PERIYAR UNIVERSITY

(NAAC A⁺⁺ GRADE–STATE UNIVERSITY–NIRF RANK 63 ARIIA 10th RANK)
Salem-636 011

DEPARTMENT OF COMPUTER SCIENCE
(Supported by UGC-SAP)

M.Sc. DEGREE

DATA SCIENCE

[Choice Based Credit System (CBCS)]

OBE REGULATIONS AND SYLLABUS

(Effective from the academic year 2022-23 and thereafter)

M. Sc. DATA SCIENCE

OBE REGULATIONS AND SYLLABUS
(With effect from the academic year 2022-2023 onwards)

1. Preamble

The **M. Sc-Data Science** programme is aimed at preparing data science professionals for research, academics, industry, design, and development in analytical technologies. It also aims to reduce the gap between academic and industry by providing industry based expertise in the field of data science.

2. General Graduate Attributes

GA 1: The fundamental knowledge

- This refers to a group of concepts associated with the discipline of computer science and how they are applied in user interfaces
- Initially, students learn about basic concepts of algorithms, programming language, and user interface
- To develop in-depth knowledge about the data analytics concept through course work.

GA 2: Design and conduct experiments

- To apply the principles of data science to real-world problems.

GA 3: Software and tools

- A proficient user of existing computing tools
- A proficient developer of new computing tools
- A competent analyzer, processor, and reporter of simple and complex data.

GA 4: Solve computer science problems

- Be equipped with a range of fundamental principles of Computer Science that will provide the basis for future learning and enable them to adapt to the constant rapid development of the field.
- Be able to apply mathematics, logic, and statistics to the design, development, and analysis of software systems.

- Be able to apply ethical practices in handling data of any kind.

GA 5: To apply algorithmic principles

- Identifies the key intellectual themes of the field as algorithmic thinking, information representation, and computer programs.

GA 6: To acquire the latest technical skills

- To enable the students to acquire the latest technical skills and build their carrier based on continuous learning and adaptability.

GA 7: Theoretical and practical knowledge

- Theoretical knowledge refers to learning things, concept, principle, information regarding anything from books. It means that the theoretical knowledge is knowledge which is basic knowledge or concept of things.
- Practical knowledge and Practical skills are two different things. Practical knowledge is the knowledge in which we can apply and implement ideas using practices.
- Be able to apply algorithms such as clustering and classification to sample dataset
- To build machine intelligence algorithm for solving complex decision-making problems

GA 8: Leadership, initiative, and teamwork:

- Ability to work effectively in a team and lead in a multidisciplinary environment

GA 9: Creative:
- Demonstrates critical thinking, imagination, and intellectual agility.
- Strives to be innovative and experimental in advancing knowledge and in creating solutions.

GA 10: Specialist:
- Possesses breadth and depth of knowledge in their specialist area.
- Applies through research and inquiry a systematic and critical approach
- Analytical approach to identifying and resolving problems.

GA 11: Intellectual Rigour

- An ability to think clearly and deeply with rigour when faced with new knowledge and arguments.

o Demonstrate an ability to apply legal research results to solve legal problems.
o Developing a shared understanding of intellectual rigour

GA 12: Knowledge of a discipline

o Develop the capability of demonstrating comprehensive and considered knowledge of the discipline.
o Enables students to evaluate and utilize information and apply their disciplinary knowledge and their professional skills in the workplace.

3. Programme Specific Qualification Attributes

Mention the programme specific qualification attributes achieved through courses in the programme in terms of

- **Knowledge and understanding level (K1 and K2)**
  o Remember or recognize a term or a basic concept
  o Select an explanation for a statement related to the question topic
  o Understand the existing problems

- **Application-level (K3)**
  o Be able to solve the problems using computing techniques.

- **Analytical level (K4)**
  o Be able to separate information related to a procedure or technique into its constituent parts for better understanding and can distinguish between facts and inferences.

- **Evaluation capability level (K5)**
  o Be able to make judgments based on criteria and standards. Detects inconsistencies or fallacies within a process or product, determines whether a process or product has internal consistency and detects the effectiveness of a procedure as it is being implemented.

- **Scientific or synthesis level (K6)**
  o A scientific way to analyze and solve the problems.

4. Vision

Achievingexcellence in Information Technology Enabled Services through Teaching, Research, Extension, Skill and Consultancy.

5. Programme Objectives and Outcomes

Spelt the PEOs (Programme Educational Objectives), Programme Specific Objectives (PSOs) and Programme Outcomes (POs).

Programme Educational Outcomes (PEOs)

PEO1: Apply algorithmic, mathematical and scientific reasoning to a variety of computational problems

PEO2: Implement software systems that meet specified design and performance requirements.

PEO3: Work effectively in teams to design and implement solutions to computational problems

PEO4: Communicate effectively, both orally and in writing. Design, correctly implement and document solutions to significant computational problems

Programme Specific Outcomes (PSOs)

PSO1: An ability to apply profound knowledge to analyze and design software and systems containing hardware and software components of varying complexity.

PSO2: An ability to apply mathematical model, algorithmic principles, and computer science theory in the design of real-time applications

Programme Outcomes (POs)

PO1: Gain the skills and knowledge required to operate your computer and perform common tasks. This basic computer skills course will allow you to gain an understanding of the most popular, current technologies used at home and in the workplace.

PO2: The objective was set as training the students to use mathematical knowledge they have to designing and conducting 'Development Experiments' as well as analyzing and interpreting data using the Scientific

Method.

PO3: Engage in effective software development practices over the entire system lifecycle including requirements, analysis, design, implementation, and testing.

PO4: Apply problem-solving skills and the knowledge of computer science to solve real problems.

PO5: Apply algorithmic reasoning to a variety of computational problems. Design, correctly implement and document solutions to significant computational problems

PO6: Learn new Programming concepts and coding skills are now essential for the high-tech industry. Everything is about new ways to build and use products and services that are more intuitive.

PO7: Theoretical knowledge and practical knowledge can often lead to a deeper understanding of a concept through seeing it in the context of a greater whole and understanding the why behind it.

PO8: Students will develop *critical* thinking skills. Students will develop an understanding of change *processes* and be able to think critically about obstacles to change.

PO9: Seeks and Progresses opportunities for change and growth. Is flexible and able to adapt to rapidly changing environments

PO10: It is a critical field that promotes systematic ways to design, evaluate, and manage computing solutions.

PO11: Intellectual rigour is encouraged for example during an assessment exercise where a debate or discussion occurs about a challenging topic.

PO12: Knowledge of a discipline is defined as "command of a discipline to enable a smooth transition and contribution to professional and community settings.

PEO-POMapping:

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
PEO1	■	■	■	■	■		■	■		■		
PEO2	■	■	■	■					■		■	■
PEO3	■	■	■		■	■	■	■		■	■	
PEO4	■	■	■		■	■	■	■		■	■	

PO-GAMAPPING:

	GA1	GA2	GA3	GA4	GA5	GA6	GA7	GA8	GA9	GA10	GA11	GA12
PO1	■											
PO2		■										
PO3			■									
PO4				■								
PO5					■							
PO6						■						
PO7							■					
PO8								■				
PO9									■			
PO10										■		
PO11											■	
PO12												■

6. Candidate's eligibility for admission

A candidate who has passed Bachelor's Degree in Electronics and Electrical Engineering/ Electronics and communication Engineering/ Computer Science and Engineering/ Information Technology/ Electronics and Instrumentation Engineering/ Computer Science/ Computer Technology/ Software Engineering/ BCA/ Mathematics/ Statistics/ Physics/ Data Science/ Data Analytics/ Electronics/ Electronics and Communications under 10+2/3+2/3/4 system of this University or any of the degree of any other University accepted by the syndicate as equivalent thereto subject to such conditions as may be prescribed therefore shall be permitted to appear and qualify for the **M.Sc. Data Science** degree examination of this University after a course of study of two academic years.

7. Duration of the programme

The Programme for the degree of **Master of Science in DATA SCIENCE** shall consist of **two Academic years** divided into four semesters. Each semester consists of 90 working days. The Period of Concluding the Programme apart from academic years should be completed within Three years.

8. CBCS- Programme structure
The programme structure comprises as follows.

Course Component	No. of Courses	No. of Courses * Hours of Learning/ Week	Marks	Credits
Part A (Credit Courses)				
Core Theory Courses	10	40	1000	40
Elective Theory Courses	03	12	300	12
Supportive Courses	1	3	100	03
Core - Practical	7	28	700	13
Elective Practical	2	8	200	04
Project	1	-	200	20
Online Courses	1	-		2
MP/Internship	1	2	100	1
Total	**26**	**93**	**2600**	**95**

9. Curriculum structure for each semester as per your courses alignment

Course	L T P	No. of Hours / Week	Number of Credits
Semester-I			
Core Course- 22UPCSC4C01 Statistics and Probability	3+1+0	4	4
Core Course- 22UPCSC4C02 Mathematical methods for Data Science	3+1+0	4	4
Core Course- 22UPCSC4C03 Design of Algorithms	3+1+0	4	4
Core Course- 22UPCSC4C04 Big Data Technologies	3+1+0	4	4
Core Course-22UPCSC4C05 Research Methodology	3+1+0	4	4
Core Course-22UPCSC4C06 R Programming – Lab	0+0+4	4	2
Core Course-22UPCSC4C07 SQL and Hadoop– Lab	0+0+4	4	2
Core Course- 22UPCSC4C08 Python Programming - Lab	0+0+2	2	1
SWAYAM /MOOC/SOFT SKILL-01 (Optional)	0+0+0		
Total		30	25

Semester-II			
Core Course- 22UPCSC4C09 Principles and Techniques of Data Science	3+1+0	4	4
Core Course- 22UPCSC4C10 Statistics Modeling	3+1+0	4	4
Core Course- 22UPCSC4C11 Machine Learning Techniques	3+1+0	4	4
22UPSOC2H01 Fundamentals of Human Rights	2+0+0	2	2
Elective-I	3+1+0	4	4
Non-major Elective / Supportive Course	2+1+0	3	3
Core Course- 22UPCSC4C12 Machine Learning -Lab	0+0+4	4	2
Core Course- 22UPCSC4C13 Essential Technologies for Data Science Lab	0+0+4	4	2
Core Course- 22UPCSC4C14 Mini Project/Summer Internship	0+0+2	2	1
SWAYAM / MOOC / SOFT SKILL-02 (Optional)	0+0+1	1	2
Total		32	28

Semester-III			
Core Course- 22UPCSC4C15 Principles of Management	3+1+0	4	4
Core Course- 22UPCSC4C16 Deep Learning Techniques	3+1+0	4	4
Elective-II	3+1+0	4	4
Elective-III	3+1+0	4	4
Core Course- 22UPCSC4C17 Deep Learning-Lab.	0+0+4	4	2
Elective Lab I	0+0+4	4	2
Elective Lab II	0+0+4	4	2
22UPCSC4C18 Soft Skill Development Lab	0+0+2	2	2
Total		30	24

Semester-IV			
Option-I			
Core Course- 22UPCSC4C19 Internship based project. Dissertation and Viva Voice (Industry/Research)	20	-	20
Total			20
Total no. of Credits **Core Theory (10)** **Elective Theory (3)** **CorePractical (7)** **Elective Practical (2)** **Supportive (1)** **Project (1)**			40 12 14 04 03 20
Grand Total			95

Elective –I

22UPCSC4E01   Optimization Techniques

22UPCSC4E02   Advanced Statistical Techniques

22UPCSC4E03   Advanced Data Visualization

22UPCSC4E04   Discrete Mathematics

Elective –II

22UPCSC4E05   Computer Vision

22UPCSC4E06   Spark Technology

Elective – III

22UPCSC4E07   Econometrics and Finance

22UPCSC4E08   Natural Language Processing

Elective Lab I

22UPCSC4E09   Computer Vision Lab

22UPCSC4E10   Spark Technology Lab

Elective Lab II

22UPCSC4E11   Econometrics and Finance Lab

22UPCSC4E12   Natural Language Processing Lab

10. CBCS – Scheme of Examinations semester wise structure

Course	Number of Credits	Hours Per Week	Examination Duration (hrs)	I. A	ESE	Total
Semester-I						
Core Course- 22UPCSC4C01 Statistics and Probability	4	4	3	25	75	100
Core Course- 22UPCSC4C02 Mathematical methods for Data Science	4	4	3	25	75	100
Core Course- 22UPCSC4C03 Design of Algorithms	4	4	3	25	75	100
Core Course- 22UPCSC4C04 Big Data Technologies	4	4	3	25	75	100
Core Course-22UPCSC4C05 Research Methodology	4	4	3	25	75	100
Core Course-22UPCSC4C06 R Programming – Lab	2	4	3	40	60	100
Core Course-22UPCSC4C07 SQL and Hadoop– Lab	2	4	3	40	60	100
Core Course- 22UPCSC4C08 Python Programming - Lab	1	2	3	40	60	100
Total	25					800
Semester-II						
Core Course- 22UPCSC4C09 Principles and Techniques of Data Science	4	4	3	25	75	100
Core Course- 22UPCSC4C10 Statistics Modeling	4	4	3	25	75	100
Core Course- 22UPCSC4C11 Machine Learning Techniques	4	4	3	25	75	100
22UPSOC2H01 Fundamentals of Human Rights	2	2	3	25	75	100
Elective-I	4	4	3	25	75	100
Non-major Elective / Supportive Course	3	3	3	25	75	100
Core Course- 22UPCSC4C12 Machine Learning -Lab	2	4	3	40	60	100

Core Course- 22UPCSC4C13 Essential Technologies for Data Science Lab	2	4	3	40	60	100
Core Course- 22UPCSC4C14 Mini Project/Summer Internship	1	2	3	40	60	100
SWAYAM/MOOC	2					100
Total	28					1000
Semester-III						
Core Course- 22UPCSC4C15 Principles of Management	4	4	3	25	75	100
Core Course- 22UPCSC4C16 Deep Learning Techniques	4	4	3	25	75	100
Elective-II	4	4	3	25	75	100
Elective-III	4	4	3	25	75	100
Core Course- 22UPCSC4C17 Deep Learning Lab	2	4	3	40	60	100
Elective Lab I	2	4	3	40	60	100
Elective Lab II	2	4	3	40	60	100
22UPCSC4C18 Soft Skill Development Lab	2	2	3	100		100
Total	24					800
Semester-IV						
Core Course-22UPCSC4C19 Dissertation and Viva Voice (Industry/Research	20	--	3	50	150	200
Total	20					200
Total no. of Credits **Core Theory (10)** **Elective Theory (3)** **CorePractical (7)** **Elective Practical (2)** **Supportive (1)** **Project (1)**	40 12 14 04 03 20					
Grand Total	95					2600

EXAMINATIONS – THEORY

EVALUATION OFINTERNAL ASSESSMENT

Test   : 10   (5+5 Marks, 5 marks from best one of Test 1 and Test 2, 5 marks from test 3 – model examinations)

Seminar   : 05 Marks

Assignment   : 05 Marks

Attendance   : 05 Marks

Total   **: 25 Marks**

EVALUATION OFEXTERNAL EXAMINATIONS

QUESTION PAPER PATTERN

Time duration: 3 Hours   Max. Marks: 75

PART- A: 20× 1 = 20

Answer all the questions

(Objective type four questions from each unit)

PART- B: 3 × 5 = 15

Answer any three questions out of five questions

(Questions must be of analytical type)

PART- C: 5× 8 = 40

Answer all the questions

(Either or type for each unit)

The Passing minimum shall be 50% out of 75 marks (38 marks)

Part	K Level	CO Coverage
A: Objective type	K1, K2	CO1-20%, CO2-20%, CO3-20%, CO4-20% and CO5-20%
B: Analytical type	K3, K4, K5	CO1-20%, CO2-20%, CO3-20%, CO4-20% and CO5-20%
C: Essay type	K1,K2, K3, K4, K5, K6	CO1-20%, CO2-20%, CO3-20%, CO4-20% and CO5-20%

PRACTICAL / SOFTWAREDEVELOPMENT
EVALUATION OFINTERNALASSESSMENT

Test 1   :   15Marks

Test 2   :   15Marks

Record   :   10 Marks

Total   :   40 Marks

EVALUATION OF EXTERNAL EXAMINATIONS

Time duration: 3 Hours   Max. Marks: 60

QUESTION PAPER PATTERN
1. One compulsory question from the given list of objectives   : 30 Marks
2. One Either/OR type question from the given list of objectives   : 30 Marks

Distribution of Marks

  Problem Understanding   : 05 Marks

  Program writing   : 10 Marks

  Debugging   : 10 Marks

  For Correct Results   : 05 Marks

  Mini-Project Viva-Voce (joint)   : 60 Marks

DISSERTATION
  Evaluation (External)   : 50 Marks

  Viva-voce (joint)   : 100 Marks

11. REGULATIONS OF PROJECT WORK

a. Students should do their five months [Dec to Apr] Project work in Company /Institutions.

b. The Candidate should submit the filled-in format as given in Annexure-I to the department for approval during the First Week of January in their Project semester.

c. Periodically the project should be reviewed a minimum three times by the advisory committee.

d. Students should prepare three copies of the dissertation and submit the same to the college on 30th April for the evaluation by examiners. After evaluation one copy is to be retained in the College Library and one copy is to be submitted to the University (Registrar) and the student can hold one copy.

e. A Sample format of the dissertation is enclosed in Annexure-II.

f. The format of the Title page and certificate are enclosed in Annexure III.

g. The Students should use OHP / PowerPoint presentations during their Project Viva-voce Examinations.

12. Examinations

Examinations are conducted in semester pattern. The examination for the Semester I & III will be held in November/December and that for the Semester II and IV will be in the month of April/May. Candidates failing in any subject (theory, practical and skill) will be permitted to appear for such failed subjects in the same syllabus structure at subsequent examinations within next 3 years after normal completion of the programme. Failing which, the candidate has to complete the course in the present existing syllabus structure.

13. Scheme for Evaluation and Attainment Rubrics

Evaluation will be done on a continuous basis and will be evaluated four times during the course work. The first evaluation will be in the 7th week, the second in the 11th week, third in the 16th week and the end – semester examination in the 19th week. Evaluation may be by objective type questions, short answers, essays or a combination of these, but the end semester examination is a University theory examination with prescribed question paper pattern.

ANNEXURE - I

PERIYAR UNIVERSITY

College Name   :

Course   :

Student Name   :

Register Number   :

Title of the Project   :

Address of Organization/Institution :

Name of the External Guide   :

Designation   :

Place :

Date :   Signature of External Guide

                                                                      (with seal)

Name of the Internal Guide   :

Qualification   :

Teaching Experience   :

Place :

Date: Signature of Internal Guide

Principal [Approved or not Approved]

[University Use]

COLLEGE BONAFIDE CERTIFICATE

COMPANY ATTENDANCE CERTIFICATE

Acknowledgment

CONTENTS

Format of the title page

TITLE OF THE DISSERTATION

A Dissertation submitted in partial fulfillment of the requirements for the degree of
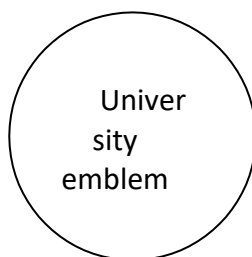
Master of Science in Data Science

to the

Periyar University, Salem - 11

By

STUDENT NAME

REG. NO.

Univer
sity
emblem

DEPARTMENT NAME (AFFILIATED TO PERIYAR UNIVERSITY)

PLACE with Pin Code

MONTH – YEAR

Format of the Certificate

DEPARTMENT NAME

(AFFILIATED TO PERIYAR UNIVERSITY)

PLACE with PIN CODE
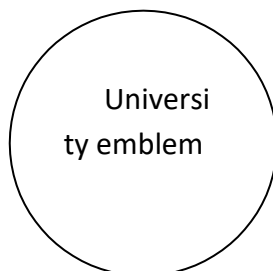
Universi
ty emblem

MONTH – YEAR

PROJECT WORK

TITLE OF THE DISSERTATION

Bonafide Work Done

by STUDENT NAME

REG.NO

A Dissertation submitted in partial fulfillment of the requirements for the degree of **Master of Science in Data Science** to the **Periyar University, Salem -11.**

INTERNAL GUIDE   HEAD OF THE DEPARTMENT

Submitted for the Viva-Voce Examination held on _____

Internal Examiner   External Examiner

M. Sc-DATASCIENCE - SYLLABUS

COURSE CODE - 22UPCSC4C01   **Credits: 4**

STATISTICS AND PROBABILITY

Course Objectives:

● The main objective of this course is to provide students with the foundations of probability and basic statistical analysis mostly used in varied applications in Business, Science etc.
● To use appropriate statistical methods to collect, organize, display, and analyze relevant data.
● To apply concepts of various probability distributions to find probabilities.

UNIT I:

Data Collection – Classification and Tabulation - Diagrammatic representation: One, Two, three dimensional diagrams - Graphical representation: Histogram - Frequency Curve - Cumulative frequency curve - Density plot - Bee swarm plot - Violin plots - drawing inference about them.

UNIT II:

Descriptive Statistics: - Measures of Central tendency - Measures of Dispersion - Measures of shape.

UNIT III:

Bivariate data: Correlation and its types - Coefficient of correlation - Rank correlation - Regression - Regression equation - Difference between correlation and regression- Partial and Multiple correlation

UNIT IV:

Combinatorial Probability: Definitions - Basic terminology - Axioms of probability - Counting principle - Addition theorem - Conditional Probability - Multiplication theorem -independent events (simple problems)

UNIT V:

Probability distributions: Random variables - Types of random variables - Distributions function and its properties - Expectation of random variable - Basic discrete and continuous distributions - Estimation of mean and variance of the distribution

Text Books:

1. Gupta S P, Statistical Methods, 46th revised edition, Sultan Chand and Sons, 2021.

Book for Reference:

1. Ehsanes, A. K.Md. and Rohatgi V. K. An Introduction to Probability and Statistics, 2ndEdition, Wiley.
2. Gupta,S.C. and Kapoor, V.K., Fundamentals of Mathematical Statistics, Sultan Chand, 2017.

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

S.NO	Course Outcomes
CO1	Understanding Data, classifying, tabulating and presenting data.
CO2	Understand the descriptive measures of data.
CO3	Observe and analyze the behavior of variables using the concept of correlation and regression
CO4	Understand the terminologies of basic probability with simple problems.
CO5	Understanding the concept of probability distributions, random variables and using these techniques to generate data from various distributions.

The mapping of course outcomes with programme outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	M	S	M	S	S	L	S	S	S	S	S	S
CO2	S	M	S	S	S	S	S	M	S	S	M	S
CO3	S	S	S	M	S	S	S	S	S	S	L	S
CO4	M	S	S	S	S	S	S	M	S	L	M	S
CO5	S	M	S	S	S	S	S	S	L	S	S	M

S- Strong; M-Medium; L-Low

COURSE CODE- 22UPCSC4C02   **Credits: 4**

MATHEMATICAL METHODS FOR DATA SCIENCE

Course Objective

- To provide suitable and effective methods called Numerical Methods

- To obtain approximate representative numerical results of the problems.

- To provide students with the foundations of essential linear algebra methods

- To solve various system equations and numerical methods

- To know Numerical methods to mostly use in varied applications in Business, Science etc.

Unit 1: Vectors &functions , Systems of linear equations, Row reduction and echelon forms, Matrix operations

Unit 2: Linear dependence and independence, Subspaces and basis and dimensions, Orthogonal basis and orthogonal projections, Linear models and least-squares problems

Unit 3: Determinants and their properties, Inverse of a matrix, Cramer's Rule

Unit 4: Eigenvalues and eigenvectors, Positive definite matrices, Linear transformations, Singular Value Decomposition, Hessian matrix

Unit 5: Numerical Analysis: Iterative method, Taylor Series, Cauchy method, Newton Raphson Method.

Textbook:

1.David C. Lay, Steven R. Lay, Judi J. McDonald, Linear Algebra and Its Applications, Pearson publication,5th edition, 2018.

 2. An Introduction to Numerical Analysis by Devi prasad

Reference Book:

Strang, Gilbert, et al. Introduction to linear algebra. Vol. 3. Wellesley, MA: Wellesley-Cambridge Press, Fifth Edition, 2016.

Course Outcomes

On the successful completion of the course, students will be able to

S.NO	Course Outcomes
CO1	Demonstrate understanding of basic mathematical concepts in data science, relating to linear algebra, probability, and calculus.
CO2	Employ methods related to these concepts in a variety of data science applications.
CO3	Analyze and evaluate the accuracy of common numerical methods.
CO4	Apply numerical methods to obtain approximate solutions to mathematical problems.
CO5	Apply Numerical analysis which has enormous application in the field of data Science

The mapping of course outcomes with program me outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	S	S	L	S	S	S	S	S	L	S	S	S
CO2	M	S	S	S	S	S	S	S	S	S	S	M
CO3	M	S	S	S	M	S	M	S	S	S	M	S
CO4	S	S	S	S	S	M	S	S	S	S	S	M
CO5	S	M	S	S	S	M	S	S	L	S	S	S

S- Strong; M-Medium; L-Low

DESIGN OF ALGORITHMS

Course Objective:

- To prepare a wiser consumer of data structures, algorithms, and heuristics.
- To explain different computational models (e.g., divideand-conquer), order notation and various complexity measures (e.g., running time, disk space) to analyse the complexity/performance of different algorithms.
- The analysis of the algorithm will be used to show the efficiency of the algorithm over the naive techniques.

Unit 1:   Introduction: Algorithm, Psuedo code for expressing algorithms, Performance Analysis-Space complexity, Time complexity, Asymptotic Notation-Big oh notation, Omega notation, Theta notation and Little oh notation Elementary Data Structures: Stacks and Queues – Trees – Dictionaries –Priority Queues – Sets and Disjoint Set Union – Graphs

Unit 2:   General Sorting method Bubble, Selection, Insertion, Divide and conquer- Merge & Quick sort, applications-Binary search, Quick sort, Merge sort, Linear Search

Unit 3:   The Greedy Method: General Method - Container Loading - Knapsack Problem – Tree Vertex Splitting – Job Sequencing with Deadlines - Minimum Cost Spanning Trees – Optimal Storage On Tapes – Optimal Merge Patterns - Single Source Shortest Paths

Unit 4:   Dynamic Programming: The General Method – Multistage Graphs – All-Pairs Shortest Paths – Single-Source Shortest Paths - Optimal Binary Search Trees - String Editing - 0/1 Knapsack - Reliability Design - The Traveling Salesperson Problem - Flow Shop Scheduling. Basic Traversal and Search Techniques: Techniques for Binary Trees – Techniques for Graphs – Connected Components and Spanning Trees – Biconnected Components and DFS

Unit 5:   Backtracking: The General Method – The 8-Queens Problem – Sum of Subsets – Graph Coloring – Hamiltonian Cycles – Knapsack Problem Branch and Bound: Least Cost searched

Text Books:

1. Ellis Horowitz, SatrajSahni and SanguthevarRajasekaran, *Fundamentals of Computer Algorithms,* Universities Press, Second Edition, 2009

Reference Books:

1. T. H. Cormen, C. E. Leiserson, R. L. Rivest. *Introduction to Algorithms,* The MIT Press, Cambridge, Massachusetts, 3rd edition,2009.
2. Skiena, Steven S. *The algorithm design manual.* Springer International Publishing, 2020.

Course Outcomes

On the successful completion of the course, students will be able to

S.NO	Course Outcomes
CO1	Ability to analyze the performance of algorithms.
CO2	Demonstrate a familiarity with major algorithms and data structures.
CO3	Describe the Greedy Methods and explain when an algorithmic design situation calls for it.Synthesize greedy algorithms, and analyze them.
CO4	Understand the Dynamic Programming problems and their applications.
CO5	Synthesize efficient algorithms in common engineering design situations.

The mapping of course outcomes with program me outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	S	S	S	S	S	S	S	S	L	M	S	M
CO2	S	S	L	S	S	S	S	S	S	M	S	M
CO3	S	S	S	S	S	S	S	S	M	S	S	S
CO4	S	S	S	S	S	S	M	S	S	S	M	S
CO5	S	M	S	S	S	S	S	S	M	L	S	S

S- Strong ; M-Medium ; L-Low

COURSE CODE-22UPCSC4C04   **Credits: 4**

BIG DATA TECHNOLOGIES

Course Objective:

- To study the basic technologies that forms the foundations of Database.
- The course provides grounding in basic and advanced methods to big data technology and tools, including MapReduce and Hadoop and its ecosystem.
- To optimize business decisions and create competitive advantage with Big Data analytics

Unit 1: Introduction to Databases and Transactions What is database system, purpose of database system, view of data, relational databases, database architecture, transaction management, Database Design, ER-Diagram overview, ER-Model, Constraints, ER-Diagrams, ERD Issues, weak entity sets, Codd's rules, Relational Schemas, Introduction to UML Relational

Unit 2: Database model: Logical view of data, keys, integrity rules. Relational Database design: features of good relational database design, atomic domain and Normalization (1NF, 2NF, 3NF, BCNF).

SQL: data definition, aggregate function, Null Values, nested sub queries, Joined relations. Triggers, Transaction management: ACID properties, serializability and concurrency

Unit 3: Introduction to Big Data, World of Big Data: Why and Where, Characteristics of Big Data and Dimensions of Scalability, Foundations for Big Data Systems and Programming,

Unit 4: The Hadoop Ecosystem: Map Reduce, HDFS, Pig, Hive, Hbase, Scoop, yarn

Unit 5: NoSQL: Overview of NoSQL Databases Definition of the Four Types of NoSQL Database, The Value of Relational Databases, Getting at Persistent Data, Concurrency, Integration, Impedance Mismatch, Application and Integration Databases, Attack of the Clusters, The Emergence of NoSQL, Key Points, Comparison of relational databases to new NoSQL stores, MongoDB, Cassandra, HBASE, Neo4j use and deployment, Application, RDBMS approach, Challenges NoSQL approach, Key-Value and Document Data Models, Column-Family Stores, Aggregate-Oriented Databases

Text Books:

1. Silberschatz, Abraham, Henry F. Korth, and ShashankSudarshan. *Database system concepts*. Vol. 4. New York: Mcgraw-hill, 1997.
2. Sadalage, Pramod J., and Martin Fowler. *NoSQL distilled: a brief guide to the emerging world of polyglot persistence.* Pearson Education, 2013.

3. White, Tom. *Hadoop: The definitive guide.* " O'Reilly Media, Inc.", 2012.

Reference Books:

1. Kenneth Cukier and Viktor Mayer-Schönberger, Big Data: A Revolution That Will Transform How We Live, Work, and Think,HodderAnd Stoughton, 2013.
2. Rob, Coronel, "Database Systems",Course Technology Inc, 7th edition, 2006.
3. Eric Redmond &Jim R. Wilson , Seven Databases in Seven Weeks: A Guide to Modern Databases and the NoSQL Movement, O'Reilly publisher, 2012.

Course Outcomes

On the successful completion of the course, students will be able to

S.NO	Course Outcomes
CO1	Illustrate the usage of data on different data ecosystems.
CO2	To design and build a simple database system and demonstrate competence with the fundamental tasks involved with modelling, designing, and implementing a DBMS.
CO3	Understand the fundamentals of various big data techniques.
CO4	Analyze the HADOOP and Map Reduce technologies associated with big data analytics.
CO5	Develop an application using different eco system tools by taking standard sample data set.

The mapping of course outcomes with program me outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	S	S	M	S	S	S	S	M	S	L	S	S
CO2	S	M	S	S	S	S	S	S	S	S	S	M
CO3	S	S	S	S	S	S	S	S	S	M	S	S

CO4	S	S	M	S	M	L	S	S	M	S	S	S
CO5	M	S	M	M	S	S	S	S	S	L	M	S

S- Strong ; M-Medium ; L-Low

COURSE CODE-22UPCSC4C05   **Credits: 4**

<div align="center">

RESEARCH METHODOLOGY

</div>

Course Objective:

- To develop a research orientation among the students and to acquaint them with fundamentals of research process.
- To identify the research problem and formulate objectives
- To develop an understanding of various research designs and techniques
- To identify various sources of information for literature review and data collection
- To learn the art of writing the reports and to be cautious of plagiarism

Unit 1: Introduction: Objectives and types of research– Definition of Research – Importance, limitations, Motivation and objectives – Research methods vs. Methodology–Types of research: Descriptive, Analytical, Applied vs. Fundamental, Quantitative vs. Qualitative, Conceptual vs. Empirical, concept of applied and basic research process– criteria of good research.

Unit 2: Research Problem and Design: Selecting the Problem - Necessity of Defining the Problem - Technique Involved in Defining a Problem - Research Design: Meaning of Research Design - Need for Research design - Features of a Good Design - Important Concepts Relating to research design - Different Research Designs - Basic Principles of Experimental Designs

Unit 3: Sampling Design: Census and Sample Survey - Implications of a Sample Design - Steps in Sampling Design - Criteria of Selecting a Sampling Procedure - Characteristics of a Good Sample Design - Different Types of Sample Designs - Random Sample Selection- Random Sample from an Infinite Universe - Complex Random Sampling Designs

Unit 4: Data Collection and AnalysisCollection of Primary Data - Observation Method - Interview Method - Questionnaires - Schedules - Difference between Questionnaires and Schedules - Collection of Secondary Data - Selection of Appropriate Method for Data Collection - Processing and Analysis of Data: Processing Operations - Problems in Processing - Elements/Types of Analysis - Statistics in Research - Measures of Central Tendency - Measures of Dispersion - Measures of Asymmetry - Measures of Relationship

Unit 5: Interpretation and Report Writing: Meaning of Interpretation - Technique of Interpretation - Precaution in Interpretation - Significance of Report Writing - Different Steps in Writing Report - Layout of the Research Report - Types of Reports - Oral Presentation - Mechanics for Writing Research Reports - Precautions for Writing Research Reports

Text Books

1. Kothari, C.R. (2019) Research Methodology Methods and Techniques. Fourth Edition, New Age International Publishers, New Delhi.

Reference Books

1. John W. Best, James V. Kahn (2005) Research in Education. Tenth Edition, Pearson Publisher
2. Fred N. Kerlinger (2017)Foundations of Behavioural Research. Surjeet Publications.

Course Outcomes

On the successful completion of the course,

S.NO	Course Outcomes
CO1	Students would have understood the concepts of research process in science.
CO2	To design the Research Proposal and decide the sample techniques and size
CO3	Students would have become familiar with the mode of collecting data and do the interpretation of the same.
CO4	Independently work in a research environment, consolidate the outcome of research and write technical papers.
CO5	Have basic knowledge on how to present research techniques.

The mapping of course outcomes with programme outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	M	S	L	S	S	S	S	S	S	S	S	M
CO2	S	S	S	S	S	M	S	S	S	L	S	S
CO3	M	S	S	S	S	S	S	S	M	S	S	S
CO4	S	M	S	S	S	S	S	S	S	L	S	S
CO5	S	L	S	S	S	M	S	M	S	S	S	S

S- Strong; M-Medium ; L-Low

COURSE CODE-22UPCSC4C06   **Credits: 2**

R Programming Lab

Implement the following in R studio.

1   Installing R on Windows, and R Studio

2   Data Types - R Objects and Attributes, Data Types - Vectors and Lists

3   Data Types - Data Frames, Matrices, Factors & Functions

4   Expression & Logical Statement in R

5   Subsetting of List, Matrices, & Dataframe

6   Dataframe functions on inbuilt Dataset

7   Dplyr Function on Dataset

8	Dplyr Function on Dataset
9	Basic Plotting with R
10	GGplots with R
11	Working with R Markdown & pushing Code to Git
12	Building R Shiny Dashboard App

COURSE CODE-22UPCSC4C07   **Credits: 2**

SQL and Hadoop Lab

Implement the following in Hadoop.

1 Viewing all databases, creating a Database, Viewing all Tables in a Database, Creating Tables (With and Without Constraints), Inserting/Updating/Deleting Records in a Table, Saving (Commit) and Undoing (rollback)

2 Altering a Table, Dropping/Truncating/Renaming Tables, Backing up / Restoring a Database

3 Real time Retail dataset Database creation

4 Queries with Aggregate functions (group by and having clause)

5 Data Analysis with SQL on Big Mart Dataset

6   Join Queries- Inner Join, Left & Right Join, Full Join, Cross Join

7   Startup VM, Hadoop 1.0 & Linux Command Basic

8   Transfer of file to HDFS & Map Reduce Program-I (Counting Problem)

9   Map Reduce Program-II (Data Analysis on Sales dataset)

10   Basic Hive Commands

11   Advance Hive Commands

12   Real time Data analysis with scoop & Hive

COURSE CODE-22UPCSC4C08   **Credits: 1**

Python Programming Lab

Implement the following in Python.

1   Introduction to Python, How to Use Shell, Spyder and Jupyter Notebook

2   Values, Data Types, Operators & Function Calls

3   Type Casting, Statements and Expressions

4   Input & Output Functions, Pythonic Functions, Exception Handling

5   Python OOPS Basic

6   Python OOPS Advance

7   Python GUI Development

8   Python Module: OS, Numpy

9   Python Module: Pandas

10   Python Module: Matplotlib and Seaborn

11   Python Module: Scipy

12   Python Module: Sklearn Pre-processing and Imputation

COURSE CODE-22UPCSC4C09   **Credits: 4**

PRINCIPLES AND TECHNIQUES OF DATA SCIENCE

Course Objective:

● To identify the scope and essentiality of Data warehousing and Data Mining.
● To develop research interest towards advances in data mining.
● To analyze the data, data science lifecycle, data collection and cleaning, exploratory data analysis and visualization, statistical inference and prediction, and decision-making algorithms for respective applications.

Unit 1: Why Data Mining? Moving toward the Information Age Data Mining as the Evolution of Information Technology, What Is Data Mining, What Kinds of Data Can Be Mined, Database Data, Data Warehouses, Transactional Data, Other Kinds of Data, OLTP & Online Analytical Processing(OLAP), Graphs Database

Unit 2: Getting to Know Your Data: Data Objects and Attribute Types, Basic Statistical Descriptions of Data, Data Visualization, Measuring Data Similarity and Dissimilarity: Euclidean, Jaccard's Index & Cosine Similarity

Unit 3: Data Pre-processing on Big Data: Data Cleaning, Data Integration, Data Reduction, Data Transformation and Data Discretization (ETL Operations)

Unit 4: Data Science Tools: One API AI toolkit

Unit 5: Data Science Applications in Uses Cases BSFI, Retail, Telecom & Healthcare

 Text Books:

1. Daimi, Kevin, et al. *Principles of Data Science.* Ed. Hamid R. Arabnia. Springer, 2020.
2. Han, Jiawei, MichelineKamber, and Jian Pei. "Data mining concepts and techniques third edition." *The Morgan Kaufmann Series in Data Management Systems* 5.4 (2011).
3. Boston Provided Material for (Unit 4 & 5)

Reference Books

1.   SinanOzdemir Principles of Data Science: Mathematical techniques and theory to succeed in data-driven industries, Packt Publishing Limited (13 December 2016)
2.   Cielen, Davy, Arno DB Meysman, and Mohamed Ali. *Introducing data science: big data, machine learning, and more, using Python tools.* Manning Publications Co., 2016.

Course Outcomes

On the successful completion of the course, students will be able to

S.NO	Course Outcomes
CO1	Understand fundamentals of data and Data Mining Principles.
CO2	To Understand importance of qualitative data, terminologies related to Data Science.
CO3	Understand and Extract knowledge using data processing concepts in data science.
CO4	Evaluate the databases, file formats and its applications using API Toolkit.
CO5	Analyze and design data mining applications.

The mapping of course outcomes with program me outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	S	S	L	S	S	S	L	S	L	S	S	M
CO2	S	S	M	M	S	S	S	M	S	S	S	S
CO3	S	S	S	S	M	S	S	S	M	L	S	S
CO4	S	S	M	S	S	S	S	S	S	S	S	S
CO5	S	M	S	S	M	S	S	M	S	S	S	S

S- Strong; M-Medium ; L-Low

COURSE CODE-22UPCSC4C10   **Credits: 4**

STATISTICAL MODELING

Course Objective:

- To understand the principles of statistical inference-Estimation theory and hypothesis testing.
- To learn about the sampling distributions-small and large samples and interpret the results.
- To understand the concept of regression and importance of R square value to interpret the results.
- To know the concept of ANOVA and its applications.
- To have the knowledge of non-parametric methods such as Kolmogrovsmirnov tests – Kruskal Wallis Test.

UNIT I:

Introduction to Statistical Inference: Estimation - Characteristics of estimation - Methods of estimation - Testing of hypothesis - Confidence interval for mean and population proportion

UNIT II:

Sampling distributions: Large sample test - Test for equality of mean and proportions - Small sample tests - Student's t test, Snedecor's F test, Chi-square test for independence of attributes - contingency table.

UNIT III:

Simple and Multiple linear regression models - Assumptions - Least square estimates - R square and adjusted R square.

UNIT IV:

Analysis of Variance - Assumptions - One way and Two-way classification (simple problems) - Introduction to Design of experiments - Principles of experimental designs - CRD, RBD, LSD (concepts only).

UNIT V:

Non-parametric methods: Advantages and disadvantages - Sign test, Run tests, Median test, Mann Whitney U test, Kolmogorov Smirnov test, Kruskal Wallis test.

TextBooks:

1.Gupta S P, Statistical Methods, 46th revised edition, Sultan Chand and Sons, 2021.

2. **Book for Reference:**

1.Gupta,S.C. and Kapoor, V.K., Fundamentals of Mathematical Statistics, Sultan Chand, 2017.

2.Gupta, S.C and Kapoor, V.K., "Fundamentals of Applied Statistics", Sultan Chand, 2018.

3.   Kothari, C.R. (2019) Research Methodology Methods and Techniques. Fourth Edition, New Age International Publishers, New Delhi.

Course Outcomes

On the successful completion of the course, students will be able to

S. No	Course Outcome
CO1	Analysis on statistical inference which arises in real life problems.
CO2	To observe and analyze the behavior of sampling distributions.
CO3	To interpret all of the results of Regression models using R square values.
CO4	To apply Analysis of Variance (ANOVA) in various society problems.
CO5	To differentiate the problems of parametric and nonparametric procedures.

The mapping of course outcomes with program me outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	M	L	S	S	M	S	S	S	S	L	S	M
CO2	S	S	S	S	S	S	M	L	S	S	S	S
CO3	S	S	S	M	S	M	S	S	S	S	S	S
CO4	S	S	S	S	S	S	M	S	L	S	M	S
CO5	M	S	S	S	L	S	S	M	S	S	M	S

S- Strong ; M-Medium ; L-Low

COURSE CODE-22UPCSC4C11   **Credits: 4**

MACHINE LEARNING TECHNIQUES

Course Objective:

- To introduce students to the basic concepts and techniques of Machine Learning.
- To understand the regression methods, regularization methods, clustering methods and metrics.
- To learn Decision trees, CART and Ensemble Techniques.
- To understand the problems using various machine learning techniques
- To study the recent machine learning software for solving practical problems.

Unit 1: Introduction of Machine Learning, Learning problem & System, Designing a learning system, concept of learning task, Hypothesis Space & Version Space

Unit 2: Introduction to Machine Linear Regression, Machine Logistics Regression, Regularization, Cross Validation & Metrics

Unit 3: Introduction to Decision Tree, CART, & Ensemble Techniques

Unit 4: Introduction to SVM, Kernels, Neural Networks (MLP), Collaborative filtering

Unit 5: Clustering Algorithms: K Means, K Mode, K Median,Hierarchical clustering, DBSCAN, PCA, TSNE

Text Books:

1. Tom M. Mitchell, *"Machine Learning"*, McGraw Hill,U.K, 2017.

2. Shalev-Shwartz, Shai, and Shai Ben-David. *Understanding machine learning: From theory to algorithms*. Cambridge university press, 2014.

Reference Books

1. Saleh, Hyatt. *Machine Learning Fundamentals: Use Python and scikit-learn to get up and running with the hottest developments in machine learning*. Packt Publishing Ltd, 2018.

Course Outcomes

On the successful completion of the course, students will be able to

S. No	Course Outcome
CO1	Apply the machine learning concepts in real life problems.

CO2	To Implement and analyze existing learning algorithms, including well-studied methods for classification, regression, clustering.
CO3	To Identify machine learning techniques suitable for a given problem.
CO4	To Design application using machine learning techniques.
CO5	To Solve the problems using various machine learning techniques

The mapping of course outcomes with program me outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	M	S	S	S	S	M	S	S	S	S	S	S
CO2	S	M	S	M	S	S	M	S	S	S	S	S
CO3	S	S	L	S	S	S	S	M	S	S	L	S
CO4	S	S	M	S	S	M	S	S	S	S	S	M
CO5	S	M	S	S	M	L	S	S	S	S	S	S

S- Strong; M-Medium; L-Low

Course Code: 22UPSOC2H01   **Credits: 2**

Fundamentals of Human Rights

Unit I: Introduction:

Meaning and Definitions of Human Rights – Characteristics and Importance of Human Rights – Evolution of Human Rights – Formation, Structure and Functions of the UNO - Universal Declaration of Human Rights – International Covenants – Violations of Human Rights in the Contemporary Era.

Unit II: Human Rights in India:

Development of Human Rights in India – Constituent Assembly and Indian Constitution – Fundamental Rights and its Classification – Directive Principles of State Policy – Fundamental Duties.

Unit III:

Rights of Marginalized and other Disadvantaged People: Rights of Women – Rights of Children – Rights of Differently Abled – Rights of Elderly - Rights of Scheduled Castes – Rights of Scheduled Tribes – Rights of Minorities – – Rights of Prisoners – Rights of Persons Living with HIVAIDS – Rights of LGBT.

Unit IV:

Human Rights Movements: Peasant Movements (Tebhaga and Telangana) – Scheduled Caste Movements (Mahar and Ad-Dharmi) – Scheduled Tribes Movements (Santhal and Munda) – Environmental Movements (Chipko and Narmada Bachao Andolan) – Social Reform Movements (Vaikom and Self Respect).

Unit V:

Redressal Mechanisms: Protection of Human Rights Act, 1993 (Amendment 2019) – Structure and Functions of National and State Human Rights Commissions – National Commission for SCs – National Commission for STs – National Commission for Women – National Commission for Minorities – Characteristics and Objectives of Human Rights Education.

References

1. Sudarshanam Gankidi, Human Rights in India: Prospective and Retrospective, Rawat Publications, Jaipur, 2019.
2. Satvinder Juss, Human Rights in India, Routledge, New Delhi, 2020.
3. Namita Gupta, Social Justice and Human Rights in India, Rawat Publications, Jaipur, 2021.
4. Mark Frezo, The Sociology of Human Rights, John Willy & Sons, U.K. 2014.
5. Chiranjivi J. Nirmal, Human Rights in India: Historical, Social and Political Perspectives, Oxford University Press, New York, 2000.
6. Dr. S. Mehartaj Begum, Human Rights in India: Issues and perspectives, APH Publishing Corporation, New Delhi, 2010.
7. Asha Kiran, The History of Human Rights, Mangalam Publications, Delhi, 2011.
8. Bani Borgohain, Human Rights, Kanishka Publishers & Distributors, New Delhi-2, 2007.
9. Jayant Chudhary, A Textbook of Human Rights, Wisdom Press, New Delhi, 2011.

COURSE CODE-22UPCSC4C12   **Credits: 2**

MACHINE LEARNING LAB

Implement the following.

1   Introduction to Python Framework on ML & significations

2   Linear Regression on Real Time Dataset

3   Logistics Regression on Real Time Dataset

4   Lasso & Ridge Regression on Real Time Dataset

5   KNN on Real Time Dataset

6   Decision Tree on Real Time Dataset

7   SVM on Real Time Dataset

8   MLP on Real Time Dataset

9   Random Forest on Real Time Dataset

10   K-Means on Real Time Dataset

11   DBSCAN on Real Time Dataset

12   TSNE & PCA on Real Time Dataset

ESSENTIAL TECHNOLOGIES FOR DATA SCIENCE LAB

Implement the following.

1   Introduction to Intel NumPy, Pandas & DAAL

2   Introduction to MKL

3   Data Engineering on Banking Dataset

4   Data Analysis on Banking Dataset

5   Building Dashboard on Sales dataset

6   Data Engineering on Banking Dataset

7   Data Analysis on Banking Dataset

8   Clustering on Sales dataset

9   Time series Data Understanding & Handling

10   Trend Analysis

11   Smoothing, Lag & Pre-processing of Time series Data

12   Forecasting Models with Time series Models, AR, MA, ARIMA

COURSE CODE-22UPCSC4C14   **Credits: 1**

MINI PROJECT AND SUMMER INTERNSHIP

Guidelines

Problem Statement on any topic from semester 2 subjects

Survey and study of published literature on the assigned topic; Working out a preliminary Approach to the Problem relating to the assigned topic; Conducting preliminary Analysis/Modelling/Simulation/Experiment/Design/Feasibility; Preparing a Written Report on the Study conducted for presentation to the Department; Final Seminar, as oral Presentation before a Departmental Committee.

PRINCIPLES OF MANAGEMENT

Course Objective:

- The course objective is to gain an understanding of business management, including historical and current management theory, types of managers and roles of managers in an organization.
- The role of a data science in organizations is complex and spread across horizontal to different functional departments of the organization such as sales and marketing, accounts, HR, facility and Maintenance.
- The course will enable the stakeholder to learn the essential knowledge identify the problem, analysis and formulate the real problem into data science problem.

Unit 1: Introduction Concept of Management

Management: Art and Science-Management Vs Administration- Levels of Management- Functions of management- Management as a Profession- Management skills- Qualities and characteristics of managers. Evolution of Management thought: Early contributions: Taylor and Scientific Management- Fayol's Administrative Management- Bureaucracy- Human Relations- and Modern Approach-Social responsibility of managers- Managerial Ethics.

Unit 2: Marketing Management for Data Scientist

Introduction- Importance and scope of Marketing- Philosophies of marketing management-Elements of marketing - Needs, Wants, Demands- Customer- Consumer-Markets and marketers; Marketing vs. Selling, Consumer markets and Industrial markets- Concept of marketing management- Marketing - Mix, Functions of marketing management- Marketing environment- Factors affecting marketing environment- Marketing information system and Marketing research- Strategic marketing planning- Marketing Myopia.

Unit 3: Banking and Finance for Data Scientist

Introduction to Banking segment- RBI -(Role and functions- Regulation of money and credit, Monetary policy- Commercial Banks: Structure and Scope NBFC's- Concept- Nature- Scope and their Implications- Securitization of Auto Loans and Housing Loans- Securitization in India- Leasing and Hire Purchase: Concepts and Scope-Parties Involved- Evaluation of Lease Transaction-Types of Lease and their Implications- Hire Purchase and Lease - Differences and Implications for the business.

Unit 4:Retail Consumer Behavior Analytics

Introduction Concept- Definition- Types of Consumers- Nature- Scope of Consumer Behavior- Four Drivers of Successful Relationships between Marketers and Customers- Consumer Behavior and Life-Style Marketing, Consumer motivation and involvement- consumer perception- Information processing- perception Learning, Consumer attitude: Family and household influences- family decision making and consumption-related roles- social class- Reference groups- Cultural and sub- cultural differences- Personality- Values and life-style.

Unit 5:Data Analytics for Manufacturing and Supply Chain Management

Introduction Manufacturing industry- sectors and segmentation- functional aspects of manufacturing industry- Inventory - management of BOM, MRP, and ERP system - analytical use cases. Supply chain Management: Introduction - significance and scope- Drivers of supply chain- supply chain case studies: DELL and Philips- different modes of transportation - price determination case study.

Text Book:

1.   Voich, Dan, and Daniel A. Wren. Principles of management: resources and systems. Ronald Press Company, 1968.

2.   Kotler, Philip. Marketing management, millenium edition. Prentice-Hall, Inc., 2001.

3.   Choudhry, M. (2018). An introduction to banking: principles, strategy and risk management. John Wiley & Sons.

4.   Schiffman, L. G., Wisenblit, J., & Kumar, S. R. (2011). Consumer Behavior by Pearson. Pearson Education India.

5.   Rahimi, I., Gandomi, A. H., Fong, S. J., &Ülkü, M. A. (Eds.). (2020). Big Data Analytics in Supply Chain Management: Theory and Applications. CRC Press.

Reference Book:

1.   Klosterman, S. (2019),"Data Science Projects with Python: A case study approach to gaining valuable insights from real data with machine learning", 2nd Edition,Packt Publisher.

2.   Kumar, U. D. (2017). Business analytics: The science of data-driven decision making. Wiley.

Course Outcomes

On the successful completion of the course, students will be able to

S. No	Course Outcomes
CO1	Compare and contrast the levels, roles, and functions of management
CO2	Distinguish between modern theories of management, including quality management and systems management theory
CO3	Understanding the Finance management theory
CO4	Break down quantitative management and the roles of branches such as operations management
CO5	Illustrate the types of planning and its function in management.

The mapping of course outcomes with program me outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	S	S	L	M	S	S	M	L	S	M	S	L
CO2	S	S	S	S	L	S	S	S	L	S	S	S
CO3	M	S	M	S	S	S	M	S	M	S	S	S
CO4	S	L	S	S	S	S	S	S	S	S	S	M
CO5	S	S	S	S	S	L	M	S	S	S	S	S

S- Strong; M-Medium; L-Low

COURSE CODE-22UPCSC4C16   **Credits: 4**

DEEP LEARNING TECHNIQUES

Course Objective:

- To understand the context of deep learning and their suitability for various data domains such as text, images, and videos.
- To improve the performance of a Deep Learning model the goal is to the reduce the optimization function which could be divided based on the classification and the regression problems.
- Students will be able to extract patterns from real-world datasets by running several classes of deep learning methods.

Unit 1:Deep Neural Networks: Introduction – Back propagation: The Gory Details -Setup and Initialization Issues -Vanishing and Exploding Gradient Problems-Gradient-Descent Strategies. **Regularization:** Penalty-Based Regularization-Ensemble Methods-Early Stopping.

Unit 2: Restricted Boltzmann Machines: Introduction -Hopfield Networks - Boltzmann Machine- Restricted Boltzmann Machines - Applications of Restricted Boltzmann Machines.

Unit 3: Recurrent Neural Networks: Introduction -Architecture of Recurrent Neural Networks -Challenges of Training Recurrent Networks-Echo-State Networks-Long Short-Term Memory (LSTM)-Gated Recurrent Units (GRUs)-Applications of Recurrent Neural Networks.

Unit 4: **Convolutional Neural Networks** : Introduction -Basic Structure of a Convolutional Network-Training a Convolutional Network-Convolutional Architectures-Visualization and Unsupervised Learning-Applications of Convolutional Networks.

Unit 5: **Autoencoder:** Under complete Auto encoders-Regularized Autoencoders-Representational Power, Layer Size and Depth-Stochastic Encoders and Decoders-Denoising Autoencoders-Learning Manifolds with Autoencoders-Contractive Autoencoders-Predictive Sparse Decomposition-Applications of Autoencoders

Text Book:

1. "Neural Networks and Deep Learning "Charu C. Aggarwal Springer, 2018.
   [Unit 1, 2 3, 4]
2. "Deep Learning (Adaptive Computation and Machine Learning series)" Ian Goodfellow, YoshuaBengio, AaronCourville , MIT Press, 2016.
   [Unit 5]

Reference Book:

1. Deep Learning with Python by Francois Chollet, Manning Publications,December 2017.
2. Online Resources: Michael Nielsen, "Neural Networks and Deep Learning", Determination Press, 2015. http://neuralnetworksanddeeplearning.com/

Course Outcomes

On the successful completion of the course, students will be able to

S. No	Course Outcomes
CO1	To know the main techniques in deep learning and the main research in this field.
CO2	Evaluate Boltzmann Machines and their types.
CO3	To understand the fundamentals and different types of RNN models.
CO4	Demonstrate an understanding of relations and functions of CNN models.
CO5	Be able to identify new application requirements.

The mapping of course outcomes with program me outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	S	S	M	S	S	L	M	S	S	M	S	S
CO2	S	S	S	S	M	S	L	S	L	S	S	M
CO3	S	M	S	S	S	L	S	M	S	S	S	S
CO4	M	S	S	S	S	S	S	S	M	S	S	M
CO5	S	S	S	S	S	M	S	M	S	S	S	S

S- Strong; M-Medium; L-Low

COURSE CODE-22UPCSC4C17   **Credits: 2**

DEEP LEARNING LAB

Implement the following

1. Introduction to DL and Framework

2. Feed Forward Network on sample dataset

3. Multi layer perceptron (MLP) on real-time dataset

4. Convolution neural network on binary classification task: Cat and Dog dataset

5. Convolution neural network on multi-classification task: Dog breed classifications

6. Transfer learning using pre trained architectures

7. Hyper parameter optimization on CNN models

8. Recurrent neural network on stock price prediction

9. Gated Recurrent neural network on Image segmentation task

10. LSTM on Price prediction

11. LSTM on Image segmentation

12. Anomaly detection using Auto encoders

Course Code: 22UPCSC4C18   **Credits: 2**

Soft Skill Development Lab

Course Objective

This course enables the students:

- To enable students to gain basic communication skills in professional and social contexts effectively.
- To acquire useful words and apply them in situational context.
- To develop listening and reading skills through comprehension passages
- To enrich the leadership qualities and interpersonal communication
- To enhance essential characteristics in writing

Exercises

1. Characteristics of Technical Writing
2. Development of Employability Skills
3. Vocabulary Development
4. Sentence Completion
5. Error Spotting
6. Interpretation of Verbal Analogy
7. Interpretation of Reading (Comprehension -Conception)
8. Interpretation of Reading (Comprehension -Reasoning)
9. Practice for writing E-mails/Technical Blogs/Forums
10. PPT Preparation / Demonstration of Technical Presentation
11. Preparation of Resume
12. Preparation for Job Interviews / Mock Interview Section
13. Group Discussion Skills
14. Developing Listening Skill(Comprehension)
15. Practice for Short Speeches / Situational Conversation
16. English through Mass Media
17. Essential Grammar
18. Communicating and collaborating with peer members
19. Team Empowerment
20. Persuasive Communication

Text Books
1. Uma Narula, "Development Communication: Theory and Practice", Revised Edition, Har-Aanad Publication, 2019.
2. Annette Capel and Wendy Sharp, "Cambridge English: Objective First", Fourth Edition, Cambridge University Press, 2013.
3. Emma Sue-Prince, "The Advantage: The 7 Soft Skills You Need to Stay One Step Ahead", First Edition, FT Press, 2013.
4. Guy Brook-Hart, "Cambridge English: Business Benchmark", Second Edition, Cambridge University Press, 2014.
5. Norman Lewis, "How to Read Better & Faster", Binny Publishing House, NewDelhi, 1978.

Reference Books

1 Michael McCarthy and Felicity O‘Dell, "English Vocabulary in Use:100 Units of Vocabulary Reference and Practice", Cambridge University Press,1996.

2. Murphy, Raymond, "Intermediate English Grammar", Second Edition, Cambridge University Press, 1999.

COURSE CODE-22UPCSC4C19   **Credits: 20**

<div align="center">

INTERNSHIP BASED PROJECT

</div>

Evaluation for Internship:

- The Internship duration for successful complete the credit should be 4 to 6 months.
- Student should have a valid letter Internship letter.
- During the Internship duration students should also complete Industry Certifications relevant to their Internship.
- Students will be evaluated in two phase presentation.
  - Phase -I: Initial presentation during 1 Month of the Internship duration
  - Phase -II: Final Presentation at the end of the semester.
- Project report need to be submit.

COURSE CODE-22UPCSC4E01   **Credits: 4**

OPTIMIZATION TECHNIQUES

Course Objective:

- To understand the fundamental knowledge of Linear Programming and Dynamic Programming problems.
- To Learn the classical optimization techniques and numerical methods of optimization.
- To study the basics of different evolutionary algorithms.
- To understand the Integer programming techniques and apply different optimization techniques
- To solve various models arising from engineering areas.

Unit 1: Linear Programming Problem Linear Programming Problem (LPP): Mathematical Formulation of Linear Programming Problem - Graphical Solution of LPP - canonical and standard forms of linear programming problem- Simplex method for solving LPP

Unit 2: Transportation Model: North West corner Method, Least cost method, and Vogel's Approximation Method. Assignment Model: Hungarian assignment model – Travelling Sales Man Problem

Unit 3: Project Scheduling PERT/CPM Networks – Fulkerson's Rule – Measure of Activity – PERT Computation – CPM Computation – Resource Scheduling.

Unit 4: Simplex Method – Gradient of function – Steepest Descent method – Conjugate Gradient method

Unit 5: Particle Swarm Optimization method – Ant Colony optimization algorithm – Fruit Fly method – Fire Fly method

Text Books:

1. Christos H. Papadimitriou and Kenneth Steiglitz, "Combinatorial Optimization -- Algorithms and Complexity", Dover Publications, 2013

2. David G. Luenberger and Yinyu Ye, "Linear and Non-linear Programming", Stanford University, Springer,4TH Edition, 2016.

3. Stephen Boyd and LievenVandenberghe, "Convex Optimization", Cambridge University Press, 2004.

Reference Books

1. Hamdy A. Taha, "Operations Research: An Introduction", Pearson, 2010

Course Outcomes

On the successful completion of the course, students will be able to

S. No	Course Outcomes
CO1	To develop the fundamental Linear Programming and Dynamic Programming problems of industrial process.
CO2	Learn efficient computational procedures to solve optimization problems.
CO3	To apply the basics of different evolutionary algorithms.
CO4	To analyze the fundamentals of Integer programming technique.
CO5	To apply different techniques to solve various optimization problems arising from business areas.

The mapping of course outcomes with program me outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	S	L	S	S	S	S	L	S	M	S	L	S
CO2	S	M	S	L	S	S	S	M	S	S	M	S
CO3	M	S	S	S	S	M	S	S	S	S	M	S
CO4	S	S	M	S	S	S	S	S	S	M	S	S
CO5	M	S	S	S	S	M	M	S	S	S	M	S

S- Strong; M-Medium; L-Low

COURSE CODE-22UPCSC4E02   **Credits: 4**

ADVANCED STATISTICAL TECHNIQUES

Course Objective:

- To understand the fundamental knowledge of Multivariate distributions, variance, covariance matrix.
- To Learn the Linear Regression Analysis and Ridge Regression.
- To study the basics of PCA and SVD algorithms.
- To understand the ECM, Linear Discriminant analysis and Logistic regression
- To solve ARIMA models for time series forecasting models.

Unit 1 (Multivariate Analysis): Notion of random vector, Mean vector , Variance-Covariance matrix and its properties – Multivariate normal distribution – sample generation from normal distribution – Computing partial and multiple correlations from variance- covariance matrix.

Unit 2 Multiple Linear Regression Analysis – Estimation and Testing – Multicollinearity – Ridge Regression – Variable selection

Unit 3 Principal Component Analysis – Interpreting PCAs – its usage- Singular Value Decomposition – Applications in Data Science

Unit 4 Classification: Minimum ECM rule – Mahalanobis D-square – Linear Discriminant analysis for two classes and more than classes problems – performance measures – Logistic regression

Unit 5 Time Series Models: Auto-regressive models – Moving average models – ARIMA models – Forecasting using time series models.

Text Books and Reference:

1. Makridakis, S, Wheelwright, S, C. and Hyndman, R.J., "Forecasting: Methods and Applications", 3rd Edition, Wiley Publications(Unit 1,3,4).

2. Montgomery, D.C., Peck, E.A. and Vining, G.G., "Introduction to Linear Regression Analysis", 5th Edition, Wiley Publications. (Unit 2)

3. Johnson, R.A and Wichern, D.W, "Applied Multivariate Statistical Analysis", 6th Edition, Pearson Publications. (Unit 5)

Course Outcomes

On the successful completion of the course, students will be able to

S. No	Course Outcome
CO1	To develop multivariate distribution and their applications in data science.

CO2	To use regression techniques and variable selection methods.	
CO3	Implement dimension reduction techniques using software on real life problems.	
CO4	To Analyze performance measure of LDA and regression technique	
CO5	To apply different techniques to solve various time series forecasting problems arising from engineering areas.	

The mapping of course outcomes with program me outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	S	S	M	S	S	L	S	S	S	M	S	S
CO2	S	S	S	S	M	S	S	S	S	S	S	S
CO3	S	M	S	S	S	S	M	S	S	S	M	S
CO4	M	S	S	S	S	S	S	S	S	M	S	M
CO5	S	S	S	S	S	M	S	S	M	S	M	S

S- Strong ; M-Medium ; L-Low

COURSE CODE-22UPCSC4E03   **Credits: 4**

ADVANCED DATA VISUALIZATION

Course Objective:

- To introduce students to the basic concepts and techniques of data visualization.
- To design and visualize various data including multivariate data, networks, text and cartography.
- An understanding of the key techniques and theory used in visualization, including data models, graphical perception and techniques for visual encoding and interaction.

Unit 1   Introduction : The case for data visualization and storytelling, Data visualization and storytelling details and best practices, Introduction to Tableau, PowerBI Importing Data / Connecting to External Sources Interface Overview Creating Sheets and Dashboard (software Demo)

Unit 2   Elements of Data Visualization: Measures & Dimension concepts, Design Fundamentals Design Principles, Colors, and "Chart Junk" Design perspectives from the experts The Shaffer 4 C's of Data Visualization Not-so-best practices (examples) Critique and redesign, Creating a good data set for analysis Data modeling fundamentals for analytics Selecting data for your KPIs,

Unit 3   Story telling with Data: Depth Design Fundamentals of Data Visulization, What are the main approaches to storytelling with data? - Dashboards vs. Storyboards vs. Info graphics - Designing with the user in mind The Duell Rules for Actionable Visualizations, Basic Charting Techniques (Hands-on Session)

Unit 4   Interactive Elements & Advance Graphics : Interactive Visualization Features – build interactive visualization Actions and filters Calculated measures Data blending, joins, and custom queries Custom Shape Files, Multiple choice and short answer on Blackboard Create graphs appropriate for data Chart critique Geocoding and Mapping, Advanced Chart types Custom Color Palettes

Unit 5   Case Studies: Compare and Contrast real-world examples Flowing Data - Nathan Yau Information is Beautiful Tableau Vizzes in the wild

Textbook

1. Yau, Nathan. *Visualize this: the FlowingData guide to design, visualization, and statistics.* John Wiley & Sons, First Edition, 2011.
2. Few, Stephen. *Information Dashboard Design: Displaying data for at-a-glance monitoring.* Vol. 5. Burlingame, CA: Analytics Press, Second Edition, 2013.

Reference Books

1. Hamilton, Max. "The visual display of quantitative information. Edward R. Tufte. Graphics Press, Cheshire, Connecticut, 5th printing (2001).
2. Knaflic, Cole Nussbaumer. *Storytelling with data: A data visualization guide for business professionals.* John Wiley & Sons, First Edition, 2015.
3. Wexler, Steve, Jeffrey Shaffer, and Andy Cotgreave. *The big book of dashboards: visualizing your data using real-world business scenarios.* John Wiley & Sons, First Edition, 2017.

Course Outcomes

On the successful completion of the course, students will be able to

S. No	Course Outcome
CO1	To demonstrate proficiency with statistical analysis of data.
CO2	To introduce the tools required to manage and analyze the elements of data visualization.
CO3	Apply visual design principles to simple and complex models that tell the stories found in data.
CO4	To understand how to work with data for real life.
CO5	To apply mathematical and statistical models and concepts to detect patterns in data, as well as draw inferences and conclusions supported by the data.

The mapping of course outcomes with program me outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	S	S	S	S	M	S	S	S	S	M	S	S
CO2	M	S	S	S	S	S	S	M	S	S	S	S
CO3	S	S	S	M	S	S	S	L	S	S	S	S
CO4	S	S	S	S	S	S	S	S	S	S	S	M
CO5	S	S	L	S	S	M	S	S	S	S	S	S

S- Strong ; M-Medium ; L-Low

COURSE CODE-22UPCSC4E04   **Credits: 4**

DISCRETE MATHEMATICS

(Theorems and Proofs are not expected)

Objectives:

- To understand the applications of functions and relations
- To understand the basic concepts of mathematical logic and predicate calculus
- To understand the concept of the method of induction
- To develop the skills in solving recurrence relations

Unit 1

Well-formed formulas – truth table of well-formed formula – a tautology, contradiction, and contingency –the equivalence of formulas. Algebra of propositions – Functionality complete sets – Normal forms of well-formed formulas- Rules of Inference for propositional calculus – well-formed formulas of predicate calculus – Rules of Inference for predicate calculus – predicate formulas involving two or more quantifiers.

Unit 2

Set theory – relations- functions – set identities – Binary relations – properties of binary relations in a set – Equivalence relations and partial orderings – Representation of a relation by a matrix representation of a relation by a digraph - Basics of Counting – Integers and Induction.

Unit 3

Formulation as Recurrence Relations-solving recurrence Relation by Iteration- solving Recurrence Relations- Solving Linear Homogeneous Recurrence Relations of Order Two-Solving Linear Nonhomogeneous Recurrence Relations. Permutations-Combinations-Permutations with repetitions-Combinations with repetition-permutations of sets with indistinguishable objects.

Unit 4

Definition and examples-properties of lattices –lattices as algebraic systems- Sub lattices and lattice Isomorphism-special classes of lattice –distributive lattices and Boolean algebras.

Unit 5

Connected Graphs-Euler Graphs-Hamiltonian circuits and paths – planar graphs – matrix representation of graphs.

Text Book:

1. N.Chandrasekaran and M.Umaparvathi, Discrete mathematics, PHI Learning Private Limited, New Delhi, 2010

References:

1. J.P.Trembley and R.Manohar, *Discrete Mathematical Structures with Applications to Computer Science,* Tata McGraw Hill, New Delhi, 2001.

2. Sengadir, T. *Discrete mathematics and combinatorics.* Pearson Education India, 2009.
3. RakeshDube, AdeshPandey, Ritu Gupta, *Discrete Structures and Automata Theory*, Narosa Publishing House, New Delhi, First Edition, 2006.

Course Outcomes

On the successful completion of the course, students will be able to

S. No	Course Outcomes
CO1	Understand the use of well-formed formula, logical connectives, predicates of formula and so on.
CO2	Evaluate Relations, functions, recurrence relations and their types.
CO3	Apply lattices with use of Hasse diagram and study of its properties and various study of different graphs.
CO4	Demonstrate an understanding of relations and functions and be able to determine their properties.
CO5	To verify whether an algorithm works well and perform analysis in terms of memory and time.

The mapping of course outcomes with program me outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	S	S	M	S	S	L	M	S	M	S	M	S
CO2	S	S	L	S	S	S	S	S	L	S	S	M
CO3	S	S	S	S	M	L	S	S	S	S	M	S
CO4	S	S	S	S	S	M	S	S	M	S	S	M
CO5	S	S	M	S	S	S	S	S	S	S	S	S

S- Strong ; M-Medium ; L-Low

COURSE CODE-22UPCSC4E05   **Credits: 3**

COMPUTER VISION

Course Objective:

● This project tests your knowledge of image processing and feature extraction techniques that allow you to programmatically represent different facial features.

● Student also use your knowledge of deep learning techniques to program a convolutional neural network to recognize facial keypoints.

- Facial keypoints include points around the eyes, nose, and mouth on any face and are used in many applications, from facial tracking to emotion recognition.

Unit 1: Introduction To Computer Vision: Geometric Camera Models - Image Formation - Intrinsic And Extrinsic Parameters - Light And Shading – Color. Filters: Linear Filters And Convolution- Shift Invariant Linear Systems- Spatial Frequency And Fourier Transforms- Sampling And Aliasing. Image Features: Gradient- SIFT-HOG- Texture.

Unit 2: Image Segmentation Applications- Segmentation By Clustering Pixels- Segmentation, Clustering, and Graphs. Grouping And Model Fitting: Hough Transform- Fitting Lines And Planes- Fitting Curved Structures- Fitting Using Probabilistic Models- Model Selection. Tracking - Simple Tracking Strategies- Tracking Using Matching.

Unit 3: Image Registration: Registering Rigid Objects- Range Data- Range Data Segmentation- Range Image Registration- Object Recognition- Kinect. Classification: Introduction- Strategies- Building Good Image Features.

Unit 4: Deep Learning: introduction – basic types- object detection- Convolutional Neural Networks (CNNs) architecture- Pedestrian Detection- Challenges of Object Detection. Face Recognition: Pose-Invariant Face Recognition- Illumination-Invariant Face Recognition- Multi-stream Convolutional Neural Networks- Face Anti-spoofing- Kinship Verification.

Unit 5: Face Recognition (FR) in Video Surveillance (VS): Introduction- Background of Video-Based FR Through Deep Learning- Deep Learning Architectures for FR in VS- Performance Evaluation. 3D Data Processing: Introduction- Circle Convolutional RBM (CCRBM). Object Instance Search: Introduction- Compact Invariant Deep Descriptors.

Textbook:

1. "COMPUTER VISION A MODERN APPROACH" by David A. Forsyth, Jean Ponce, Second Edition, Pearson publication, 2012.(Unit-I,II,III)

2. "Deep Learning in Object Detection and Recognition" by Editors Xiaoyue Jiang, AbdenourHadid, Yanwei Pang, Eric Granger, XiaoyiFeng, Springer, 2019.(Unit- IV,V)

Reference Book:

1. Hands-On Computer Vision with TensorFlow 2: Leverage Deep Learning to Create Powerful Image Processing Apps with TensorFlow 2.0 and Keras by Benjamin Planche and Eliot Andres.

Course Outcomes

On the successful completion of the course, students will be able to

S. No	Course Outcomes
CO1	Learn where computer vision techniques are used in industry.

	Identify basic concepts, terminology, theories, models and methods in the field of computer vision
CO2	Identify basic concepts, terminology, theories, models and methods in the field of computer vision
CO3	Demonstrate awareness of the current key research issues in computer vision
CO4	Build a CNN-based image classifier.
CO5	Suggest a design of a computer vision system for a specific problem.

The mapping of course outcomes with program me outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	S	S	M	S	M	L	S	S	M	S	M	S
CO2	S	S	S	S	S	S	M	S	L	S	S	S
CO3	S	S	S	S	S	L	S	S	S	S	M	S
CO4	S	S	S	M	S	S	S	S	M	S	S	M
CO5	S	S	S	S	S	S	S	M	S	S	S	S

S- Strong ; M-Medium ; L-Low

COURSE CODE-22UPCSC4E06   **Credits: 3**

SPARK TECHNOLOGY

Course Objective:

- This course is for novice programmers or business people who would like to understand the core tools used to wrangle and analyze big data.
- Student will have the opportunity to walk through hands-on examples with Hadoop and Spark frameworks, two of the most common in the industry.

Unit 1: Introduction to SPARK: - purpose- Components of the Spark - unified stack Resilient Distributed Dataset (RDD)-Downloading and installing Spark standalone- Scala and Python overview Launching and using Spark's Scala and Python shell.

Unit 2: Distributed Dataset and Data Frames: create parallelized collections and external datasets-Work with Resilient Distributed Dataset (RDD) operations- Utilize shared variables and key-value pairs.

Unit 3: Spark Application: Understand the purpose and usage of the Spark Context- Initialize Spark with the various programming languages-Describe and run some Spark examples- Pass functions to Spark, Create and run a Spark standalone application- Submit applications to the cluster.

Unit 4: Spark configuration: Understand components of the Spark cluster Configure Spark to modify the Spark properties- environmental variables- or logging properties Monitor Spark using the web UIs- metrics- and external instrumentation Understand performance tuning considerations.

Unit 5: Introduction to Spark libraries: Understanding the different library and their usage.

Text Book:

1. Karau, H., Konwinski, A., Wendell, P., &Zaharia, M. Learning spark: lightning-fast big data analysis. " O'Reilly Media, Inc."2015.
2. Zecevic, P., &Bonaci, M. Spark in action. Manning.2017

Reference Book:

1. Chambers, B., &Zaharia, M. Spark: The definitive guide: Big data processing made simple. O'Reilly Media, Inc.",2018

Course Outcomes

On the successful completion of the course, students will be able to

S. No	Course Outcomes
CO1	To provide insight into Big Data Hype, its technologies opportunities, and challenges.
CO2	To design Hadoop stack ranging from the basic HDFS components to application execution frameworks, and languages, services.
CO3	To developed detailed look at the Hadoop Distributed File System (HDFS) and understand the read/write process to HDFS.
CO4	To learn about the big idea of Map/Reduce and you will learn how to design, implement, and execute tasks in the map/reduce framework.
CO5	To learn the trade-offs in map/reduce and how that motivates other tools.

The mapping of course outcomes with program me outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	S	S	L	M	S	S	M	L	S	S	S	M
CO2	S	S	S	S	L	S	S	S	L	M	S	S
CO3	s	S	S	S	S	S	S	S	S	S	S	S
CO4	S	S	S	S	S	S	S	M	S	S	S	S
CO5	S	M	S	S	M	S	S	S	S	M	S	S

S- Strong ; M-Medium ; L-Low

COURSE CODE-22UPCSC4E07   **Credits: 4**

ECONOMETRICS AND FINANCE

Course Objective:

- The objective is to equip students with a firm grounding in the methods and practice of estimation and inference in econometric models.
- Forecast financial data using high-level econometric techniques and measure their effectiveness.
- To extend knowledge and equipment with methods and techniques that allow you to analyse these finance-related issues.

Unit 1:   Introduction to Econometric: - econometrics- Types of data, Steps involved in formulating an econometric model- Econometric packages for modelling financial data.

Unit 2:   Classical Linear Regression Model assumptions-Estimation of the regression model- Properties of Ordinary Least Square estimators.

Unit 3: Regression analysis: Objective- Statistical Analysis and Interpretation of results- Hypothesis Testing-Types of Hypothesis- Test statistic- Critical Region- Multiple Linear Regression Model: Interpretation of the model, Statistical Analysis, Interpretation of the results.

Unit 4: Hypothesis testing: Level of significance and confidence interval approach- Goodness of Fit(R^2)-Concepts of Explained Sum of Squares (ESS)-Residual Sum of Squares -Total Sum of Squares.

Unit 5: Time Series Analysis: Autocorrelation- Stationarity and non-stationarity – reasons- concepts and testing- Random walk with and without drift- Unit-root concept and testing- Co-integration-Granger Causality- Box-Jenkins ARIMA- ARCH and GARCH; Error Correction Model (ECM)- Vector Auto Regressive (VAR).

Text Book:

1. Chris, B. (2014). Introductory econometrics for finance, 3rd Edition.
2. Stock, J. H., & Watson, M. W. (2012). Introduction to econometrics, New York: Pearson.

Reference Book:

1. William H.Greene, Econometric Analysis Second Edition, International publisher.

Course Outcomes

On the successful completion of the course, students will be able to

S. No	Course Outcomes
CO1	To Identify key macroeconomic indicators and measures of economics change, growth, and development.
CO2	Explain the function of market and prices as allocative mechanisms.
CO3	To Discuss the application of analysis.
CO4	To Calculate, present, and discuss descriptive statistics.
CO5	To Demonstrate the ability to collect, process, and interpret data, including statistical inference.

The mapping of course outcomes with program me outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	S	S	L	M	S	S	M	L	S	S	S	L

CO2	S	S	S	S	L	S	S	S	M	S	S	S
CO3	M	S	M	S	S	S	S	S	S	S	M	S
CO4	S	S	S	S	S	S	S	S	S	S	S	M
CO5	S	M	S	S	M	S	S	S	M	S	S	S

S- Strong ; M-Medium ; L-Low

COURSE CODE-22UPCSC4E08   **Credits: 4**

<div align="center">

NATURAL LANGUAGE PROCESSING

</div>

Course Objective:

● NLP relates our perceptions, personal beliefs, emotions, behaviours, attitudes and communication styles
● Resulting the outcome we experience on a day-to-day basis.
● You will progress from understanding the basics, to confidently applying the different NLP principles (and practices) in all aspects of your personal or professional life.

Unit 1:   Introduction: NLP tasks in syntax- semantics- and pragmatics-Applications such as information extraction- question answering- and machine translation-The role of machine learning- Brief history of the field.

Unit 2:   N-gram Language Models: The role of language models- Simple N-gram models-Estimating parameters and smoothing-Evaluating language models.

Unit 3: **Semantic Analysis I:** Lexical semantics and word-sense disambiguation-Compositional semantics-Semantic Role Labeling and Semantic Parsing.

Unit 4:   Semantic   Analysis   II: Lexical   semantics   and   word-sense disambiguation. Compositional semantics-Semantic Role Labeling and Semantic Parsing- Named entity recognition and relation extraction-IE using sequence labeling.

Unit 5:   Text Classification Model: Supervised Classification, Document Classification, Part-of-Speech Tagging-Exploiting Context- Sequence Classification- Sentence Segmentation- Identifying Dialogue Act Types- Recognizing Textual Entailment- Scaling Up to Large Datasets.

Text Book:

1. Elhadad, M. (2010). Book Review: Natural Language Processing with Python by Steven Bird, Ewan Klein, and Edward Loper. Computational Linguistics, 36(4).
2. Bender, E. M. (2013). Linguistic fundamentals for natural language processing: 100 essentials from morphology and syntax. Synthesis lectures on human language technologies, 6(3), 1-184.

Reference Book:

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

Course Outcome:

On the successful completion of the course, students will be able to

S.NO	Course Outcomes
CO1	A language that motivates and creates change – presuppositions.
CO2	To assist other people in an NLP based coaching relationship
CO3	Break out of unhelpful habits and generate the change you want.
CO4	To understand how the mind works and the psychology of mind and behavior.
CO5	Enhance your mental wellbeing, overall assertiveness and resilience. Understand the principal motivations behind human performance

The mapping of course outcomes with programme outcomes is tabulated as follows.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	M	S	M	S	S	S	M	S	S	S	S	M
CO2	S	M	S	S	L	S	S	S	S	S	S	S

CO3	S	S	S	S	M	S	S	S	S	S	L	S
CO4	S	S	S	S	S	S	S	S	S	S	S	M
CO5	S	M	S	M	S	S	S	S	S	M	S	S

S- Strong; M-Medium; L-Low

COURSE CODE-22UPCSC4E09   **Credits: 2**

COMPUTER VISION LAB

Implement the following

1.   Fundamental of computer vision

2.   Introduction to open cv

3.   How to Read, Write, Show Images in Open CV

4.   Image processing using open CV

5.   Image transformation I

6.   Image transformation II

7.   Feature detection and description

8.   Feature matching and model fitting

9.   Colour fundamentals

10.   Clustering and classification

11.   Dimensional reduction and sparse representation

12.   Evaluating the learning models

COURSE CODE-22UPCSC4E10   **Credits: 2**

SPARK TECHNOLOGY LAB

Implement the following

1   Fundamental of spark

2   Introduction to different library mllib,pyspark

3   Software installation

4   working on spark shell

5   Working on spark session and data sources

6   Basic intro to Rdd

7   Working on Dataset and Dataframe

8   Using spark with hadoop

9   Spark Core

10   Spark Streaming

11   Spark SQL

12   MLlib (Machine Learning)

COURSE CODE-22UPCSC4E11   **Credits: 2**

ECONOMETRICS AND FINANCE LAB

Implement the following

1   Fundamental of Econometric

2   Autoregression (AR)

3   Moving Average (MA)

4   Autoregressive Moving Average (ARMA)

5   Autoregressive Integrated Moving Average (ARIMA)

6   Seasonal Autoregressive Integrated Moving-Average (SARIMA)

7   Seasonal Autoregressive Integrated Moving-Average with Exogenous

Regressors (SARIMAX)

8   Vector Autoregression (VAR)

9   Vector Autoregression Moving-Average (VARMA)

10   Vector Autoregression Moving-Average with Exogenous Regressors (VARMAX)

11   Simple Exponential Smoothing (SES)

12   Holt Winter's Exponential Smoothing (HWES)

COURSE CODE- 22UPCSC4E12   **Credits: 2**

NATURAL LANGUAGE PROCESSING LAB

Implement the following

1. Fundamentals of NLP I Tokenization & Lemmatization

2. Fundamentals of NLP II Stemming & Sentence Segmentation

3. NLP Using Scikit Library

4. NLP using Spacy library

5. Working with TF-IDF

6. Naïve Bayes Classifier

7. Word cloud using python

8. Python key word extraction

9. Named entity recognition

10. LATENT semantic analysis

11. Determine optimum number of topics in a document

12. Fundamental of topic modelling