JSS COLLEGE OF ARTS, COMMERCE AND SCIENCE

Affiliated & Autonomous under University of Mysore

Accredited with Grade A+ NAAC CGPA: 3.30

Recognized by UGC as College with Potential for Excellence

Master of Science in Data Science

(Choice Based Credit System)

2026 -2028

PG Department of Computer science

JSS COLLEGE OF ARTS, COMMERCE AND SCIENCE

(Autonomous)

MYSURU

1	**Title of the Course**	**M.Sc. (Data Science)**
2	Duration of the course	2 years (Four Semesters)
3	Eligibility criteria for admission	The candidate seeking admission in to M.Sc. (Data Science) should have • Bachelors Degree in mathematics/Computer Science/Computer applications/Statistics or • B.Voc with Computer Science/Computer Applications/Information Technology/Web Technologies, B.Sc (Data Science) & B.Sc(Artificial Intelligence/Machine Learning) (or) • Bachelors Degree in Engineering or Technology in CSE/ECE/EEE/E&I/IT with 50% of marks (inclusive all subjects)
4	Level of the Course	Post Graduate
5	Mode of Admission	The mode of admission is through PGCET conducted by University of Mysore.
6	Objectives of the course	The Objective of M.Sc., Data Science course is to impart knowledge and skill-oriented training in the recent advancements in Data Science with an aim to develop research and innovations.
7	Course Requirement	The course shall include Theory papers, Labs, Assignments, Tests, Seminars and Project Work.
8	Number of working days	In each semester at least ninety working days must be dedicated for theory classes, practical classes and seminars.

JSS COLLEGE OF ARTS, COMMERCE AND SCIENCE

M.Sc. (Data Science)

(CBCS)

COURSE STRUCTURE

I SEMESTER

Course Code	Course Name	Teaching Hours/Week			CORE / IDC/DSE/ SEC/OEC/MOOCS	Internal Marks	External Marks	No. of Credits
		Lecture	Tutorial	Practical				
DS101	Programming with R and Python	3	0	0	Core	30	70	3
DS102	Data Structures	3	0	0	Core	30	70	3
DS103	Database Management Systems	4	0	0	Core	30	70	4
DS104	Probability and Statistics	4	0	0	Core	30	70	4
DS105	Operating Systems	4	0	0	Core	30	70	4
DS106	Python & R Lab	0	0	2	Core	15	35	1
DS107	Data Structures Lab	0	0	2	Core	15	35	1
TOTAL FOR FIRST SEMESTER						**180**	**420**	**20**

II SEMESTER

Course Code	Course Name	Teaching Hours/Week			CORE / IDC/DSE/ SEC/OEC/MOOCS	Internal Marks	External Marks	No. of Credits
		Lecture	Tutorial	Practical				
DS201	Data Processing and Visualization using Tableau	3	0	0	Core	30	70	3
DS202	Data Mining & Data warehousing	3	0	0	Core	30	70	3
DS203	Data Science with R Programming	4	0	0	Core	30	70	4
DS204	Optimization Techniques	4	0	0	SEC	30	70	4
DOMAIN SPECIFIC ELECTIVE COURSES (CHOOSE ANY ONE)								
DS205E1	Linear Regression Models	4	0	0	DSE	30	70	4
DS205E2	Analysis and Design of Algorithms	4	0	0	DSE	30	70	4
DS205E3	Time Series Analysis	4	0	0	DSE	30	70	4
LAB PRACTICALS								
DS206	Data Processing and Visualization Lab	0	0	2	Core	15	35	1
DS207	Data Mining and Data Warehousing Lab	0	0	2	Core	15	35	1
TOTAL FOR SECOND SEMESTER						**180**	**420**	**20**

III SEMESTER

Course Code	Course Name	Teaching Hours/Week			CORE / IDC/DSE/ SEC/OEC/MOOCS	Internal Marks	External Marks	No. of Credits
		Lecture	Tutorial	Practical				
DS301	Data Science using Python	3	0	0	Core	30	70	3
DS302	MongoDB	3	0	0	Core	30	70	3
DOMAIN SPECIFIC ELECTIVE COURSES (CHOOSE ANY TWO)								
DS302E1	React JS	4	0	0	DSE	30	70	4
DS302E2	**Cryptography and Network Security**	4	0	0	DSE	30	70	4
DS302E3	Social Network Analysis	4	0	0	DSE	30	70	4
DS302E4	Applied Data Analytics	4	0	0	DSE	30	70	4
DS302E5	Block Chain Technology	4	0	0	DSE	30	70	4
LAB PRACTICALS								
DS303	Data Science using Python Lab	0	0	2	Core	15	35	1
DS304	Mongo DB Lab	0	0	2	Core	15	35	1
ENTREPRENURAL & INNOVATION/IT SKILL RELATED TO DOMAIN SPECIFIC ELECTIVE COURSES (CHOOSE ANY ONE)								
OE305	**Cyber Security**	4	0	0	OEC	30	70	4
OE306	Bio Informatics	4	0	0	OEC	30	70	4
OE307	Web Designing	4	0	0	OEC	30	70	4
TOTAL FOR THIRD SEMESTER						**180**	**420**	**20**

IV SEMESTER

Course Code	Course Name	Teaching Hours/ Week			CORE / IDC/DSE/ SEC/OEC/MOOCS	Internal Marks	External Marks	No. of Credits
		Lecture	tutorial	Practical				
DS401	Dynamic Web Programming using Django and Flask	3	0	0	Core	30	70	3
DOMAIN SPECIFIC ELECTIVE COURSES (CHOOSE ANY ONE)								
DS402E1	Web Scraping with Python	4	0	0	DSE	30	70	4
DS402E2	Deep Learning with PyTorch	4	0	0	DSE	30	70	4
DS402E3	**Machine Learning**	4	0	0	DSE	30	70	4
DS402E4	Design Thinking	4	0	0	DSE	30	70	4
DS402E5	**Big Data Analytics**	4	0	0	DSE	30	70	4
LAB PRACTICALS								
DS403	Dynamic Web Programming using Django and Flask Lab	0	0	2	Core	15	35	1
ENTREPRENURAL & INNOVATION/IT SKILL RELATED TO DOMAIN SPECIFIC ELECTIVE COURSES (CHOOSE ANY ONE)								
DS404E1	Quantum Computing	4	0	0	SEC	30	70	4
DS404E2	**AI Concepts and Techniques With Python**	4	0	0	SEC	30	70	4
DS404E3	Data Wrangling with Java Script	4	0	0	SEC	30	70	4
DS405	**Project**	**0**	**0**	**4**	**Core**	**30**	**70**	**4**
TOTAL FOR IV SEMESTER						**135**	**415**	**16**

Note

1. At the end of 2nd semester, every student must undergo Summer Internship/Apprenticeship/Project Work/Industrial Training/Research based Project Work for Six Weeks and must prepare a report concerned as per approved project guidelines, and submit the same to the college 14 days before the commencement of third semester end examinations.

2. Students may be allowed to register and appear for MOOCS from the third semester itself. However, students are to complete the MOOCS successfully and submit pass certificate of the same to the University through the Principal of the College concerned for approval and endorsement of the same on grade cards and PCs and ODs as per the regulations of the University.

- **Total number of credits at the end of course:**

S.No	Semester	Credits
1	I Semester	20
2	II Semester	20
3	III Semester	20
4	IV Semester	16
TOTAL		**76**

SCHEME OF EXAMINATION AND ASSESMENT:

In view of the CBCS syllabus, Each Course is Assess with Components . Component 1 (C1), Component 2 (C2), and Component 3 (C3), The following is the scheme which will be followed for the assessment of marks for theory (Core/ DSE/ SEC/PROJECT) irrespective of the Credits associated with each Course. Thirty percent of the marks will be assessed for the internals (C1 and C2) and remaining seventy percent will be for the semester end examinations (C3). Each Course carries 100 marks and hence thirty marks for internal assessment and remaining seventy marks will be for Semester End Examinations. Out of thirty marks for internals, fifteen marks will be allotted to each C1 and C2 components.

Each Course (HC/ SC/ OE) consists of three components namely C1, C2 and C3. C1 and C2 are designated as Internal Assessment (IA) and C3 as Semester End Examination. Each Course (Core/ DSE/ SEC/PROJECT) carries 100 Marks and hence the allotment of marks to C1, C2 and C3 Components will be fifteen, fifteen and seventy marks respectively. i.e.,

C1 Component : 15 Marks C2 Component : 15 Marks	Internal Assessment Marks
C3 Component : 70 Marks	Semester End Examination
Total : 100 Marks	

The above will be followed in common for all (Core/ DSE/ SEC/PROJECT) Courses in all the four semesters.

Practicals:

C1 and C2	Semester End Exam	Total
15 marks	35 marks	50

*****Continuous assessment sheet given below.

Note: For practical courses, there shall be a continuous evaluation during the semester for 15 sessional marks and end examination shall be for 35 marks. Day-to-day work in the laboratory shall be evaluated for 15 marks by the concerned laboratory teacher based on the regularity/record/viva. The end examination shall be conducted by the internal and external examiner from home University / Other University.

I SEMESTER

Programming with R and Python 3:0:0

Course Outcomes:

At the end of this course the students should be able to:

CO1: Understand the basic concepts python programming including syntax, data types and operators.

CO2: Learn basic algorithmic problem, solving techniques (decision structures, loops, functions). Modules and packages

CO3: Use and understand objects used in programming and data structures

CO4: Design, document, implement and test solutions to programming problems. Related to R includes vector matrices and arrays

Course Content:

UNIT -I

Introduction: History of Python, features of python, Reading Input from the Console, Variables, data types, Assignment Statements, Operators, Control Statements in Python, Looping Statements, Break, continue , pass, Lists, Tuples and Dictionaries and Sets in Python.

UNIT-II

Python functions: Defining a Function – Calling a Function –Functions with/without Return Values – Positional and Keyword Arguments –Passing Arguments by Reference Values, lambda expressions.

UNIT-III

Classes and objects: Introduction to Object – Oriented Programming – Basic principles of Object – Oriented Programming in Python – Classes and Objects, Inheritance and Polymorphism, Abstract Classes and Interfaces, Exception Handling.

UNIT-IV

R Programming: Introduction to R Programming, Data Types, Vectors, Lists, Matrices, Arrays, Data Frames, Merging Data Frames, Variables, Functions in R, R - CSV Files, R - Excel File, R - Pie Charts, R - Bar Charts, R – Histograms, R - Line Graphs, R – Histograms, R – Scatterplots

Text Books:

- Core Python Black book Dream Tech Publishers Dr R. Nageswara Rao
- Reema Thareja, Python Programming using Problem Solving Approach, Oxford University Press(2017)
- Hands on Programming with R from Oreilly Publications by GarettGrolemund

Data Structures 3:0:0

Course Outcomes:

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

CO1: Understand the importance of various types of data structures in solving a problem through programming

CO2: Able to identify the suitability of a particular data structure to solve a problem

CO3: Critically evaluate the efficient representation of data structures in the memory

CO4: Understand the importance of indexing and how it is achieved through a particular data structure.

Course Content

UNIT-I

Data Structures: Array as an ADT & types of array, pointers, Stack: Definition and examples, Primitive operations, Example, The stack as an ADT, Representing stacks, Implementing the pop, push operations, Applications of Stacks: Expression Evaluations, Expression conversion, Recursion as application of stack, Properties of recursive definition or algorithm, Towers of Hanoi problem

UNIT-II

Queue: Sequential representation, the queue as an ADT, Priority queue, Array implementation of a priority queue, Linked lists, inserting and removing nodes from a list, Linked implementations of stacks ,Linked list implementation of queues. Circular Lists, Stack as a Circular List, Queue as a Circular List, Primitive Operations on Circular Lists, Double Linked Lists.

UNIT-III

,Binary Tree: ,Binary Tree Representations, Node Representation of Binary Trees, Internal and External Nodes. Implicit Array Representation of Binary Trees, Binary Tree Traversals.

Searching: Sequential Searching, efficiency of Sequential Searching ,Binary Search, General search tree – B tree , Algorithm for B-tree insertion &Sorting : Insertion Sort, Shell Sort, Radix Sort

UNIT-IV

Hashing: Resolving Hash Clashes by Open Addressing, Deleting items from a Hash Table, Separate Chaining. Dynamic Memory management: Compaction of Blocky of Storage First Fit, Best Fit and Worst Fit, Improvements in the First-Fit Method, Freeing Storage Blocks, Boundary Tag Method.

Text Books :

Text Books:

- Data Structures UsingCandC++byAaronMTenenbaum,YedidyahLangsamand Moshe J Augustine,PHI, Edition,2011
- Data Structures, Algorithms and Applications in C++, Sahani, University Press(India) Pvt Ltd, 2ndEdition
- The complete reference C,Herbert Schildt, Fifth Edition,Tata McGraw Hill

Database Management System 4:0:0

Course Outcomes

CO1: Understand the significance of databases, types of databases, merits and limitations ofdifferent DBMS

CO2: Explain and apply the concept of normalization for database design

CO3: Understand and apply concurrency control and transaction processing mechanisms

CO4: Learn the characteristics implementation of object oriented and distributed database management systems and their architecture

Course Content

UNIT-I:

Comparison between different databases: Significance of Databases, Database System Applications, Advantages and Disadvantages of different Database Management systems, Comparison between DBMS, RDBMS, Distributed and Centralized DB.

Normalization**:** Functional Dependency, Anomalies in a Database, The normalization process: Conversion to first normal form, Conversion to second normal form, Conversion to third normal form, The boyce-code normal form (BCNF), Fourth Normal form and fifth normal form, normalization and database design, Denormalization

UNIT-II:

Introduction of transaction processing, advantages and disadvantages of transaction processing system, online transaction processing system, serializability and recoverability, view serializability, resolving deadlock, distributed locking Transaction management in multi-database system, long duration transaction, high-performance transaction system.

UNIT–III:

Oriented paradigm, OODBMS architectural approaches, Object identity, procedures and encapsulation, Object oriented data model: relationship, identifiers, Basic OODBMS terminology, Inheritance , Basic interface and class structure, Type hierarchies and inheritance, Type extents and persistent programming languages, OODBMS storage issues

UNIT -IV:

Introduction of DDB, DDBMS architectures, Homogeneous and Heterogeneous databases, Distributed data storage, Advantages of Data Distribution, Disadvantages of Data Distribution Distributed transactions, Commit protocols, Availability.

design techniques used in RDBMS, extension techniques in RDBMS, standards for OODBMS ,Products and applications:

Text Books:

- Henry F Korth, Abraham Silberschatz and S Sudharshan, "Database System Concepts", Sixth Edition, McGraw Hill, 2011
- C J Date, A Kannan and S Swamynathan,"An Introduction to Database Systems", EighthEdition, Pearson Education, 2006
- R Elmasri, S B Navathe, "Fundamentals of Database Systems", Fifth Edition, PearsonEducation/Addison Wesley, 2007
- Thomas Cannolly and Carolyn Begg, "Database Systems, A Practical Approach to Design,Implementation and Management", Third Edition, Pearson Education, 2007
- Subramaniam, " Multimedia Databases", Morgan Kauffman Publishers, 2008

Probability and Statistics 4:0:0

Course Outcomes:

After the completion of the course, student will be able to

CO1: Apply the basic rules and theorems of probability theory such as Baye's Theorem, determine probabilities that help to solve engineering problems and to determine the expectation and variance of a random variable from its distribution.

CO52: Able to perform and analyze of sampling, means, proportions, variances and estimates the maximum likelihood based on population parameters.

CO3: Learn how to formulate and test hypotheses about sample means, variances and proportions and to draw conclusions based on the results of statistical tests.

CO4: Design various ciphers using number theory.

Course Content:

UNIT-I

Axioms of Probability, Conditional Probability, Independence of the Multiplication Rule, Bayes' theorem

Random Variables, Discrete Probability Densities, Expectation and distribution parameters, Binomial distribution, Poisson distribution and its applications

UNIT-II

Hypothesis Testing, significance testing, Hypothesis and significance test on the mean, Hypothesis tests on the Variance

Estimating proportions, testing hypothesis on a proportion, Comparing two proportions: estimation, comparing two proportions: hypothesis testing**.**

UNIT-III

Independent samples, Comparing variances: the F-distribution, Comparing means, variances equal

One-way classification fixed effects model, comparing variances, pair wise comparisons, randomized complete block design.

UNIT-IV

Model and parameter estimation, inferences about slope, Inferences about intercept, Co-efficient of determination, Least square procedures for model fitting, a matrix approach to least squares, interval estimation

Text Books

- Susan Milton and Jesse C. Arno ld, Introduction to Probability and Statistics, Fourth edition.
- William Mendenhall, Robert J Beaver, Barbara M Beaver, Introduction to Probability and Statistics,Twelth edition, Thomson.

Operating Systems 4:0:0

Course Outcomes:

At the end of this course students will able to understand

CO1: Understand fundamental operating system abstractions such as processes, threads, files, semaphores, IPC abstractions, shared memory regions, etc.,
CO2: Analyze important algorithms eg. Process scheduling and memory management algorithms.
CO3: Categorize the operating system's resource management techniques, dead lock management techniques, memory management techniques
CO4: Understanding the Distributed Operating Systems

Course Content:

UNIT- I

Introduction to Operating System Concept, Types of Operating Systems, Operating Systems Concepts, Operating System Operations. Operating Systems Structures- Operating System Services, User Operating- System Interface, Introduction to System calls, Types of System Calls. Processes Management: Process Management: Process concept, Process State Diagram, Process control block, Process Scheduling, Inter process Communication, Threads- Threading Issues, Scheduling- Basic Concepts, Scheduling Criteria, Scheduling Algorithms.

UNIT- II

Process Synchronization,The Critical-Section Problem, Peterson's Solution, Synchronization Hardware, Semaphores, Classic Problems of Synchronization, Monitors. Principles of deadlock: System Model, Deadlock characterization, Deadlock handling, Deadlock Prevention, Detection and Avoidance, Recovery Starvation, Critical Regions form Deadlock.

UNIT –III

Memory Management,Swapping, Contiguous Memory Allocation, Paging, structure of the Page Table, Segmentation. Virtual Memory Management- Demand Paging, Page-Replacement Algorithms.

UNIT- IV

File-System Interface**,** File Concept, Access Methods, Directory structure, File-System mounting, Files Sharing, Protection,File-System implementation.

Distributed Operating Systems, Types of network based Operating systems, Network Structure, Network Topology, Communication Structure, Communication Protocols, Robustness, Design Issues .A review of Mobile Operating Systems, Features of Android Operating Systems.

Text Books

- Abraham Silberschatz,& Peter Baer Galvin, Gagne, Operating System Concepts, Ninth Edition, Wiley, 2015
- William Stallings, Operating Systems-Internals and Design Principles, Fifth Edition, Pearson Education, 2007
- Achyut S Godbole, Operating Systems, Second Edition, TMH, 2007 Flynn/McHoes, Operating Systems, Cengage Learning, 2008.
- Deitel & Deitel, Operating System, Third Edition, Pearson Education, 2008

PYTHON PROGRAMMING LAB 0:0:2

Course Outcome:

CO1: Understand the fundamental concepts of python and its main components.
CO2: Develop (Read and Write) python programs using variables, assignments, and conditional statements using functions.
CO3: Illustrate and implement different data structures.
CO4: Demonstrate Object-oriented concepts and file handling.
CO5: Analyze and plot data using python visualization libraries.

List of Indicative Experiments

1.Test and Debug simple Python programs
2.Different data types in python(variables constants and strings)
3.Programs on different operators
4.Control statements and Loops
5.Working on Functions
6.Data structures in python (List, Tuple, Dictionary and Set)
7.Objects and Classes manipulation using python
8.Open, Read and write data from/to files in Python
9.Different plots using Mat plot lib
10Visualization of data using sea born

Data Structures And Algorithms Lab 0:0:2

Course Outcome:

CO1: Understand the fundamental concepts of python and its main components.
CO2: Develop (Read and Write) python programs using variables, assignments, and conditional Statements using functions.
CO3: Illustrate and implement different data structures.
CO4: Demonstrate Object-oriented concepts and file handling.
CO5: Analyze and plot data using python visualization libraries.

List of Indicative Experiments

1. Stacks and Queues
2. Lists
3. Linear Search and Binary Search
4. Sorting Algorithms
5. Graph Traversal Algorithms
6. Tree Traversal Algorithms
7. Shortest Path Algorithms
8. Knapsack Problem
9. Travelling Sales man Problem
10 N-Queens's Problem

II SEMESTER

Data Processing and Visualization using Tableau 3:0:0

Course Outcomes:

At the end of the course, student will be able to:

CO1: Create basic calculations, including arithmetic calculations, custom aggregations, and ratios; and use quick table calculations.

CO2: Apply analytics using reference lines and box plots.

CO3: Create parameters to help users dynamically modify values.

CO4: Build interactive dashboards and stories to reveal data insights.

UNIT-I

Connect to and Customize Data: Connecting to Data, customizing a Data Source, working with a Data Extract, Organize Data and Create Filters: Creating Groups in Your Data, CreatingHierarchies in Your Data, Understanding Filtering in Tableau, Filtering Your Data, Sorting Your Data, Using Sets to Highlight Data.

UNIT-II

Build Common Views: Working with Dates to Visualize Time-Based Data, Creating Custom Date Fields and Hierarchies, Comparing Multiple Measures in Views, Using Scatter Plots To Show Relationships Between Measures, Creating Spreadsheet-like Views Using Text Tables, Using a Highlight Table to Show Specific Values, Showing Breakdowns of the Whole Using Pie Charts, Showing Breakdowns of the Whole Using Tree Maps, Using Bar-in-Bar Charts and Bullet Graphs to Compare Measures, Map Geographic Data: Creating Symbol and Filled Maps, Creating a Density Map

UNIT-III

Create Calculated Fields: Creating Calculated Fields for Deeper Analysis, Working with String and Type Conversion Functions, Working with Date Functions, Working with Aggregate Functions,Apply Table Calculations: Using Quick Table Calculations to Analyse Data, Create Level of Detail (LOD) Expressions: Using Level of Detail Expressions, Using Filters with LOD Expressions, Using Fixed LOD Expressions for Cohort Analysis

UNIT-IV

Apply Analytics: Highlighting Values with Reference Lines and Bands, Using Parameters to Control Data in the View, Using Histograms and Box & Whisker Plots to Show Distribution,

Use a Background Image: Module Overview, Concept: Plotting Data on a Custom Background, Demo: Adding an Image and Plotting Data, Activity: Plot Data on an Image, Create Dashboards and Stories.

References:

- Jones, Ben. Communicating data with Tableau: Designing, developing, and delivering datavisualizations. " O'Reilly Media, Inc.", 2014.
- Murray, Daniel G. Tableau your data!: fast and easy visual analysis with tableau software. JohnWiley & Sons, 2013.
- Baldwin, David. Mastering Tableau. Packt Publishing Ltd, 2016.
- urray, Daniel G. Tableau your data!: fast and easy visual analysis with tableau software. JohnWiley & Sons, 2013.

Data Mining & Data Warehousing 3:0:0

Course outcomes:
After the completion of the course, student will be able to

CO1: Acquire the knowledge of data preprocessing and data quality modeling and design of data warehouses and algorithms for data mining.

CO2: Be able to design data warehouses and apply acquired knowledge for understanding data and select suitable methods for data analysis.

CO3: Apply data mining tools and techniques for solving real-world problems and decision making.

CO4: Analyze and implement clustering techniques.

Course Content

UNIT-I

Introduction to data mining and Data Warehousing, Modeling: Data Cube and OLAP, Data Warehouse Implementation, Data Mining – types of data, types of patterns, Data cleaning, Data integration

UNIT-II

Data Reduction, Wavelet Transforms, Attribute Subset Selection, Histogram, Clustering, Sampling, Data Cube Aggregation Data Transformation: Strategies Overview, Data Transformation by Normalization

UNIT-III

Mining Frequent Patterns, Associations & Correlations: pattern evaluation methods Classification, Decision tree Induction, Attribute Selection Measures, Tree Pruning, Bayes Classification Methods

Unit-IV

Cluster Analysis: Requirement for Cluster Analysis, clustering methods Data Mining Applications & Trends:

Mining Sequence Data; Time Series, Symbolic, Statistical Data Mining, Visual Data Mining, Data Mining Applications

Text Books:

- Jiawei Micheline Kamber, 'Data Mining Concepts and Techniques', Morgan Kauf MannPublishers
- George M Marakas, 'Modern Data Warehousing, Mining and Visualization', PearsonEducation, 2003
- W H Inmon, 'Building the Data Warehouse', Wiley dreamtech, 3rd Edition
- Mastering Data Mining – Michael J A Berry & Gordon S Linoff (Wiley Pub)
- Data Warehousing (Pearson Ed) – Sam Anahory & Dennis Murray

Data Science with R Programming 4:0:0

Course outcomes:

After the completion of the course, student will be able to

CO1: Understand the processes of data science identifying the problem to be solved, data collection, preparation, modeling, evaluation and visualization.

CO2: Be aware of the challenges that arise in Data Sciences.

CO3: Be able to identify the application of the type of algorithm based on the type of the problem.

CO4: Be comfortable using commercial and open source tools such as the R/Python language and its associated libraries for data analytics and Visualization.

UNIT I:

Introduction to Data Science and Big data, Benefits and Uses, facts of Data, Data Science Process. History and Overview of R. Overview of the Data Science Process-Setting the research goal, Retrieving Data, Data Preparation, Exploration, Modeling, data Presentation and Automation. Getting Data in and out of R, Using reader package, Interfaces to the outside world.

UNIT II:

Machine Learning: Understanding why data scientists use machine learning-What is machine learning and why we should care about, Applications of machine learning in data science, The modeling process, Types of Machine Learning- Supervised and Unsupervised.

UNIT III:

Handling large Data on a Single Computer: The problems we face when handling large data, General Techniques for handling large volumes of data, Generating programming tips for dealing with large datasets.

UNIT IV:

Sub setting R objects, Vectorised Operations, Managing Data Frames with the dplyr, Control structures, functions, Scoping rules of R, Coding Standards in R, Loop Functions, Debugging, Simulation. Case studies on preliminary data analysis.

References:

- Davy Cielen, Arno.D.B.Maysman, MohamedAli, "Introducing Data Science" Manning Publications, 2016.
- Roger D. Peng, "R Programming for DataScience" Lean Publishing, 2015.
- Nina Zumel, John Mount, "Practical Data Science with R", Manning Publications, 2014.
- Tony Ojeda, Sean Patrick Murphy, Benjamin Bengfort, AbhijitDasgupta, "PracticalData Science Cookbook", Packt Publishing Ltd., 2014.

Linear Regression Models 4:0:0

Course Outcomes:

Having completed this course, the learner will be able to

CO1: Distinguish between General Linear Models and Regression Analysis
CO2: Estimate parameters of General Linear Model using least square method under various conditions on data matrix, conditions on parameters and testing of hypothesis about parameters
CO3: Derive the distributions of sum of squares due to regressors, errors for ANOVA and ANCOVA model
CO4: Fit simple and multiple linear regression to data with step by step procedure

UNIT I

Introduction of models based on nature of data; need of estimation of linear parametric functions; Gauss-Markov set-up, Normal Equations and Least Square Estimators, Error and estimation spaces, variances and covariance of least square estimators

UNIT II

Estimation of error variance, Distribution of sums of squares due to regression and errors. Least square with correlated observations, least square estimators with restriction on parameters, simultaneous estimators of linear parametric functions.

UNIT III

Tests of hypotheses for one and more than one linear parametric functions, confidence intervals and regions, Analysis of Variance, Analysis of Covariance.Multiple comparison tests due to Tukey and Scheffe, simultaneous estimation of linear parametric functions.

UNIT IV

Introduction to one-way random effects linear models and estimation variance components.Multiple Linear Regression, Regression model building – best subset selection, stepwise regression, residual analysis, Model selection criteria.

Text Books:

- Weisberg, S. (2005). Residual and Influence in Regression. Wiley Series in
- Probability and Statistics, Wiley
- Draper, N. R. and Smith, H. (1998).Applied Regression Analysis. Third Edition, Wiley
- Gunst, R. F. and Mason, R. L. (1980). Regression Analysis and its Applications –A Data
- Oriented Approach. Macel and Dekker.
- Rao, C. R. (2001) Linear Statistical Inference and its Applications. Ed. II, Wiley Eastern.
- Weisberg's. (1985). Applied Linear Regression. Wiley Series in Probability and Statistics, Wiley
- Gujarathi, D. N. and Sangeetha (2007). Basic Econometrics, Ed. IV , Tata MacGraw Hill

Design and Analysis Algorithms 4:0:0

Course outcomes:

After the completion of the course, student will be able to

CO1: Analyze different scenarios for running time of algorithms using asymptotic notations and Design using Recursion.

CO2: Apply divide and conquer strategy for design of various algorithms.

CO3: Describe and apply dynamic-programming approach for designing graph and matrix based algorithms

CO4: Implementation of graph and matching algorithms Course s Content.

UNIT- I

Introduction: Algorithm specification, pseudo code conventions

Performance Analysis: Space Complexity, Time Complexity, Asymptotic Notation,

Mathematical Analysis: Recursive and Non recursive algorithms

BRUTE FORCE – Bubble Sort, Selection Sort, Sequential Search, String Matching

UNIT-II

DIVIDE- AND – CONQUER: General Method, Binary Search, Finding the Maximum and Minimum, Merge Sort, Quick Sort, Strassen's Matrix Multiplication

THE GREEDY METHOD: The General Method, Tree Vertex Splitting, Job Sequencing with Deadlines, Minimum-Cost Spanning Trees - Prim's Algorithm, Kruskal's Algorithm, Optimal Storage on Tapes, Optimal Merge Patterns.

UNIT-III

DYNAMIC PROGRAMMING: The General Method, Multistage Graphs, All Pairs Shortest Paths Single-Source Shortest Paths: General Weights, String Editing, 0/1 Knapsack, the Traveling Salesperson Problem
BACKTRACKING: The General Method, the 8-Queens Problem, Sum of Subsets, Graph Coloring, Hamiltonian Cycles

UNIT-IV

Depth First search and Breadth First Search, Topological sorting, Pre sorting, Balanced Search Trees, Heaps and Heap sort
NP-Hard and NP-Complete Problems: Basic Concepts, Nondeterministic Algorithms, The Classes NP-Hard And NP-Complete ,NP-Hard Graph Problems

References:

- Herbert Scheldt: The Complete Reference C++, 6th Edition, Tata McGraw Hill2013
- Anany Levitin: Introduction to the Design and Analysis of Algorithms, PearsonEducation,2003
- CoremenT H ,LeisersonC E ,andRivestR L :IntroductiontoAlgorithms,PHI,1998
- Horowitz E , Sahani S , Rajasekharan S : Computer Algorithms, GalgotiaPublications, 2001

Time Series Analysis 4:0:0

Course outcomes:
After the completion of the course, student will be able to
CO1: Understand the fundamentals of time series forecasting and explain the concepts of MMSE forecasts, ARIMA forecasting, and forecast functions.
CO2: Compute and interpret one-step and multi-step forecasts using appropriate time series models and update forecasts as new observations become available.
CO3: Perform diagnostic checking of fitted time series models using residual analysis and statistical tests to assess model adequacy.
CO4: Analyze real-world time series data, develop forecasting models, and interpret the results for decision-making and research purposes.

UNIT-I:

Stochastic processes , The Autocovariance and Autocorrelation Function ,The Partial Autocorrelation Function ,White Noise Processes , Estimation of the Mean , Autocovariance and Autocorrelation , Moving Average and Autoregressive Representations of Time Series Processes , Linear Difference Equations .

UNIT-II

Autoregressive Processes , Moving Average processes , The Dual Relationship Between AR(p) and MA(q) Processes , Autoregressive Moving Average ARMA(p,q) Processes Nonstationarity in the mean , Autoregressive Integrated Moving Average (ARIMA) Models , Nonstationarity in the variance and the Autocovariance .

UNIT-III

Introduction , Minimum Mean Square Error Forecasts , Computation of Forecasts, The ARIMA Forecast as a Weighted Average of Previous observations , updating Forecasts , Eventual Forecasts Functions ,A Numerical Example.

UNIT-IV

The Method of Moments , Maximum Likelihood Method , Nonlinear Estimation , Ordinary least Squares(OLS) Estimation in Time Series Analysis , Diagnostic Checking ,Empirical Examples for Series w1-w7 , Model Selection Criteria .

Textbook :

- William W.S.Wei –Department of Statistics , the Fox School of Business and Mangement , temple university - Time Series Analysis –Univariate and Multivariate Methods –Second edition

III Semester

Block chain Technology 4:0:0

Course outcomes:

After the completion of the course, student will be able to

CO1: To Understand Blockchain terminologies with its applications. Design

CO2: To learn working principles of Blockchain and methodologies used in Bitcoin

CO3: To gain knowledge on Ethereum Network, Wallets, Nodes, Smart contract & DApps

CO4: To learn blockchain Based Application Architecture using Hyperledger and the Smart Contract Lifecycle

UNIT-I

Distributed systems, CAP theorem, Byzantine Generals problem, Consensus. The history of blockchain, Introduction to blockchain, Various technical definitions of blockchains, Generic elements of a blockchain, Features of a blockchain, Applications of blockchain technology, Tiers of blockchain technology, Consensus in blockchain, CAP theorem and blockchain, Benefits and limitations of blockchain.

UNIT -II

Decentralization using blockchain, Methods of decentralization, Blockchain and full ecosystem decentralization, Smart contract, Decentralized organizations, Decentralized autonomous organizations, Decentralized autonomous corporations, Decentralized autonomous societies Decentralized applications, Platforms for decentralization. Cryptographic primitives: Symmetric cryptography, Asymmetric cryptography, Public and private keys, Hash functions: Compression of arbitrary messages into fixed length digest, Easy to compute, Pre-image resistance, Second pre-image resistance, Collision resistance, Message Digest (MD),Secure Hash Algorithms (SHAs), Merkle trees, Patricia trees, Distributed hash tables (DHTs), Digital signatures, Elliptic Curve Digital signature algorithm (ECDSA).

UNIT -III

Bitcoin, Bitcoin definition, Transactions, The transaction life cycle, The transaction structure, Types of transaction, The structure of a block , The structure of a block header, The genesis block, The bitcoin network, Wallets, Smart Contracts-History, Definition, Ricardian contracts, Smart contract templates, Oracles, Smart Oracles, Deploying smart contracts on a blockchain, The DAO.

UNIT -IV

Ethereum 101, Introduction, Ethereum clients and releases, The Ethereum stack, Ethereum blockchain, Currency (ETH and ETC), Forks, Gas, The consensus mechanism, The world state, Transactions, Contract creation transaction, Message call transaction, Elements of the Ethereum blockchain , Ethereum virtual machine (EVM), Accounts, Block, Ether, Messages, Mining, The Ethereum network. Hands-on: Clients and wallets –Geth.

References

- Imran Bashir. “Mastring BlockChain”, Third Edition, Packt – 2020.
- Andreas M. , Mastering Bitcoin: Programming the Open Blockchain – O’rielly – 2017.

Cyber Security 4:0:0

Course outcomes:

After the completion of the course, student will be able to

CO1: Explain the fundamentals of information systems, information security, cyber security, information assurance, and security risk analysis.

CO2: Apply security technologies and mechanism and data protection techniques to secure information assets.

CO3: Evaluate and implement organizational security policies, standards, and best practices for maintaining confidentiality, integrity, and availability of information.

CO4: Interpret and apply legal, ethical, and regulatory requirements related to information security.

UNIT-I

Introduction to information systems, Types of information Systems, Development of Information Systems, Introduction to information security, Need for Information security, Threats to Information Systems, Information Assurance, Cyber Security, and Security Risk Analysis.

UNIT -II

Application security (Database, E-mail and Internet), Data Security Considerations-Backups, Archival Storage and Disposal of Data, Security Technology-Firewall and VPNs, Intrusion Detection, Access Control.

Security Threats -Viruses, Worms, Trojan Horse, Bombs, Trapdoors, Spoofs, E-mail viruses, Macro viruses, Malicious Software, Network and Denial of Services Attack, Security Threats to E-Commerce.

UNIT-III

Developing Secure Information Systems, Application Development Security, Information Security Governance & Risk Management, Security Architecture & Design Security Issues in Hardware, Data Storage & Downloadable Devices, Physical Security of IT Assets, Access Control, CCTV and intrusion Detection Systems, Backup Security Measures.

UNIT-IV

Security Policies, WWW policies, Email Security policies, Information Security Standards-ISO, IT Act, Copyright Act, Patent Law, IPR. Cyber Laws in India; IT Act 2000 Provisions, Intellectual Property Law: Copy Right Law, Software License.

Text Books:

1. Charles P. Pfleeger, Shari Lawerance Pfleeger, "Analysing Computer Security ", Pearson Education India.

2. V.K. Pachghare, "Cryptography and information Security", PHI Learning Private Limited, Delhi India. 3.Dr. Surya Prakash Tripathi, Ritendra Goyal, Praveen kumar Shukla ,"Introduction to Information Security and Cyber Law" Willey Dreamtech Press.

4. CHANDER, HARISH," Cyber Laws And It Protection " , PHI Learning Private Limited ,Delhi ,India

Network Security And Cryptography 4:0:0

Course outcomes:

After the completion of the course, student will be able to

CO1: Explain the fundamental concepts of network security.

CO2: Analyze and implement authentication mechanisms to ensure data integrity and authenticity.

CO3: Examine web security technologies to secure web-based applications and electronic commerce.

CO4: Design and assess system security solutions to protect networks from cyber threats

UNIT-I

Introduction : OSI Security Architecture, Security Attacks, Security Services, Security Mechanism, Model for Network Security. Classical Encryption Technique: Symmetric Cipher Model, Substitution Techniques, Transposition Techniques.

Block Ciphers, Data Encryption Standard and Advanced Encryption Standard: Block Cipher Principles, The Data Encryption Standard, Block Cipher Design Principles and Modes of operation,

UNIT-II

Public Key Cryptography and Key Management: Principles of Public Key Cryptosystem, RSA algorithm, Key management, Diffie Hellman Key exchange Message Authentication and Hash Function: Authentication Requirement, Authentication Functions, Message Authentication Code, Hash Functions, Digital Signatures, Digital Signature Standard. Authentication Applications : Kerberos, X.509 Authentication Service.

UNIT-III

Electronic Mail Security : Pretty Good Privacy (PGP), S/MIME; IP Security: IP Security Overview; IP Security Architecture; Authentication Header; Encapsulating Security Payload; Combining Security Associations; Key Management.

UNIT-IV

Web Security : Web security Considerations; Secure Socket layer (SSL) and Transport layer Security (TLS); Secure Electronic Transaction (SET) System Security: Intruders, Intrusion Detection, Firewall Design Principles- Characteristics, Types of Firewall and Firewall Configuration.

References:

1. William Stallings, “Cryptography and Network Security – Principles and Practices”, 4th Edition, Pearson Education, 2009.

Data Science Using Python 0:2:0

Course outcomes:

After the completion of the course, student will be able to

CO1: Demonstrate proficiency in Python programming by applying control structures, functions, object-oriented programming concepts, and built-in data structures .

CO2: Explain the fundamentals of data science, data analytics, and descriptive statistics, and perform statistical analysis.

CO3: Apply data preprocessing techniques to prepare datasets for analysis.

CO4: Develop data-driven solutions by integrating statistical analysis, database management, and basic machine learning functionalities using Python.

UNIT I

Getting started with Python: Basics of Python including data types, operators, variables, expressions, control structures, conditional statements, looping structure, functions, OOP concepts. Python sequence data structures including String, Array, List, Tuple, Set, and Dictionary and associated operations

UNIT II

Data Science and Descriptive Statistics: Understanding data, its type and data science, Considering the emergence and fundamentals of data science, Data science life cycle, Data analysis (Univariate, bivariate and Multivariate), Data analytics, Data measurement scale, Data descriptive statistics (Measures of central tendency, dispersion/variation, measure of location, Shape and symmetry), Understanding Python's role in data science.

UNIT III

Exploring Python libraries: Reading from a text file, CSV file, excel file, Streaming, Sampling and uploading data, Managing data with relational database (MySQL), Exploring statistics of data using stats package from SciPy, Numerical computing and working with multidimensional array through NumPy, Data manipulation through Pandas, Basic functionalities of machine learning using Scikit-learn, Extracting components from HTML document through beautiful soup.

UNIT IV

Data preparation: Data loading, Data cleaning, dealing with missing data, removing duplicates, Slicing and Dicing, Filtering, and selecting data, Concatenating and transforming, adding new cases and variables, removing data, Sorting and shuffling, Aggregating data, Handling outliers, Data wrangling, Data Normalization.

Data Visualization: Visualizing data through figure, subplot and its properties, graphs, plots in matplotlib, basics of advanced visualizations with Seaborn.

Text Books:

- Python for data science for dummies, John Paul Mueller, Luca Massaron, and Wiley
- Programming through Python, M. T. Savaliya, R. K. Maurya, G. M. Magar, STAREDU Solutions
- Pandas for everyone: Python Data Analysis, Daniel Y. Chen, Pearson
- Introducing Data Science: Big Data, Machine Learning, and More, Using Python Tools, Davy Cielen, Arno D.B. Meysman, et a1., Minning
- Applied Data Science with Python and Jupyter, Alex Galea,
- Business Analytics: The Science of Data - Driven Decision Making, U Dinesh Kumar, Wiley India.

MongoDB 3:0:0

Course outcomes:

After the completion of the course, student will be able to

CO1: Understand the fundamentals of NoSQL databases and MongoDB and CRUD operations.

CO2: Apply MongoDB query techniques, data modeling principles, schema design practices, indexing methods, and aggregation framework for efficient data retrieval and analysis.

CO3: Analyze and implement MongoDB replication and sharding mechanisms to ensure high availability, fault tolerance, and horizontal scalability in distributed database environments.

CO4: Demonstrate MongoDB administration skills by implementing security mechanisms and performing backup and restore operations for database management and recovery.

UNIT-I

Overview of MongoDB and NoSQL databases, Installing and setting up MongoDB, Introduction to BSON (Binary JSON) data format, CRUD Operations in MongoDB Performing CRUD operations (Create, Read, Update, Delete) in MongoDB, Using the MongoDB shell for interacting with databases, Working with collections and documents

UNIT-II

Querying data using find() method, Query operators and expressions in MongoDB, Working with nested documents and arrays, Data Modeling and Schema Design, Understanding data modeling concepts in MongoDB, Schema design best practices.

Introduction to indexing in MongoDB, Types of indexes, Performance optimization techniques, Aggregation Framework Overview of MongoDB aggregation framework, Using aggregation pipeline for data transformation and analysis, Aggregation operators and stages.

UNIT-III

Understanding replica sets and high availability in MongoDB, Configuring and managing replica sets, Automatic failover and recovery, Sharding Introduction to sharding in MongoDB for horizontal scaling, Configuring and managing sharded clusters, Shard key selection and distribution.

UNIT -IV

Securing MongoDB deployments, Authentication and authorization mechanism, Role-based access control (RBAC), Backup and Restore Backup strategies for MongoDB deployments, Using mongodump and mongorestore utilities, Configuring automated backups

References:

"MongoDB: The Definitive Guide" by Shannon Bradshaw, Kristina Chodorow, and Michael Dirolf

React JS 4:0:0

Course outcomes:

After the completion of the course, student will be able to

CO1: Explain the fundamentals of React JS, ES6 features, JSX, Virtual DOM, and develop reusable React components using modern JavaScript concepts.

CO2: Design interactive user interfaces by implementing React components, managing data.

CO3: Develop React applications using Hooks, forms, routing, and integrate external APIs..

CO4: Build, manage, and deploy complete React applications by utilizing Context API for state management.

UNIT-I

Introduction to React and its features, Component-based UI architecture, Virtual DOM and React DOM, JSX syntax and expressions, Setting up React using Create React App,ES6 fundamentals: let and const, Arrow functions, Classes, Modules, Destructuring, Introduction to Functional and Class Components,Component rendering and structure

UNIT-II

Props and data passing between components, State management in components, Event handling in React, Conditional rendering, Rendering lists and using keys, Dynamic UI updates, Component communication, Best practices for reusable components

UNIT -III

React Hooks such as use State and use Effect, Forms and controlled components, Form validation techniques, React Router and navigation between pages, Introduction to API integration Fetch API and Axios, Async/Await concepts, Handling loading states and errors.

UNIT-IV

Introduction to global state management. Context API and state sharing, Project structure and organization, UI styling techniques in React, Integrating API data into applications, Development of a simple React mini project, Testing and debugging basics, Deployment of React applications using: Netlify, Other cloud deployment platforms.

References

- Stoyan Stefanov, *React Up & Running*, O'Reilly.
- Alex Banks & Eve Porcello, *Learning React*, O'Reilly.
- Kirupa Chinnathambi, *Learning React: A Hands-On Guide*, Addison-Wesley.
- Robin Wieruch, *The Road to React* (Online Book).
- Dave Ceddia, *Pure React* (Online Book).
- Official Documentation: React.js – react.dev

SOCIAL NETWORK ANALYSIS 4:0:0

Course outcomes:

After the completion of the course, student will be able to

CO1: Explain the foundations of Social Network Analysis and levels of network analysis in social and behavioral sciences.

CO2: Identify, collect, measure, and evaluate different types of social network data, applying appropriate methods for data collection, sampling, validity, reliability, and longitudinal analysis.

CO3: Apply graph-theoretic, sociometric, and algebraic notations to represent social network data and analyze network .

CO4: Analyze cohesive subgroups in social networks using clique-based and connectivity-based approaches, interpret subgroup structures.

UNIT-I

The Social Networks Perspective, Historical and Theoretical Foundations, Empirical Motivations, Theoretical Motivations, Mathematical Motivations, Fundamental Concepts in Network, Distinctive Features, Complexity, Descriptive and Statistical Methods, Theory Driven Methods, Chronology, Levels of Analysis.

UNIT-II

Introduction to Network Data, Structural and Composition Variables, Modes, Affiliation Variables , Boundary Specification and Sampling, Types of Networks, Network Data, Measurement and Collection, Measurement, Collection, Longitudinal Data Collection, Measurement Validity, Reliability, Accuracy.

UNIT-III:

Data , Graph Theoretic Notation , A Single Relation, Multiple Relations, Sociometric Notation, Single Relation, Multiple Relation, Algebraic Notation, Two Sets of Actors, Different Types of Pairs, Sociometric Notation,Prominence: Centrality and Prestige , Actor Centrality, Actor Prestige, Group Centralization and Group Prestige, Non directional Relations, Degree Centrality, Closeness Centrality,

UNIT-IV:

Background, Social Group and Subgroup, Notation, Subgroups Based on Complete Mutuality, Definition of a Clique, An Example. Considerations, Reachability and Diameter, n-cliques , An Example. Considerations, n-clans and n-clubs, Subgroups Based on Nodal Degree, k-plexes, k-cores, Comparing Within to Outside Subgroup Ties, LS Sets, Lambda Sets, Measures of Subgroup Cohesion, Directional Relations, Cliques Based on Reciprocated Ties, Connectivity in Directional Relations, n-cliques in Directional Relations, Valued Relations, Cliques, n-cliques, and k-plexes . Other Approaches for Valued Relations, Interpretation of Cohesive Subgroups, Other Approaches, Matrix Permutation Approaches, Multidimensional Scaling, Factor Analysis.

References:

- Social network analysis – methods and applications., Stanley Wasserman and Katherine Faust

APPLIED DATA ANALYTICS 4:0:0

Course outcomes:

After the completion of the course, student will be able to

CO1: Apply descriptive and inferential statistical techniques to analyze data.

CO2: Create and interpret effective data visualizations and dashboards using appropriate charts, plots, and visualization techniques.

CO3: Implement and evaluate data mining techniques for prediction and decision-making tasks.

CO4: Analyze datasets using clustering and association rule mining techniques, including hierarchical clustering, k-means clustering, Apriori algorithm, and measures of cluster and rule quality.

UNIT-I

Basic Concepts ,Descriptive Statistics, Statistical Inferences Data Measurement Measures of Central Tendency and Dispersion Common Statistical Graphs Determination of Outliers Statistical Inferences, Point Estimation and Required Properties of Point Estimators, Interval Estimations for Mean, Proportion and Variance of Population Sample Size Determination Hypothesis Testing 1. Hypothesis Testing for Mean, Proportion and Variance of Population – Single Sample Test, Hypothesis Testing for Mean, Proportion and Variance of Population – Two Samples Test Type I and Type II Errors – Power of the Test Observed Significance Level

UNIT-II

Data Visualization, Introduction to Data Visualization, Basic Charts for Numerical Data and Categorical Data, Distribution Plots, Multivariate Charts: Combo Chart, Combination Chart, Stacked Column Chart, Data Dashboard, What is a Data Dashboard, Applications and Benefits of Data Dashboard, Design and Construct a Data Dashboard

.

UNIT-III

Regression Analysis, Linear Regression and Least Square Method, Residual Analysis , Multiple Regression, Goodness of Fit Tests, Data Classification, k-Nearest Neighbor Algorithm for Estimation and Prediction, Distance Functions: Euclidian, Manhattan, Minkowski, Min-Max Normalization, Z-Score Standardization, Logistics Regression , Bayesian Networks, Model Evaluation Measures for Classification Task VIII.

UNIT-IV

Hierarchical Clustering Method , k-Means Clustering, Measuring Cluster Goodness: The Silhouette Method and The Pseudo-F Statistic , Association Rules, Affinity Analysis, The a Priori Algorithm – Generating Frequent Itemsets , The a Priori Algorithm – Generating Association Rules, Measure the Usefulness of Associate Rules

References:

- Intelligent Data Analysis by Michael Berthold and David J. Hand.
- Data Science and Big data Analytics by David Dietrich, Barry Heller, and Beibei Yang.

BIO INFORMATICS DATABASES 4:0:0

Course Outcome:

CO1: Understand biological data types and bioinformatics databases.

CO2: Analyze genome databases and model organism data.

CO3: Apply sequence and protein database concepts.

CO4: Use bioinformatics search tools and resources effectively.

UNIT-I

Types of biological data – Genomic DNA, recombinant DNA, ESTs, STS – Primary databases – GenBank, EMBL, DDBJ – Composite databases – UniProt – Literature databases – PubMed, PLoS – Bioinformatics resources – NCBI, EBI, ExPASy, RCSB.

UNIT-II

Viral genome database:-ICTVdb; Bacterial Genomes database:- Genomes OnLine Database – GOLD, Microbial Genome Database-MBGD; Genome Browsers:- Ensembl, VEGA genome browser, NCBI-NCBI map viewer, KEGG, MIPS, UCSC Genome Browser; Archeal Genomics, Eukaryotic genomes with special reference to model organisms:- Yeast(SGD), Drosophila (FlyBase), C.elegans (WormBase), Rat, Mouse, Human (OMIM / OMIA), plants – Arabidopsis thaliana (TAIR), Rice, PlasmodiumDB, etc.

UNIT-III

Nucleotide sequence Databases:- GenBank, EMBL, DDBJ; Protein sequences Databases:- Swiss-Prot, TrEMBL, UniProt, UniProtKB, UniParc, UniRef, UniMES; Sequence motifs Databases:- Prosite, ProDom, Pfam, InterPro, Gene Ontology; Sequence file formats:- GenBank, FASTA, PIR, ALN/ClustalW2.

UNIT-IV

Structure databases – PDB, SCOP, CATH – Protein interaction databases – STRING – Gene expression databases – GEO – Bioinformatics search tools – BLAST, FASTA – Motif search – ScanProsite – Structural search tools – VAST, DALI – Proteomics tools – ExPASy, EMBOSS.

Text Books:

- Bioinformatics: Sequence and Genome Analysis by Mount D., Cold Spring Harbor Laboratory Press, New York. 2004
- Bioinformatics- a Practical Guide to the Analysis of Genes and Proteins by Baxevanis, A.D. and Francis Ouellellette, B.F., Wiley India Pvt Ltd. 2009
- Introduction to bioinformatics by Teresa K. Attwood, David J. Parry-Smith. Pearson Education. 1999

Web Designing 4:0:0

Course Outcome:

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

CO1: Apply a structured approach to identifying needs, interests, and functionality of a website,design dynamic websites that meet specified needs and interests.

CO2: Write well-structured, easily maintained, standards-compliant CSS code to present HTML pages in different ways.

CO3: Use JavaScript to add dynamic content to pages, select appropriate HTML, CSS.

CO4: Write well-structured, easily maintained JavaScript code following accepted good practice.Course Content:

UNIT-I

Fundamentals of Web: Internet, WWW, Web Browsers, and Web Servers, URLs, MIME, HTTP, Security, the Web Programmers Toolbox. HTML: Origins and evolution of HTML, Basic syntax, Standard HTML document structure, Basic text markup, Hypertext Links, Lists,Tables, Forms

UNIT-II

Cascading Style Sheets (CSS): Introduction, Levels of style sheets, Style specification Formats, Selector forms, Property value forms, Font properties , List properties , Color , Alignment of text, The Box model,Background images, The<space> and tags

UNIT-III

JavaScript: Overview of JavaScript, Object orientation and JavaScript, General syntactic Characteristics, Primitives, operations, and expressions, Screen output and keyboard input,Control statements, Object creation and modification, Arrays, Functions,Constructor, Pattern Matching using regular expressions, Errors in scripts, Examples.

UNIT-IV

Introduction to PHP : Origins and uses of PHP, Overview of PHP, Arrays, Functions, Pattern matching, Form handling, Files, Cookies, Session tracking .

XML: Introduction, Syntax, Uses, Document structure, Namespaces, XML schemas, displaying raw XML documents, Displaying XML documents with CSS, XSLT style sheets, XML processors, Web services.

Text Books:

- Programming the World Wide Web – Robert W. Sebesta, 4th Edition, Pearson Education, 2008.
- Internet & World Wide Web How to program – M. Deitel, P.J. Deitel, A. B. Goldberg, 3rd Edition, Pearson Education / PHI, 2004.
- Web Programming Building Internet Applications – Chris Bates, 3rd Edition, Wiley India, 2006.
- The Web Warrior Guide to Web Programming – XueBai et al, Thomson, 2003.

IV Semester

Dynamic Web Programming using Django and Flask 0:2:0

Course Outcomes: On successful completion of course, students should be able to

CO1: Recall Evolution of Web Development, Basics of HTML, CSS, and JavaScript, Server-Side Development Basics

CO2: Demonstrate advanced JavaScript and Frontend Frameworks, Database Integration

CO3: Apply and integrate server-side frameworks, Security in Web Development.

CO4: Analyze Security in Web Development, showcasing synthesis and evaluation in web development. CO5: Evaluate RESTful APIs, emerging trends in dynamic web development and DOM manipulation.

UNIT-I

Evolution of Web Development-Static vs. Dynamic Websites, Emergence of Dynamic Content, Client-Side vs. Server-Side Programming- Roles and Responsibilities, Communication between Client and Server, Basics of HTML, CSS, and JavaScript- HTML Structure and Tags, CSS Styling and Layout, JavaScript Fundamentals ,Setting up a Development Environment - Code Editors (e.g., Visual Studio Code) ,Local Web Servers (e.g., Node.js). Server-Side Development Basics: Server-Side Scripting Languages-Overview of PHP, Python.

UNIT-II

Document Object Model (DOM)-Selecting and Modifying DOM Elements, Creating and Appending Elements, Introduction to Frontend Frameworks: Overview of Frontend Frameworks- React, Angular, Vue.js.

Server-Side Frameworks:Introduction to Server-Side Frameworks-Express.js, Django, Flask,Routing and Middleware in Server-Side Frameworks-Defining Routes and Handling HTTP Methods,Implementing Middleware for Request Processing-Handling Requests and Responses Processing Client Requests, Constructing Server Responses, Building a Basic Server-Side Application

UNIT-III

Introduction to RESTful Architecture-Principles and Constraints, RESTful API Design Best Practices, Creating APIs with Server-Side Frameworks- Defining Endpoints and Methods ,Handling API Requests and Responses, Consuming APIs on the Client Side-Making API Requests from Frontend Applications, Handling API Responses and Errors, Authentication and Authorization in APIs- Token-Based Authentication.

UNIT-IV

Progressive Web Apps (PWAs)- Offline Capabilities, Push Notifications, Web Assembly and its Applications-Running Native Code in Browsers, Serverless Architecture-Functions as a Service (FaaS),Benefits and Use Cases,Future Trends in Dynamic Web Development- Voice Interfaces, Artificial Intelligence in Web Development

Text Books

- Steve Holden Publisher Python Web Programming" New Riders,2022
- Miguel Grinberg, Flask Web Development; Developing Web Applications with Python O'Reilly Media.2020 2 William S. Vincent,
- Django for Beginners: Build websites with Python and Django" William S. Vincent.2022

Web Scraping with Python 4:0:0

Course Outcomes

CO1: Apply Python programming concepts for data handling.

CO2: Use data preprocessing and wrangling techniques.

CO3: Extract data from web sources using scraping tools.

CO4: Analyze numerical data using Python libraries.

UNIT-I :

Parts of Python Programming Language, Identifiers, Keywords, Statements and Expressions, Variables, Operators, Precedence and Associativity, Data Types, Indentation, Comments, Program Execution, Reading Input, Print Output, Type Conversions, Control Flow Statements, The if Decision Control Flow Statement, The if…else Decision Control Flow Statement, The if…elif…else Decision Control Statement, Nested if Statement, The while Loop, The for Loop, The continue and break Statements

UNIT-II :

Creating and Storing Strings, Basic String Operations, Accessing Characters in String by Index Number, String Slicing and Joining, String Methods, Formatting Strings, Lists-Creating Lists, Basic List Operations, Indexing and Slicing in Lists, Built-In Functions Used on Lists, List Methods. Sets, Tuples and Dictionaries. Files: reading and writing files. Class Definition – Constructors – Inheritance – Overloading

UNIT-III :

Loading from CSV files, Accessing SQL databases. Cleansing Data with Python: Stripping out extraneous information, Normalizing data AND Formatting data. Combining and Merging Data Sets – Reshaping and Pivoting – Data Transformation – String Manipulation, Regular Expressions.

UNIT-IV :

Data Acquisition by Scraping web applications – Submitting a form - Fetching web pages – Downloading web pages through form submission – CSS Selectors. NumPy Essentials: TheNumPy, Matplotlib package – Plotting Graphs – Controlling Graph – Adding Text – More Graph Types – Getting and setting values – Patches.

References:

- Allen B. Downey, "Think Python: How to Think Like a Computer Scientist,,,,, 2nd edition.
- Guido van Rossum and Fred L. Drake Jr, ―An Introduction to Python – Revised and updated for Python 3.2, Network Theory Ltd., 2011.
- Jake Vander plas, "Python Data Science Handbook: Essential tools for working with data", O,,Reilly Publishers
- Mark Lutz, "Programming Python", O'Reilly Media, 4th edition, 2010.
- Tim Hall and J-P Stacey, "Python 3 for Absolute Beginners", Apress, 1st edition, 2009.
- Shai Vaingast, "Beginning Python Visualization Crafting Visual Transformation Scripts", Apress, 2nd edition,
- Wes Mc Kinney, "Python for Data Analysis", O'Reilly Media, 2012.

Deep Learning with PyTorch 4:0:0

Course Outcome:

CO1: Understand deep learning concepts and PyTorch framework.

CO2: Apply tensor operations for data representation.

CO3: Use pretrained models and deep learning architectures.

CO4: Develop and train neural network models using PyTorch.

UNIT-I :

The deep learning revolution , PyTorch for deep learning , Why PyTorch, The deep learning competitive landscape, An overview of how PyTorch supports deep learning projects, Hardware and software requirements.

UNIT-II :

A pretrained network that recognizes the subject of an image, Obtaining a pretrained network for image recognition , AlexNet , ResNet, Ready, set, almost run Run!, A pretrained model that fakes it until it makes it, The GAN game, CycleGAN, A network that turns horses into zebras, A pretrained network that describes scenes, NeuralTalk, Torch Hub

UNIT-III :

Multidimensional arrays, From Python lists to PyTorch tensors, Constructing our first tensors, The essence of tensors, Indexing tensors, Named tensors, Tensor element types . Specifying the numeric type with d type A d type for every occasion, Managing a tensor's dtype attribute, The tensor API Tensors: Scenic views of storage.

Real-world data representation using tensors Working with images , Adding color channels Loading an image file, Changing the layout, Normalizing the data, 3D images: Volumetric data, Loading a specialized format, Representing tabular data, Using a real-world dataset, Loading a wine data tensor, Representing scores, One-hot encoding

UNIT-IV :

The case for convolutions, What convolutions do, Convolutions in action, Padding the boundary, Detecting features with convolutions, Looking further with depth and pooling, Putting it all together for our network, Subclassing nn.Module, Our network as an nn.Module. How PyTorch keeps track of parameters and submodules, The functional API Training our convent, Measuring accuracy, Saving and loading our model, Training on the GPU. Model design Adding memory capacity: Width, Helping our model to converge and generalize: Regularization, Going deeper to learn more complex structures: Depth.

References:

- TextBook Author: Eli Stevens Luca Antiga Thomas Viehmann Foreword by Soumith Chintala

Machine Learning 4:0:0

Course Outcome

CO1: Gain knowledge about basic concepts of Machine Learning
CO2:Identify machine learning techniques suitable for a given problem
CO3: Solve the problems using various machine learning techniques
CO4: Design and implement machine learning solutions to classification, regression, and clustering problems; and be able to evaluate and interpret the results of the algorithms

UNIT-I

Introduction, Machine learning definition, importance of machine learning, machine learning framework, types of machine learning, relation to other fields, examples of machine learning applications, designing a learning system, issues in machine learning

.

UNIT-II

Introduction to Supervised Learning, Decision tree based classifier, Bayesian theory based classifier, Neural network based classifier, Nearest neighbour classifier, Support vector classifier, performance evaluation

UNIT-III

Introduction to Unsupervised Learning, Clustering methods, Criteria functions for clustering, Similarity measures, Component analysis, Low dimensional analysis and multidimensional scaling.

Unit-IV

Additional topics, Reinforcement learning, Genetic algorithms, Analytical learning, Ensemble of classifiers, Design and analysis of machine learning experiments

References

- Machine Learning: a Probabilistic Perspective by Kevin Patrick Murphy, MIT Press, March 2014
- Introduction to Machine Learning by Alex Smola and S V N Vishwanathan, Cambridge University Press
- Understanding Machine Learning: From Theory to Algorithms by Shai Shalev-Shwartzand Shai Ben-David

Design Thinking 4:0:0

Course Outcome:

CO1: Understand principles and stages of design thinking.
CO2: Apply observation and problem definition techniques.
CO3: Use ideation and creativity tools for solution generation.
CO4: Apply design thinking to real-world case studies**.**

UNIT-I

Introduction to design thinking for engineering, essential design thinking skills, core principles of design thinking, foundation of design thinking, building an effective design thinking team, discussion of problem from IT sector.

UNIT-II

Stages of design thinking, roles of observation in understanding product and process challenges, preparation of AEIOU and other canvas with defined problem statement, exploring design thinking in innovation – understanding the role of design thinking in product and process innovation, differentiating engineering design and design thinking.

UNIT-III

Understanding listening and empathizing techniques, exploring observation methods, utilizing a structured, open – ended approach for effective communication, Ideation Tools, Brainstroming techniques for idea generation, applying innovation to foster creativity, application of SCAMPER technique to an approved problem statement.

UNIT-IV

Methods and tools for design thinking practice, empathy map, storytelling, critical items diagram, mind map, journey map. Prototype- exploration map, minimum viable product(MVP), feasibility testing, viability testing, A/B Testing, Sustainability Testing, collect feedback, iterate and improve the ideas. PDC canvas verification.

Quick-Commerce Applications – food and grocery delivery (eg, swiggy, instamart), Ride sharing(eg. Uber, Ola, RApido). E-Commerce Applications – Online Shopping(eg. Amazon).

Text Books:

- Soni, Pavan. Design your Thinking : The Mindsets, Toolsets and skill sets for creative problem – solving Penguin Random House India Private Limited, 2020
- Tim Brown, Change by Design: How Design thinking Transforms organizations and inspires innovation, HarperCollins e- books 2009
- Lockwood, Thomas. Design Thinking: Integrating innovation, customer experience and brand value simon and Schuster, Allworth Press 2010

Big Data Analytics 4:0:0

Course Outcome

CO1: Understand the key issues in big data management and its associated applications inintelligent business and scientific computing

CO2: Acquire fundamental enabling techniques and scalable algorithms like Hadoop, MapReduce and NO SQL in big data analytics

CO3: Interpret business models and scientific computing paradigms, and apply software tools for big data analytics

CO4: Achieve adequate perspectives of big data analytics in various applications like recommender systems, social media applications etc

UNIT I

Importance of Big data, Risks of big data, structure of big data- exploring big data, most big data doesn't matter- filtering big data effectively -mixing big data with traditional data- the need for standards-today's big data is not tomorrow's big data Web data: the original big data-web data overview.Data Analysis-Evolution of analytic scalability – convergence–parallel processing systems –cloud computing–grid computing–map reduce–enterprise analytics and box–analytic data sets

UNIT II

Introduction to streams concepts, stream data model and architecture, stream computing, sampling data in a stream, filtering streams, counting distinct elements in a stream, estimating moments, counting oneness in a window, decaying window

UNIT III

Mining frequent item sets -– apriori algorithm –clustering techniques – hierarchical – k- means – clustering high dimensional data –frequent pattern based clustering methods – clustering in non-euclidean space –clustering for streams and parallelism

UNIT IV

Map reduce – hadoop, hive, mapr – sharding – nosql databases - s3 - hadoop distributed file systems –visualizations - visual data analysis techniques, interaction techniques; systems and applications

References:

- Bill Franks, Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with advanced analytics, John Wiley&sons, 2013
- Anand Rajaraman and Jeffrey David Ullman, Mining of Massive Data sets, Cambridge University Press, 2014
- PaulZ ikopoulos, Chris Eaton, Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data: Analytics for Enterprise Class Hadoop and Streaming Data, McGraw Hill Professional, 2012
- Glenn J Myatt, Making Sense of Data, John Wiley & Sons, Pete Warden, Big Data Glossary,O ‟Reilly

Quantum Computing 4:0:0

Course Outcome:

CO1: Understand key concepts and challenges in big data analytics.
CO2: Apply scalable techniques such as Hadoop and MapReduce.
CO3: Analyze large datasets using appropriate tools and methods.
CO4: Perform data mining tasks such as clustering and pattern analysis.

Unit I

Introduction: Elementary I quantum mechanics:linear, algebra for quantum mechanics,Quantum states in Hilbert space,The Bloch sphere, Density operators, generalized measurements, no-cloning theorem.

Unit II

Quantum correlations,Bell inequalities and entanglement, Schmidt decomposition, super dense coding, teleportation,Quantum cryptography-quantum key distribution

Unit III

Quantum gates and algorithms, Universal set of gates, quantum circuits, Solovay-Kitaev theorem, Deutsch-Jozsa algorithm, factoring.

Unit IV

Programming a quantum computer,The IBMQ, coding a quantum computer using a simulator to carry out basic quantum measurement and state analysis.

Text-books

(1) Phillip Kaye, Raymond Laflamme et. al., An introduction to Quantum Computing, Oxford University press, 2007.
(2) Chris Bernhardt, Quantum Computing for Everyone, The MIT Press,Cambridge, 2020
(3) David McMahon-Quantum Computing Explained-Wiley-Interscience , IEEE Computer Society (2008)

AI Concepts and Techniques 4:0:0

Course Outcome

CO1: Explain what constitutes "Artificial" Intelligence and how to identify systems with Artificial Intelligence

CO2: Explain how Artificial Intelligence enables capabilities that are beyond conventional technology, for example, chess-playing computers, self-driving cars, robotic vacuum cleaners

CO3: Use classical Artificial Intelligence techniques, such as search algorithms, minimax algorithm, neural networks, tracking, robot localization

CO4: Ability to apply Artificial Intelligence techniques for problem solving

CO5: Explain the limitations of current Artificial Intelligence techniques

UNIT: I

Intelligent Agents: Agents and environment; the nature of environment; the structure of agents. Problem-solving: Problem solving agents; uninformed search strategies. Informed search strategies,Constraint satisfaction problems; Backtracking search for CSPs. Adversial search: Games; Optimal decisions in games; Alpha-Beta pruning.Knowledge-based agents,Effective propositional inference; Agents based on propositional logic.

UNIT: II

Syntax and semantics of first-order logic; Using first-order logic; Knowledge engineering in first-order logic. Propositional versus first-order inference; Unification and lifting; Forward chaining; Backward chaining; Resolution.

UNIT: III

Ontological engineering; Categories and objects; Actions, situations, and events; Mental events and mental objects; The Internet shopping world; Reasoning systems for categories; Reasoning with default information; Truth maintenance systems. Learning: Forms of Learning; Inductive learning; Learning decision trees.

UNIT: IV

Introduction; Robot Hardware: sensors and Effectors; Robotic Perception: localization, mapping, other types of perception; Planning to Move: configuration space, cell decomposition methods and skeletonization methods; Planning uncertain movements: robust methods; Moving: dynamics and control, potential field control and reactive control; Robotic Software: architectures, sub sumption architecture , three-layer architecture and robotic programming languages ; Application domains

References:

- Artificial Intelligence, Elaine Rich, Kevin Knight, Shivashankar Nair, Tata McGraw Hill
- Artificial Intelligence, Patrick Henry Winston, AWL
- Artificial Intelligence and Expert Systems, Dan W Patterson, PHI
- Artificial Intelligence, Nils J Nilson, Elsevier, Morgan Kaufmann

\

Data Wrangling 4:0:0

Course Outcome:

CO1: Understand the process and importance of data wrangling.
CO2: Handle and process structured and unstructured data.
CO3: Apply data cleaning and transformation techniques.
CO4: Analyze and visualize datasets using appropriate tools.

UNIT-I

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

UNIT-II

Installing Python Packages-Parsing Excel Files-Parsing Excel Files -Getting Started with Parsing-PDFs and Problem Solving in Python- Programmatic Approaches to PDF Parsing-Converting PDF to Text-Parsing PDFs Using pdf miner- Acquiring and Storing Data-Databases: A Brief Introduction Relational Databases: MySQL and PostgreSQL-Non-Relational Databases: NoSQL When to Use a Simple File-Alternative Data Storage

UNIT-III

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

UNIT-IV

Exploring Data-Importing Data-Exploring Table Functions- Analysing Data-Separating and Focusing the Data Presenting Data-Visualizing the Data- Charts-Time-Related Data-Maps-Interactives -Words Images, Video, and Illustrations-Presentation Tools-Publishing the Data-Open Source Platforms, Web scraping: What to Scrape and How-Analyzing a Web Page-Network/Timeline Interacting with JavaScript-In-Depth Analysis of a Page- Getting Pages-Reading a Web Page-Reading a Web Page with LXML-XPath-Advanced Web Scraping

Text books

- Principles of Data Wrangling: Practical Techniques for Data Preparation
- Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython
- Data Wrangling with Python: Creating actionable data from raw sources
- Data Wrangling with Python: Tips and Tools to Make Your Life Easier