and perform the following Simple Queries, Simple Queries with Aggregate functions, Queries with Aggregate functions (group by and having clause), Queries involving- Date Functions, String Functions, Math Functions Join Queries- Inner Join, Outer Join Subqueries- With IN clause, With EXISTS clause

8. For a given set of relation tables perform the following a. Creating Views (with and without check option), Dropping views, Selecting from a view

9. Write a PL/SQL program using FOR loop to insert ten rows into a database table.

10. Given the table EMPLOYEE (EmpNo, Name, Salary, Designation, DeptID) write a cursor to select the five highest paid employees from the table.

11. Illustrate on usage of procedures and functions.

12. Illustrate on usage of triggers.

13. Given an integer i, write a PL/SQL procedure to insert the tuple (i, 'xxx') into a given relation.

14. Write a PL/SQL block to satisfy some conditions by accepting input from the user.

15. Write a PL/SQL block that handles all types of exceptions.

Second Semester			
Title of the Paper: Machine Learning			
Paper	CPT 2.1	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Couse Outcomes:

- To understand the concepts of machine learning.

- To appreciate supervised and unsupervised learning and their applications.

- To appreciate the concepts and algorithms of learning.

UNIT I   **9 Hours**

Introduction: Definition-Examples of machine learning applications –Well posed learning problems- Designing a learning system- Perspectives and issues Concept learning and general to specific ordering: Inductive learning hypothesis- Concept learning as search – candidate elimination algorithm-inductive bias.

UNIT II   **10 Hours**

Regression and classification:

Regression: Linear Regression-Simple-Multiple Decision Tree-Pruning: Introduction – Representation-Algorithm- issues Classification: Support Vector machine – Naïve Bayes-Applications.

UNIT III   **10 Hours**

Clustering and Learning:

Clustering: k-Means clustering– adaptive Hierarchical clustering –Applications- Neural network: Perceptron, multilayer network- back propagation- introduction to deep neural network Instance based learning: k-NN– Radial basis functions Case based reasoning-Reinforcement learning -Applications.

UNIT IV **10 Hours**

Probabilistic graphical models:

Graphical Models: Undirected graphical models - Markov Random Fields - Directed Graphical Models -Bayesian Networks - Conditional independence properties - Inference – Learning Generalization - Hidden Markov Models - Conditional random fields(CRFs).

UNIT V **9 Hours**

Machine learning experiments:

Design-Cross validation - Measuring Performance -Hypothesis testing- Assessing Performance -Comparison of algorithms, Datasets-Case study.

Text Books:

1. Tom M. Mitchell, Machine learning, McGraw-Hill, 1997.

2. Ethem Alpaydin, Introduction to Machine Learning, MIT Press, Third Edition, 2014.

Reference Books:

1. Stephen Marsland, Machine Learning – An Algorithmic Perspective, Chapman and Hall, CRC Press, Second Edition, 2014.

 2. Shalev-Shwartz, Shai, Shai Ben-David, Understanding Machine Learning: From theory to algorithms, Cambridge University Press, 2014.

Second Semester			
Title of the Paper: Statistics with R Programming			
Paper	**CPT 2.2**	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcomes:

- be able to use and program in the programming language R
- be able to use R to solve statistical problems
- be able to implement and describe Monte Carlo the technology
- be able to minimize and maximize functions using R

UNIT I   8 Hours

Introduction, How to run R, R Sessions and Functions, Basic Math, Variables, Data Types, Vectors, Conclusion, Advanced Data Structures, Data Frames, Lists, Matrices, Arrays, Classes.

UNIT II   10 Hours

R Programming Structures, Control Statements, Loops, - Looping Over Non vector Sets,- If-Else, Arithmetic and Boolean Operators and values, Default Values for Argument, Return Values, Deciding Whether to explicitly call return- Returning Complex Objects, Functions are Objective, No Pointers in R, Recursion, A Quicksort Implementation-Extended Extended Example: A Binary Search Tree.

UNIT III   10 Hours

Doing Math and Simulation in R, Math Function, Extended Example Calculating Probability- Cumulative Sums and Products-Minima and Maxima- Calculus, Functions Fir Statistical Distribution, Sorting, Linear Algebra Operation on Vectors and Matrices, Extended Example: Vector cross Product- Extended Example: Finding Stationary Distribution of Markov Chains, Set Operation, Input /out put, Accessing the Keyboard and Monitor, Reading and writer Files,

UNIT IV   **8 Hours**

Graphics, Creating Graphs, The Workhorse of R Base Graphics, the plot() Function –
Customizing Graphs, Saving Graphs to Files.

UNIT V   **12 Hours**

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- other Generalized Linear Models-Survival
Analysis, Nonlinear Models, Splines- Decision- Random Forests,

Text Books:

1. The Art of R Programming, Norman Matloff, Cengage Learning

2. R for Everyone, Lander, Pearson

Reference Books:

1. R Cookbook, PaulTeetor, Oreilly.

2. R in Action,Rob Kabacoff, Manning

Second Semester			
Title of the Paper: Multivariate Data Analysis			
Paper	SPT 2.3 A	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcomes:

- Inculcate the knowledge on various multivariate statistical techniques and its applications
- Know the usage of dependence and interdependence multivariate methods
- Know the statistics associated with principal component and factor analysis
- Impart the regression and classification techniques
- Learn the various clustering methods

UNIT I 8 Hours

Introduction to Multivariate Analysis:

Meaning of Multivariate Analysis – Multivariate Analysis in Statistical Terms – Basic concepts: Variate, Measurement Scales, Measurement Error, Multivariate Measurement, Statistical Significance and Statistical Power. Classification of Multivariate Techniques: Dependence and Independence Techniques – Applications of Multivariate Techniques.

UNIT II 10 Hours

Multiple Regression Analysis

Concept of Simple and Multiple Regressions – Illustrations. Prediction using Single and Several Independent Variables – Decision Process in Multiple Regression Analysis: Objectives, Research Design, Assumptions, Estimation of Regression Model – Assessing Model Fit – Interpretation of Regression Variate using Regression Coefficients and Assessing Multi collinearity.

UNIT III 10 Hours

Factor Analysis

Notion of Principal Components and Factors – Concept of Data Summarization and Data Reduction - Introduction to Principal Component Analysis and Factor Analysis – Illustrations.

Decision Process in Factor Analysis: Objectives, Design, Assumptions, Deriving Factors, Interpretation of Factors, Validation of Factors – Illustrations.

UNIT IV   **10 Hours**

Discriminant Analysis

Concept of Discriminant Function – Meaning of Discriminant Analysis – Decision Process in Discriminant Analysis: Objectives, Research Design, Assumptions, Estimation of Discriminant Model, Assessing Model Fit, Interpretation.

UNIT V   **10 Hours**

Cluster Analysis

Meaning and Conceptual Development of Cluster Analysis – Decision Process in Cluster Analysis: Objectives, Research Design, Assumptions, Deriving Clusters, Interpretation of Clusters, 23 Validation and Profiling of Clusters – Illustrations – Basic Notion of Hierarchical and Nonhierarchical Clusters.

Reference Books:

1. Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2018). Multivariate Data Analysis, Eighth Edition, Pearson.

2. Johnson, R. A., and Wichern, D. W. (2015). Applied Multivariate Statistical Analysis, Sixth Edition, Pearson.

3. Johnson, D. E. (1998). Applied Multivariate Methods for Data Analysts, First Edition, Duxbury Press.

Second Semester			
Title of the Paper: Probability Distribution			
Paper	**SPT 2.3 B**	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcomes:

- Develop problem-solving techniques needed to calculate probability and conditional probability.
- Formulate fundamental probability distribution and density functions, as well as functions of random variables, derive the probability density function of transformations.
- Derive the expectation and conditional expectation, and describe their properties.
- Understand various types of generating functions used in statistics.
- Describe commonly used univariate discrete and continuous probability distributions.
- Apply sampling distributions to testing of hypotheses.
- Translate and correlate the statistical problems into Statistical analysis

UNIT I   **8 Hours**

Descriptive Statistics:

Raw Data – Graphical Plots and Charts - Frequency Distribution – Histogram and Frequency Polygons – Relative Frequency Distributions – Cumulative Frequency Distributions – Frequency Curves and Their Types - Measures of Central Tendency: Mean, Median, Mode, Trimmed Mean – Measures of Dispersion: Range, Standard Deviation, Quartile Deviation, Mean and Median Absolute Deviation – Moments - Measures of Skewness and Kurtosis – Notion of Linear Correlation and Linear Regression – Simple Problems.

UNIT II   **10 Hours**

Basic Probability, Random Variables and Probability Distributions:

Concept of Probability – Axioms of Probability - Conditional Probability – Simple Problems - Independent Events - Bayes' Rule (without proof) and Simple Applications. Discrete and Continuous Random Variables, Probability Distributions for Discrete and Continuous Random Variables – Distribution Functions for Discrete and Continuous Random Variables - Joint Distributions - Independent Random Variables - Probability Distributions of Functions of Random Variables – Marginal and Conditional Distributions – Mathematical Expectation.

UNIT III   **10 Hours**

Special Probability Distributions:

Notions of Binomial, Poisson Distribution and Normal Distributions – Properties – Relationship Between Binomial and Normal Distributions, Poisson and Normal Distributions – Uniform, Exponential, Gamma Distributions, t, Chi-square and F Distributions - Bivariate Normal Distribution – Simulation: Random Number Generation from Exponential, Gamma and Normal Distributions.

UNIT IV   **10 Hours**

Sampling Theory and Statistical Estimation Theory:

Population and Sample - Random Samples – Sampling With and Without Replacement, Sampling Distributions, Sampling distributions of Mean, Proportion and Difference of Means, Standard 10 Error. Estimation of Parameters, Properties of Estimators: Unbiasedness, Consistency, Efficiency, Sufficiency. Point and Interval Estimates and Their Reliability, Confidence Interval Estimates of Population Parameters Based on Normal, t and Chi-square Distributions.

UNIT V   **10 Hours**

Statistical Decision Theory:

Statistical Decisions, Statistical Hypothesis, Tests of Hypothesis and Significance, One-tail and Two-tail Tests. Parametric Tests: Tests Involving Normal, t, Chi-square and F Distributions - Test for Goodness of Fit, Contingency Tables, Tests for Independence of Attributes, One-way and Two-way Analysis of Variance. Non-parametric Tests: Sign Test, Run Test, Wilcoxon Signed Rank Test, Mann-Whitney U test, Kruskal-Wallis Test.

Reference Books:

1. Montgomery, D. C., and Runger, G. C. (2018). Applied Statistics and Probability for Engineers, Seventh Edition, John Wiley & Sons, Inc.

2. Bruce, P., Bruce, A., and Gedeck, P. (2020). Practical Statistics for Data Scientists, Second Edition, O'Reilly Media, Inc.

3. Spiegel, M. R., Schiller, J. J., and Alu Srinivasan, R. (2013). Probability and Statistics, Fourth Edition, Schaum's Outline Series, McGraw Hill Companies, Inc.

Second Semester			
Title of the Paper: Information and Communication Technology			
Paper	OET 2.4	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcomes:

- Introduction to computers, classification of computers, anatomy of computer, constituents and architecture, microcontrollers

- Operating systems, functions of operating systems, classification of operating systems, kernel, shell, basics of Unix, shell programming, booting

- Prepare documents using work processing tools. Interpret data and represent it graphically using spread sheet. Prepare professional presentations.

- Internet basics, features, applications, services, internet service providers, browsing, email, searching and ICT tools.

UNIT I   **8 Hours**

Introduction to Computers and Windows:

Introduction, History of computer, Block diagram of a computer, Generation of computer, Classification of computers, Characteristics of computer, Applications of computer.

Windows concepts, general features of windows, different parts of windows screen, Setting the date and time, Icon explanation, paint, notepad, calculator, control panel, mouse properties, multitasking, my computer, folder creation, use of recycle bin, task bar.

UNIT II   **10 Hours**

Word Processing – MS-Word:

Introduction to MS Word, parts of MS- word, Parts of Ms-word screen, Backstage view, editing the text, formatting the text, Text effects, Bullets add numbering, Paragraph formatting, Borders & Shadings, Text styles, Table formatting, Picture, clipart and shapes adding, SmartArt representing, Screenshot explanation, Header & Footer, Word art, Drop cap, Equations & Symbols, Page setup,

Watermark, Envelopes and labels, Mail merge, Auto correct, Word count, Spell & grammar check, Commenting, Restrict editing, Document views, Zoom options, Navigation Pane, Arranging & Splitting word screen, Printing document, Exiting word.

UNIT III   8 Hours

Worksheet – MS- Excel:

Introduction to MS-excel screen, Parts of MS-excel screen, Worksheet basic, Creating worksheet, Entering text, Dates, alphanumeric and values, conditional formatting, table formatting, Cell styles, Auto sum, Sorting & Filtering, Editing the table, Illustrations, Explaining charts, Page setup, Print area, Sheet options, Arranging text & Images, Applying different Formulas, Insert functions, protecting sheet & workbook, workbook views, exiting excel.

UNIT IV   8 Hours

Presentation Graphics – MS-Power Point

Introduction to MS-Power point screen, Parts of MS-Power point screen, News slide & Layout options, editing text, Images & Illustrations, Photo album, Page setup, Slide Themes, Background styles & Graphics, Slide Transition effects, Sound effects for text and images, New animation effects, Order & Reordering animation, Previewing, Slide show from beginning & Custom slide shows, Rehearse timings, Media controls, Reviewing, Slide sorter and other Views, Color/Grayscale, Window options, exiting PowerPoint.

UNIT V   14 Hours

Fundamentals of Internet:

What is Internet?, Internet applications, Internet Addressing – Entering a Web Site Address, Intranet, Search engine, video conference, online forms, drive, URL–Components of URL, Searching the Internet, Browser –Types of Browsers, Introduction to Social Networking: Twitter, Tumblr, LinkedIn, Facebook, flickr, Skype, yahoo, YouTube, WhatsApp .

E-mail: Definition of E-mail -Advantages and Disadvantages –User Ids, Passwords, Email Addresses, Domain Names, Mailers, Message Components, Message Composition, Mail Management.

G-Suite: Google drive, Google documents, Google spread sheets, Google Slides and Google forms.

Overview of Internet security, E-mail threats and secure E-mail, Viruses and antivirus software, Firewalls, Cryptography, Digital signatures, Copyright issues. What are GOI digital initiatives in higher education? (SWAYAM, Swayam Prabha, National Academic Depository,

National Digital Library of India, E-Sodh-Sindhu, Virtual labs, eacharya, e-Yantra and NPTEL).

Reference Books:

1. Microsoft Office 2010: John Walkenbach, Herb Tyson, Michael R Groh, Faithe Wempen.

2. Microsoft Office 2010 for windows: Steve Schwartz.

3. Microsoft office 2010 The complete reference- by Virginia Andersen.

4. Microsoft Office 2010 Certification Prep By Laura Story, Dawna Walls

5. In-line/On-line : Fundamentals of the Internet and the World Wide Web, 2/e – byRaymond   Greenlaw and Ellen Hepp, Publishers : TMH

6. Internet technology and Web design, ISRD group, TMH.

7. Information Technology – The breaking wave, Dennis P.Curtin, Kim Foley, Kunai Sen and Cathleen Morin, TMH.

Second Semester			
Title of the Paper: Machine Learning Lab			
Paper	CPP 2.5	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	48	Credits	04

List of Practical Programs:

1. The probability that it is Friday and that a student is absent is 3 %. Since there are 5 school days in a week, the probability that it is Friday is 20 %. What is the probability that a student is absent given that today is Friday? Apply Baye's rule in python to get the result. (Ans: 15%) 13

2. Extract the data from database using python

3. Implement k-nearest neighbor's classification using python

4. Given the following data, which specify classifications for nine combinations of VAR1 and VAR2 predict a classification for a case where
   VAR1=0.906 and VAR2=0.606, using the result of kmeans clustering with 3 means (i.e., 3 centroids) periments

VAR1	VAR2	CLAS
1.713	1.586	0
0.180	1.786	1
0.353	1.240	1
0.940	1.566	0
1.486	0.759	1
1.266	1.106	0
1.540	0.419	1
0.459	1.799	1
0.773	0.186	1

5. The following training examples map descriptions of individuals onto high, medium and low credit-worthiness.

   Medium skiing design single twenties no -> highRisk
   high golf trading married forties yes -> lowRisk
   low speedway transport married thirties yes -> medRisk
   medium football banking single thirties yes -> lowRisk
   high flying media married fifties yes -> highRisk
   low football security single twenties no -> medRisk
   medium golf media single thirties yes -> medRisk
   medium golf transport married forties yes -> lowRisk
   high skiing banking single thirties yes -> highRisk
   low golf unemployed married forties yes -> highRisk

Input attributes are (from left to right) income, recreation, job, status, age-group, home-owner. Find the unconditional probability of `golf' and the conditional probability of `single' given `medRisk' in the dataset?

6. Implement linear regression using python.

7. Implement Naïve Bayes theorem to classify the English text

8. Implement an algorithm to demonstrate the significance of genetic algorithm

9. Implement the finite words classification system using Back-propagation algorithm

Second Semester			
Title of the Paper: Statistics with R Programming Lab			
Paper	**CPP 2.6**	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

List of Practical Programs:

1. Installation of R in Windows and Linux environment

2. Write a R program to find measures of central tendency

3. Write a R program to perform different operations on Matrices

4. Write a R program to store data into a List and perform different operations

5. Write a R program to store data into Data frame and perform different operations

6. Write a R program to find biggest of three elements

7. Write a R program to find roots of a quadratic equation

8. Write a R program to find sum of elements of vector

9. Write a R program to find factorial of a number using recursion

10. Write a R program to find gcd of two numbers using recursion

11. Write a R program to mean, variance, standard deviation for the given discrete probability distribution

12. Write a R program to mean, variance, standard deviation for the given continuous probability distribution

13. Write a R program to represent the given data in the form of graphs using built in functions

14. Write a R program to fit Binomial distribution to the given data

15. Write a R program for Z test /t-test/F test/ Chi-square test/ Z test

16. Write a R program to fit a linear regression/multiple linear regression.

Second Semester			
Title of the Paper: Multivariate Data Analysis Lab			
Paper	SPP 2.7 A	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

List of Practical Programs:

1. Computation of Mean vector and covariance matrix for multivariate data set

2. Generation of multivariate data using multivariate normal distribution

3. Fitting of linear, quadratic, exponential and logistic models

4. Principal Component analysis and factor analysis

5. Linear and quadratic discriminant analysis with classification of two and three groups.

6. Cluster analysis with hierarchical clustering (single linkage, average linkage, Wards method) and non-hierarchical clustering (k-means)

Second Semester			
Title of the Paper: Data Analysis Using Excel Lab			
Paper	SPP 2.7. B	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	52	Credits	02

List of Practical Programs:

.ABOUT EXCEL

1. Introduction, Uses of Excel, New functions and features of excel 2007

2. Getting started with excel: Opening a blank or new workbook, general organization

3. Highlights and main functions: Home, Insert, page layout, formulas

4. Highlights and main functions: Data, review, view, add-inns

5. Using the Excel help function

II. GENERAL EXCEL LESSONS

6. Customizing the Quick Access Toolbar

7. Creating and Using Templates

8. Working with Data: Entering, Editing, Copy, Cut, Paste, Paste Special

9. Formatting Data and Using the Right Mouse Click

10. Saving, Page Setup, and Printing

11. Using Headers and Footers

12. Manipulating Data, using Data Names and Ranges, Filters and Sort and Validation Lists

13. Data from External Sources

14. Using and Formatting Tables

15. Basic Formulas and Use of Functions

16.   Data Analysis Using Charts and Graphs

17.   Managing, Inserting, and Copying Worksheets

18.   Securing the Excel Document (Protect Cells and Workbook)

III. ADVANCED EXCEL LESSONS

19.   Advanced Formulas and Functions

20.   Advanced Worksheet Features

21.   Advanced Data Analysis using PivotTables and Pivot Charts

PART B – STATISTICAL TOOLS FOR EXECUTION USING EXCEL

1.   Tabulation, bar diagram, Multiple Bar diagram, Pie diagram,

2.   Measure of central tendency: Mean, median, mode,

3.   Measure of dispersion: variance, standard deviation, Coefficient of variation. Correlation, regression lines.

4.   t-test, F-test,

5.   ANOVA one way classification,

6.   chi square test, independence of attributes.

7.   Time series: forecasting Method of least squares, moving average method.

8.   Inference and discussion of results. Shortcut Keys

Second Semester			
Title of the Paper: Information and Communication Technology Lab			
Paper	OEP 2.8	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	52	Credits	02

List of Practical Programs:

1. Create a Letter format by using MS-word

2. Create a company letter head by using MS-word.

3. Create your own resume by using MS-word

4. Create a greeting card

5. Create a Cover Page of project by using MS-word.

6. Create a Mail merge letter.

7. Create a Spreadsheet in MS-EXCEL and enter the marks of a student, calculate total and Print grade if the student has passed in all subjects.

8. Enter the following details for 10 employee's employee code, name, basic salary, DA, HRA, Loans, total salary and tax.

9. Create a Spreadsheet in MS-EXCEL Implement five functions each for Arithmetic, date and time, financial, logical and statistical functions. Write the syntax, example and output for simple programs

10. A bank offers loan for housing and vehicle at an interest of 10.25% for housing and 14.2% for vehicle. For a loan applicants compute the monthly premium (EMI), given total installments as 24 months. Also compute the monthly interest and monthly principal amount and total amount of principal and interest using financial library functions in a spreadsheet.

11. Create a simple bar chart to high light the sales of a company for 3 different periods:

12. Create a pie chart for a sample data and give legends.

13. Create a worksheet importing data from database and calculate sum of all the columns.

14. Create text and images with effects.

15. Create a Power-point presentation with animation & sound effects consist of at least 10 slides.

16. Create Google sheet to collect the marks scored in previous examination and share the link to your friends to gather the data.

17. Design your own Google form and share with your friends to gather the data.

18. Schedule a Google meet, share the link to invite your friends and present the document as a host (use different layouts).

Third Semester			
Title of the Paper: Advanced Software Engineering			
Paper	CPT 3.1	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcomes:

- Categorize problems based on their characteristics and practical importance.
- Apply the correct process models for software development.
- Apply the techniques, skills, and modern engineering tools necessary for engineering practice.
- Define, formulate and analyze a problem as per the testing techniques.
- Apply new Generation of Software Engineering Technology to Meet Current and Future Industrial Challenges of Emerging Software Trends

UNIT I   10 Hours

Software process Models and lifecycle: Software Product, Product, Software Processes, Evolving Role of Software, Software: A Crisis on the Horizon and Software Myths, Software Engineering: A Layered Technology, Study of different Software Process Models, The Linear Sequential Model, The Prototyping Model, The RAD Model, Evolutionary Process Models, Component-Based Development, Process, Product and Process, Object Oriented Software Engineering.

Project Management Concepts & Project Metrics: The Management Spectrum, People, Product, Process, Project, the W5HH Principle, Metrics in the Process and Project Domains (FP & LOC), Software Measurement, Metrics for Project and Software Quality

UNIT II   10 Hours

Software Project Planning, Scheduling and Tracking: Project Planning Objectives, Software Project Estimation using COCOMO Model, Software Scope and Resources, Empirical Estimation Models, Automated Estimation Tools, Basic Concepts and Relationship between People and Effort, Defining a Task Set for the Software Project, Selecting Software Engineering Tasks, Defining a Task Network and Scheduling, Earned Value Analysis and Error Tracking

Software Requirements Specification: Requirement Gathering and Analysis, Software Requirement Specification (SRS), Formal requirements specification and verification - axiomatic and algebraic specifications

UNIT III   10 Hours

Analysis Modeling, Software Design Concepts and Principles: The Elements of the Analysis Model, Data Modeling, Functional Modeling and Information Flow, Behavioral Modeling and Structured Analysis, Software Design and Software Engineering, The Design Process, Design Principles, Design Concepts, Modular Design, Design Heuristics for Effective Modularity, The Design Model ,Design Documentation, Function oriented v/s object-oriented design, Object Modeling using UML, Software Architecture and Data Design, Architectural Styles, Analyzing Alternative Architectural Designs, Mapping Requirements into a Software Architecture

User Interface Design, Component Level Design: User Interface Design, Task Analysis and Modeling, Interface Design Activities and Implementation Tools, Design Evaluation, Structured Programming and Comparison of Design Notation.

Risk Analysis & Management: Reactive versus Proactive Risk Strategies, Software Risks (Risk Identification, Risk Projection, Risk Refinement, Risk Mitigation), Risks Monitoring and Management

UNIT IV   10 Hours

Coding, Software Testing Techniques & Software Testing Strategies: Software Testing Fundamentals and Test Case Design, White-Box Testing and Black-Box Testing, ISO/IEC/IEEE Software Testing standards, Testing for Specialized Environments, A Strategic Approach to Software Testing and Issues, Unit Testing, Integration and Validation Testing, System Testing, Software Documentation and Debugging Techniques

Software Quality Assurance and Configuration Management - Quality Concepts and Software Quality Assurance, Quality Planning and Control, Software Reviews (Formal Technical Reviews), Software Reliability and Fault Tolerance, The ISO 9000 Quality Standards, The SCM Process, Identification of Objects in the Software Configuration, Six Sigma, Version Control and Change Control

UNIT V   8 Hours

Agile Software Development

Agile Software Development Values, principles, stakeholders, Agile Principles: Twelve Practices of Agile process, Communication, Planning, Agile Product Management: Estimation, Quality, Risk, Metrics and Measurements, Agile Requirements: User Stories, Backlog Management, Agile Architecture: Feature-Driven Development, Agile Risk Management: Risk and Quality Assurance, Agile Testing: Test-Driven Development, User Acceptance Test.

Text Book:

1. Roger S. Pressman, Software Engineering: A practitioner's approach, McGraw Hill.

2. Pankaj Jalote, An integrated approach to Software Engineering by Springer.

3. Ian Sommerville, Software Engineering, Addision and Wesley.

Third Semester			
Title of the Paper: Artificial Intelligence			
Paper	**CPT 3.2**	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcomes:

- Explain what constitutes "Artificial" Intelligence and how to identify systems with Artificial Intelligence.
- Explain how Artificial Intelligence enables capabilities that are beyond conventional technology,
- Use classical Artificial Intelligence techniques, such as search algorithms, minimax algorithm, neural networks, tracking, and robot localization.
- Ability to apply Artificial Intelligence techniques for problem-solving.
- Explain the limitations of current Artificial Intelligence techniques.

UNIT I   **8 Hours**

PROBLEM SOLVING

Introduction to AI- Foundations of AI – Risks and benefits of AI - Agents and Environments – Structure of Agents - Uninformed Search Strategies- Informed Search Strategies- Heuristic functions - Local Search Algorithm.

UNIT II   **10 Hours**

SEARCH IN COMPLEX ENVIRONMENT, GAMES AND KNOWLEDGE REPRESENTATION

Introduction to Game Playing-Alpha Beta Pruning- Constraint Satisfaction Problems - Knowledge Representation using First order logic- Knowledge Engineering in First Order Logic Proportional vs First Order Logic.

UNIT III   **10 Hours**

KNOWLEDGE REASONING AND PLANNING

Inference- Forward and Backward Chaining-Unification-Uncertainty-Inference in Bayesian Network – Inference in Temporal models – Hidden Markov Models – Kalman Filters –

Dynamic Bayesian Networks – Combining Beliefs and desires under uncertainty – Decision Networks.

UNIT IV   **10 Hours**

PRODUCTION SYSTEM AND PLANNING

Introduction to Production system-control strategies-Rete Algorithm-Planning-STRIPS-Planning with state space search-Partial Order Planning-Planning Graphs-Planning, acting in the real world.

UNIT V   **10 Hours**

EXPERT SYSTEM

Expert System- Architecture and Roles of Expert System-Typical Expert System-MYCINXOON-DART Case Study-Construction of simple reflex agent with sensor and actuator using Arduino.

Text Books:

1. Stuart Russell, Peter Norvig, "Artificial Intelligence – A Modern Approach", 3rd Edition, Pearson Education / Prentice Hall of India, 2010.

2. Joseph C. Giarratano, Gary D. Riley," Expert Systems: Principles and Programming", 4 th Edition, 2015.

Reference Books:

1. Nils J. Nilsson, "Artificial Intelligence: A new Synthesis", Harcourt Asia Pvt. Ltd., 2000.

2. Kevin Night and Elaine Rich, Nair B., "Artificial Intelligence (SIE)", Mc Graw Hill2008.

3. W. Patterson, 'Introduction to Artificial Intelligence and Expert Systems', Prentice Hall of India, 2007

4. Prateek Joshi, "Artificial Intelligence with Python", Packt Publishing, 2017.

Third Semester			
Title of the Paper: Deep Learning			
Paper	**SPT 3.3 A**	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcomes

- The students should be able to precisely state the classical algorithms, models, and theories in the area.
- Students should be able to identify appropriate algorithm given a practical task.
- Students should also be able to implement and solve the tasks using deep learning techniques.

UNIT I   **8 Hours**

INTRODUCTION:

Introduction – Definition of learning systems – Goals and applications of Machine Learning – Types of Machine Learning – Machine Learning process – Hypothesis space and Version space

UNIT II   **10 Hours**

SUPERVISED LEARNING

Linear models for Regression – Linear models for Classification – Decision Tree Learning – Bayesian Learning – Naïve Bayes – Ensemble Methods – Bagging – Boosting – Support Vector Machines.

UNIT III   **10 Hours**

EVALUATION Performance Evaluation metrics – ROC Curves – Validation methods – Bias-variance decomposition – Model complexity

UNIT IV   **10 hours**

UNSUPERVISED LEARNING Clustering – K-means – K-mode- K-median – Hierarchical clustering – DBSCAN – Principal Component Analysis – Independent Component Analysis

UNIT V   **10 hours**

ADVANCED LEARNING

Sampling – Basic sampling methods – Monte Carlo – Gibbs Sampling – Computational Learning theory – Reinforcement learning – Markov Decision Processes.

Text Books:

1. Tom Mitchell, Machine Learning, McGraw-Hill, UK, 2017

2. Ethem Alpaydin, Introduction to Machine Learning, MIT Press, Third Edition, 2014.

Reference Books:

1. Stephen Marsland, Machine Learning – An Algorithmic Perspective, Chapman and Hall, CRC Press, Second Edition, 2014.

2. Shalev-Shwartz, Shai, Shai Ben-David, Understanding Machine Learning: From theory to algorithms, Cambridge University Press, 2014.

Third Semester			
Title of the Paper: Cloud Computing			
Paper	SPT 3.3 B	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcomes:

- Define Cloud Computing and memorize the different Cloud service and deployment models.

- Describe importance of virtualization along with their technologies.

- Use and Examine different cloud computing services.

- Analyze the components of open stack & Google Cloud platform and understand Mobile Cloud Computing.

- Design & develop backup strategies for cloud data based on features.

UNIT I   **8 Hours**

Introduction; Principles of Parallel and Distributed Computing

Cloud computing at a glance; Historical Developments; building Cloud computing environment; computing platforms and Technologies Principles of Parallel and Distributed Computing: Eras of Computing; parallel Vs. distributed computing; elements of distributed computing; technologies of Distributed computing

UNIT II   **10 Hours**

Virtualization and Cloud Computing Architecture Characteristics of virtualized environments; virtualization techniques; virtualization and cloud computing; pros and cons of virtualization; examples. Cloud Reference model; Types of clouds; cloud economics; open challenges

Aneka: Cloud application Platform Overview; anatomy of the Aneka container; building Aneka clouds; cloud programming and management

UNIT III   10 Hours

Concurrent Computing and High-Throughput Computing and Map Reduce Programming
Introducing parallelism; programming with threads; multithreading with Aneka; applications;
Task Computing; task based Application Model; Task based Programming; Data Intensive
Computing; Technologies; Aneka Map Reduce Programming

UNIT IV   10 Hours

Cloud Platforms in Industry and Cloud Applications Amazon Web services; Google App
Engine; Microsoft Azure; Cloud scientific Applications; Business and Consumer Applications

UNIT V   10 Hours

Advanced Topics in Cloud Computing and Cloud Security Energy Efficiency Clouds; Market
based management clouds; Federated Clouds; Third Party Cloud Services; Infrastructure
Security: Network level security, Host level security, and Application level security; Data
security and Storage.

Text Books:

1. Mastering Cloud Computing, Editors: Rajkumar Buyya, Christian Vecchiola, S. Thamarai Selvi, MGH2013 Chapters: 1 ,2,3,4,5,6,7,8,9,10,11

Reference Books:

1. Cloud Computing: Principles and Paradigms, Editors: RajkumarBuyya, James Broberg, Andrzej M. Goscinski, Wile, 2011
2. Cloud Computing Bible, Barrie Sosinsky, Wiley-India, 2010
3. Cloud Computing: Principles, Systems and Applications, Editors: Nikos Antonopoulos, Lee Gillam, Springer, 2012

4. Cloud Security: A Comprehensive Guide to Secure Cloud Computing, Ronald L. Krutz, Russell Dean Vines, Wiley-India, 2010

Third Semester			
Title of the Paper: Web Programming			
Paper	OET 3.4	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcomes:

- To use the syntax and semantics of HTML and DHTML
- To develop different parts of a web page
- To understand how CSS can enhance the design of a webpage.
- To create and apply CSS styling to a webpage
- To get familiarity with the JavaScript language and understand Document Object Model handling of Java Script

UNIT I   8 Hours

Basics of Internet and Web The basics of Internet, World Wide Web, Web page, Home page, Web site, Static, Dynamic and Active web page, Overview of Protocols – Simple Mail Transfer Protocol, Gopher, Telnet, Emails, TFTP, Simple Network Management Protocol, Hyper Text Transfer Protocol, Client server computing concepts.

Web Client and Web Sever Web Browser, Browsers e.g., Netscape navigator, Internet Explorer, Mozilla Firefox, Client Side Scripting Languages- VB Script and Java Script, Active X control and Plug-ins; Web Server Architecture, Image maps, CGI, API web database connectivity-DBC, ODBC

UNIT II   10 Hours

Introduction to HTML Introduction to HTML, Essential Tags, Tags and Attributes, Text Styles and Text Arrangements, Text, Effects, Exposure to Various Tags (DIV, MARQUEE, NOBR, DFN, HR, LISTING, Comment, IMG), Color and Background of Web Pages, Lists and their Types, Attributes of Image Tag, Hypertext, Hyperlink and Hypermedia, Links, Anchors and URLs, Links to External Documents, Different Section of a Page and Graphics, Footnote and eMailing, Creating Table, Frame, Form and Style Sheet.

UNIT III   **10 Hours**

 DHTML Dynamic HTML, Document Object Model, Features of DHTML, CSSP (Cascading Style Sheet Positioning) and JSSS (JavaScript assisted Style Sheet), Layers of Netscape, The ID Attribute, DHTML Events.

UNIT IV   **8 Hours**

Java Script Objects, Methods, Events and Functions, Tags, Operators, Data Types, Literals and Type Casting in JavaScript, Programming Construct, Array and Dialog Boxes, Relating JavaScript to DHTML, Dynamically Changing Text, Style, Content.

UNIT V   **12 Hours**

Introduction to PHP Server Side Programming , Introduction to PHP, Basic Programming Concepts of PHP: Variables, Data-types, Constants, Scope of Variables, Type of Variables, Type Casting, Operators, Operators Precedence, References, Arrays; Control Structures: Branching, If statement, Switch statement; Looping: for Loop, while Loop, do while Loop, for each Loop; Functions: User Defined Functions, Built-in Function, Functions for Variables; Script Controlling Functions, Array Functions, Date and Time Functions, Mathematical Functions, String Functions, PHP Server Variables; Working with form, Uploading files to Web Server using PHP

Reference Books:

1. Burdman, "Collaborative Web Development", Addison Wesley.

2. Sharma &Sharma, "Developing E-Commerce Sites", Addison Wesley

3. Ivan Bayross, "Web Technologies Part II", BPB Publications.

Third Semester			
Title of the Paper: Minor Project			
Paper	**CPP 3.5**	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	48	Credits	

The Mini-project is an application that should be formally initiated and should be developed and to be implemented by the respective team. The team coordination and proper communication should be promptly maintained by the team leader. Based on the ability/abilities of the students and recommendations of the mentor, a single discipline or a multidisciplinary Minor-project can be assigned to the students.

Note:

1. It is team project with a maximum of three in a team
2. Students shall select a domain and develop an application with social relevance
3. Documentation is to be based on the standards
4. Evaluation pattern is like Lab examination
5. Need to submit a report, presentation with demo.
6. During the examination, each student must demonstrate the project individually.

Semester End Examination Marks Distribution:

Sl.No	Description	Marks
1.	Project Demonstration	20
2.	Viva-Voce	10
3.	Project Report	05
	Total:	**35 Marks**

Third Semester			
Title of the Paper: Artificial Intelligence Lab			
Paper	**CPP 3.6**	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	48	Credits	02

List of Practical Programs:

1. Write a program to implement DFS and BFS

2. Write a Program to find the solution for travelling salesman Problem

3. Write a program to implement Simulated Annealing Algorithm

4. Write a program to find the solution for wampus world problem

5. Write a program to implement 8 puzzle problem

6. Write a program to implement Towers of Hanoi problem

7. Write a program to implement A* Algorithm
8. Build a bot which provides all the information related to you in college.

9. Build a virtual assistant for Wikipedia using Wolfram Alpha and Python

10. Write a program to implement Hill Climbing Algorithm

11. Implement Mini-Max algorithm for game playing (Alpha-Beta pruning)

12. Propositional Model Checking Algorithms

13. Implement Forward Chaining Algorithm

14. Implement backward Chaining Algorithm

15. Implement Naïve Bayes Models

16. Implement Bayesian Networks and perform inferences

Third Semester			
Title of the Paper: Deep Learning Lab			
Paper	**SPP 3.7 A**	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	48	Credits	02

List of Practical Programs:

1. Style transfer for an images

2. Build multi-layer neural networks

3. Convolutional neural networks

4. Recurrent neural networks

5. Generative adversarial networks

6. Deploying a sentiment analysis model

Third Semester			
Title of the Paper: Cloud Computing Lab			
Paper	SPP 3.7 B	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	48	Credits	

List of Practical Programs:

1. Install Virtual box / VMware Workstation with different flavors of Linux or windows OS on top of windows7 or 8.

2. Install a C compiler in the virtual machine created using virtual box and execute Simple Programs

3. Install Google App Engine. Create hello world app and other simple web applications using python/java.

4. Use GAE launcher to launch the web applications.

5. Simulate a cloud scenario using Cloud Sim and run a scheduling algorithm that is not present in Cloud Sim.

6. Find a procedure to transfer the files from one virtual machine to another virtual machine.

7. Find a procedure to launch virtual machine using try stack (Online Open stack Demo Version)

8. Install Hadoop single node cluster and run simple applications like word count.
   Cloud computing (Programs on SAAS)

9. Create an word document of your class time table and store locally and on cloud with doc and pdf format.

10. Create a spread sheet which contains employee salary information and calculate gross and total salary using formula DA=10%OF BASIC,HRA=30%OF BASIC PF= 10% OF BASIC IF BASIC<=3000 12% OF BASIC IF BASIC>3000 TAX=10% OF BASIC IF BASIC<=1500 11% OF BASIC IF BASIC>1500 AND BASIC<=2500 12% OF BASIC IF BASIC>2500 NET_SALARY=BASIC_SALARY+DA+HRA-PF-TAX 57

11. Prepare a ppt on cloud computing- introduction, models, services and architecture

12. Create your resume in a neat format using Google and Zoho cloud

Cloud computing (Programs on SAAS)

13. Write a Google app engine program to generate n even numbers and deploy it to Google cloud

14. Google app engine program multiply two matrices

15. Google app engine program to validate user; create a database login(username, password)in mysql and deploy to cloud

16. Google app engine program to validate the user use mysql to store user info and deploy on to cloud

Third Semester			
Title of the Paper: Web Programming Lab			
Paper	**OEP 3.8**	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	48	Credits	02

List of Practical Programs:

1. Write a HTML code to create simple Web page to display your college information.

2. Write a HTML code to create different types of ordered lists.

3. Write a HTML code to insert an image of 'Tim Berners Lee' and display his Bio data.

4. Write a HTML code to illustrate text formatting tags

5. Write a HTML code to illustrate Img tag

6. Write a HTML code to illustrate Hyper Link tag (Anchor tag)

7. Write a HTML code to illustrate Frame tag

8. Write a HTML code to embed multimedia on to a web page.

9. Write a HTML code to create your class time table.

10. Write a HTML code to create a form to accept student details. When the form runs in the Browser fill the textboxes with data. Write JavaScript code that verifies that all textboxes has been filled. If a textboxes has been left empty, popup an alert indicating which textbox has been left empty

11. Write a PHP program to create and manage a database using SQL commands.

12. Write a PHP program to create and validate a email id.

13. Using PHP and SQL, create and validate a sample login form.

14. Using PHP and SQL, develop a program to accept student information viz. Accession name, regno, course, sem, marks in 3 subjects from a web page and store the information in a database and to calculate total and average and to display the results with proper headings.

15. Using PHP and SQL, develop a program to accept employee information viz. Accession no, name and basic from a web page and store the Information in a database and to calculate DA, HRA, PF, TAX, Gross and net pay and to display the details with proper headings.

Fourth Semester			
Title of the Paper: Research Methodology			
Paper	**CPT 4.1**	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcomes:

- Understand the psychology of research which includes different perspectives and necessity of research.
- Apply the research knowledge to formulate a suitable problem statement by adopting different research methods and models.
- Analyze the research outcome by using suitable statistical tool.
- Write or present a scientific report and research proposal by adopting copyright based ethical values.

UNIT 1   **8 Hours**

Research Methodology: Introduction, Meaning of Research, Objectives of Research, Motivation in Research, Types of Research, Research Approaches, Significance of Research, Research Methods versus Methodology, Research and Scientific Method, Importance of Knowing How Research is Done, Research Process, Criteria of Good Research, and Problems Encountered by Researchers in India.

UNIT II   **10 Hours**

Defining the Research Problem: Research Problem, Selecting the Problem, Necessity of Defining the Problem, Technique Involved in Defining a Problem, an Illustration. Reviewing the literature: Place of the literature review in research, Bringing clarity and focus to your research problem, Improving research methodology, Broadening knowledge base in research area, Enabling contextual findings, How to review the literature, searching the existing literature, reviewing the selected literature, Developing a theoretical framework, Developing a conceptual framework, Writing about the literature reviewed.

UNIT III   **10 Hours**

Research Design: Meaning of Research Design, Need for Research Design, Features of a Good Design, Important Concepts Relating to Research Design, Different Research Designs, Basic Principles of Experimental Designs, Important Experimental Designs. Design of Sample Surveys: Introduction, Sample Design, Sampling and Non-sampling 76 Errors, Sample Survey versus Census Survey, Types of Sampling Designs.

UNIT IV   **10 Hours**

Data Collection: Experimental and Surveys, Collection of Primary Data, Collection of Secondary Data, Selection of Appropriate Method for Data Collection, Case Study Method. Interpretation and Report Writing: Meaning of Interpretation, Technique of Interpretation, Precaution in Interpretation, Significance of Report Writing, Different Steps in Writing Report, Layout. Types of Reports, Oral Presentation, Mechanics of Writing a Research Report, Precautions for Writing Research Reports.

UNIT V   **10 Hours**

Intellectual Property (IP) Acts: Introduction to IP: Introduction to Intellectual Property (IP), different types of IPs and its importance in the present scenario, Patent Acts: Indian patent acts 1970.Design Act: Industrial Design act 2000. Copy right acts: Copyright Act 1957. Trade Mark Act, 1999.

Text books:

1. Research Methodology: Methods and Techniques, C.R. Kothari, Gaurav GargNew Age International 4th Edition, 2018.

2. Research Methodology a step-by-step guide for beginners. (For the topic Reviewing the literature under UNIT 2) Ranjit Kumar SAGE Publications Ltd 3rd Edition, 2011 Study Material. 3. Intellectual property, Debirag E. Bouchoux, Cengage learning, 2013.

Reference Books:

1. Research Methods: the concise knowledge baseTrochim, Atomic Dog Publishing, 2005.

2. Conducting Research Literature Reviews: From the Internet to Paper FinkAge Publications, 2009.

Fourth Semester			
Title of the Paper: Data Visualization			
Paper	**CPT 4.2**	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcomes:

- To interpret data plots and understand core data visualization concepts such asrrelation, linear relationships, and log scales.

- To explore the relationship between two continuous variables using scatter plots and line plots.

- To translate and present data and data correlations in a simple way, data analysts use a wide range of techniques — charts, diagrams, maps, etc.

UNIT I   **8 Hours**

Introduction of visual perception, visual representation of data, Gestalt principles, information overloads.

UNIT II   **10 Hours**

Creating visual representations, visualization reference model, visual mapping, visual analytics, Design of visualization applications.

UNIT III   **10 Hours**

Classification of visualization systems, Interaction and visualization techniques misleading, Visualization of one, two and multi-dimensional data, text and text documents.

UNIT IV   **10 Hours**

Visualization of groups, trees, graphs, clusters, networks, software, Metaphorical visualization

UNIT V   **10 Hours**

Visualization of volumetric data, vector fields, processes and simulations, Visualization of maps, geographic information, GIS systems, collaborative visualizations, evaluating visualizations.

Textbooks:

1. Ward, Grinstein Keim, Interactive Data Visualization: Foundations, Techniques, and Applications. Natick: A K Peters, Ltd, 1st Edition, 2010.

2. Kieran Healy, Data Visualization: A Practical Introduction, 1st Edition, 2018.

3. Andy Krik, Data Visualization: a successful design process 1st Edition, 2012.

4. Corey Lanum, Visualizing Graph Data 1st Edition, 2016.

Reference Books:

1. E. Tufte, The Visual Display of Quantitative Information, Graphics Press.2nd Edition, 2001.

2. Ward, Grinstein Keim, Interactive Data Visualization: Foundations, Techniques, and Applications. Natick: A K Peters, Ltd. 1st Edition, 2014.

3. Kieran Healy, Data Visualization: A Practical Introduction, sew Edition, 2013.

4. Andy Krik, Data Visualization: a successful design process 1st Edition, 2016.

Fourth Semester			
Title of the Paper: Natural Language Processing			
Paper	SPT 4.3 A	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcomes:

- NLP relates our perceptions, personal beliefs, emotions, behavior's, attitudes and communication styles
- Resulting the outcome we experience on a day-to-day basis.
- Understanding the basics, to confidently applying the different NLP principles (and practices) in all aspects of personal or professional life.

UNIT I   8 Hours

Introduction:

NLP tasks in syntax- semantics- and pragmatics Applications such as information extraction- question answering- and machine translation-The role of machine learning- Brief history of the field.

UNIT II   10 Hours

N-gram Language Models:

The role of language models- Simple N-gram models-Estimating parameters and smoothing-Evaluating language models.

UNIT III   10 Hours

Semantic Analysis I:

Lexical semantics and word-sense disambiguation Compositional semantics-Semantic Role Labeling and Semantic Parsing.

UNIT IV   10 Hours

Semantic Analysis II:

Lexical semantics and word-sense disambiguation. Compositional semantics-Semantic Role Labeling and Semantic Parsing- Named entity recognition and relation extraction-IE using sequence labeling.

UNIT V   **10 Hours**

Text Classification Model:

Supervised Classification, Document Classification, Part-of-Speech Tagging-Exploiting Context- Sequence Classification Sentence Segmentation- Identifying Dialogue Act Types-Recognizing Textual Entailment- Scaling Up to Large Datasets.

Text Book:

1. Elhadad, M. (2010). Book Review: Natural Language Processing with Python by Steven Bird, Ewan Klein, and Edward Loper. Computational Linguistics, 36(4).

2. Bender, E. M. (2013). Linguistic fundamentals for natural language processing: 100 essentials from morphology and syntax. Synthesis lectures on human language technologies, 6(3), 1-184.

Reference Book:

3. Hapke, H. M., Lane, H., & Howard, C. (2019). Natural language processing in action.

Fourth Semester			
Title of the Paper: Big Data Analytics			
Paper	SPT 4.3 B	CIE Marks	30
No of Hours/Week:	4	SEE Marks	70
Total Hours:	48	Credits	04

Course Outcomes:

- Identify Big Data and its Business Implications.
- List the components of Hadoop and Hadoop Eco-System
- Access and Process Data on Distributed File System
- Manage Job Execution in Hadoop Environment
- Develop Big Data Solutions using Hadoop Eco System

UNIT I   8 Hours

Introduction to Big Data:

Introduction: What is big data – why big data – convergence of key trends - unstructured data – industry examples of big data – Web analytics - big data and marketing – fraud and big data - risk and big data – credit risk management – big data and algorithmic trading - big data and healthcare – big data in medicine – advertising and big data – big data technologies - cloud and big data– mobile business intelligence – crowd sourcing analytics.

UNIT II   10 Hours

Hadoop:

History of Hadoop - The Hadoop Distributed File System – components of Hadoop - Analyzing the Data with Hadoop - Design of HDFS – HDFS concepts - Hadoop I/O – data integrity – compression – serialization – Avro – file-based data structures.

UNIT III   10 Hours

Map Reduce:

Map Reduce: MapReduce workflows – unit tests with MRUnit – test data and local tests – anatomy of MapReduce job run – classic Map-reduce – YARN – failures in classic Map-reduce and YARN – job scheduling – shuffle and sort – task execution –MapReduce types – input formats – output formats.

UNIT IV   10 Hours

Hadoop Eco System:

HBase – data model and implementations – HBase clients – HBase examples. Cassandra – Cassandra data model –Cassandra examples – Cassandra clients –Hadoop integration. Pig – Grunt – pig data model – Pig Latin – developing and testing Pig Latin scripts. Hive – data types and file formats – HiveQL data definition – HiveQL data manipulation –HiveQL queries-case study.

UNIT V   **10 Hours**

Graph Databases:

Introduction - Neo4J - Key concept and characteristics -Modeling data for neo4j - Importing data into neo4j - visualizations - neo4j - Cypher Query Language –data visualization.

Text Books

1. Daimi, Kevin, Hamid R. Arabnia, Principles of Data Science. Ed. Springer, 2020.

2. SinanOzdemir, Principles of Data Science: Mathematical Techniques and Theory to Succeed in Data-Driven Industries, Packt Publishing Limited, 2016.

3. Rik Van Bruggen, Learning Neo4j, Second Edition, PacktPubishers, 2014.

4. Michael Minelli, Michelle Chambers, Ambiga Dhiraj, Big Data, Big Analytics: Emerging

Business Intelligence and Analytic Trends for Today's Businesses, Wiley, 2013.

5. Tom White, "Hadoop: The Definitive Guide", Fourth Edition, O'Reilly Publishers, 2012.

Reference Books:

1.Andreas Francois Vermeulen, Ankurgupta, Cindy Gross, David Kjerrumgaard and Scott Shaw, Practical Hive: A Guide to Hadoop's Data Warehouse System, Apress Media publishers, 2016.

2. Eric Lubow and Russell Baradberry, Practical Cassandra: A Developer's Approach, Addison Wesley publishers, 2014.

3. EMC Education Services, Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data, Wiley publishers, 2015.

4. Bart Baesens, Analytics in a Big Data World: The Essential Guide to Data Science and its Applications, Wiley Publishers, 2015.

5. Kim H. Pries and Robert Dunnigan, Big Data Analytics: A Practical Guide for Managers, CRC Press, 2015.

Fourth Semester			
Title of the Paper: Mobile App Development Lab			
Paper	CPP 4.4	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	48	Credits	02

List of Practical Programs:

1. Develop an application that uses GUI components, Font and Colours

2. Develop an application that uses Layout Managers and event listeners.

3. Write an application that draws basic graphical primitives on the screen.

4. Develop an application that makes use of databases.

5. Develop an application that makes use of Notification Manager.

6. Implement an application that uses multi-threading.

7. Develop a native application that uses GPS location information

8. Implement an application that writes data to the SD card.

9. Implement an application that creates an alert upon receiving a message

10. Write a mobile application that makes use of RSS feed

11. Develop a mobile application to send an email.

12. Develop a Mobile application for simple needs

13. Develop an application that uses GUI components, Font and Colours

Fourth Semester			
Title of the Paper: Data Visualization Lab			
Paper	CPP 4.5	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	48	Credits	02

List of Practical Programs:

1. Loading and Distinguishing Dependent and Independent parameters

2. Exploring Data Visualization tools.

3. Drawing Charts

4. Drawing Graphs

5. Data mapping

6. Creating Scatter Plot maps

7. Using BNF Notations

8. Working with REGEX

9. Visualize Network Data

10. Understanding Data Visualization frameworks

Fourth Semester			
Title of the Paper: Natural Language Processing Lab			
Paper	SPP 4.6 A	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	48	Credits	02

List of Practical Programs:

1.   Fundamentals of NLP I Tokenization & Lemmatization

2.   Fundamentals of NLP II Stemming & Sentence Segmentation

3.   NLP Using Scikit Library

4.   NLP using Spacy library

5.   Working with TF-IDF

6.   Naïve Bayes Classifier

7.   Word cloud using python

8.   Python key word extraction

9.   Named entity recognition

10.   LATENT semantic analysis

11.   Determine optimum number of topics in a document

12.   Fundamental of topic modelling

Fourth Semester			
Title of the Paper: Big Data Analytics Lab			
Paper	SPP 4.6 B	CIE Marks	15
No of Hours/Week:	4	SEE Marks	35
Total Hours:	48	Credits	02

List of Practical Programs:

1. Installing Hadoop; Understanding different Hadoop modes. Startup scripts, Configuration files.

2. Hadoop Implementation of file management tasks, such as adding files and directories, retrieving files and deleting files.

3. Run a basic Word Count Map Reduce program to understand Map Reduce Paradigm.

4. Hive Installation and Table Operations.

5. Hive Databases, Tables, Views, Functions and Indexes.

6. Neo4j - Crud operations using datasets; Find a relationship between datasets; Construct a graph; String and aggregation operations.

7. Pig Latin scripts - sort, group, join, project, and filter operations.

8. Installation of Cassandra and perform key space and table operation; Crud operations

9. Installation of Hbase and simple operations

Fourth Semester			
Title of the Paper: Major Project			
Paper	PROJECT 4.7	CIE Marks	15 + 30
No of Hours/Week:	4	SEE Marks	Viva - Vove 35 Dissertation 70
Total Hours:	48	Credits	06

General Guidelines for Project to be done in Fourth Semester

The syllabus proposes the introduction of a project to be done by students in Semester IV. It is an individual student project with the allocation of 150 Marks.

The objective of introducing the Project is to introduce a Project Based Learning which helps the student

1. To explore the important core and applications areas of Computer Science.

2. To know about innovations, technological developments and new research initiatives in various areas of Computer Science.

3. To motivate students to write a research or technical paper on the project undertaken by them. This makes the course learner centered and helps them to understand the concepts covered in the syllabus and how to apply to real life situations.

Working on a project is expected to increase the problem solving ability and analytical thinking, thus helping them to face the industrial and professional demands (at least partially) once he or she completes the course.

- The projects shall be undertaken by the students under the guidance of the teacher teaching course or the experts approved by the teacher In charge.
- Each student can chose a topic with the approval of the teacher In charge.
- The topic selected should be related to the topics covered in the syllabus or any other allied area of Computer Science.
- The project work should be spread over a period of at least 16 weeks.
- The project should cover problem solving using the concepts mentioned in the syllabus, and approved by the teacher.
- Weightage shall be given to research projects, live projects and the projects with new concepts.

Theory Question paper pattern:

Instructions to Candidates: Answer all the Questions

Time: 3 Hours   **Max. Marks = 70**

I. Answer any Ten Questions of the following

1. Answer in Brief   **(10 X 2 = 20)**

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

II. Write Short notes on the following (Answer any FOUR)   **(4 x 5 = 20)**

2.

3.

4.

5.

6.

III. Answer any three questions of the following

7. Essay type questions   (Answer any THREE)   **(3 x 10 = 30)**

(a)

(b)

(c)

(d)

(e)

PRACTICAL EXAMINATION PATTERN:

1)	Writing and Execution of the Programs	25
2)	Viva - Voce	5
3)	Records	5
	Total Marks	35
4)	**Internal marks for practical**	15
Total:	**Grand**	**50 Marks**