MAHATMA GANDHI INSTITUTE OF TECHNOLOGY (Autonomous)
B.Tech. in Computer Science & Engineering (Data Science)
Scheme of Instruction and Examination
(Choice Based Credit System)
Applicable from the Academic Year 2022-23
V SEMESTER

Sl. No.	Course Code	Course Title	Instruction Hours Per Week			Examination Max. Marks		Duration of SEE in Hours	Credits
			L	T	P/D	CIE	SEE		
1.	CS501PC	Algorithm Design and Analysis	3	0	0	40	60	3	3
2.	CS502PC	Computer Networks	3	0	0	40	60	3	3
3.	CD501PC	Introduction to Data Science	3	1	0	40	60	3	4
4.		**Professional Elective – I**	3	0	0	40	60	3	3
5.		**Professional Elective – II**	3	0	0	40	60	3	3
6.	CS551PC	Computer Networks Lab	0	0	2	40	60	3	1
7.	CD551PC	R-Programming Lab	0	0	2	40	60	3	1
8.	CD552PC	ETL-Kafka/Talend	0	0	2	40	60	3	1
9.	EN551HS	Advanced English Communication Skills Laboratory	0	0	2	40	60	3	1
10.	MC502ES	Artificial Intelligence	3	0	0	40	60	3	0
11.	MC501HS	Intellectual Property Rights	3	0	0	40	60	3	0
		Total	21	1	8	440	660	33	20

L: Lecture   T: Tutorial   D: Drawing   P: Practical   CIE - Continuous Internal Evaluation   SEE - Semester End Examination

VI SEMESTER

Sl. No.	Course Code	Course Title	Instruction Hours Per Week			Examination Max. Marks		Duration of SEE in Hours	Credits
			L	T	P/D	CIE	SEE		
1.	CS601PC	Machine Learning	3	0	0	40	60	3	3
2.	IT601PC	Automata Theory and Compiler Design	3	0	0	40	60	3	3
3.	CD601PC	Big Data Analytics	3	0	0	40	60	3	3
4.		**Professional Elective – III**	3	0	0	40	60	3	3
5.		**Open Elective – I**	3	0	0	40	60	3	3
6.	CS651PC	Machine Learning Lab	0	0	2	40	60	3	1
7.	CD651PC	Big Data Analytics Lab	0	0	2	40	60	3	1
8.		**Professional Elective – III Lab**	0	0	2	40	60	3	1
9.	CD652PC	Industrial Oriented Mini Project/ Summer Internship/ Skill Development Course (UI design- Flutter)	0	0	4	-	100	-	2
10.	MC602ES	Cyber Security	3	0	0	40	60	3	0
		Total	18	0	10	360	640	27	20
11.	MC601BS	Environmental Science	3	0	0	40	60	3	0

L: Lecture   T: Tutorial   D: Drawing   P: Practical   CIE - Continuous Internal Evaluation   SEE - Semester End Examination

VII SEMESTER

Sl. No.	Course Code	Course Title	Instruction Hours Per Week			Examination Max. Marks		Duration of SEE in Hours	Credits
			L	T	P/D	CIE	SEE		
1.	CD701PC	Predictive Analytics	3	0	0	40	60	3	3
2.	CD702PC	Web and Social Media Analytics	3	0	0	40	60	3	3
3.		**Professional Elective – IV**	3	0	0	40	60	3	3
4.		**Professional Elective – V**	3	0	0	40	60	3	3
5.		**Open Elective – II**	3	0	0	40	60	3	3
6.	CD751PC	Predictive Analytics Lab	0	0	2	40	60	3	1
7.	CD752PC	Web and Social Media Analytics Lab	0	0	2	40	60	3	1
8.	CD753PC	Project Stage – I	0	0	6	100	-	-	3
		Total	**15**	**0**	**10**	**380**	**420**	**21**	**20**

L: Lecture   T: Tutorial   D: Drawing   P: Practical   CIE - Continuous Internal Evaluation   SEE - Semester End Examination

VIII SEMESTER

Sl. No.	Course Code	Course Title	Instruction Hours Per Week			Examination Max. Marks		Duration of SEE in Hours	Credits
			L	T	P/D	CIE	SEE		
1.	MS801HS	Organizational Behavior	3	0	0	40	60	3	3
2.		**Professional Elective – VI**	3	0	0	40	60	3	3
3.		**Open Elective – III**	3	0	0	40	60	3	3
4.	CD851PC	Project Stage – II including Seminar	0	0	22	40	60	-	11
		Total	**9**	**0**	**22**	**160**	**240**	**9**	**20**

L: Lecture   T: Tutorial   D: Drawing   P: Practical   CIE - Continuous Internal Evaluation   SEE - Semester End Examination

Professional Elective – I

CM511PE	Data Ware Housing and Business Intelligence
CM512PE	Web Programming
CD511PE	Graph Theory
CM514PE	Image Processing

Professional Elective – II

CD521PE	Spatial and Multimedia Databases
CD522PE	Information Retrieval Systems
IT521PE	DevOps
CD523PE	Computer Vision and Robotics

Professional Elective – III

CS631PE	Full Stack Development
CS633PE	Mobile Application Development
CD631PE	Cryptography and Network Security
CD632PE	Data Visualization Techniques

Professional Elective – III-Lab

CS611PE	Full Stack Development Lab
CS613PE	Mobile Application Development Lab
CD611PE	Cryptography and Network Security Lab
CD612PE	Data Visualization Techniques Lab

Professional Elective – IV

CD741PE	Natural Language Processing
CD742PE	Internet of Things
CD743PE	Cloud Computing
CD744PE	Information Storage Management

Professional Elective – V

CD751PE	Privacy Preserving in Data Publishing
CD752PE	Database Security
CD753PE	Data Science Applications
CD754PE	Mining Massive Datasets

Professional Elective – VI

CM851PE	Web Security
CD861PE	Data Stream Mining
CD862PE	Video Analytics
CD863PE	Block Chain Technology

Open Elective – I

CM621OE	Advanced Data Structures
CD621OE	Database Management Systems

Open Elective – II

CD721OE	Fundamentals of Data Science
CD722OE	R Programming

Open Elective – III

CD821OE	Data Analytics
CD822OE	Data Mining

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester
CS501PC: ALGORITHM DESIGN AND ANALYSIS
(Common to CSE, CSE (Data Science))

Prerequisites:
- A course on "Computer Programming and Data Structures"
- A course on "Advanced Data Structures"

Course Objectives:

- Introduces the notations for analysis of the performance of algorithms.
- Introduces the data structure disjoint sets.
- Describes major algorithmic techniques (divide-and-conquer, greedy, Dynamic Programming, backtracking and branch and bound methods) and mention problems for which each technique is appropriate
- Describes how to evaluate and compare different algorithms using worst-, average-, and best-case analysis.
- Explains the difference between tractable and intractable problems, and introduces the Problems that are P, NP and NP complete.

Course Outcomes:

Student will be able to:
- Acquire the knowledge of algorithm analysis and its notations that are applied on the problems solved by divide and conquer paradigm.
- Use greedy approach to solve an appropriate problem for optimal solution.
- Apply dynamic programming approach to solve suitable problems
- Apply the concept of back tracking, branch and bound paradigm for real time problems.
- Analyze the complexity of problems and differentiate that in terms of P and NP problems with examples.

Unit I:
Introduction: Algorithm, Pseudo code for expressing algorithms, Performance Analysis-Space complexity, Time complexity, Asymptotic Notation- Big oh notation, Omega notation, Theta notation and Little oh notation.
Divide and conquer: General method, applications-Binary search, Quick sort, Merge sort, Stassen's matrix multiplication.

Unit II:
Sets and Disjoint Set Unions: Introduction, Union and Find Operations with algorithms.
Greedy method: General method, applications-Job sequencing with deadlines, knapsack problem, Minimum cost spanning trees-Prim and rascal's Algorithm, Single source shortest path problem.

Unit III:
Dynamic Programming: General method, applications- Multistage Graphs, Optimal binary search trees, 0/1 knapsack problem, All pairs shortest path problem, The Traveling Sales Person problem, Reliability design.

Unit IV:
Backtracking: General method, applications- The 8-Queens problem, sum of subsets problem, graph coloring, Hamiltonian cycles.
Branch and Bound: General method, applications - 0/1 knapsack problem- LC Branch and Bound solution, FIFO Branch and Bound solution, Travelling sales person problem.

Unit V:
NP-Hard and NP-Complete problems: Basic concepts, Nondeterministic algorithms, The classes NP - Hard and NP-Complete, Cook's theorem, NP-Hard Graph Problems-Clique Decision Problem (CDP), Node cover decision problem.

TEXT BOOKS:
1. Ellis Horowitz, Satraj Sahni and Sanguthevar Rajasekharam, Fundamentals of Computer Algorithms, Galgotia publications Pvt. Ltd, Second Edition, 2007.
2. Aho, Ullman and Hopcroft, Design and Analysis of algorithms, Pearson education, Reprint 2002.

REFERENCE BOOKS:
1. R.C.T.Lee, S.S.Tseng, R.C.Chang and T.T Sai, Introduction to Design and Analysis of Algorithms A strategic approach, Mc Graw Hill, 2005.
2. Algorithm Design: Foundations, Analysis and Internet Examples, M.T. Goodrich and R.Tamassia, John Wiley and sons.
3. Introduction to Algorithms, 3rd Ed, T. H. Cormen, C.E. Leiserson, R. L. Rivest, and Clifford Stein, PHI Pvt. Ltd., Pearson Education.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester
CS502PC: COMPUTER NETWORKS
(Common to CSE, CSBS, CSE (AI&ML), CSE (Data Science))

Course Objectives:

- The objective of the course is to equip the students with a general overview of the concepts and fundamentals of computer networks.
- Familiarize the students with the standard models for the layered approach to communication between machines in a network and the protocols of the various layers.

Course Outcomes:

- Gain the knowledge of the basic computer network technology.
- Gain the knowledge of the functions of each layer in the OSI and TCP/IP reference model.
- Identify and analyze various routing algorithms, congestion control algorithms.
- Outline the transport layer protocols like TCP and UDP.
- List and examine the applications of HTTP, WWW, DNS, Email, FTP and the underlying protocols.

UNIT - I
Network Hardware: Local Area Networks, Metropolitan Area Networks, Wide Area Networks, Internetwork. Network software: Protocol Hierarchies, Design Issues for the Layers, Connection-Oriented and Connectionless Services. **Reference Models:** OSI, TCP/IP Reference models, Comparison of OSI and TCP/IP Models Example Networks: ARPANET, Internet. **Physical Layer: Guided Transmission media:** Magnetic Media, Twisted pairs, Coaxial Cable, Fiber Optics, Unguided Transmission Media: Radio waves, Microwaves, Infrared.

UNIT - II
Data link layer: Design issues, Framing: Character Count, Character Stuffing, Bit Stuffing Error Detection and Correction: Block Codes, Simple Parity Check, LRC, Hamming Distance, Checksum, Hamming Code, CRC, **Elementary data link protocols:** simplex protocol, A simplex stop and wait protocol for an error-free channel, A simplex stop and wait protocol for noisy channel. **Sliding Window protocols:** Pipelining, Piggybacking, A one-bit sliding window protocol, A protocol using Go-Back-N, A protocol using Selective Repeat, **Medium Access sub layer:** The channel allocation problem, Multiple access protocols: ALOHA, Carrier sense multiple access protocols, collision free protocols. Wireless LANs, Connecting devices at the data link layer.

UNIT - III
Network Layer: Design issues, Routing algorithms: shortest path routing, Flooding, Hierarchical routing, Broadcast, Multicast, distance vector routing, Link State Routing, **Congestion Control Algorithms:** Approaches to Congestion Control, Traffic aware routing, Admission Control, Traffic throttling, Load shedding. **Quality of Service:** Traffic Shaping, Packet scheduling, Admission Control, Integrated services, Differentiated Services, Internetworking, **Network layer in the Internet:** IPv4 protocol, IP Addresses, IPv6 protocol, Internet Control Protocols: ICMP, ARP, RARP, BOOTP, and DHCP.

UNIT - IV
Transport Layer: Transport Services: Services provided to Upper layer, Transport service primitives, Berkeley Sockets Elements of Transport protocols: Addressing, Error and Flow Control, Multiplexing, Crash Recovery, Connection management. **Internet Transport Protocols: TCP:** Service Model, TCP Protocol, Segment header, TCP Connection establishment and Release, TCP Connection management, TCP Sliding Window, Timer management, TCP Congestion Control. **UDP:** Protocol, UDP Header

UNIT - V
Application Layer: Domain name system- DNS Name Space, Resource records, Name Servers. SNMP, **Electronic Mail:** Architecture and Services, User Agent, Message Formats, Message Transfer, Final Delivery. SMTP, POP3, ICMP. **File transfers:** FTP, TFTP, Telnet.

TEXT BOOKS:
1. Computer Networks -- Andrew S Tanenbaum, David. j. Wetherall, 5th Edition. Pearson Education/PHI, 2011.
2. Data Communications and Networking – Behrouz A. Forouzan. Fifth Edition TMH, 2017.

REFERENCE BOOKS:

1. An Engineering Approach to Computer Networks-S. Keshav, 2nd Edition, Pearson Education.

2. "Data and Computer Communications" by William Stallings 9th Edition, Pearson Education 2010.TCP/IP Illustrated" by W. Richard Stevens, Addison-Wesley Professional 2nd edition 2011.

L	T	P	C
3	1	0	4

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester
CD501PC: INTRODUCTION TO DATA SCIENCE

Course Objectives

- Learn concepts, techniques and tools they need to deal with various facets of data science practice, including data collection and integration.
- Understand the basic types of data and basic statistics.
- Identify the importance of data reduction and data visualization techniques.

Course Outcomes

After completion of the course, the student should be able to
- Understand basic terms of Data science and Statistical Inference means.
- Identifying different types of data and statistical descriptions of data.
- Utilize R elements for data handling.
- Understanding conditionals and control flow through R elements.
- Perform data reduction and apply visualization techniques.

UNIT-I:
Introduction: Definition of Data Science- Big Data and Data Science hype – and getting past the hype- Datafication - Current landscape of perspectives - Statistical Inference - Populations and samples - Statistical modeling, probability distributions, fitting a model – Over fitting.
Basics of R: Introduction, R- Environment Setup, Programming with R, Basic Data Types.

UNIT-II:
Data Types & Statistical Description
Types of Data: Attributes and Measurement, Attribute The Type of an Attribute, The Different Types of Attributes, Describing Attributes by the Number of Values, Asymmetric Attributes, Binary Attribute, Nominal Attributes, Ordinal Attributes, Numeric Attributes, Discrete versus Continuous Attributes. Basic Statistical Descriptions of Data: Measuring the Central Tendency: Mean, Median, and Mode, Measuring the Dispersion of Data: Range, Quartiles, Variance, Standard Deviation, and Interquartile Range, Graphic Displays of Basic Statistical Descriptions ofData.

UNIT-III:
Vectors: Creating and Naming Vectors, Vector Arithmetic, Vector sub setting
Matrices: Creating and Naming Matrices, Matrix Sub setting, Arrays, Class.
Factors and Data Frames: Introduction to Factors: Factor Levels, Summarizing a Factor, Ordered Factors, Comparing Ordered Factors, Introduction to Data Frame, subsetting of Data Frames, Extending Data Frames, Sorting Data Frames.
Lists: Introduction, creating a List: Creating a Named List, Accessing List Elements, Manipulating List Elements, Merging Lists, Converting Lists to Vectors

UNIT-IV:
Conditionals and Control Flow: Relational Operators, Relational Operators and Vectors, Logical Operators, Logical Operators and Vectors, Conditional Statements.
Iterative Programming in R: Introduction, While Loop, For Loop, Looping Over List.
Functions in R: Introduction, writing a Function in R, Nested Functions, Function Scoping, Recursion, Loading an R Package, Mathematical Functions in R.

UNIT-V:
Charts and Graphs: Introduction, Pie Chart: Chart Legend, Bar Chart, Box Plot, Histogram, Line Graph: Multiple Lines in Line Graph, Scatter Plot.
Regression: Linear Regression Analysis, Multiple Linear regression.

Text Books:
1. Doing Data Science, Straight Talk from The Frontline. Cathy O'Neil and Rachel Schutt, O'Reilly,2014
2. Jiawei Han, Micheline Kamber and Jian Pei. Data Mining: Concepts and Techniques, 3rd ed. The Morgan Kaufmann Series in Data Management Systems.
3. K G Srinivas, G M Siddesh, "Statistical programming in R", Oxford Publications.

Reference Books:
1. Introduction to Data Mining, Pang-Ning Tan, Vipin Kumar, Michael Steinbanch, PearsonEducation.
2. Brain S. Everitt, "A Handbook of Statistical Analysis Using R", Second Edition, 4 LLC, 2014.
3. Dalgaard, Peter, "Introductory statistics with R", Springer Science & Business Media, 2008.
4. Paul Teetor, "R Cookbook", O'Reilly, 2011.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester
CM511PE: DATA WAREHOUSING AND BUSINESS INTELLIGENCE (PE-1)
(Common to CSE (AI &ML) & CSE (Data Science))

Course Objectives

The Student should be able to:
- This course is concerned with extracting data from the information systems that deal with the day-to-day operations and transforming it into data that can be used by businesses to drive high-level decision making.
- Students will learn how to design and create a data warehouse, and how to utilize the process of extracting, transforming, and loading (ETL) data into data warehouses..

Course Outcomes

Students will be able to:
- Understand architecture of data warehouse and OLAP operations.
- Understand Fundamental concepts of BI and Analytics
- Application of BI Key Performance indicators
- Design of Dashboards, Implementation of Web Analytics
- Understand Utilization of Advanced BI Tools and their Implementation.
- Implementation of BI Techniques and BI Ethics.

UNIT – I
DATA WAREHOUSE: Data Warehouse-Data Warehouse Architecture- Multidimensional Data Model-Data cube and OLAP Technology-Data Warehouse Implementation -DBMS schemas for Decision support - Efficient methods for Data cube computation.

UNIT - II
Business Intelligence: Introduction – Definition, Leveraging Data and Knowledge for BI, BI Components, BI Dimensions, Information Hierarchy, Business Intelligence and Business Analytics. BI Life Cycle. Data for BI - Data Issues and Data Quality for BI.

UNIT - III
BI Implementation - Key Drivers, Key Performance Indicators and Performance Metrics, BI Architecture/Framework, Best Practices, Business Decision Making, Styles of BI-vent-Driven alerts-A cyclic process of Intelligence Creation. The value of Business intelligence -Value driven and Information use.

UNIT - IV
Advanced BI – Big Data and BI, Social Networks, Mobile BI, emerging trends, Description of different BI-Tools (Pentaho, KNIME)

UNIT - V
Business intelligence implementation -Business Intelligence and integration implementation-connecting in BI systems- Issues of legality- Privacy and ethics- Social networking and BI.

TEXT BOOKS:

1. Data Mining – Concepts and Techniques - JIAWEI HAN & MICHELINE KAMBER, Elsevier.
2. Rajiv Sabherwal "Business Intelligence" Wiley Publications, 2012.

REFERENCE BOOKS:

1. Efraim Turban, Ramesh Sharda, Jay Aronson, David King, Decision Support and Business Intelligence Systems,9th Edition, Pearson Education, 2009.
2. David Loshin, Business Intelligence - The Savy Manager's Guide Getting Onboard with Emerging IT, Morgan Kaufmann Publishers, 2009.
3. Philo Janus, Stacia Misner, Building Integrated Business Intelligence Solutions with SQLServer, 2008 R2 & Office 2010, TMH, 2011.
4. Business Intelligence Data Mining and Optimization for decision making [Author: Carlo-Verellis] [Publication:(Wiley)]
5. Data Warehousing, Data Mining & OLAP- Alex Berson and Stephen J. Smith- Tata McGraw-Hill Edition, Tenth reprint 2007
6. Building the Data Warehouse- W. H. Inmon, Wiley Dreamtech India Pvt. Ltd.
7. Data Mining Introductory and Advanced topics –MARGARET H DUNHAM, PEA.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester
CM512PE: WEB PROGRAMMING (PE-I)
(Common to CSE (AI &ML) & CSE (Data Science))

Course Objectives

The student should be able to:
- Understand the technologies used in Web Programming.
- Know the importance of object-oriented aspects of Scripting.
- Understand creating database connectivity using JDBC.
- Learn the concepts of web-based application using sockets

Course Outcomes

Upon Completion of the course, the students will be able to

- Understand the technologies used in web programming to design web pages.
- Apply the Object-oriented aspects of scripting
- Create databases with connectivity using JDBC.
- Understanding server-side programming using Servlets.
- Build web-based applications using sockets.

UNIT – I

SCRIPTING: Web page Designing using HTML, Scripting basics- Client side and server side scripting. Java Script- Object, names, literals, operators and expressions- statements and features- events - windows -documents - frames - data types - built-in functions- Browser object model - Verifying forms.-HTML 5-CSS3- HTML 5 canvas - Web site creation using tools.

UNIT – II

JAVA: Introduction to object-oriented programming-Features of Java – Data types, variables and arrays–Operators – Control statements – Classes and Methods – Inheritance. Packages and Interfaces – Exception Handling – Multithreaded Programming – Input/ Output – Files – Utility Classes – String Handling.

UNIT – III

JDBC: JDBC Overview – JDBC implementation – Connection class – Statements - Catching DatabaseResults, handling database Queries. Networking– Inet Address class – URL class- TCP sockets – UDPsockets, Java Beans –RMI.

UNIT – IV

APPLETS: Java applets- Life cycle of an applet – Adding images to an applet – Adding sound to an applet. Passing parameters to an applet. Event Handling. Introducing AWT: Working with Windows Graphics and Text. Using AWT Controls, Layout Managers and Menus. Servlet – life cycle of a servlet.The Servlet API, Handling HTTP Request and Response, using Cookies, Session Tracking. Introduction to JSP.

UNIT – V

XML AND WEB SERVICES: Xml – Introduction-Form Navigation-XML Documents- XSL – XSLT- Web services- UDDI- WSDL-Java web services – Web resources.

TEXT BOOKS:

1. Harvey Deitel, Abbey Deitel, Internet and World Wide Web: How To Program 5th Edition.
2. Herbert Schildt, Java - The Complete Reference, 7th Edition. Tata McGraw- Hill Edition.
3. Michael Morrison XML Unleashed Tech media SAMS.

REFERENCE BOOKS:

1. John Pollock, Javascript - A Beginners Guide, 3rd Edition –- Tata McGraw-Hill Edition.
2. Keyur Shah, Gateway to Java Programmer Sun Certification, Tata McGraw Hill, 2002.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester
CD511PE: GRAPH THEORY (PE-I)

Prerequisites

- An understanding of Mathematics in general is sufficient

Course Outcomes

- Know some important classes of graph theoretic problems;
- Be able to formulate and prove central theorems about trees, matching, connectivity, colouring and planar graphs;
- Be able to describe and apply some basic algorithms for graphs;
- Be able to use graph theory as a modelling tool.

UNIT-I:
Introduction-Discovery of graphs, Definitions, Subgraphs, Isomorphic graphs, Matrix representations of graphs ,Degree of a vertex, Directed walks, paths and cycles, Connectivity in digraphs, Eulerian and Hamilton digraphs, Eulerian digraphs, Hamilton digraphs, Special graphs, Complements, Larger graphs from smaller graphs, Union, Sum, Cartesian Product, Composition, Graphic sequences, Graph theoretic model of the LAN problem, Havel-Hakimi criterion,Realization of a graphic sequence.

UNIT-II:
Connected graphs and shortest paths - Walks, trails, paths, cycles, Connected graphs, Distance, Cut-vertices and cut edges, Blocks, Connectivity, Weighted graphs and shortest paths, Weighted graphs, Dijkstra‟s shortest path algorithm,Floyd-Warshall shortest path algorithm.

UNIT-III:
Trees- Definitions and characterizations, Number of trees, Cayley‟s formula, Kircho-matrix-tree theorem, Minimum spanning trees, Kruskal‟s algorithm, Prim‟s algorithm, Special classes of graphs, Bipartite Graphs, Line Graphs, Chordal Graphs, Eulerian Graphs, Fleury‟s algorithm, Chinese Postman problem, Hamilton Graphs, Introduction, Necessary conditions and sufficient conditions

UNIT-IV:
Independent sets coverings and matchings – Introduction, Independent sets and coverings: basic equations Matchings in bipartite graphs, Hall‟s Theorem, K¨onig‟s Theorem, Perfect matchings in graphs, Greedy and approximation algorithms.

UNIT-V:
Vertex Colorings- Basic definitions, Cliques and chromatic number, Mycielski‟s theorem, Greedy coloring algorithm,Coloring of chordal graphs, Brooks theorem, Edge Colorings, Introduction and Basics, Gupta-Vizing theorem, Class-1and Class-2 graphs, Edge-coloring of bipartite graphs, Class-2graphs, Hajos union and Class-2 graphs, A scheduling problem and equitable edge-coloring.

TEXT BOOKS:
1. J. A. Bondy and U. S. R. Murty. Graph Theory, volume 244 of Graduate Texts in Mathematics. Springer, 1stedition, 2008.
2. J. A. Bondy and U. S. R. Murty. Graph Theory with Applications.

REFERENCE BOOKS:
1. Lecture Videos: http://nptel.ac.in/courses/111106050/13.
2. Introduction to Graph Theory, Douglas B. West, Pearson.
3. Schaum's Outlines Graph Theory, Balakrishnan, TMH.
4. Introduction to Graph Theory, Wilson Robin j, PHI.
5. Graph Theory with Applications to Engineering and Computer Science, Narsing Deo, PHI.
6. Graphs - An Introductory Approach, Wilson and Watkins

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester
CM514PE: IMAGE PROCESSING (PE –I)
(Common to CSE (AI &ML), CSE (Data Science))

Prerequisites

- Students are expected to have knowledge in linear signals and systems, Fourier Transform, basic linear algebra, basic probability theory and basic programming techniques; knowledge of Digital Signal Processing is desirable.
- A course on "Computational Mathematics"
- A course on "Computer Oriented Statistical Methods"

Course Objectives

The student should be able to:
- Provide a theoretical and mathematical foundation of fundamental Digital Image Processing concepts.
- The topics include image acquisition; sampling and quantization; preprocessing; enhancement; restoration; segmentation; and compression.

Course Outcomes

Upon Completion of the course, the students will be able to
- Demonstrate the knowledge of the basic concepts of two-dimensional signal acquisition, sampling, and quantization.
- Demonstrate the knowledge of filtering techniques.
- Demonstrate the knowledge of 2D transformation techniques.
- Demonstrate the knowledge of image enhancement, segmentation, restoration and compression techniques.

UNIT - I
Digital Image Fundamentals: Digital Image through Scanner, Digital Camera. Concept of Gray Levels. Gray Level to Binary Image Conversion, Sampling and Quantization, Relationship between Pixels, Imaging Geometry. 2D
Transformations-DFT, DCT, KLT and SVD.

UNIT - II
Image Enhancement in Spatial Domain Point Processing, Histogram Processing, Spatial Filtering, Enhancement in Frequency Domain, Image Smoothing, Image Sharpening.

UNIT - III
Image Restoration Degradation Model, Algebraic Approach to Restoration, Inverse Filtering, Least Mean Square Filters, Constrained Least Squares Restoration, Interactive Restoration.

UNIT - IV
Image Segmentation Detection of Discontinuities, Edge Linking and Boundary Detection, Thresholding, Region Oriented Segmentation.

UNIT - V
Image Compression Redundancies and their Removal Methods, Fidelity Criteria, Image Compression Models, Source Encoder and Decoder, Error Free Compression, Lossy Compression.

TEXT BOOK:

1. Digital Image Processing: R.C. Gonzalez & R. E. Woods, Addison Wesley/ Pearson Education,2nd Ed, 2004.
2. Rafael C G., Woods R E. and Eddins S L, Digital Image Processing, Prentice Hall, 3rd edition, 2008.

REFERENCE BOOKS:

1. Fundamentals of Digital Image Processing: A. K. Jain, PHI.
2. Digital Image Processing using MAT LAB: Rafael C. Gonzalez, Richard E. Woods, Steven L.Eddins: Pearson Education India, 2004.
3. Digital Image Processing: William K. Pratt, John Wilely, 3rd Edition, 2004.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester

CD521PE: SPATIAL AND MULTIMEDIA DATABASES (PE-II)

Course Objectives

- Introduce the basic concepts, data models and indexing structures for spatial data, multimedia data.

Course Outcomes

- Understand data models, storage, indexing and design of spatial databases.
- Evaluate multidimensional data structures.
- Represent image database with R-tree.
- Store and retrieve multimedia data.

UNIT-I:
Introduction to Spatial Databases: Overview, beneficiaries, GIA and SDBMS, users, Space taxonomy, query language, query processing, query optimization.
Spatial Concepts and Data Models: Models of Spatial information, three step database design, Extending the ER model with spatial concept, object- oriented data modeling. Spatial Query Languages.

UNIT-II:
Spatial Storage and Indexing: Storage-disks and files, spatial indexing, TR*, spatial join index. Query processing and optimization – Evaluation of Spatial operations, query optimization, Analysis of Spatial index structures, distributed and parallel spatial database system.
Multidimensional Data Structures: k- d Trees, Point Quadtrees, The MX-Quadtree, R-Trees, comparison of Different Data Structures.

UNIT-III:
Image Databases: Raw Images, Compressed Image Representations, Image Processing: Segmentation, Similarity-Based Retrieval, Alternative Image DB Paradigms, Representing Image DBs with Relations, Representing Image DBs with R-Trees, Retrieving Images By Spatial Layout, Implementations.
Text/Document Databases: Precision and Recall, Stop Lists, Word Stems, and Frequency Tables, Latent Semantic Indexing, TV-Trees, Other Retrieval Techniques.

UNIT-IV:
Video Databases : Organizing Content of a Single Video, Querying Content of Video Libraries, Video Segmentation, video Standards.
Audio Databases : A General Model of Audio Data, Capturing Audio Content through Discrete Transformation, Indexing Audio Data.
Multimedia Databases : Design and Architecture of a Multimedia Database, Organizing Multimedia Data Based on The Principle of Uniformity, Media Abstractions, Query Languages for Retrieving Multimedia Data, Indexing SMDSs with Enhanced Inverted Indices, Query Relaxation/Expansion.

UNIT-V:
Creating Distributed Multimedia Presentations: Objects in Multimedia Presentations, Specifying Multimedia Documents with Temporal Constraints, Efficient Solution of Temporal Presentation Constraints, Spatial Constraints.
Distributed Media Servers: Distributed multimedia server architecture, distributed retrieval plans, optimal distributed retrieval plans.

TEXT BOOKS:
1. Shashi Shekhar, Sanjiv Chawla, Spatial Databases-A Tour, Pearson Education.
2. V.S. Subrahmanian, Principles of Multimedia Database Systems, Morgan Kauffman

REFERENCE BOOKS:
1. Multimedia Databases: An object relational approach, Lynne Dunckley, Pearson Education.
2. Multimedia Database Systems, Prabhakaran, Springer.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester

CD522PE: INFORMATION RETRIEVAL SYSTEMS (PE –II)

Prerequisites

- Data Structures

Course Objectives

- To learn the important concepts and algorithms in IRS
- To understand the data/file structures that are necessary to design, and implement information retrieval (IR) systems.

Course Outcomes

- Ability to apply IR principles to locate relevant information large collections of data
- Ability to design different document clustering algorithms
- Implement retrieval systems for web search tasks.
- Design an Information Retrieval System for web search tasks

UNIT-I:
Introduction to Information Retrieval Systems: Definition of Information Retrieval System, Objectives of Information Retrieval Systems, Functional Overview, Relationship to Database Management Systems, Digital Libraries and Data Warehouses Information Retrieval System Capabilities: Search Capabilities, Browse Capabilities, Miscellaneous Capabilities.

UNIT-II:
Cataloging and Indexing: History and Objectives of Indexing, Indexing Process, Automatic Indexing, Information Extraction. **Data Structure:** Introduction to Data Structure, Stemming Algorithms, Inverted File Structure, N-Gram Data Structures, PAT Data Structure, Signature File Structure, Hypertext and XML Data Structures, Hidden Markov Models

UNIT-III:
Automatic Indexing: Classes of Automatic Indexing, Statistical Indexing, Natural Language, Concept Indexing, Hypertext Linkages. Document and Term Clustering: Introduction to Clustering, Thesaurus Generation, Item Clustering, Hierarchy of Clusters.

UNIT-IV:
User Search Techniques: Search Statements and Binding, Similarity Measures and Ranking, Relevance Feedback, Selective Dissemination of Information Search, Weighted Searches of Boolean Systems, Searching the INTERNET and Hypertext. **Information Visualization:** Introduction to Information Visualization, Cognition and Perception, Information Visualization Technologies.

UNIT-V:
Text Search Algorithms: Introduction to Text Search Techniques, Software Text Search Algorithms, Hardware Text Search Systems .**Multimedia Information Retrieval:** Spoken Language Audio Retrieval, Non-Speech Audio Retrieval, Graph Retrieval, Imagery Retrieval, Video Retrieval.

TEXT BOOKS:

1. Information Storage and Retrieval Systems – Theory and Implementation, Second Edition, Gerald J. Kowalski, Mark T. Maybury, Springer

REFERENCE BOOKS:

1. Frakes, W.B., Ricardo Baeza-Yates: Information Retrieval Data Structures and Algorithms, Prentice Hall, 1992.
2. Information Storage & Retrieval By Robert Korfhage – John Wiley & Sons.
3. Modern Information Retrieval By Yates and Neto Pearson Education.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester
IT521PE : DevOps
(Professional Elective – II)
(Common to IT & CSE [Data Science])

Prerequisites:
- Software Engineering
- Software Project Management

Course Objectives:

- Describe the agile relationship between development and IT operations.
- Understand the skill sets and high-functioning teams involved in DevOps and related methods to reach a continuous delivery capability.
- Implement automated system update and DevOps lifecycle.

Course Outcomes:

- Identify components of DevOps environment.
- Describe Software development models and architectures of DevOps.
- Apply different project management, integration, testing and code deployment tool.
- Investigate different DevOps Software development models.
- Assess various DevOps practices.
- Collaborate and adopt DevOps in real-time projects.

UNIT – I : Introduction:
Introduction, Agile development model, DevOps, and ITIL. DevOps process and Continuous Delivery, Release management, Scrum, Kanban, delivery pipeline, identifying bottlenecks.

UNIT – II : Software development models and DevOps :
DevOps Lifecycle for Business Agility, DevOps, and Continuous Testing. DevOps influence on Architecture: Introducing software architecture, The monolithic scenario, Architecture rules of thumb, The separation of concerns, Handling database migrations, Microservices, and the data tier, DevOps, architecture, and resilience.

UNIT – III : Introduction to project management:
The need for source code control, The history of source code management, Roles and code, source code management system and migrations, Shared authentication, Hosted Git servers, Different Git server implementations, Docker intermission, Gerrit, The pull request model, GitLab.

UNIT – IV : Integrating the system :
Build systems, Jenkins build server, Managing build dependencies, Jenkins plugins, and file system layout, The host server, Build slaves, Software on the host, Triggers, Job chaining and build pipelines, Build servers and infrastructure as code, Building by dependency order, Build phases, Alternative build servers, Collating quality measures.

UNIT – V : Testing Tools and automation:
Various types of testing, Automation of testing Pros and cons, Selenium - Introduction, Selenium features, JavaScript testing, Testing backend integration points, Test-driven development, REPL-driven development. Deployment of the system: Deployment systems, Virtualization stacks, code execution at the client, Puppet master and agents, Ansible, Deployment tools: Chef, SaltStackand Docker.

TEXT BOOKS

1. Joakim Verona., Practical DevOps, Packt Publishing, 2016.

REFERENCE BOOKS:

1. Deepak Gaikwad, Viral Thakkar. DevOps Tools from Practitioner's Viewpoint.Wiley publications.
2. Len Bass, Ingo Weber, Liming Zhu. DevOps: A Software Architect's Perspective.Addison Wesley.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester
CD523PE: COMPUTER VISION AND ROBOTICS (PE-II)

Pre-Requisites

- Linear Algebra and Probability

Course Objectives

- To understand the Fundamental Concepts Related To sources, shadows and shading.
- To understand The Geometry of Multiple Views.

Course Outcomes

- Implement fundamental image processing techniques required for computer vision.
- Implement boundary tracking techniques.
- Apply chain codes and other region descriptors, Hough Transform for line, circle, and ellipse detections.
- Apply 3D vision techniques and Implement motion related techniques.
- Develop applications using computer vision techniques

UNIT-I:
CAMERAS: Pinhole Cameras.
Radiometry – Measuring Light: Light in Space, Light Surfaces, Important Special Cases.
Sources, Shadows, And Shading: Qualitative Radiometry, Sources and Their Effects, Local Shading Models, Application: Photometric Stereo, Interreflections: Global Shading Models.
Color: The Physics of Color, Human Color Perception, Representing Color, A Model for Image Color, Surface Color from Image Color.

UNIT-II:
Linear Filters: Linear Filters and Convolution, Shift Invariant Linear Systems, Spatial Frequency and Fourier Transforms, Sampling and Aliasing, Filters as Templates.
Edge Detection: Noise, Estimating Derivatives, Detecting Edges.
Texture: Representing Texture, Analysis (and Synthesis) Using Oriented Pyramids, Application: Synthesis by Sampling Local Models, Shape from Texture.

UNIT-III:
The Geometry of Multiple Views: Two Views
Stereopsis: Reconstruction, Human Stereposis, Binocular Fusion, Using More Cameras Segmentation by Clustering:
Segmentation by Clustering: Segmentation, Human Vision: Grouping and Getstalt, Applications: Shot Boundary Detection and Background Subtraction, Image Segmentation by Clustering Pixels, Segmentation by Graph-Theoretic Clustering.

UNIT-IV:
Segmentation by Fitting a Model: The Hough Transform, Fitting Lines, Fitting Curves, Fitting as a Probabilistic Inference Problem, Robustness
Geometric Camera Models: Elements of Analytical Euclidean Geometry, Camera Parameters and the Perspective Projection, Affine Cameras and Affine Projection Equations
Geometric Camera Calibration: Least-Squares Parameter Estimation, A Linear Approach to Camera Calibration, Taking Radial Distortion into Account, Analytical Photogrammetry, An Application: Mobile Robot Localization.

UNIT-V:
Introduction to Robotics: Social Implications of Robotics, Brief history of Robotics, Attributes of hierarchical paradigm, Closed world assumption and frame problem, Representative Architectures, Attributes of Reactive Paradigm, Subsumption Architecture, Potential fields and Perception
Common sensing techniques for Reactive Robots: Logical sensors, Behavioural Sensor Fusion, Pro- prioceptive sensors, Proximity Sensors, Topological Planning and Metric Path Planning.

TEXT BOOKS:
1. David A. Forsyth and Jean Ponce: Computer Vision – A Modern Approach, PHI Learning (Indian Edition), 2009.
2. Robin Murphy, Introduction to AI Robotics, MIT Press

REFERENCE BOOKS:
1. E. R. Davies: Computer and Machine Vision – Theory, Algorithms and Practicalities, Elsevier (Academic Press), 4th edition, 2013.
2. The Robotics premier, Maja J Matari, MIT Press
3. Richard Szeliski "Computer Vision: Algorithms and Applications" Springer-Verlag London Limited 2011.

L	T	P	C
0	0	2	1

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester
CS551PC: COMPUTER NETWORKS LAB
(Common to CSE, CSBS, CSE (AI&ML), CSE (Data Science))

Course Objectives:

- To understand the working principle of various communication protocols.
- To understand the network simulator environment and visualize a network topology and observe its performance
- To analyze the traffic flow and the contents of protocol frames

Course Outcomes:

- Implement data link layer farming methods
- Analyze error detection and error correction codes.
- Implement and analyze routing and congestion issues in network design.
- Implement Encoding and Decoding techniques used in presentation layer.
- To be able to work with different network tools.

List of Experiments:

1. Write a program to implement framing.

i. Fixed framing, ii. Variable framings

2. Write a program to implement stuffing techniques.

i. Bit-stuffing ii. Character stuffing

3. Write a program to implement checksum

4. Write a program to compute CRC code for the polynomial CRC-12

5. Develop a simple data link layer that performs the flow control using the sliding window protocol, and loss recovery using the Go-Back-N mechanism.

6. Write a program to implement classful addressing.

7. Write a program to implement classless addressing.

8. Write a program to implement Dijsktra's algorithm to compute the shortest path through a network

9. Write a program to divide a given network into n-sub networks using CISCO Packet Tracer.

10. Implement distance vector routing algorithm for obtaining routing tables at each node.

11. Write a program to implement Link state routing

12. Write a program for congestion control using Leaky bucket algorithm

13. Implement data encryption and data decryption.

14. Basic Trouble shooting commands ipconfig,ping ,tracert,nslookup,arp etc.

TEXT BOOK:

1. Computer Networks—Andrew S Tanenbaum, David. j. Wetherall, 5th Edition. Pearson Education/PHI,2011.

L	T	P	C
0	0	2	1

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester
CD551PC: R-PROGRAMMING LAB

Pre-requisites

- Any programming language

Course Objectives

- Familiarize with R basic programming concepts, various data structures for handling datasets, various graph representations and Exploratory Data Analysis concepts

Course Outcomes

- Setup R programming environment.
- Understand and use R – Data types and R – Data Structures.
- Develop programming logic using R – Packages.
- Analyze data sets using R – programming capabilities

List of Programs:

1. Download and install R-Programming environment and install basic packages using install. packages() command in R.
2. Learn all the basics of R-Programming (Data types, Variables, Operators etc,.)
3. Write R command to
   i.   Illustrate summation, subtraction, multiplication, and division operations on vectors using vectors.
   ii. Enumerate multiplication and division operations between matrices and vectors in R console
4. Write R command to
   1. Illustrates the usage of Vector subsetting and Matrix subsetting ii) Write a program to create an array of 3×3 matrixes with 3 rows and 3 columns.
5. Write an R program to draw i) Pie chart ii) 3D Pie Chart, iii) Bar Chart along with chart   legend by considering suitable CSV file
6. Create a CSV file having Speed and Distance attributes with 1000 records. Write R program to draw i) Box plots ii) Histogram iii) Line Graph iv) Multiple line graphs v) Scatter plot to demonstrate the relation between the cars speed and the distance.
7. Implement different data structures in R (Vectors, Lists, Data Frames)
8. Write an R program to read a csv file and analyze the data in the file using EDA (Explorative Data Analysis) techniques.
9. Write an R program to illustrate Linear Regression and Multi linear Regression considering suitable CSV file

TEXT BOOKS:

1. R Programming for Data Science by Roger D. Peng
2. The Art of R Programming by Norman Matloff Cengage Learning India

REFERENCE BOOKS:

1. Hadley Wickham, Garrett Grolemund, R for Data Science: Import, Tidy, Transform, Visualize, and Model Data 1st Edition, O'Reilly.
2. Tilman M. Davies, The book of R a first course in programming and statistics, no starch press.

L	T	P	C
0	0	2	1

B.Tech. V Semester

CD552PC: ETL-KAFKA/TALEND

Course Objectives

- Develop a comprehensive understanding of Extract, Transform, Load (ETL) processes using Apache Kafka and Talend.
- Understand how to scale Kafka clusters seamlessly to handle growing data volumes, ensuring optimal performance for ETL operations.

Course Outcomes

- Learn to design and deploy fault-tolerant Kafka clusters, ensuring data integrity and availability in real-world scenarios.
- Gain practical experience in cluster management, topic creation, and basic operations such as producing and

List of Programs

1. Install Apache Kafka on a single node.
2. Demonstrate setting up a single-node, single-broker Kafka cluster and show basic operations such as creating topics and producing/consuming messages.
3. Extend the cluster to multiple brokers on a single node.
4. Write a simple Java program to create a Kafka producer and Produce messages to a topic.
5. Implement sending messages both synchronously and asynchronously in the producer.
6. Develop a Java program to create a Kafka consumer and subscribe to a topic and consume
7. Write a script to create a topic with specific partition and replication factor settings.
8. Simulate fault tolerance by shutting down one broker and observing the cluster behavior.
9. Implement operations such as listing topics, modifying configurations, and deleting topics.
10. Introduce Kafka Connect and demonstrate how to use connectors to integrate with external systems.
11. Implement a simple word count stream processing application using Kafka Stream
12. Implement Kafka integration with the Hadoop ecosystem.

Text Books:

1. Neha Narkhede, Gwen Shapira, Todd Palino, Kafka – The Definitive Guide: Real-time data and stream processing at scale, O′Reilly.

L	T	P	C
0	0	2	1

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester
EN551HS: Advanced English Communication Skills Laboratory
(Common to CSE, IT, CSBS, CSE (AI&ML) & CSE (DS))

Introduction:

The introduction of the Advanced English Communication Skills Lab is considered essential at the B.Tech 3rd year level. At this stage, the students need to prepare themselves for their career which may require them to listen to, read, speak and write in English both for their professional and interpersonal communication in the globalized context.

The proposed course should be a laboratory course to enable students to use appropriate English and perform the following:

1. Gathering ideas and information to organize ideas relevantly and coherently.
2. Making oral presentations.
3. Writing formal letters.
4. Transferring information from non-verbal to verbal texts and vice-versa.
5. Writing project/research reports/technical reports.
6. Participating in group discussions.
7. Engaging in debates.
8. Facing interviews.
9. Taking part in social and professional communication.

Course Objectives:
This Lab focuses on using multi-media instruction for language development to meet the following targets:
1. Improve the students' fluency in English, with a focus on vocabulary.
2. Enable them to listen to English spoken at normal conversational speed by educated English speakers.
3. Respond appropriately in different socio-cultural and professional contexts.
4. Communicate their ideas relevantly and coherently in writing.
5. Prepare the students for placements.

Course Outcomes:
     Students will be able to:
1. Enhance listening proficiency and reading comprehension and cultivate critical thinking ability.
2. Acquire essential vocabulary and develop strategic planning skills for effective technical writing and gain expertise in E-Correspondence and (N) etiquette.
3. Understand the nuances of oral skills (Speaking skills), gain competence in delivering effective presentations, employing suitable language and body language.
4. Communicate confidently in group discussions and enhance the employability skills of students.
5. Apply effective techniques and strategies for successful job interviews.

Syllabus:
The following course content to conduct the activities is prescribed for the Advanced English Communication Skills (AECS) Lab:

1. **Activities on Listening and Reading Comprehension:** Active Listening – Development of Listening Skills Through Audio clips - Benefits of Reading – Methods and Techniques of Reading – Basic Steps to Effective Reading – Common Obstacles – Discourse Markers or Linkers - Sub-skills of reading - Reading for facts, negative facts and Specific Details- Guessing Meanings from Context, Inferring Meaning - Critical Reading — Reading Comprehension – Exercises for Practice.

2. **Activities on Writing Skills**: Vocabulary for Competitive Examinations - Planning for Writing – Improving Writing Skills - Structure and presentation of different types of writing – Free Writing and Structured Writing - Letter Writing – Writing a Letter of Application –Resume vs. Curriculum Vitae – Writing a Résumé **–** Styles of Résumé - e-Correspondence – Emails – Blog Writing - (N)etiquette – Report Writing – Importance of Reports – Types and Formats of Reports– Technical Report Writing– Exercises for Practice.

3. **Activities on Presentation Skills –** Dealing with Glossophobia or stage fear, starting a conversation – responding appropriately and relevantly – using the right language and body language – Role Play in different situations including Seeking Clarification, Making a Request, Asking for and Refusing Permission, Participating in a Small Talk – Oral presentations (individual and group) through JAM sessions- PPTs – Importance of Presentation Skills – Planning, Preparing, Rehearsing and Making a Presentation - Understanding Nuances of Delivery - Presentations through Posters/Projects/Reports – Checklist for Making a Presentation and Rubrics of Evaluation.

4. **Activities on Group Discussion (GD):** Types of GD and GD as a part of a Selection Procedure - Dynamics of Group Discussion - myths and facts (Dos and Don'ts) of GD - Intervention, Summarizing - Modulation of Voice, Body Language, Relevance, Fluency and Organization of Ideas - GD Strategies – Exercises for Practice.

5. **Activities on Interview Skills**: Concept and Process - Interview Preparation Techniques - Types of Interview Questions – Pre-interview Planning, Opening Strategies, Answering Strategies - Interview Through Tele-conference & Video-conference - Mock Interviews.

Text Books:

1. Effective Technical Communication by M Ashraf Rizvi. McGraw Hill Education (India) Pvt. Ltd. 2nd Edition.
2. Academic Writing: A Handbook for International Students by Stephen Bailey, Routledge, 5th Edition.

Reference Books:

1. Rizvi, M. Ashraf (2018). *Effective Technical Communication*. (2nd ed). McGraw Hill Education (India) Pvt. Ltd.
2. Suresh Kumar, E. (2015). *Engineering English.* Orient BlackSwan Pvt. Ltd.
3. Bailey, Stephen. (2018). Academic *Writing: A Handbook for International Students.* (5th Edition). Routledge.
4. Koneru, Aruna. (2016). *Professional Communication*.   McGraw Hill Education (India) Pvt. Ltd.
5. Raman, Meenakshi & Sharma, Sangeeta. (2015). *Technical Communication, 3E: Principles and Practice.* Oxford University Press.
6. Anderson, Paul V. (2007). *Technical Communication***.** Cengage Learning Pvt. Ltd. New Delhi.
7. McCarthy, Michael; O'Dell, Felicity & Redman, Stuart. (2017). *English Vocabulary in Use* Series. Cambridge University Press.
8. Sen, Leela. (2009). *Communication Skills.* PHI Learning Pvt Ltd., New Delhi.
9. Elbow, Peter. (1998). *Writing with Power.* Oxford University Press.
10. Goleman, Daniel. (2013). *Emotional Intelligence*: *Why it can matter more than IQ.* Bloomsbury Publishing.
11. Professional Communication by Aruna Koneru, McGraw Hill Education (India) Pvt. Ltd, 2016.
12. Technical Communication by Meenakshi Raman & Sangeeta Sharma, Oxford University Press 2009.
13. How to Write and Speak Better, Reader's Digest, 2003.
14. TOEFL Reading & Writing Workout, The Princeton Review.
15. How to prepare for Group Discussions and Interviews by Harimohan Prasad and Rajneesh Prasad,TataMcgrawHill.
16. Keep Talking, Frederick Klippel, Cambridge University Press, South Asian edition (6 May 2010).
17. Objective English, Edgar Thorpe & Showick Thorpe, Pearson; 5th edition (1 August 2013).
18. Communication Skills for Engineers, Sunitha Mishra, C.Murali Krishna, Pearson; 4th Edition.

****

L	T	P	C
3	0	0	0

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester
MC502ES: Artificial Intelligence
(Common to all branches except CSE, IT, CSBS, CSE (AI&ML))

Course Objectives:

- To train the students to understand different types of AI agents, various AI search algorithms, fundamentals of knowledge representation, building of simple knowledge-based systems and to apply knowledge representation, reasoning.
- Study of Markov Models enable the student ready to step into applied AI.

UNIT - I
Introduction: AI problems, Agents and Environments, Structure of Agents, Problem Solving Agents **Basic Search Strategies**: Problem Spaces, Uninformed Search (Breadth-First, Depth-First Search, Depth-first with Iterative Deepening), Heuristic Search (Hill Climbing, Generic Best-First, A*), Constraint Satisfaction (Backtracking, Local Search)

UNIT - II
Advanced Search: Constructing Search Trees, Stochastic Search, A* Search Implementation, Minimax Search, Alpha-Beta Pruning
Basic Knowledge Representation and Reasoning: Propositional Logic, First-Order Logic, Forward Chaining and Backward Chaining, Introduction to Probabilistic Reasoning, Bayes Theorem

UNIT - III
Advanced Knowledge Representation and Reasoning: Knowledge Representation Issues, Non- monotonic Reasoning, Other Knowledge Representation Schemes
Reasoning Under Uncertainty: Basic probability, Acting Under Uncertainty, Bayes' Rule, Representing Knowledge in an Uncertain Domain, Bayesian Networks

UNIT – IV
Learning: What Is Learning? Rote Learning, Learning by Taking Advice, Learning in Problem Solving, Learning from Examples, Winston's Learning Program, Decision Trees.

UNIT - V
Expert Systems: Representing and Using Domain Knowledge, Shell, Explanation, Knowledge Acquisition.

TEXT BOOK:

1. Artificial Intelligence, Elaine Rich, Kevin Knight, Shivasankar B. Nair, The McGraw Hill publications, Third Edition, 2009

REFERENCE BOOKS:
1. Russell, S. and Norvig, P, Artificial Intelligence: A Modern Approach, Third Edition, Prentice-Hall, 2010.
2. George F. Luger, Artificial Intelligence: Structures and Strategies for Complex Problem Solving, Pearson Education, 6th ed., 2009.

L	T	P	C
3	0	0	0

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. V Semester
MC501HS: Intellectual Property Rights
(Common to EEE, CSE, IT, CSBS, CSE (AI&ML) & CSE (DS))

Course Objectives: The objectives of the course are:

1. To enable the students to have an overview of Intellectual Property Rights.
2. To provide comprehensive knowledge to the students regarding Trademarks Registration process and law related to it.
3. To disseminate knowledge on Copyrights, its related rights and recent developments.
4. To make the students understand Patent Regime in India and abroad.
5. To understand the framework of Trade secrets.

Course Outcomes: By the end of the course students shall:

1. Gain knowledge on Intellectual property rights and their importance.
2. Understand Indian and International Trademark Law and procedure for registration of Trademarks.
3. Acquire knowledge on Copyright Law, and the privileges awarded to the copyright owners.
4. Familiarized with the process of acquiring the patent and relevant laws.
5. Learn the importance of trade secrets for business sustainability.

UNIT – I: INTRODUCTION TO INTELLECTUAL PROPERTY
Introduction of IPR-Meaning of intellectual property, types of intellectual property-trademarks, copyrights, patents, trade secrets, importance of intellectual property rights, International organizations-WTO-WIPO-USPTO-INTA, International Conventions, agencies and treaties- Paris Convention-Berne Convention- Madrid Protocol-NAFTA-PCT-GATT-TRIPS.

UNIT – II: TRADEMARKS
Trademarks: Purpose and functions of Trademarks-Categories of marks, acquisition of trademark rights - Protectable matter - Selecting and evaluating Trademark- Trademark registration process – Trademark Infringement - Remedies for infringement of Trademarks-New developments in Trademark Law- International Trademarks Law.

UNIT III: COPYRIGHT
Copyrights-Fundamentals of Copyright Law - Requirements of Copyrightability - Originality of material, fixation of material, Authorship works, exclusions from copyright protection- Rights of Copyright Owner-Right of reproduction of copyrighted work, right to do derivative works ,right to distribute copies of the copyrighted work, right to perform the work publicly, right to display the copyrighted work, – Copyright Ownership issues – Joint Works, Works made for Hire, Specially commissioned works, Copyright Registration - Notice of Copyright – Copyright Infringement - Remedies for infringement in Copyrights-New developments in Copyright Law- International Copyright Law.

UNIT IV: PATENTS
Concept of Patent - Classification – Utility Patents – Design Patents and Plant Patents, Patent searching process-Types of Patent Applications-Patent Registration Process, Ownership, Transfer, Assignment and Licensing of Patent-Patent Infringement, Remedies for Infringement of Patents, New developments in Patent Law- International Patent Law.

UNIT – V: TRADE SECRETS & LAW OF UNFAIR COMPETITION
Trade Secrets: Trade secret law, determination of trade secret status, measures for protecting trade secret status-Liability for misappropriation of trade secrets, protection for submissions, trade secret litigation. New developments in Trade secrets Law - International Trade Secret law.
Law of Unfair Competition: Passing off, Misappropriation, Right of publicity, Dilution of trademarks, Product disparagement, False advertising, Internet Piracy.

TEXT BOOKS:

1. Deborah. E.Bouchoux, Intellectual property, Cengage learning India Pvt.Ltd., 4th edition, 2013.
2. Prabuddha Ganguli, Intellectual property right, Tata McGraw Hill Publishing Company, 8th edition, 2016.

REFERENCES

1. Richard Stim, Intellectual Property, Cengage learning India Pvt. Ltd. 3rd edition, 2017
2. Vinod.V. Sope, Managing Intellectual Property, Asoka K. Ghosh, 2nd edition, 2010.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. VI Semester
CS601PC - Machine Learning
(Common to CSE, CSE (Data Science))

Prerequisites: 1. Data Structures 2. Knowledge on statistical methods

Course Objectives:

- This course explains machine learning techniques such as decision tree learning, Bayesian learning etc.
- To understand computational learning theory.
- To study the pattern comparison techniques.

Course Outcomes:

- Understand the concepts of computational intelligence like machine learning
- Ability to get the skill to apply machine learning techniques to address the real time problems in different areas
- Understand the Neural Networks and its usage in machine learning application.

Unit - I

Introduction: Learning- types of machine learning, Perspectives and issues in machine learning, bias variance trades off, testing machine learning algorithms, confusion matrix, accuracy metrics, turning data into probabilities. Dimensionality reduction- Linear discriminant analysis, principal component analysis, Factor analysis, independent component analysis.

Unit - II

Supervised Learning: Decision Tree Learning – Introduction, decision tree representation, appropriate problems for decision tree learning, the basic decision tree learning algorithm, Information Gain, Gini Index. issues in decision tree learning, overfitting, pruning.Instance-Based Learning- Introduction, k-nearest neighbours' algorithm, remarks on lazy and eager learning.

Unit - III

Neural Networks: The perceptron- learning rate, bias, perceptron learning algorithm, linear separability- Exclusive-OR function, backpropagation algorithm, initialization of weights, MLP in practice, deriving back propagation, radial basis functions. An illustrative example: face recognition.

Unit - IV

Bayesian learning – Introduction, Conditional probability, Joint probability, Bayes theorem, maximum likelihood hypotheses for predicting probabilities, Bayes optimal classifier, Naïve Bayes classifier, an example: learning to classify text, Bayesian belief networks. Support vector machines, optimal separation kernels, SVM algorithm.

Unit - V

Unsupervised learning- Clustering, Hierarchical and partitional clustering, k means algorithm, ensemble learning-Boosting Bagging. Reinforcement Learning – Introduction, the learning task, Q–learning, non- deterministic, rewards and actions, temporal difference learning.

TEXT BOOKS:

1. Machine Learning – Tom M. Mitchell, - MGH
2. Ethem Alpaydin, Introduction to Machine Learning (Adaptive Computation and Machine Learning), The MIT Press, 2015

REFERENCE BOOKS:

1.Machine Learning: An Algorithmic Perspective, Stephen Marshland, Taylor & Francis.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. VI Semester
IT601PC : Automata Theory and Compiler Design
(Common *to IT, CSBS & CSE (Data Science)*)

Course Objectives:

1. To introduce the fundamental concepts of formal languages, grammars and automata theory.
2. To understand deterministic and non-deterministic machines and the differences betweendecidability and undecidability.
3. Introduce the major concepts of language translation and compiler design and impart theknowledge of practical skills necessary for constructing a compiler.
4. Topics include phases of compiler, parsing, syntax directed translation, type checking use ofsymbol tables, intermediate code generation.

Course Outcomes:

1. Able to employ finite state machines for modeling and solving computing problems.
2. Able to design context free grammars for formal languages.
3. Able to distinguish between decidability and undecidability.
4. Demonstrate the knowledge of patterns, tokens & regular expressions for lexical analysis.
5. Acquire skills in using lex tool and design LR parsers

UNIT – I:
Introduction to Finite Automata: Structural Representations, Automata and Complexity, the Central Concepts of Automata Theory – Alphabets, Strings, Languages, Problems.
Nondeterministic Finite Automata: Formal Definition, an application, Text Search, Finite Automata with Epsilon-Transitions.
Deterministic Finite Automata: Definition of DFA, How A DFA Process Strings, The language of DFA, Conversion of NFA with €transitions to NFA without €transitions. Conversion of NFA to DFA

UNIT – II:
Regular Expressions: Finite Automata and Regular Expressions, Applications of Regular Expressions, Algebraic Laws for Regular Expressions, Conversion of Finite Automata to Regular Expressions.
Pumping Lemma for Regular Languages: Statement of the pumping lemma, Applications of the Pumping Lemma.
Context-Free Grammars: Definition of Context-Free Grammars, Derivations Using a Grammar, Leftmost and Rightmost Derivations, the Language of a Grammar, Parse Trees, Ambiguity in Grammars and Languages.

UNIT – III :
Push Down Automata: Definition of the Pushdown Automaton, the Languages of a PDA,Equivalence of PDA's and CFG's, Acceptance by final state
Turing Machines: Introduction to Turing Machine, Formal Description, Instantaneous description, The language of a Turing machine
Undecidability: Undecidability, A Language that is Not Recursively Enumerable, An Undecidable Problem That is RE, Undecidable Problems about Turing Machines

UNIT – IV :

Introduction: The structure of a compiler

Lexical Analysis: The Role of the Lexical Analyzer, Input Buffering, Recognition of Tokens, The Lexical- Analyzer Generator Lex

Syntax Analysis: Introduction, Context-Free Grammars, Writing a Grammar, Top-Down Parsing, Bottom- Up Parsing, Introduction to LR Parsing: Simple LR, More Powerful LR Parsers

UNIT – V:

Syntax-Directed Translation: Syntax-Directed Definitions, Evaluation Orders for SDD's,Syntax- Directed Translation Schemes, Implementing L-Attributed SDD's.

Intermediate-Code Generation: Variants of Syntax Trees, Three-Address Code

Run-Time Environments: Stack Allocation of Space, Access to Nonlocal Data on the Stack,Heap Management

TEXT BOOKS :

1. Introduction to Automata Theory, Languages, and Computation, 3nd Edition, John E. Hopcroft, Rajeev Motwani, Jeffrey D. Ullman, Pearson Education.
2. Theory of Computer Science – Automata languages and computation, Mishra and Chandrashekaran, 2nd Edition, PHI.

REFERENCE BOOKS :

1. Compilers: Principles, Techniques and Tools, Alfred V. Aho, Monica S. Lam, Ravi Sethi,Jeffry D. Ullman, 2nd Edition, P earson.
2. Introduction to Formal languages Automata Theory and Computation, Kamala Krithivasan,Rama R, Pearson.
3. Introduction to Languages and The Theory of Computation, John C Martin, TMH.
4. lex & yacc – John R. Levine, Tony Mason, Doug Brown, O'reilly Compiler Construction, Kenneth C. Louden, Thomson. Course Technology.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. VI Semester
CD601PC: BIG DATA ANALYTICS

Course Objectives

- Provide the knowledge of principles and techniques for Big data Analytics and give an exposure of the frontiers of Big data Analytics.

Course Outcomes

- Understand the importance of big data analytics and its types
- Perform analytics on big data
- Proficiency in big data storage and processing in Hadoop
- Data analytics through MongoDB
- Data analytics through R.

UNIT-I:
Types of Digital data: Classification of Digital Data,
Introduction to Big Data: Evolution of Big Data, definition of big data, Traditional Business Intelligence vs BigData, Coexistence of Big Data and Data Warehouse.
Big Data Analytics: introduction to Big Data Analytics, What Big Data Analytics Isn't, Sudden Hype Around Big Data Analytics, Classification of Analytics, Greatest Challenges that Prevent Business from Capitalizing Big Data, Top Challenges Facing Big Data, Big Data Analytics Importance, Data Science, Terminologies used in Big Data Environments.

UNIT-II:
Hadoop: Features of Hadoop, Key advantages of hadoop, versions of hadoop, overview of hadoop ecosystem, Hadoop distributions. Need of hadoop, RDBMS vs Hadoop, Distribution computing challenges, History of hadoop, Hadoop overview, HDFS.

UNIT-III:
Processing data with hadoop, introduction to mapreduce programming, mapper, reducer, combiner, partitioner
NoSQL: Types of NoSQL Databases, advantages of NoSQL, Use of NoSQL in industry, SQL vs
NoSQL, newSQL, comparison of Nosql, sql and newsql.

UNIT-IV:
MongoDB, necessity of mongodb, terms used in mongodb and RDBMS, datatypes in mongoDB, mongodb query language.

UNIT-V:
Introduction to R programming, operators, control statements and functions, interfacing with R, vectors, matrices, lists, data frames, factors and tables, accessing input and output, graphs in R, R apply family.

Text Books:
1. Big Data Analytics, Seema Acharya, Subhashini Chellappan, Wiley 2015.
2. R programming for beginners, sandhya arora, latesh malik, university press

Reference Books:
1. Chandramouli Subramanian, Asha A Geroge, C R Rene Robin, big data analytics, University press.
2. Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Business, Michael Minelli, Michehe Chambers, 1st Edition, Ambiga Dhiraj, Wiley CIO Series, 2013.
3. Hadoop: The Definitive Guide, Tom White, 3rd Edition, O'Reilly Media, 2012.
4. Big Data Analytics: Disruptive Technologies for Changing the Game, Arvind Sathi, 1st Edition, IBM Corporation, 2012.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. VI Semester
CS631PE – Full Stack Development
Professional Elective – III
(Common to CSE, IT, CSE(AI&ML) (PE-II), CSE (Data Science)

Pre-Requisites:
- Object Oriented Programming
- Web Technologies

Course Objectives:

- Students will become familiar to implement fast, efficient, interactive and scalable web applications using run time environment provided by the full stack components.

Course Outcomes:

- Understand Full stack components for developing web application.
- Apply packages of NodeJS to work with Data, Files, Http Requests and Responses.
- Use MongoDB data base for storing and processing huge data and connects with NodeJS application.
- Design faster and effective single page applications using Express and Angular.
- Create interactive user interfaces with react components.

UNIT-I
Introduction to Full Stack Development:
Understanding the Basic Web Development Framework- User, Browser, Webserver, Backend Services, Full Stack Components - Node.js, MongoDB, Express, React, Angular. Java Script Fundamentals, NodeJS- Understanding Node.js, Installing Node.js, Working with Node Packages, creating a Node.js Application, Understanding the Node.js Event Model, Adding Work to the Event Queue, Implementing Callbacks.

UNIT-II
Node.js:
Working with JSON, Using the Buffer Module to Buffer Data, Using the Stream Module to Stream Data, Accessing the File System from Node.js- Opening, Closing, Writing, Reading Files and other File System Tasks. Implementing HTTP Services in Node.js- Processing URLs, Processing Query Strings and Form Parameters, Understanding Request, Response, and Server Objects, Implementing HTTP Clients and Servers in Node.js, Implementing HTTPS Servers and Clients. Using Additional Node.js Modules-Using the os Module, Using the util Module, Using the dns Module, Using the crypto Module.

UNIT-III
MongoDB:
Need of NoSQL, Understanding MongoDB, MongoDB Data Types, Planning Your Data Model, Building the MongoDB Environment, Administering User Accounts, Configuring Access Control, Administering Databases, Managing Collections, Adding the MongoDB Driver to Node.js, Connecting to MongoDB from
Node.js, Understanding the Objects Used in the MongoDB Node.js Driver, Accessing and Manipulating Databases, Accessing and Manipulating Collections

UNIT-IV
Express and Angular:
Getting Started with Express, Configuring Routes, Using Requests Objects, Using Response Objects. Angular: importance of Angular, Understanding Angular, creating a Basic Angular Application, Angular Components, Expressions, Data Binding, Built-in Directives, Custom Directives, Implementing Angular Services in Web Applications.

UNIT-V
React:
Need of React, Simple React Structure, The Virtual DOM, React Components, Introducing React Components, Creating Components in React, Data and Data Flow in React, Rendering and Life Cycle Methods in React, Working with forms in React, integrating third party libraries, Routing in React.

TEXT BOOKS:
1. Brad Dayley, Brendan Dayley, Caleb Dayley., Node.js, MongoDB and Angular Web Development, 2nd Edition, Addison-Wesley, 2019.
2. Mark Tielens Thomas, React in Action, 1st Edition, Manning Publications.

REFERENCE BOOKS:
1. Vasan Subramanian, Pro MERN Stack, Full Stack Web App Development with Mongo, Express, React, and Node, 2nd Edition, Apress, 2019.
2. Chris Northwood, The Full Stack Developer: Your Essential Guide to the Everyday Skills Expected of a Modern Full Stack Web Developer', 1st edition, Apress, 2018.
3. Kirupa Chinnathambi, Learning React: A Hands-On Guide to Building Web Applications Using React and Redux, 2nd edition, Addison-Wesley Professional, 2018.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. VI Semester
CS633PE – Mobile Application Development
Professional Elective –III
(Common to CSE, IT, CSE (Data Science))

Course Objectives:

- To demonstrate their understanding of the fundamentals of Android operating systems.
- To improves their skills of using Android software development tools.
- To demonstrate their ability to develop software with reasonable complexity on mobile platform.
- To demonstrate their ability to deploy software to mobile devices.
- To demonstrate their ability to debug programs running on mobile devices

Course Outcomes:

- Student understands the working of Android OS Practically.
- Student will be able to develop Android user interfaces
- Student will be able to develop, deploy and maintain the Android Applications

UNIT – I:
Introduction to Android Operating System: Android OS design and Features – Android development framework, SDK features, Installing and running applications on Android Studio, Creating AVDs, Types of Android applications, Best practices in Android programming, Android tools Android application components – Android Manifest file, Externalizing resources like values, themes, layouts, Menus etc, Resources for different devices and languages, Runtime Configuration Changes Android Application Lifecycle – Activities, Activity lifecycle, activity states, monitoring state changes

UNIT – II:
Android User Interface: Measurements – Device and pixel density independent measuring units Layouts – Linear, Relative, Grid and Table Layouts
User Interface (UI) Components – Editable and non-editable TextViews, Buttons, Radio and Toggle Buttons, Checkboxes, Spinners, Dialog and pickers
Event Handling – Handling clicks or changes of various UI components
Fragments – Creating fragments, Lifecycle of fragments, Fragment states, Adding fragments to Activity, adding, removing and replacing fragments with fragment transactions, interfacing between fragments and Activities, Multi-screen Activities

UNIT – III:
Intents and Broadcasts: Intent – Using intents to launch Activities, Explicitly starting new Activity, Implicit Intents, Passing data to Intents, Getting results from Activities, Native Actions, using Intent to dial a number or to send SMS
Broadcast Receivers – Using Intent filters to service implicit Intents, Resolving Intent filters, finding and using Intents received within an Activity
Notifications – Creating and Displaying notifications, Displaying Toasts

UNIT – IV:
Persistent Storage: Files – Using application specific folders and files, creating files, reading data from files, listing contents of a directory Shared Preferences – Creating shared preferences, saving and retrieving data using Shared Preference

UNIT – V:
Database – Introduction to SQLite database, creating and opening a database, creating tables, inserting retrieving and etindelg data, Registering Content Providers, Using content Providers (insert, delete, retrieve and update)

TEXT BOOKS:

1.Professional Android 4 Application Development, Reto Meier, Wiley India, (Wrox), 2012.
2.Android Application Development for Java Programmers, James C Sheusi, Cengage Learning, 2013.

REFERENCE BOOK:

1.   Beginning Android 4 Application Development, Wei-Meng Lee, Wiley India (Wrox), 2013.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. VI Semester

CD631PE: CRYPTOGRAPHY AND NETWORK SECURITY (PE-III)

Course Objectives

- Explain the objectives of information security.
- Explain the importance and application of each of confidentiality, integrity, authentication and availability.
- Understand various cryptographic algorithms.
- Understand the basic categories of threats to computers and networks.
- Describe public-key cryptosystem.
- Describe the enhancements made to IPv4 by IPSec.
- Understand Intrusions and intrusion detection.
- Discuss the fundamental ideas of public-key cryptography.
- Generate and distribute a PGP key pair and use the PGP package to send an encrypted e mail message.
- Discuss Web security and Firewalls.

Course Outcomes

- Student will be able to understand basic cryptographic algorithms, message and web authentication and security issues.
- Ability to identify information system requirements for both of them such as client and server.
- Ability to understand the current legal issues towards information security.

UNIT-I:
Security Concepts: Introduction, The need for security, Security approaches, Principles of security, Types of Security attacks, Security services, Security Mechanisms, A model for Network Security.
Cryptography Concepts and Techniques: Introduction, plain text and cipher text, substitution techniques, transposition techniques, encryption and decryption, symmetric and asymmetric key cryptography, steganography, key range and key size, possible types of attacks.

UNIT-II:
Symmetric key Ciphers: Block Cipher principles, DES, AES, Blowfish, RC5, IDEA, Block cipher operation, Stream ciphers, RC4.
Asymmetric key Ciphers: Principles of public key cryptosystems, RSA algorithm, Elgamal Cryptography, Diffie-Hellman Key Exchange, Knapsack Algorithm.

UNIT-III:
Cryptographic Hash Functions: Message Authentication, Secure Hash Algorithm (SHA-512).
Message authentication codes: Authentication requirements, HMAC, CMAC, Digital signatures, Elgamal Digital Signature Scheme.
Key Management and Distribution: Symmetric Key Distribution Using Symmetric & Asymmetric Encryption, Distribution of Public Keys, Kerberos, X.509 Authentication Service, Public – Key Infrastructure.

UNIT-IV:
Transport-level Security: Web security considerations, Secure Socket Layer and Transport Layer Security, HTTPS, Secure Shell (SSH).
Wireless Network Security: Wireless Security, Mobile Device Security, IEEE 802.11 Wireless LAN, IEEE 802.11i Wireless LAN Security.

UNIT-V:
E-Mail Security: Pretty Good Privacy, S/MIME IP Security: IP Security overview, IP Security architecture, Authentication Header, Encapsulating security payload, Combining security associations, Internet Key Exchange.
Case Studies on Cryptography and security: Secure Multiparty Calculation, Virtual Elections, Single sign On, Secure Inter-branch Payment Transactions, Cross site Scripting Vulnerability.

Text Books:
1. Cryptography and Network Security - Principles and Practice: William Stallings, Pearson Education, 6th Edition.
2. Cryptography and Network Security: Atul Kahate, Mc Graw Hill, 3rd Edition.

Reference Books:
1. Cryptography and Network Security: C K Shyamala, N Harini, Dr T R Padmanabhan, Wiley India, 1st Edition.
2. Cryptography and Network Security: Forouzan Mukhopadhyay, Mc Graw Hill, 3rd Edition.
3. Information Security, Principles, and Practice: Mark Stamp, Wiley India.
4. Principles of Computer Security: WM. Arthur Conklin, Greg White, TMH.
5. Introduction to Network Security: Neal Krawetz, CENGAGE Learning.
6. Network Security and Cryptography: Bernard Menezes, CENGAGE Learning.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. VI Semester
CD632PE: DATA VISUALIZATION TECHNIQUES (PE-III)

Course Objectives

- To understand various data visualization techniques.

Course Outcomes

- Know the historical development and evolution of data visualization techniques.
- Analyze and visualize high-dimensional datasets using appropriate techniques.
- Visualize large multidimensional datasets using appropriate methods.
- Create insightful visual representations for diverse types of data.

UNIT-I:
Introduction, A Brief History of Data Visualization, Good Graphics, Static Graphics.

UNIT-II:
Data Visualization Through Their Graph Representations, Graph-theoretic Graphics, High-dimensional Data Visualization, Multivariate Data Glyphs: Principles and Practice, Linked Views for Visual Exploration, Linked Data Views, Visualizing Trees and Forests.

UNIT-III:
Multidimensional Scaling, Huge Multidimensional Data Visualization, Multivariate Visualization by Density Estimation, Structured Sets of Graphs, Structural Adaptive Smoothing by Propagation– Separation Methods, Smoothing Techniques for Visualization.

UNIT-IV:
Data Visualization via Kernel Machines, Visualizing Cluster Analysis and Finite Mixture Models, Visualizing Contingency Tables, Mosaic Plots and their Variants.

UNIT-V:
Parallel Coordinates: Visualization, Exploration and Classification of High- Dimensional Data, Matrix Visualization, Visualization in Bayesian Data Analysis.

TEXT BOOKS:

1. Handbook of Data Visualization by Chun-houh Chen, 2008.
2. Matthew Ward, Georges Grinstein and Daniel Keim, "Interactive Data Visualization Foundations, Techniques, Applications", 2010.
3. Colin Ware, "Information Visualization Perception for Design", 2nd edition, Margon Kaufmann Publishers, 2004

REFERENCE BOOKS:

1. Robert Spence "Information visualization – Design for interaction", Pearson Education, 2nd Edition, 2007.
2. Alexandru C. Telea, "Data Visualization: Principles and Practice," A. K. Peters Ltd, 2008.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. VI Semester
CM621OE: ADVANCED DATA STRUCTURES
(Open Elective – I)

Prerequisites: Programming for Problem Solving

Course Objectives

- Exploring basic data structures such as stacks and queues.
- Introduces a variety of data structures such as hash tables, search trees, tries, heaps, graphs.
- Introduces sorting and pattern matching algorithms.

Course Outcomes

- Ability to select the data structures that efficiently model the information in a problem.
- Ability to assess efficiency trade-offs among different data structure implementations or   combinations.
- Implement and know the application of algorithms for sorting and pattern matching.
- Design programs using a variety of data structures, including hash tables, binary and general tree structures, search trees, tries, heaps, graphs, and AVL-trees.

UNIT - I

Introduction to Data Structures, abstract data types, Linear list – singly linked list implementation, insertion, deletion and searching operations on linear list, Double Linked list implementation, insertion, deletion and searching operations, Circular Linked List implementation, insertion, deletion and searching operations.

UNIT - II

Stacks- Operations, array and linked representations of stacks, stack applications, Queues- operations, array and linked representations.
Hash Table Representation: hash functions, collision resolution-separate chaining, open addressing- linear probing, quadratic probing, double hashing, rehashing.

UNIT - III

Search Trees: Binary Search Trees, Definition, Implementation, Operations- Searching, Insertion and Deletion, B- Trees, AVL Trees, Definition, Height of an AVL Tree, Operations – Insertion, Deletion and Searching.

UNIT - IV

Graphs: Graph Implementation Methods. Graph Traversal Methods.Sorting: Quick Sort, Heap Sort, Merge Sort.

UNIT - V

Pattern Matching and Tries: Pattern matching algorithms-Brute force, the Boyer –Moore algorithm, the Knuth-Morris-Pratt algorithm, Standard Tries, Compressed Tries, Suffix tries.

TEXT BOOKS:

1. Fundamentals of Data Structures in C, 2 nd Edition, E. Horowitz, S. Sahni and Susan Anderson Freed, Universities Press.
2. Data Structures using C – A. S.Tanenbaum, Y. Langsam, and M.J. Augenstein, PHI/PearsonEducation.

REFERENCE BOOK:

1. Data Structures: A Pseudocode Approach with C, 2 nd Edition, R. F. Gilberg and B.A.Forouzan, Cengage Learning.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. VI Semester

CD621OE: DATABASE MANAGEMENT SYSTEMS (OE-I)

Prerequisites

- A course on "Data Structures".

Course Objectives

- To understand the basic concepts and the applications of database systems.
- To master the basics of SQL and construct queries using SQL.
- Topics include data models, database design, relational model, relational algebra, transaction control, concurrency control, storage structures and access techniques.

Course Outcomes

- Gain knowledge of fundamentals of DBMS, database design and normal forms
- Master the basics of SQL for retrieval and management of data.
- Be acquainted with the basics of transaction processing and concurrency control.
- Familiarity with database storage structures and access techniques

UNIT-I:
Database System Applications: A Historical Perspective, File Systems versus a DBMS, the Data Model, Levels of Abstraction in a DBMS, Data Independence, Structure of a DBMS.
Introduction to Database Design: Database Design and ER Diagrams, Entities, Attributes, and Entity Sets, Relationships and Relationship Sets, Additional Features of the ER Model, Conceptual Design With the ER Model

UNIT-II:
Introduction to the Relational Model: Integrity constraint over relations, enforcing integrity constraints, querying relational data, logical data base design, introduction to views, destroying/altering tables and views.Relational Algebra,

UNIT-III:
SQL: QUERIES, CONSTRAINTS, TRIGGERS: form of basic SQL query, UNION, INTERSECT, and EXCEPT, Nested Queries, aggregation operators, NULL values, complex integrity constraints in SQL, triggers and active databases.
Schema Refinement: Problems caused by redundancy, decompositions, problems related to decomposition, reasoning about functional dependencies, FIRST, SECOND, THIRD normal forms, BCNF, lossless join decomposition, multi-valued dependencies, FOURTH normal form, FIFTH normal form.

UNIT-IV:
Transaction Concept, Transaction State, Implementation of Atomicity and Durability, Concurrent Executions, Serializability, Recoverability, Implementation of Isolation, Testing for serializability, Lock Based Protocols, Timestamp Based Protocols, Validation- Based Protocols.

UNIT-V:
Data on External Storage, File Organization and Indexing, Cluster Indexes, Primary and Secondary Indexes, Index data Structures, Hash Based Indexing, Tree base Indexing, Comparison of File Organizations, Indexes-Intuitions for tree Indexes, Indexed Sequential Access Methods (ISAM), B+ Trees: A Dynamic Index Structure.

TEXT BOOKS:

1. Database Management Systems, Raghurama Krishnan, Johannes Gehrke, Tata Mc Graw Hill 3rd Edition
2. Database System Concepts, Silberschatz, Korth, Mc Graw hill, V edition

REFERENCE BOOKS:

1. Database Systems design, Implementation, and Management, Peter Rob & Carlos Coronel 7th Edition.
2. Fundamentals of Database Systems, Elmasri Navrate, Pearson Education
3. Introduction to Database Systems, C. J. Date, Pearson Education
4. Oracle for Professionals, The X Team, S.Shah and V. Shah, SPD.
5. Database Systems Using Oracle: A Simplified guide to SQL and PL/SQL,Shah, PHI.
6. Fundamentals of Database Management Systems, M. L. Gillenson, Wiley Student Edition.

L	T	P	C
0	0	2	1

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. VI Semester
CS651PC – Machine Learning Lab
(Common to CSE, CSE (Data Science))

Pre-Requisites: Python programming

Course Objectives:

- To get an overview of the various machine learning techniques and can able to demonstrate them using python.

Course Outcomes:

After completion of course, students would be able to:
- Understand and apply data preprocessing techniques
- Understand complexity of Machine Learning algorithms and their limitations.
- Understand modern notions in data analysis-oriented computing.
- Be capable of confidently applying common Machine Learning algorithms in practice and implementing their own.
- Be capable of performing experiments in Machine Learning using real-world data.

1. Write a program to demonstrate the following in python Numpy: Indexing, Array, Slicing.
2. Write a program to demonstrate the following in python Pandas: Series, DataFrames, read .csv files, analysing data.
3. Write a program to demonstrate the following in python Matplotlib: pyplot, markers, line, labels, grid, subplots.
4. Write a program to demonstrate the working of the decision tree based ID3 algorithm. Use an appropriate data set for building the decision tree and apply this knowledge to classify a new sample.
5. Implement Random Forest classification in python.
6. Implement k-nearest neighbour classification using python.
7. Implement linear regression using python.
8. Implement Naïve Bayes theorem to classify English text
9. Implement an algorithm to demonstrate the significance of genetic algorithm.
10. Implement the finite words classification system using Back-propagation algorithm.

TEXT BOOKS:
1. Python Programming: A Modern Approach, Vamsi Kurama, Pearson.
2. Python Machine Learning, Sebastian Raschka, PACKT publishing, 2015

REFERENCES:
1. Machine Learning – Tom M. Mitchell, - MGH, 1997
2. Pattern recognition and Machine Learning, C. M. Bishop, Springer, 2007.
3. Introduction to Machine Learning, E. Alpaydin, MIT Press, 3rd Edition, 2014.

L	T	P	C
0	0	2	1

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. VI Semester
CD651PC: BIG DATA ANALYTICS LAB

Course Objectives

- Provide knowledge of Big data Analytics principles and techniques.
- Designed to give an exposure of the frontiers of Big data Analytics.

Course Outcomes

- Use Excel as an Analytical tool and visualization tool.
- Ability to program using HADOOP and Map reduce
- Ability to perform data analytics using ML in R.
- Use MongoDB to perform data analytics.

LIST OF EXPERIMENTS:
1. Create a Hadoop cluster
2. Implement a simple map-reduce job that builds an inverted index on the set of input documents (Hadoop)
3. Process big data in HBase
4. Store and retrieve data in Pig
5. Perform data analysis using MongoDB
6. Using Power Pivot (Excel) Perform the following on any dataset
   a) Big Data Analytics
   b) Big Data Charting

Text Books:
1. Big Data Analytics, Seema Acharya, Subhashini Chellappan, Wiley 2015.
2. Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Business, Michael Minelli, Michehe Chambers, 1st Edition, Ambiga Dhiraj, Wiley CIO Series, 2013.
3. Hadoop: The Definitive Guide, Tom White, 3rd Edition, O"Reilly Media, 2012.
4. Big Data Analytics: Disruptive Technologies for Changing the Game, Arvind Sathi, 1st Edition, IBM Corporation, 2012

Reference Books:
1. Big Data and Business Analytics, Jay Liebowitz, Auerbach Publications, CRC press (2013)
2. Using R to Unlock the Value of Big Data: Big Data Analytics with Oracle R Enterprise and Oracle R Connector for Hadoop, Tom Plunkett, Mark Hornick, McGraw-Hill/Osborne Media (2013), Oracle press.
3. Professional Hadoop Solutions, Boris lublinsky, Kevin t. Smith, Alexey Yakubovich, Wiley, ISBN: 9788126551071, 2015.
4. Understanding Big data, Chris Eaton, Dirk deroos et al., McGraw Hill, 2012.
5. Intelligent Data Analysis, Michael Berthold, David J. Hand, Springer, 2007.
6. Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics, Bill Franks, 1st Edition, Wiley and SAS Business Series, 2012.

L	T	P	C
0	0	2	1

B.Tech. in Computer Science & Engineering (Data Science)
B. Tech. VI Semester
CS611PE–Full Stack Development Lab
Professional Elective – III Lab
(Common to CSE, IT, CSE (Data Science))

Pre-Requisites:

1. Object Oriented Programming
2. Web Technologies

Course Objectives:

- Introduce fast, efficient, interactive and scalable web applications using run time environment provided by the full stack components.

Course Outcomes:

- Design flexible and responsive Web applications using Node JS, React, Express and Angular.
- Perform CRUD operations with MongoDB on huge amount of data.
- Develop real time applications using react components.
- Use various full stack modules to handle http requests and responses.

List of Experiments

1. Create an application to setup node JS environment and display "Hello World".
2. Create a Node JS application for user login system.
3. Write a Node JS program to perform read, write and other operations on a file.
4. Write a Node JS program to read form data from query string and generate response using NodeJS
5. Create a food delivery website where users can order food from a particular restaurant listed inthe website for handling http requests and responses using NodeJS.
6. Implement a program with basic commands on databases and collections using MongoDB.
7. Implement CRUD operations on the given dataset using MongoDB.
8. Perform Count, Limit, Sort, and Skip operations on the given collections using MongoDB.
9. Develop an angular JS form to apply CSS and Events.
10. Develop a Job Registration form and validate it using angular JS.
11. Write an angular JS application to access JSON file data of an employee from a server using $http service.
12. Develop a web application to manage student information using Express and Angular JS.
13. Write a program to create a simple calculator Application using React JS.
14. Write a program to create a voting application using React JS
15. Develop a leave management system for an organization where users can apply different types of leaves such as casual leave and medical leave. They also can view the available number of days using react application

.

16. Build a music store application using react components and provide routing among the web pages.
17. Create a react application for an online store which consist of registration, login, product information pages and implement routing to navigate through these pages.

TEXT BOOKS:

1. Brad Dayley, Brendan Dayley, Caleb Dayley., Node.js, MongoDB and Angular Web Development, 2nd Edition, Addison-Wesley,2019.
2. Mark Tielens Thomas., React in Action, 1st Edition, Manning Publications.

REFERENCE BOOKS:

1. Vasan Subramanian, Pro MERN Stack, Full Stack Web App Development with Mongo, Express, React, and Node, 2nd Edition, Apress,2019.
2. Chris Northwood, The Full Stack Developer: Your Essential Guide to the Everyday Skills Expected of a Modern Full Stack Web Developer', 1st edition, Apress, 2018.
3. Brad Green& Seshadri. Angular JS. 1st Edition. O'Reilly Media, 2013.
4. Kirupa Chinnathambi, Learning React: A Hands-On Guide to Building Web Applications Using React and Redux, 2nd edition, Addison-Wesley Professional, 2018.

L	T	P	C
0	0	2	1

B.Tech. in Computer Science & Engineering (Data Science)
VI Semester Syllabus
CS613PE–Mobile Application Development Lab
Professional Elective – III Lab
(Common to CSE, IT, CSE (Data Science))

Course Objectives:

- To learn how to develop Applications in android environment.
- To learn how to develop user interface applications.
- To learn how to develop URL related applications.

Course Outcomes:

- Student understands the working of Android OS Practically.
- Student will be able to develop Android user interfaces
- Student will be able to develop, deploy and maintain the Android Applications.

List of Experiments

1.   (a) Create an Android application that shows Hello + name of the user and run it on an emulator.

i.   Create an application that takes the name from a text box and shows hello message along with the name entered in text box, when the user clicks the OK button.

2.   Create a screen that has input boxes for User Name, Password, Address, Gender (radio buttons for male and female), Age (numeric), Date of Birth (Date Picket), State (Spinner) and a Submit button. On clicking the submit button, print all the data below the Submit Button. Use (a) Linear Layout (b) Relative Layout and (c) Grid Layout or Table Layout.

3.   Develop an application that shows names as a list and on selecting a name it should show the details of the candidate on the next screen with a "Back" button.
If the screen is rotated to landscape mode (width greater than height), then the screen should show list on left fragment and details on right fragment instead of second screen with back button. Use Fragment transactions and Rotation event listener.

4.   Develop an application that uses a menu with 3 options for dialing a number, opening a website and to send an SMS. On selecting an option, the appropriate action should be invoked using intents.

5.   Develop an application that inserts some notifications into Notification area and whenever a notification is inserted, it should show a toast with details of the notification.

6.   Create an application that uses a text file to store user names and passwords (tab separated fields and one record per line). When the user submits a login name and password through a screen, the details should be verified with the text file data and if they match, show a dialog saying that login is successful. Otherwise, show the dialog with Login Failed message.

7.   Create a user registration application that stores the user details in a database table.

8.   Create a database and a user table where the details of login names and passwords are stored. Insert some names and passwords initially. Now the login details entered by the user should be verified with the database and an appropriate dialog should be shown to the user

9.   Create an admin application for the user table, which shows all records as a list and the admin can select any record for edit or modify. The results should be reflected in the table

10. Develop an application that shows all contacts of the phone along with details like name, phone number, mobile number etc.

11. Create an application that saves user information like name, age, gender etc. in shared preference and retrieves them when the program restarts.

12. Create an alarm that rings every Sunday at 8:00 AM. Modify it to use a time picker to set alarm time.

13. Create an application that shows the given URL (from a text field) in a browser.

TEXT BOOKS:

1. Professional Android 4 Application Development, Reto Meier, Wiley India, (Wrox), 2012
2. Android Application Development for Java Programmers, James C Sheusi, Cengage Learning, 2013

REFERENCE BOOK:
1. Beginning Android 4 Application Development, Wei-Meng Lee, Wiley India (Wrox), 2013.

L	T	P	C
0	0	2	1

B.Tech. in Computer Science & Engineering (Data Science)
VI Semester Syllabus
CD611PE: CRYPTOGRAPHY AND NETWORK SECURITY LAB (PE-III)

Course Objectives

- Explain the objectives of information security
- Explain the importance and application of each of confidentiality, integrity, authentication and availability
- Understand various cryptographic algorithms.

Course Outcomes

- Understand basic cryptographic algorithms, message and web authentication and security issues.
- Identify information system requirements for both of them such as client and server.
- Understand the current legal issues towards information security.

LIST OF EXPERIMENTS:

1. Write a C program that contains a string (char pointer) with a value 'Hello world'. The program should XOR each character in this string with 0 and displays the result.
2. Write a C program that contains a string (char pointer) with a value 'Hello world'. The program should AND or and XOR each character in this string with 127 and display the result.
3. Write a Java program to perform encryption and decryption using the following algorithms a. Ceaser cipher b. Substitution cipher c. Hill Cipher
4. Write a C/JAVA program to implement the DES algorithm logic.
5. Write a C/JAVA program to implement the Blowfish algorithm logic.
6. Write a C/JAVA program to implement the Rijndael algorithm logic.
7. Write the RC4 logic in Java Using Java cryptography; encrypt the text "Hello world" using Blowfish. Create your own key using Java key tool.
8. Write a Java program to implement RSA algorithm.
9. Implement the Diffie-Hellman Key Exchange mechanism using HTML and JavaScript.
10. Calculate the message digest of a text using the SHA-1 algorithm in JAVA.
11. Calculate the message digest of a text using the MD5 algorithm in JAVA.

TEXT BOOKS:

1. Cryptography and Network Security - Principles and Practice: William Stallings, Pearson Education, 6th Edition.
2. Cryptography and Network Security: Atul Kahate, Mc Graw Hill, 3rd Edition

REFERENCE BOOKS:

1. Cryptography and Network Security: C K Shyamala, N Harini, Dr T R Padmanabhan, Wiley India, 1st Edition.
2. Cryptography and Network Security: Forouzan Mukhopadhyay, Mc Graw Hill, 3rd Edition.
3. Information Security, Principles, and Practice: Mark Stamp, Wiley India.
4. Principles of Computer Security: WM. Arthur Conklin, Greg White, TMH.
5. Introduction to Network Security: Neal Krawetz, CENGAGE Learning.
6. Network Security and Cryptography: Bernard Menezes, CENGAGE Learning

L	T	P	C
0	0	2	1

B.Tech. in Computer Science & Engineering (Data Science)
VI Semester Syllabus
CD612PE: DATA VISUALIZATION TECHNIQUES LAB(PE-III)

Course Objectives

- Understand the various types of data, apply and evaluate the principles of data visualization.
- Acquire skills to apply visualization techniques to a problem and its associated dataset.

Course Outcomes

- Identify the different data types, visualization types to bring out the insight.
- Relate the visualization towards the problem based on the dataset to analyze and bring out valuable insight on a large dataset.
- Demonstrate the analysis of a large dataset using various visualization techniques and tools.
- Identify the different attributes and showcasing them in plots. Identify and create various visualizations for geospatial and table data.
- Ability to create and interpret plots using R/Python

LIST OF EXPERIMENTS:
1. Acquiring and plotting data.
2. Statistical Analysis – such as Multivariate Analysis, PCA, LDA, Correlation regression and analysis of variance
3. Financial analysis using Clustering, Histogram and HeatMap
4. Time-series analysis – stock market
5. Visualization of various massive dataset - Finance - Healthcare - Census - Geospatial
6. Visualization on Streaming dataset (Stock market dataset, weather forecasting)
7. Market-Basket Data analysis-visualization
8. Text visualization using web analytics.

TEXT BOOKS:
1. Matthew Ward, Georges Grinstein and Daniel Keim, "Interactive Data Visualization Foundations, Techniques, Applications", 2010.
2. Colin Ware, "Information Visualization Perception for Design", 2nd edition, Margon Kaufmann Publishers, 2004

REFERENCE BOOKS:
1. Robert Spence "Information visualization – Design for interaction", Pearson Education, 2 nd Edition, 2007.
2. Alexandru C. Telea, "Data Visualization: Principles and Practice," A. K. Peters Ltd, 2008.

L	T	P	C
0	0	4	2

B.Tech. in Computer Science & Engineering (Data Science)
VI Semester Syllabus
CD652PC: Industrial Oriented Mini Project/ Summer Internship/ Skill Development Course (UI design- Flutter)

Course Objectives

- Learns to Implement Flutter Widgets and Layouts
- Understands Responsive UI Design and with Navigation in Flutter
- Knowledge on Widges and customize widgets for specific UI elements, Themes
- Understand to include animation apart from fetching data.

Course Outcomes

- Implements Flutter Widgets and Layouts
- Responsive UI Design and with Navigation in Flutter
- Create custom widgets for specific UI elements and also Apply styling using themes and custom styles.
- Design a form with various input fields, along with validation and error handling
- Fetches data and write code for unit Test for UI components and also animation.

List of Experiments: Students need to implement the following experiments
1. a) Install Flutter and Dart SDK.
    b) Write a simple Dart program to understand the language basics.
2. a) Explore various Flutter widgets (Text, Image, Container, etc.).
    b) Implement different layout structures using Row, Column, and Stack widgets.
3. a) Design a responsive UI that adapts to different screen sizes.
    b) Implement media queries and breakpoints for responsiveness.
4. a) Set up navigation between different screens using Navigator.
    b) Implement navigation with named routes.
5. a) Learn about stateful and stateless widgets.
    b) Implement state management using set State and Provider.
6. a) Create custom widgets for specific UI elements.
    b) Apply styling using themes and custom styles.
7. a) Design a form with various input fields.
    b) Implement form validation and error handling.
8. a) Add animations to UI elements using Flutter's animation framework.
    b) Experiment with different types of animations (fade, slide, etc.).
9. a) Fetch data from a REST API.
    b) Display the fetched data in a meaningful way in the UI.
10. a) Write unit tests for UI components.
    b) Use Flutter's debugging tools to identify and fix issues.

TEXT BOOKS:

1. Marco L. Napoli, Beginning Flutter: A Hands-on Guide to App Development.

L	T	P	C
3	0	0	0

B.Tech. in Computer Science & Engineering (Data Science)
VI Semester Syllabus
MC602ES – Cyber Security
(Common to all branches except CSE, IT, CSBS)

Course objectives:

- To familiarize various types of cyber-attacks and cyber-crimes
- To give an overview of the cyber laws
- To study the defensive techniques against these attacks

Course Outcomes:

- The students will be able to understand cyber-attacks, types of cybercrimes, cyber laws and how to protect them self and ultimately the entire Internet community from such attacks.

UNIT - I

Introduction to Cyber Security: Basic Cyber Security Concepts, layers of security, Vulnerability, threat, Harmful acts, Internet Governance – Challenges and Constraints, Computer Criminals, CIA Triad, Assets and Threat, motive of attackers, active attacks, passive attacks, Software attacks, hardware attacks, Spectrum of attacks, Taxonomy of various attacks, IP spoofing, Methods of defence, Security Models, risk management, Cyber Threats-Cyber Warfare, Cyber Crime, Cyber terrorism, Cyber Espionage, etc., Comprehensive Cyber Security Policy.

UNIT - II

Cyberspace and the Law & Cyber Forensics: Introduction, Cyber Security Regulations, Roles of International Law. The INDIAN Cyberspace, National Cyber Security Policy. Introduction, Historical background of Cyber forensics, Digital Forensics Science, The Need for Computer Forensics, Cyber Forensics and Digital evidence, Forensics Analysis of Email, Digital Forensics Lifecycle, Forensics Investigation, Challenges in Computer Forensics, Special Techniques for Forensics Auditing.

UNIT - III

Cybercrime: Mobile and Wireless Devices: Introduction, Proliferation of Mobile and Wireless Devices, Trends in Mobility, Credit card Frauds in Mobile and Wireless Computing Era, Security Challenges Posed by Mobile Devices, Registry Settings for Mobile Devices, Authentication service Security, Attacks on Mobile/Cell Phones, Mobile Devices: Security Implications for Organizations, Organizational Measures for Handling Mobile, Organizational Security Policies and Measures in Mobile Computing Era, Laptops.

UNIT- IV

Cyber Security: Organizational Implications: Introduction cost of cybercrimes and IPR issues, web Threats for organizations, security and privacy implications, social media marketing: security risks and perils for organizations, social computing, and the associated challenges for organizations.

Cybercrime and Cyber terrorism: Introduction, intellectual property in cyberspace, the ethical dimension of cybercrimes the psychology, mindset and skills of hackers and other cyber criminals.

UNIT - V

Privacy Issues: Basic Data Privacy Concepts: Fundamental Concepts, Data Privacy Attacks, Data linking and profiling, privacy policies and their specifications, privacy policy languages, privacy in different domains- medical, financial, etc.

Cybercrime: Examples and Mini-Cases Examples: Official Website of Maharashtra Government Hacked, Indian Banks Lose Millions of Rupees, Parliament Attack, Pune City Police Bust Nigerian Racket, e-mail spoofing instances.

Mini-Cases: The Indian Case of online Gambling, An Indian Case of Intellectual Property Crime, Financial Frauds in Cyber Domain.

TEXT BOOKS:

1. Nina Godbole and Sunit Belpure, Cyber Security Understanding Cyber Crimes, Computer Forensics and Legal Perspectives, Wiley

2. B. B. Gupta, D. P. Agrawal, Haoxiang Wang, Computer and Cyber Security: Principles, Algorithm, Applications, and Perspectives, CRC Press, ISBN 9780815371335, 2018.

REFERENCES:

1. Cyber Security Essentials, James Graham, Richard Howard and Ryan Otson, CRC Press.

2. Introduction to Cyber Security, Chwan-Hwa(john) Wu,J. David Irwin, CRC Press T&F Group.

L	T	P	C
3	0	0	0

B.Tech. in Computer Science & Engineering (Data Science)
VI Semester Syllabus
MC601BS: Environmental Science

Course Objectives:.

- To understand the importance of ecosystem and ecological balance.
- To understand the natural resources and their conservation.
- To understand the importance of biodiversity and its values.
- To gain knowledge about environmental pollution, effects and controlling measures.
- To study about global environmental problems and global issues.
- To understand the environmental policies, regulations, and sustainable development.

Course Outcomes :

After completing the course, the student will be able to:

- Get the information about ecosystem and their usage and conservation.
- Learn different types of natural resources and take up the measures to protect the resources.
- Get the information about biodiversity and their usage and conservation.
- Get the information about the types of pollution, understand their effects and controlling measures.
- Gain the knowledge about current global environmental issues and initiations to be taken to protect the environment.
- Gain the knowledge about environmental acts, EIA, sustainable development and follow the rules and regulations.

UNIT - I
Ecosystems: Definition, scope, and importance of ecosystem. Classification, structure, and function of an ecosystem, food chains, food webs, and ecological pyramids. Flow of energy, biogeochemical cycles, bioaccumulation, biomagnification, ecosystem value, services and carrying capacity. Field visits.

UNIT - II
Natural Resources: Classification of resources: Living and non-living resources. **Water resources:** Use and over utilization of surface and ground water, floods and droughts. Dams: Benefits and problems. **Mineral resources:** Use and exploitation, environmental effects of extracting and using mineral resources. **Land resources** and **Forest resources**. **Energy resources:** Growing energy needs, renewable and non-renewable energy resources, use of alternate energy resources. Case studies.

UNIT - III
Biodiversity and Biotic Resources: Introduction, definition, genetic, species and ecosystem diversity. Values of biodiversity: Consumptive use, productive use, social, ethical, aesthetic andoptional values. India as a mega diversity nation, hot spots of biodiversity. Field visit. Threats to biodiversity: Habitat loss, poaching of wildlife, man-wildlife conflicts. Conservation of biodiversity: In - Situ and Ex -situ conservation. National biodiversity act.

UNIT - IV

Environmental Pollution and Control Technologies:
Classification of pollution. **Air pollution:** Primary and secondary pollutants, automobile and industrial pollution. Ambient air quality standards. **Soil pollution:** Sources and types, impacts of modern agriculture, degradation of soil. **Solid waste:** Municipal solid waste management, composition and characteristics of e-waste and its management.
Pollution control technologies: Wastewater treatment methods: Primary, secondary and tertiary. Overview of air pollution control technologies, concepts of bioremediation.
Global Environmental issues and Global efforts: Climate change and impacts on human environment. Ozone depletion and Ozone depleting substances (ODS). Deforestation and desertification. International conventions /Protocols: Earth summit, Kyoto protocol and Montréal protocol. NAPCC-GoI Initiatives.

UNIT - V

Environmental Policy, Legislation & EIA: Environmental Protection Act, legal aspects Air Act- 1981, Water Act, biomedical waste management and handling rules, hazardous waste management and handling rules.
EIA: EIA structure, methods of baseline data acquisition. Concepts of environmental management plan
(EMP).
Towards Sustainable Future: Concept of sustainable development goals, population and its explosion. Crazy consumerism, Environmental education, Human health, Environmental ethics. Concept of green building, Principles of green chemistry, Ecological footprint, Life cycle assessment (LCA), Low carbon life style.

TEXT BOOKS:
1. Textbook of Environmental Studies for Undergraduate Courses by Erach Bharucha for University Grants Commission.
2. Environmental Studies by R. Rajagopalan, Oxford University Press.

REFERENCE BOOKS:
1. Environmental Science: Towards a sustainable future by Richard T. Wright. 2008 PHL Learning Private Ltd. New Delhi.
2. Environmental Engineering and science by Gilbert M. Masters and Wendell P. Ela. 2008, PHI Learning Pvt. Ltd.
3. Environmental Science by Daniel B. Botkin & Edward A. Keller, Wiley INDIA edition.
4. Environmental Studies by Anubha Kaushik, 4th Edition, New Age International publishers.
5. Text book of Environmental Science and Technology - Dr. M. Anji Reddy 2007, BS Publications.
6. Introduction to Environmental Science by Y. Anjaneyulu, BS Publications.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VII Semester Syllabus
CD701PC: PREDICTIVE ANALYTICS

Prerequisite

- Data Science/Data analytics.

Course Objectives

- To learn the basics and applications of predictive analytics using different techniques.

Course Outcomes

- Understand the processing steps for predictive analytics
- Construct and deploy prediction models with integrity
- Explore various techniques (machine learning/data mining, ensemble) for predictive analytics.
- Apply predictive analytics to real world examples.

UNIT-I:
Introduction – types of analytics, applications of predictive analytics, overview of predictive analytics. Setting up the problem - processing steps, business understanding, objectives, data for predictive modeling, columns as measures, target variables, measures of success for predictive models.

UNIT-II:
Prediction effect, deployment of prediction model, ethics and responsibilities The Data effect.

UNIT-III:
Machine Learning for prediction
Predictive modeling – decision trees, logistic regression, neural network, kNN, Bayesian method,
Regression model
Assessing Predictive models - Batch Approach to Model Assessment, Percent Correct Classification, Rank-Ordered Approach to Model Assessment, Assessing Regression Models.

UNIT-IV:
Ensemble effect
Model ensembles – motivation, wisdom of crowds, Bagging, Boosting, Random forests, stochastic gradient boosting, heterogeneous ensembles.

UNIT-V:
Case studies: Survey analysis, question answering– challenges in text mining, persuasion by the numbers.

Text Books:
1. Eric Siegel, Predictive analytics- the power to predict who will Click, buy, lie, or die, John Wiley & Sons, 2013.
2. Dean Abbott, Applied Predictive Analytics - Principles and Techniques for the Professional Data Analyst, 2014.

Reference Books:
1. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning-Data Mining, Inference, and Prediction, Second Edition, Springer Verlag, 2009.
2. G. James, D. Witten, T. Hastie, R. Tibshirani-An introduction to statistical learning with applications in R, Springer, 2013.
3. E. Alpaydin, Introduction to Machine Learning, Prentice Hall of India, 2010

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VII Semester Syllabus
CD702PC: WEB AND SOCIAL MEDIA ANALYTICS

Course Objectives

- Exposure to various web and social media analytic techniques.

Course Outcomes

- Knowledge on decision support systems.
- Apply natural language processing concepts on text analytics.
- Understand sentiment analysis.
- Knowledge on search engine optimization and web analytics.

UNIT-I:
An Overview of Business Intelligence, Analytics, and Decision Support: Analytics to Manage a Vaccine Supply Chain Effectively and Safely, Changing Business Environments and Computerized Decision Support, Information Systems Support for Decision Making, The Concept of Decision Support Systems (DSS), Business Analytics Overview, Brief Introduction to Big Data Analytics.

UNIT-II:
Text Analytics and Text Mining: Machine Versus Men on Jeopardy: The Story of Watson, Text Analytics and Text Mining Concepts and Definitions, Natural Language Processing, Text Mining Applications, Text Mining Process, Text Mining Tools.

UNIT-III:
Sentiment Analysis: Sentiment Analysis Overview, Sentiment Analysis Applications, Sentiment Analysis Process, Sentiment Analysis and Speech Analytics.

UNIT-IV:
Web Analytics, Web Mining: Security First Insurance Deepens Connection with Policyholders, Web Mining Overview, Web Content and Web Structure Mining, Search Engines, Search Engine Optimization, Web Usage Mining (Web Analytics), Web Analytics Maturity Model and Web Analytics Tools.

UNIT-V:
Social Analytics and Social Network Analysis: Social Analytics and Social Network Analysis, Social Media Definitions and Concepts, Social Media Analytics.
Prescriptive Analytics - Optimization and Multi-Criteria Systems: Multiple Goals, Sensitivity Analysis, What-If Analysis, and Goal Seeking.

Text Books:
1. Ramesh Sharda, Dursun Delen, Efraim Turban, BUSINESS INTELLIGENCE AND ANALYTICS: SYSTEMS FOR DECISION SUPPORT, Pearson Education

Reference Books:
1. Rajiv Sabherwal, Irma Becerra-Fernandez," Business Intelligence – Practice, Technologies and Management", John Wiley 2011.
2. Lariss T. Moss, ShakuAtre, "Business Intelligence Roadmap", Addison-Wesley It Service.
3. Yuli Vasiliev, "Oracle Business Intelligence: The Condensed Guide to Analysis and Reporting", SPD Shroff, 2012**.**

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VII Semester Syllabus
CD741PE: NATURAL LANGUAGE PROCESSING (PE-IV)
(Common to CSBS & CSE (Data Science))

Prerequisites

- • Data structures and compiler design.

Course Objectives

- Introduction to some of the problems and solutions of NLP and their relation to linguistics and statistics.

Course Outcomes

- Show sensitivity to linguistic phenomena and an ability to model them with formal grammars.
- Understand and carry out proper experimental methodology for training and evaluating empirical NLP systems
- Able to manipulate probabilities, construct statistical models over strings and trees, and estimate parameters using supervised and unsupervised training methods.
- Able to design, implement, and analyze NLP algorithms; and design different language modeling Techniques.

UNIT-I:
Finding the Structure of Words: Words and Their Components, Issues and Challenges, Morphological Models
Finding the Structure of Documents: Introduction, Methods, Complexity of the Approaches, Performances of the Approaches, Features.

UNIT-II:
Syntax I: Parsing Natural Language, Treebanks: A Data-Driven Approach to Syntax, Representation of Syntactic Structure, Parsing Algorithms.

UNIT-III:
Syntax II: Models for Ambiguity Resolution in Parsing, Multilingual Issues
Semantic Parsing I: Introduction, Semantic Interpretation, System Paradigms, Word Sense.

UNIT-IV:
Semantic Parsing II: Predicate-Argument Structure, Meaning Representation Systems.

UNIT-V:
Language Modeling: Introduction, N-Gram Models, Language Model Evaluation, Bayesian parameter estimation, Language Model Adaptation, Language Models- class based, variable length, Bayesian topic based, Multilingual and Cross Lingual Language Modeling.

Text Books:
1. Multilingual natural Language Processing Applications: From Theory to Practice –Daniel M. Bikel and Imed Zitouni, Pearson Publication

Reference Books:
1. Speech and Natural Language Processing - Daniel Jurafsky& James H Martin, Pearson Publications.
2. Natural Language Processing and Information Retrieval: Tanvier Siddiqui, U.S. Tiwary..

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VII Semester Syllabus
CD742PE: INTERNET OF THINGS (PE-IV)
(Common to CSE (AI&ML)(PE-III) & CSE (Data Science))

Prerequisites

- Computer organization, Computer Networks.

Course Objectives

- To introduce the terminology, technology and its applications
- To introduce the concept of M2M (machine to machine) with necessary protocols
- To introduce the Python Scripting Language which is used in many IoT devices
- To introduce the Raspberry PI platform, that is widely used in IoT applications
- To introduce the implementation of web-based services on IoT devices.

Course Outcomes

- Interpret the impact and challenges posed by IoT networks leading to new architectural models.
- Compare and contrast the deployment of smart objects and the technologies to connect them to network.
- Appraise the role of IoT protocols for efficient network communication.
- Identify the applications of IoT in Industry.

UNIT-I:
Introduction to Internet of Things –Definition and Characteristics of IoT, Physical Design of IoT,
Logical Design of IoT, IoT Enabling Technologies, IoT Levels and Deployment Templates
Domain Specific IoTs – Home automation, Environment, Agriculture, Health and Lifestyle.

UNIT-II:
IoT and M2M – M2M, Difference between IoT and M2M, SDN and NFV for IoT,
IoT System Management with NETCOZF, **YANG-** Need for IoT system Management, Simple Network management protocol, Network operator requirements, NETCONF, YANG, IoT Systems Management with NETCONF-YANG.

UNIT-III:
IoT Systems – Logical design using Python-Introduction to Python – Python Data types & Data structures, Control flow, Functions, Modules, Packaging, File handling, Data/Time operations, Classes, Exception, Python packages of Interest for IoT.

UNIT-IV:
IoT Physical Devices and Endpoints - Raspberry Pi, Linux on Raspberry Pi, Raspberry Pi Interfaces, Programming Raspberry PI with Python, Other IoT devices.
IoT Physical Servers and Cloud Offerings – Introduction to Cloud Storage models and communication APIs, WAMP-AutoBahn for IoT, Xively Cloud for IoT, Python web application framework –Django, Designing a RESTful web API.

UNIT-V:
Case studies- Home Automation, Environment-weather monitoring-weather reporting- air pollution monitoring, Agriculture.

TEXT BOOKS: 64

1. Internet of Things - A Hands-on Approach, Arshdeep Bahga and Vijay Madisetti, Universities Press, 2015, ISBN: 9788173719547

REFERENCE BOOKS:

1. Getting Started with Raspberry Pi, Matt Richardson & Shawn Wallace, O'Reilly (SPD), 2014, ISBN: 9789350239759.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VII Semester Syllabus
CD743PE: CLOUD COMPUTING (PE-IV)

Prerequisites

- A course on "Computer Networks".
- A course on "Operating System"

Course Objectives

- This course provides an insight into cloud computing.
- Topics covered include- Cloud Computing Architecture, Deployment Models, Service Models, Technological Drivers for Cloud Computing, Networking for Cloud Computing and Security in Cloud Computing.

Course Outcomes

- Understand different computing paradigms and potential of the paradigms and specifically cloud computing
- Understand cloud service types, cloud deployment models and technologies supporting and driving the cloud
- Acquire the knowledge of programming models for cloud and development of software application that runs the cloud and various services available from major cloud providers
- Understand the security concerns and issues in cloud computing.
- Acquire the knowledge of advances in cloud computing.

UNIT-I:

Computing Paradigms, Cloud Computing Fundamentals, Cloud Computing Architecture and Management.

UNIT-II:

Cloud Deployment Models, Cloud Service Models, Technological Drivers for Cloud Computing:

SOA and Cloud, Multicore Technology, Web 2.0 and Web 3.0, Pervasive Computing, Operating System, Application Environment.

UNIT-III:

Virtualization, Programming Models for Cloud Computing: MapReduce, Cloud Haskell, Software Development in Cloud.

UNIT-IV:

Networking for Cloud Computing: Introduction, Overview of Data Center Environment, Networking Issues in Data Centers, Transport Layer Issues in DCNs, Cloud Service Providers.

UNIT-V:

Security in Cloud Computing, and Advanced Concepts in Cloud Computing.

TEXT BOOKS:

1. Chandrasekaran, K. Essentials of cloud computing. CRC Press, 2014.

REFERENCE BOOKS:

1. Cloud Computing: Principles and Paradigms, Editors: Rajkumar Buyya, James Broberg, Andrzej M. Goscinski, Wiley, 2011
2. Enterprise Cloud Computing - Technology, Architecture, Applications, Gautam Shroff, Cambridge University Press, 2010
3. Cloud Computing Bible, Barrie Sosinsky, Wiley-India, 2010.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VII Semester Syllabus
CD744PE: INFORMATION STORAGE MANAGEMENT (PE-IV)

Course Objectives

- To understand various segments of storage technology and architectures
- To explore the inherent power of information
- To describe the different backup, recovery and replication strategies.

Course Outcomes

- Understand the evolution of storage technology and Intelligent Storage Systems
- Explore the key concepts of various Storage Networking Technologies - DAS, SANs, NAS and CAS
- Understand the basics of Storage Virtualization
- Understand the concepts of Storage security and Storage Infrastructure Management
- Analyze the purpose of backup, recovery and replication Strategies.

UNIT-I:
Introduction to Information Storage and Management: Information Storage, Evolution of Storage Technology and Architecture, Data Center Infrastructure, Key Challenges in Managing Information, Information Lifecycle. Storage System Environment - Data Protection: RAID - Intelligent Storage System.

UNIT-II:
Direct-Attached Storage and Introduction to SCSI
Types of DAS, DAS Benefits and Limitations, Disk Drive Interfaces, Introduction to Parallel SCSI, SCSI
Command Model,
Storage Area Networks
Fibre Channel: Overview, The SAN and Its Evolution, Components of SAN, FC Connectivity, Fibre Channel Ports, Fibre Channel Architecture, Zoning, Fibre Channel Login Types, FC Topologies,
Concepts in Practice: EMC Connectrix
Network-Attached Storage
General-Purpose Servers vs. NAS Devices, Benefits of NAS, NAS File I/O, Components of NAS, NAS Implementations, NAS File-Sharing Protocols, NAS I/O Operations, Factors Affecting NAS Performance and Availability, Concepts in Practice: EMC Celerra.

UNIT-III:
Content-Addressed Storage
Fixed Content and Archives, Types of Archives, Features and Benefits of CAS, CAS Architecture,
Object Storage and Retrieval in CAS, CAS Examples, Concepts in Practice: EMC Centera
Storage Virtualization
Forms of Virtualization, SNIA Storage Virtualization Taxonomy, Storage Virtualization Configurations,
Storage Virtualization Challenges, Types of Storage Virtualization, Concepts in Practice.

UNIT-IV:
Backup and Recovery
Backup Purpose, Backup Considerations, Backup Granularity, Recovery Considerations, Backup Methods, Backup Process, Backup and Restore Operations, Backup Topologies, Backup in NAS
Environments, Backup Technologies, Concepts in Practice: EMC NetWorker
Local Replication
Local Replication, Source and Target, Uses of Local Replicas, Data Consistency, Local Replication Technologies, Restore and Restart Considerations, Creating Multiple Replicas, Management Interface,
Concepts in Practice: EMC TimeFinder and EMC SnapView

Remote Replication

Modes of Remote Replication, Remote Replication Technologies, Network Infrastructure, Concepts in Practice: EMC SRDF, EMC SAN Copy, and EMC MirrorView.

UNIT-V:
Securing the Storage Infrastructure
Storage Security Framework, Risk Triad, Storage Security Domains, Security Implementations in
Storage Networking
Managing the Storage Infrastructure
Monitoring the Storage Infrastructure, Storage Management Activities, Storage Infrastructure
Management Challenges, Developing an Ideal Solution, Concepts in Practice: EMC ControlCenter.

TEXT BOOKS:

1. Marc Farley Osborne, "Building Storage Networks", Tata McGraw Hill, 2001.
2. Robert Spalding and Robert Spalding, "Storage Networks: The Complete Reference", Tata McGraw Hill, 2003.
3. Meeta Gupta, "Storage Area Network Fundamentals", Pearson Education Ltd., 2002

REFERENCE BOOKS:

1. Gerald J Kowalski and Mark T Maybury," Information Storage Retrieval Systems theory & Implementation", BS Publications, 2000.
2. Thejendra BS, "Disaster Recovery & Business continuity", Shroff Publishers & Distributors, 2006.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VII Semester Syllabus
CD751PE: PRIVACY PRESERVING IN DATA PUBLISHING (PE-V)

Prerequisites

- A course on "Data Mining".

Course Objectives

- The aim of the course is to introduce the fundamentals of Privacy Preserving Data Mining Methods
- The course gives an overview of - Anonymity and its Measures, Multiplicative Perturbation for Privacy-Preserving Data Mining, techniques for Utility-based Privacy Preserving Data.

Course Outcomes

- Understand the concepts of Privacy Preserving Data Mining Models and Algorithms.
- Demonstrate a comprehensive understanding of different tasks associated in Inference Control Methods for Privacy-Preserving Data Mining.
- Understand the concepts of Data Anonymization Methods and its Measures.
- Evaluate and appraise the solution designed for Multiplicative Perturbation.
- Formulate, Design and Implement the solutions for Utility-based Privacy Preserving Data

UNIT-I:
Introduction, Privacy-Preserving Data Mining Algorithms, The Randomization Method, Group Based Anonymization, Distributed Privacy-Preserving Data Mining.

UNIT-II:
Interface Control Methods Introduction, A Classification of Microdata Protection Methods, Perturbative Masking Methods, Nonperturbative Masking Methods, Synthetic Microdata Generation, Trading off Information Loss and Disclosure Risk.

UNIT-III:
Measure of Anonymity Data Anonymization Methods, A Classification of Methods, Statistical Measure of Anonymous, Probabilistic Measure of Anonymity, Computational Measure of Anonymity, reconstruction Methods for Randomization, Application of Randomization.

UNIT-IV:
Multiplicative Perturbation Definition of Multiplicative Perturbation, Transformation Invariant Data Mining Models, Privacy Evaluation for Multiplicative Perturbation, Attack Resilient Multiplicative Perturbation, Metrics for Quantifying Privacy Level, Metrics for Quantifying Hiding Failure, Metrics for Quantifying Data Quality.

UNIT-V:
Utility-Based Privacy-Preserving Data Types of Utility-Based Privacy Preserving Methods, Utility-Based Anonymization Using Local Recording, The Utility-Based Privacy Preserving Methods in Classification Problems, Anonymization Merginal: Injection Utility into Anonymization Data Sets.

Text Books:
1. Privacy – Preserving Data Mining: Models and Algorithms Edited by Charu C. Aggarwal and S. Yu, Springer

Reference Books:
1. Charu C. Agarwal, Data Mining: The Textbook, 1st Edition, Springer.
2. J. Han and M. Kamber, Data Mining: Concepts and Techniques, 3rd Edition, Elsevier.
3. Privacy Preserving Data Mining by Jaideep Vaidya, Yu Michael Zhu and Chirstopher W. Clifton, Springer.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VII Semester Syllabus
CD752PE: DATABASE SECURITY (PE-V)

Course Objectives

- To learn the security of databases
- To learn the design techniques of database security
- To learn the security software design.

Course Outcomes

- Identify database security problems
- Implement different security models
- Provide security for software design
- Protect object-oriented systems
- Handle security issues for active databases

UNIT-I:
Introduction: Introduction to Databases Security, Problems in Databases Security, Controls, Conclusions
Security Models -1: Introduction Access Matrix Model, Take-Grant Model, Acten Model, PN Model.

UNIT-II:
Security Models -2: Hartson and Hsiao's Model, Fernandez's Model, Bussolati and Martella's Model for Distributed databases, Bell and LaPadula's Model, Biba's Model, Dion's Model, Sea View Model, Jajodia and Sandhu's Model, The Lattice Model for the Flow Control.

UNIT-III:
Security Mechanisms: Introduction User Identification/Authentication, Memory Protection, Resource Protection, Control Flow Mechanisms, Isolation Security Functionalities in Some Operating Systems Trusted Computer System Evaluation Criteria.
Security Software Design: Introduction A Methodological Approach to Security Software Design
Secure Operating System Design Secure DBMS Design Security Packages Database Security Design

UNIT-IV:
Statistical Database Protection & Intrusion Detection Systems: Introduction Statistics Concepts and Definitions Types of Attacks Inference Controls Evaluation Criteria for Control Comparison. Introduction IDES System RETISS System ASES System Discovery.

UNIT-V:
Models For the Protection of New Generation Database Systems: Introduction A Model for the Protection of Frame Based Systems A Model for the Protection of Object-Oriented Systems SORION Model for the Protection of Object-Oriented Databases, The Orion Model, Jajodia and Kogan's Model- A Model for the Protection of Active Databases Conclusions.

Text Books:
1. Database Security by Castano, Pearson Edition(1/e)

Reference Books:
1. Database security by alfred basta, melissa zgola, CENGAGE learning
2. Database Security and Auditing: Protecting Data Integrity and Accessibility, 1st Edition, Hassan Afyouni, THOMSON Edition.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VII Semester Syllabus
CD753PE: DATA SCIENCE APPLICATIONS (PE – V)

COURSE OBJECTIVES

- To give deep knowledge of data science and how it can be applied in various fields to make life easy.

COURSE OUTCOMES

- To correlate data science and solutions to modern problems.
- To decide when to use which type of technique in data science.
- To understand applications of Data science in Education and social media..
- To explore data science in Healthcare and Bioinformatics.
- To demonstrate various case studies of data optimization using python.

UNIT-I:
Data Science Applications in various domains, Challenges and opportunities, tools for data scientists, Recommender systems – Introduction, methods, application, challenges.

UNIT-II:
Time series data – stock market index movement forecasting. Supply Chain Management – Real world case study in logistics.

UNIT-III:
Data Science in Education, Social media.

UNIT-IV:
Data Science in Healthcare, Bioinformatics.

UNIT-V:
Case studies in data optimization using Python

TEXT BOOKS:
1. Aakanksha Sharaff, G.K.Sinha , "Data Science and its applications ", CRC Press, 2021.
2. Q. A. Menon, S. A. Khoja, "Data Science: Theory, Analysis and Applications", CRC Press, 2020

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VII Semester Syllabus
CD754PE: MINING MASSIVE DATASETS (PE-V)

Prerequisites

- Students should be familiar with Data mining, algorithms, basic probability theory and Discrete math.

Course Objectives

- This course will cover practical algorithms for solving key problems in mining of massive datasets.
- This course focuses on parallel algorithmic techniques that are used for large datasets.
- This course will cover stream processing algorithms for data streams that arrive constantly, page ranking algorithms for web search, and online advertisement systems that are studied in detail..

Course Outcomes

- Handle massive data using MapReduce.
- Develop and implement algorithms for massive data sets and methodologies in the context of data mining.
- Understand the algorithms for extracting models and information from large datasets
- Develop recommendation systems.
- Gain experience in matching various algorithms for particular classes of problems

UNIT-I:

Data Mining-Introduction-Definition of Data Mining-Statistical Limits on Data Mining.

MapReduce and the New Software Stack-Distributed File Systems, MapReduce, Algorithms Using MapReduce.

UNIT-II:

Similarity Search: Finding Similar Items-Applications of Near-Neighbour Search, Shingling of Documents, Similarity-Preserving Summaries of Sets, Distance Measures.

Streaming Data: Mining Data Streams-The Stream Data Model, Sampling Data in a Stream, Filtering Streams.

UNIT-III:

Link Analysis-PageRank, Efficient Computation of PageRank, Link Spam.

Frequent Item sets - Handling Larger Datasets in Main Memory, Limited-Pass Algorithms, Counting Frequent Items in a Stream.

 Clustering-The CURE Algorithm, Clustering in Non-Euclidean Spaces, Clustering for Streams and Parallelism.

UNIT-IV:

Advertising on the Web-Issues in On-Line Advertising, On-Line Algorithms, The Matching Problem, The AdWords Problem, AdWords Implementation.

Recommendation Systems - A Model for Recommendation Systems, Content-Based Recommendations, Collaborative Filtering, Dimensionality Reduction, The Netflix Challenge.

UNIT-V:

Mining Social-Network Graphs-Social Networks as Graphs, Clustering of Social-Network Graphs, Partitioning of Graphs, Simrank, Counting Triangles.

TEXT BOOKS:

1.   Jure Leskovec, Anand Rajaraman, Jeff Ullman, Mining of Massive Datasets, 3rd Edition.

REFERENCE BOOKS:

1.   Jiawei Han & Micheline Kamber, Data Mining – Concepts and Techniques 3rd Edition Elsevier.
2.   Margaret H Dunham, Data Mining Introductory and Advanced topics, PEA.
3.   Ian H. Witten and Eibe Frank, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VII Semester Syllabus
CD721OE: FUNDAMENTALS OF DATA SCIENCE (OE-II)

Course Objectives

- Learn concepts, techniques and tools they need to deal with various facets of data science practice, including data collection and integration
- Understand the basic types of data and basic statistics
- Identify the importance of data reduction and data visualization techniques.

Course Outcomes

- Understand basic terms of statistical modeling and data science
- Implementation of R programming concepts
- utilize R elements for data visualization and prediction

UNIT-I:
Introduction
Definition of Data Science- Big Data and Data Science hype – and getting past the hype - Datafication - Current landscape of perspectives - Statistical Inference - Populations and samples - Statistical modeling, probability distributions, fitting a model – Over fitting.
Basics of R: Introduction, R-Environment Setup, Programming with R, Basic Data Types.

UNIT-II:
Data Types & Statistical Description
Types of Data: Attributes and Measurement, Attribute, The Type of an Attribute, The Different Types of Attributes, Describing Attributes by the Number of Values, Asymmetric Attributes, Binary Attribute, Nominal Attributes, Ordinal Attributes, Numeric Attributes, Discrete versus Continuous Attributes.
Basic Statistical Descriptions of Data: Measuring the Central Tendency: Mean, Median, and Mode, Measuring the Dispersion of Data: Range, Quartiles, Variance, Standard Deviation, and Interquartile Range, Graphic Displays of Basic Statistical Descriptions of Data.

UNIT-III:
Vectors: Creating and Naming Vectors, Vector Arithmetic, Vector sub setting.
Matrices: Creating and Naming Matrices, Matrix Sub setting, Arrays, Class.
Factors and Data Frames: Introduction to Factors: Factor Levels, Summarizing a Factor, Ordered Factors, Comparing Ordered Factors, Introduction to Data Frame, subsetting of Data Frames, Extending Data Frames, Sorting Data Frames.
Lists: Introduction, creating a List: Creating a Named List, Accessing List Elements, Manipulating List Elements, Merging Lists, Converting Lists to Vectors.

UNIT-IV:
Conditionals and Control Flow: Relational Operators, Relational Operators and Vectors, Logical Operators, Logical Operators and Vectors, Conditional Statements.
Iterative Programming in R: Introduction, While Loop, For Loop, Looping Over List.
Functions in R: Introduction, writing a Function in R, Nested Functions, Function Scoping, Recursion, Loading an R Package, Mathematical Functions in R.

UNIT-V:
Charts and Graphs: Introduction, Pie Chart: Chart Legend, Bar Chart, Box Plot, Histogram, Line Graph: Multiple Lines in Line Graph, Scatter Plot.
Regression: Linear Regression Analysis, Multiple Linear regression.

TEXT BOOKS:
1. Doing Data Science, Straight Talk from The Frontline. Cathy O'Neil and Rachel Schutt, O'Reilly, 2014.
2. K G Srinivas, G M Siddesh, "Statistical programming in R", Oxford Publications.

REFERENCE BOOKS:
1. Jiawei Han, Micheline Kamber and Jian Pei. Data Mining: Concepts and Techniques, 3rd ed. The Morgan Kaufmann Series in Data Management Systems.
2. Introduction to Data Mining, Pang-Ning Tan, Vipin Kumar, Michael Steinbanch, Pearson Education.
3. Brain S. Everitt, "A Handbook of Statistical Analysis Using R", Second Edition, 4 LLC, 2014.
4. Dalgaard, Peter, "Introductory statistics with R", Springer Science & Business Media, 2008.
5. Paul Teetor, "R Cookbook", O'Reilly, 2011

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VII Semester Syllabus
CD722OE: R PROGRAMMING (OE-II)

Prerequisites

- Basic programming

Course Objectives

- Understanding and being able to use basic programming concepts
- Automate data analysis
- Working collaboratively and openly on code
- Knowing how to generate dynamic documents
- Being able to use a continuous test-driven development approach.

Course Outcomes

- Understand to use and program in the programming language R
- Understand to use R to solve statistical problems
- Implement and describe Monte Carlo the technology
- Implement minimize and maximize functions using R.

UNIT-I:
Introduction: Overview of R, R data types and objects, reading and writing data, sub setting R Objects, Essentials of the R Language, Installing R, Running R, Packages in R, Calculations, Complex numbers in R, Rounding, Arithmetic, Modulo and integer quotients, Variable names and assignment, Operators, Integers, Factors, Logical operations.

UNIT-II:
Control structures, functions, scoping rules, dates and times, Introduction to Functions, preview of Some Important R Data Structures, Vectors, Character Strings, Matrices, Lists, Data Frames, Classes Vectors: Generating sequences, Vectors and subscripts, Extracting elements of a vector using subscripts, Working with logical subscripts, Scalars, Vectors, Arrays, and Matrices, Adding and Deleting Vector Elements, Obtaining the Length of a Vector, Matrices and Arrays as Vectors Vector Arithmetic and Logical Operations, Vector Indexing, Common Vector Operations.

UNIT-III:
Lists: Creating Lists, General List Operations, List Indexing Adding and Deleting List Elements, Getting the Size of a List, Extended Example: Text Concordance Accessing List Components and Values Applying Functions to Lists, DATA FRAMES, Creating Data Frames, Accessing Data Frames, Other
Matrix-Like Operations.

UNIT-IV:
Factors And Tables: Factors and Levels, Common Functions Used with Factors, Working with Tables, Matrix/Array-Like Operations on Tables, Extracting a Subtable, Finding the Largest Cells in a Table, Math Functions, Calculating a Probability, Cumulative Sums and Products, Minima and Maxima, Calculus, Functions for Statistical Distributions.

UNIT-V:
Graphics: Creating Graphs, Customizing Graphs, Saving Graphs to Files, Customizing Graphs, Creating Three-Dimensional Plots.
Debugging: Fundamental Principles of Debugging, Why Use a Debugging Tool?, Using R Debugging Facilities, Moving Up in the World: More Convenient Debugging Tools, Ensuring Consistency in Debugging Simulation Code, Syntax and Runtime Errors, Running GDB on R Itself.

TEXT BOOKS:

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

REFERENCE BOOKS:

1. R Programming for Data Science by Roger D. Peng
2. Hadley Wickham, Garrett Grolemund, R for Data Science: Import, Tidy, Transform, Visualize, and Model Data 1st Edition, O'Reilly
3. Tilman M. Davies, The book of R a first course in programming and statistics, no starch press.

L	T	P	C
0	0	2	1

B.Tech. in Computer Science & Engineering (Data Science)
VII Semester Syllabus
CD751PC: PREDICTIVE ANALYTICS LAB

Course Objectives

- To learn the basics and applications of predictive analytics using different techniques

Course Outcomes

- Understand the processing steps for predictive analytics
- Construct and deploy prediction models with integrity
- Explore various techniques (machine learning/data mining, ensemble) for predictive analytics.
- Apply predictive analytics to real world examples

List of Experiments:

Following experiments to be carried out using Python/SPSS/SAS/R/Power BI

1. Simple Linear regression
2. Multiple Linear regression
3. Logistic Regression
4. CHAID
5. CART
6. ARIMA – stock market data
7. Exponential Smoothing
8. Hierarchical clustering
9. Ward's method of clustering
10. Crowdsource predictive analytics- Netflix data

TEXT BOOKS:

1. Eric Siegel, Predictive analytics- the power to predict who will Click, buy, lie, or die, John Wiley & Sons, 2013.
2. Dean Abbott, Applied Predictive Analytics - Principles and Techniques for the Professional Data Analyst, 2014.

REFERENCE BOOKS:

1. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning-Data Mining, Inference, and Prediction, Second Edition, Springer Verlag, 2009.
2. G. James, D. Witten, T. Hastie, R. Tibshirani-An introduction to statistical learning with applications in R, Springer, 2013
3. E. Alpaydin, Introduction to Machine Learning, Prentice Hall of India, 2010

L	T	P	C
0	0	2	1

B.Tech. in Computer Science & Engineering (Data Science)
VII Semester Syllabus
CD752PC: WEB AND SOCIAL MEDIA ANALYTICS LAB

Course Objectives

* Exposure to various web and social media analytic techniques.

Course Outcomes

* Knowledge on decision support systems.
* Apply natural language processing concepts on text analytics.
* Understand sentiment analysis.
* Knowledge on search engine optimization and web analytics.

List of Programs

1. Preprocessing text document using NLTK of Python
     a. Stopword elimination
     b. Stemming
     c. Lemmatization
     d. POS tagging
     e. Lexical analysis
2. Sentiment analysis on customer review on products
3. Web analytics
     a. Web usage data (web server log data, clickstream analysis)
     b. Hyperlink data
4. Search engine optimization- implement spamdexing
5. Use Google analytics tools to implement the following
     a. Conversion Statistics
     b. Visitor Profiles
6. Use Google analytics tools to implement the Traffic Sources.

Resources:
1. Stanford core NLP package
2. GOOGLE.COM/ANALYTICS

TEXT BOOKS:

1. Ramesh Sharda, Dursun Delen, Efraim Turban, BUSINESS INTELLIGENCE AND ANALYTICS: SYSTEMS FOR DECISION SUPPORT, Pearson Education

REFERENCE BOOKS:

1.   Rajiv Sabherwal, Irma Becerra-Fernandez," Business Intelligence – Practice, Technologies and Management", John Wiley 2011.
2.   Lariss T. Moss, ShakuAtre, "Business Intelligence Roadmap", Addison-Wesley It Service.
3.   Yuli Vasiliev, "Oracle Business Intelligence: The Condensed Guide to Analysis and Reporting", SPD Shroff, 2012.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VIII Semester Syllabus
MS801HS: ORGANIZATIONAL BEHAVIOUR
(Common to CSE, IT & CSE (DS))

Course Objective: The objectives of the course are:

- To understand the perception and attribution concepts in organizational and external environmental Context.
- To discuss the concepts and theories of personality, attitudes and motivation.
- To understand the decision making and communication process in interpersonal and their intrapersonal context.
- To familiarize the students with the basic understanding of individual behavior and explore issues of power and empowerment.
- To discuss the concepts of High performance, Learning and Leadership.

Course Outcomes: Upon the completion of the subject, the student will be able to

- Analyze the concepts of perception and attribution of individuals and groups in organizations in terms of the key factors that influence organizational behavior.
- Acquire knowledge of personality and attitudes of individuals in applying motivational theories to resolve problems of employees.
- Apply group dynamics and skills required for working in groups and identify the processes used in developing communication and decision making to overcome stress and conflicts.
- Analyze organizational behavioral issues in the context of, power, and empowerment issues.
- Acquire the knowledge of learning concepts and leadership styles to achieve high performance levels.

UNIT I: INTRODUCTION, PERCEPTION AND ATTRIBUTION
Introduction to OB - Definition, Nature and Scope – Environmental and organizational context –Impact of IT, globalization, Diversity, Ethics, culture, reward systems and organizational design on Organisational Behaviour.
Cognitive Process I: Perception and Attribution: Meaning of Perception, Nature and importance of Perception — Sensation versus Perception - Perceptual selectivity and organization – Social perception.
Attribution: Meaning - Attribution Theories – Fritz heider's theory of attribution; Jones & Davis correspondent inference theory; Kelley's Covariation Model; and Weiner's Three-Dimensional Model - Locus of control –Attribution Errors – Impression Management – Types - Strategies of Impression Management.

UNIT II: PERSONALITY, ATTITUDES AND MOTIVATION
Cognitive Process II: Personality and Attitudes - Personality as a continuum –Meaning of personality, Personality Traits - Johari Window and Transactional Analysis .Attitudes: Nature and Dimension of Attitudes - Job satisfaction and organizational commitment.
Motivation: Motivational needs and processes- Theories of Motivation– the content theories of motivation: Maslow's hierarchy of needs; Herzberg's two factor theory of motivation; Alderfer's ERG theory – The process theories of work motivation: Vroom's expectancy theory of motivation; the Porter-Lawler's model.

UNIT III: COMMUNICATION, DECISION-MAKING, STRESS AND CONFLICT
Dynamics of OB - I: Communication: Meaning and types - interactive communication in organizations – barriers to communication and strategies to improve the flow of communication.
Decision Making: Participative decision-making techniques – creativity and group decision making.
Dynamics of OB – II: Stress and Conflict: Meaning and types of stress –Meaning and types of conflict - Effect of stress and intra-individual conflict -strategies to cope with stress and conflict.

UNIT IV: POWER AND EMPOWERMENT
Dynamics of OB - III Power: Meaning - Types of power – Legitimate power, Coercive power, Expert power, Informal power and Referent power.
Empowerment – Benefits, Process, Determinants, Techniques of Empowerment; Groups vs. Teams –Nature of groups – dynamics of informal groups – Dysfunctions of groups and teams –Teams in modern work place.

UNIT V: HIGH PERFORMANCE, LEARNING AND LEADERSHIP
Leading High performance: Job design and Goal setting for High Performance - Quality of Work Life - Socio technical Design and High-performance work practices - Behavioural performance management.
Learning: Reinforcement and Punishment as principles of Learning – Process of Behavioural modification.
Leadership: Definition of leadership – Traits of effective leaders - Leadership behaviour Vs Traits – Leadership skills – Leadership theories: Trait theories; Behavioural theories; Contingency approaches to leadership, Leadership Styles: Autocratic Leadership style, Democratic Leadership style, Free rein Leadership style.

TEXT BOOKS:
1. Luthans, Fred, Organizational Behavior, McGraw-Hill, 10th Edition, 2009.
2. Robbins, P. Stephen, Timothy A. Judge, Organizational Behavior, PHI/Pearson, 18th Edition, 2022.
3. Pareek Udai, Behavioral Process at Work: Oxford & IBH, New Delhi, 2009.

REFERENCE BOOKS:
1.   Schermerhorn, Organizational Behavior Wiley, 9th Edition, 2008.
2. Michael A Hitt, Organizational Behavior, Wiley, 2008.
3. Aswathappa, Organizational Behavior, Himalaya, 2009.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VIII Semester Syllabus
CM851PE: WEB SECURITY
(Common to CSE (AI&ML) PE-V, CSE(Data Science) PE – VI)

Course Objectives:

- Give an Overview of information security.
- Give an overview of Access control of relational databases

Course Outcomes: Students should be able to

- Understand the Web architecture and applications
- Understand client side and service side programming
- Understand how common mistakes can be bypassed and exploit the application
- Identify common application vulnerabilities

UNIT - I

The Web Security, The Web Security Problem, Risk Analysis and Best Practices.
Cryptography and the Web: Cryptography and Web Security, Working Cryptographic Systems and Protocols, Legal Restrictions on Cryptography, Digital Identification.

UNIT - II

The Web's War on Your Privacy, Privacy-Protecting Techniques, Backups and Antitheft, Web Server Security, Physical Security for Servers, Host Security for Servers, Securing Web Applications.

UNIT - III

Database Security: Recent Advances in Access Control, Access Control Models for XML, Database Issues in Trust Management and Trust Negotiation, Security in Data Warehouses and OLAP Systems.

UNIT - IV

Security Re-engineering for Databases: Concepts and Techniques, Database Watermarking for Copyright Protection, Trustworthy Records Retention, Damage Quarantine and Recovery in Data Processing Systems, Hippocratic Databases: Current Capabilities and Future Trends.

UNIT - V

Privacy in Database Publishing: A Bayesian Perspective, Privacy-enhanced Location-based Access Control, Efficiently Enforcing the Security and Privacy Policies in a Mobile Environment.

TEXT BOOKS:

1. Web Security, Privacy and Commerce Simson G Arfinkel, Gene Spafford, O'Reilly.

2. Handbook on Database security applications and trends Michael Gertz, Sushil Jajodia

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VIII Semester Syllabus
CD861PE: DATA STREAM MINING (PE – VI)

Prerequisites

- A basic knowledge of "Data Mining"

Course Objectives

- The aim of the course is to introduce the fundamentals of Data Stream Mining.
- The course gives an overview of – Mining Strategies, methods and algorithms for data stream mining.

Course Outcomes

- Understand how to formulate a knowledge extraction problem from data streams.
- Ability to apply methods / algorithms to new data stream analysis problems.
- Evaluate the results and understand the functioning of the methods studied.
- Demonstrate decision tree and adaptive Hoeffding Tree concepts

UNIT-I:
MOA Stream Mining, Assumptions, Requirements, Mining Strategies, Change Detection Strategies, MOA Experimental Settings, Previous Evaluation Practices, Evaluation Procedures for Data Streams, Testing Framework, Environments, Data Sources, Generation Speed and Data Size, Evolving Stream Experimental Setting.

UNIT-II:
Hoeffding Trees, The Hoeffding Bound for Tree Induction, The Basic Algorithm, Memory Management, Numeric Attributes, Batch Setting Approaches, Data Stream Approaches.

UNIT-III:
Prediction Strategies, Majority Class, Naïve Bayes Leaves, Adaptive Hybrid, Hoeffding Tree Ensembles, Data Stream Setting, Realistic Ensemble Sizes.

UNIT-IV:
Evolving Data Streams, Algorithms for Mining with Change, A Methodology for Adaptive Stream Mining, Optimal Change Detector and Predictor, Adaptive Sliding Windows, Introduction, Maintaining Updated Windows of Varying Length.

UNIT-V:
Adaptive Hoeffding Trees, Introduction, Decision Trees on Sliding Windows, Hoeffding Adaptive Trees, Adaptive Ensemble Methods, New methods of Bagging using trees of different size, New method of bagging using ADWIN, Adaptive Hoeffding Option Trees, Method performance

TEXT BOOKS:
1. DATA STREAM MINING: A Practical Approach by Albert Bifet and Richard Kirkby

REFERENCE BOOKS:
1. Knowledge discovery from data streams by Gama João. ISBN: 978-1-4398-2611-9.
2. Machine Learning for Data Streams by Albert Bifet, Ricard Gavalda; MIT Press, 2017.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VIII Semester Syllabus
CD862PE: VIDEO ANALYTICS (PE-VI)

Course Objectives

- To know the fundamental concepts of big data and analytics
- To learn various techniques for mining data streams
- To acquire the knowledge of extracting information from surveillance videos
- To learn Event Modelling for different applications.
- To understand the models used for recognition of objects in videos.

Course Outcomes

- Understand the basics of video – signals and systems.
- Able to estimate motion in a video.
- Able to detect the objects and track them.
- Recognize activity and analyze behavior.
- Evaluate face recognition technologies.

UNIT-I:
INTRODUCTION: Multidimensional signals and systems: signals, transforms, systems, sampling theorem. Digital Images and Video: human visual system and color, digital video, 3Dvideo, digital-video applications, image and video quality.

UNIT-II:
MOTION ESTIMATION Image formation, motion models, 2D apparent motion estimation, differential methods, matching methods, non-linear optimization methods, transform domain methods, 3D motion and structure estimation.

UNIT-III:
VIDEO ANALYTICS: Introduction-Video Basics-Fundamentals for Video Surveillance- Scene Artifacts- Object Detection and Tracking: Adaptive Background Modelling and Subtraction- Pedestrian Detection and Tracking Vehicle Detection and Tracking- Articulated Human Motion Tracking in Low-Dimensional Latent Spaces.

UNIT-IV:
BEHAVIORAL ANALYSIS & ACTIVITY RECOGNITION Event Modelling- Behavioral Analysis-Human Activity Recognition- Complex Activity Recognition Activity modelling using 3D shape, Video summarization, shape- based activity models- Suspicious Activity Detection.

UNIT-V:
HUMAN FACE RECOGNITION & GAIT ANALYSIS: Introduction: Overview of Recognition algorithms– Human Recognition using Face: Face Recognition from still images, Face Recognition from video, Evaluation of Face Recognition Technologies- Human Recognition using gait: HMM Framework for Gait Recognition, View Invariant Gait Recognition, Role of Shape and Dynamics in Gait Recognition.

TEXT BOOKS:

1. Murat Tekalp, ―Digital Video Processing‖, second edition, Pearson, 2015.
2. Rama Chellappa, AmitK. Roy- Chowdhury, Kevin Zhou. S, ―Recognition of Humans and their Activities using Video‖, Morgan & Claypool Publishers, 2005.
3. Yunqian Ma, Gang Qian, ―Intelligent Video Surveillance: Systems and Technology‖, CRC Press (Taylor and Francis Group), 2009.

Reference Books:
1. Richard Szeliski, ―Computer Vision: Algorithms and Applications‖, Springer, 2011.
2. Yao Wang, Jorn Ostermann and Ya-Qin Zhang, ―Video Processing and Communications‖, Prentice Hall, 2001.
3. Thierry Bouwmans, FatihPorikli, Benjamin Höferlinand Antoine Vacavant, ―Background Modeling and Foreground Detection for Video Surveillance: Traditional and Recent Approaches, Implementations, Benchmarking and Evaluation‖, CRC Press, TaylorandFrancisGroup,2014.
4. Md. Atiqur Rahman Ahad,―Computer Vision and Action Recognition- A Guide for Image Processing and Computer Vision Community for Action Understanding‖, AtlantisPress,2011.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VIII Semester Syllabus

CD863PE: BLOCKCHAIN TECHNOLOGY (PE-VI)

Prerequisites

- Knowledge in information security and applied cryptography.
- Knowledge in Computer Networks

Course Objectives

- To learn the fundamentals of Blockchain and various types of block chain and consensus mechanisms.
- To understand the public block chain system, Private block chain system and consortium blockchain.
- Able to know the security issues of blockchain technology.

Course Outcomes

- Understanding concepts behind crypto currency
- Applications of smart contracts in decentralized application development
- Understand frameworks related to public, private and hybrid blockchain
- Create blockchain for different application case studies.

UNIT-I:
Fundamentals of Blockchain: Introduction, Origin of Blockchain, Blockchain Solution, Components of Blockchain, Block in a Blockchain, The Technology and the Future.
Blockchain Types and Consensus Mechanism: Introduction, Decentralization and Distribution, Types of Blockchain, Consensus Protocol.
Cryptocurrency – Bitcoin, Altcoin and Token: Introduction, Bitcoin and the Cryptocurrency, Cryptocurrency Basics, Types of Cryptocurrencies, Cryptocurrency Usage.

UNIT-II:
Public Blockchain System: Introduction, Public Blockchain, Popular Public Blockchains, The Bitcoin Blockchain, Ethereum Blockchain.
Smart Contracts: Introduction, Smart Contract, Characteristics of a Smart Contract, Types of Smart Contracts, Types of Oracles, Smart Contracts in Ethereum, Smart Contracts in Industry.

UNIT-III:
Private Blockchain System: Introduction, Key Characteristics of Private Blockchain, Need of Private Blockchain, Private Blockchain Examples, Private Blockchain and Open Source, E- commerce Site Example, Various Commands (Instructions) in E-commerce Blockchain, Smart Contract in Private Environment, State Machine, Different Algorithms of Permissioned Blockchain, Byzantine Fault, Multichain.
Consortium Blockchain: Introduction, Key Characteristics of Consortium Blockchain, Need of Consortium Blockchain, Hyperledger Platform, Overview of Ripple, Overview of Corda.
Initial Coin Offering: Introduction, Blockchain Fundraising Methods, Launching an ICO, Investing in an ICO, Pros and Cons of Initial Coin Offering, Successful Initial Coin Offerings, Evolution of ICO, ICO Platforms.

UNIT-IV:
Security in Blockchain: Introduction, Security Aspects in Bitcoin, Security and Privacy Challenges of Blockchain in General, Performance and Scalability, Identity Management and Authentication, Regulatory Compliance and Assurance, Safeguarding Blockchain Smart Contract (DApp), Security Aspects in Hyperledger Fabric.
Applications of Blockchain: Introduction, Blockchain in Banking and Finance, Blockchain in Education, Blockchain in Energy, Blockchain in Healthcare, Blockchain in Real-estate, Blockchain In Supply Chain, The Blockchain and IoT. Limitations and Challenges of Blockchain.

UNIT-V:
Blockchain Case Studies: Case Study 1 – Retail, Case Study 2 – Banking and Financial Services, Case Study 3 – Healthcare, Case Study 4 – Energy and Utilities.
Blockchain Platform using Python: Introduction, Learn How to Use Python Online Editor, Basic Programming Using Python, Python Packages for Blockchain.
Blockchain platform using Hyperledger Fabric: Introduction, Components of Hyper ledger Fabric Network, Chain codes from Developer.ibm.com, Blockchain Application Using Fabric Java SDK.

TEXT BOOKS:
1. Blockchain Technology, Chandramouli Subramanian, Asha A. George, Abhilasj K A and Meena Karthikeyan, Universities Press.

REFERENCE BOOKS:
1. Michael Juntao Yuan, Building Blockchain Apps, Pearson, India.
2. Blockchain Blueprint for Economy, Melanie Swan, SPD O'reilly.
3. Blockchain for Business, Jai Singh Arun, Jerry Cuomo, Nitin Gaur, Pearson.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VIII Semester Syllabus
CD821OE-DATA ANALYTICS (OE-III)

Prerequisites

- A course on "Database Management Systems".
- Knowledge of probability and statistics.

Course Objectives

- To explore the fundamental concepts of data analytics.
- To learn the principles and methods of statistical analysis
- Discover interesting patterns, analyze supervised and unsupervised models and estimate the accuracy of the algorithms.
- To understand the various search methods and visualization techniques..

Course Outcomes

- Understand the impact of data analytics for business decisions and strategy
- Carry out data analysis/statistical analysis
- To carry out standard data visualization and formal inference procedures
- Design Data Architecture
- Understand various Data Sources

UNIT-I:
Data Management: Design Data Architecture and manage the data for analysis, understand various sources of Data like Sensors/Signals/GPS etc. Data Management, Data Quality(noise, outliers, missing values, duplicate data) and Data Processing & Processing.

UNIT-II:
Data Analytics: Introduction to Analytics, Introduction to Tools and Environment, Application of Modeling in Business, Databases & Types of Data and Variables, Data Modeling Techniques, Missing Imputations etc. Need for Business Modeling.

UNIT-III:
Regression – Concepts, Blue property assumptions, Least Square Estimation, Variable Rationalization, and Model Building etc.
Logistic Regression: Model Theory, Model fit Statistics, Model Construction, Analytics applications to various Business Domains etc.

UNIT-IV:
Object Segmentation: Regression Vs Segmentation – Supervised and Unsupervised Learning, Tree Building – Regression, Classification, Overfitting, Pruning and Complexity, Multiple Decision Trees etc.
Time Series Methods: Arima, Measures of Forecast Accuracy, STL approach, Extract features from generated model as Height, Average Energy etc and Analyze for prediction

UNIT-V:
Data Visualization: Pixel-Oriented Visualization Techniques, Geometric Projection Visualization Techniques, Icon-Based Visualization Techniques, Hierarchical Visualization Techniques, Visualizing Complex Data and Relations.

TEXT BOOKS:
1. Student's Handbook for Associate Analytics – II, III.
2. Data Mining Concepts and Techniques, Han, Kamber, 3rd Edition, Morgan Kaufmann Publishers

REFERENCE BOOKS:
1. Introduction to Data Mining, Tan, Steinbach and Kumar, Addision Wisley, 2006.
2. Data Mining Analysis and Concepts, M. Zaki and W. Meira
3. Mining of Massive Datasets, Jure Leskovec Stanford Univ. Anand Rajaraman Milliway Labs Jeffrey D Ullman Stanford Univ.

L	T	P	C
3	0	0	3

B.Tech. in Computer Science & Engineering (Data Science)
VIII Semester Syllabus
CD822OE-DATA MINING (OE-III)

Prerequisites

- Database Management System
- Probability and Statistics

Course Objectives

- Students will become acquainted with both the strengths and limitations of various data mining techniques like Association, Classification, Cluster and Outlier analysis.

Course Outcomes

- Understand the need of data mining and pre-processing techniques.
- Perform market basket analysis using association rule mining.
- Utilize classification techniques for analysis and interpretation of data.
- Identify appropriate clustering and outlier detection techniques to handle complex data.
- Understand the mining of data from web, text and time series data.

UNIT-I:

Introduction to Data Mining:
What Data mining? Kinds of Data, Knowledge Discovery process, Data Mining Functionalities, Kinds of Patterns, Major Issues in Data Mining. Data Objects and Attribute Types, Basic Statistical Descriptions of Data, Data Visualization, Measuring Data Similarity and Dissimilarity, Data Pre-processing: Major Tasks in Data Pre-processing, Data Cleaning, Data Integration, Data Reduction, Data Transformation and Data Discretization.

UNIT-II:

Association Analysis: Basic Concepts, Market Basket Analysis, Apriori Algorithm, FP-growth, From Association Analysis to Correlation Analysis, Pattern Mining in Multilevel Associations and Multidimensional Associations.

UNIT-III:

Classification: Basic Concepts, Decision Tree Induction, Bayes Classification Methods, Rule-Based Classification, Metrics for Evaluating Classifier Performance, Ensemble Methods, Multilayer Feed Forward Neural Network, Support Vector Machines, k-Nearest-Neighbor Classifiers.

UNIT-IV:

Cluster Analysis: Requirements for Cluster Analysis, Overview of Basic Clustering Methods,
Partitioning Methods-k-Means, k-Medoids, Hierarchical Methods-AGENES, DIANA, BIRCH, Density Based Method-DBSCAN, Outlier Analysis: Types of Outliers, Challenges of Outlier Detection, and
Overview of Outlier Detection Methods.

UNIT-V:

Advanced Concepts: Web Mining- Web Content Mining, Web Structure Mining, Web Usage Mining, Spatial Mining- Spatial Data Overview, Spatial Data Mining Primitives, Spatial Rules, Spatial Classification Algorithm, Spatial Clustering Algorithms, Temporal Mining- Modeling Temporal Events, Time Series, Pattern Detection, Sequences, Temporal Association Rules.

TEXT BOOKS:
1. Jiawei Han, Micheline Kamber, Jian Pei., Data Mining: Concepts and Techniques, 3rd Edition, Morgan Kaufmann/Elsevier, 2012.
2. Margaret H Dunham, Data Mining Introductory and Advanced Topics, 2nd Edition, Pearson Education, India, 2006

REFERENCE BOOKS:
1. Data Mining Techniques, Arun K Pujari, 3rd Edition, Universities Press.
2. Pang-Ning Tan, Michael Steinbach, Anuj Karpatne and Vipin Kumar, Introduction to Data Mining, 2nd Edition, Pearson Education India, 2021.
3. Amitesh Sinha, Data Warehousing, Thomson Learning, India, 2007.