



KANNUR UNIVERSITY
കണ്ണൂർ സർവകലാശാല

(Abstract)

M.Sc. Data Science - Scheme & Syllabus - Approved and implemented with effect from the Academic Year 2026-2027 - Order issued.

ACADEMIC J SECTION

ACAD J/ACAD J1/13187/2026

Dated: 13.10.2026

- Read:-1. Letter No. G/A/95/1/2026 Dtd: 03.07.2026 from the Principal, Mahatma Gandhi College, Iritty.
2.Email dated 17.07.2026 to the Dean, Faculty of Technology.
3.Email dated 13.08.2026 from the Dean, Faculty of Technology.
4. Orders of the Vice Chancellor in the file No. ACAD J/ACAD J1/13187/2026 Dtd : 17.08.2026.

ORDER

1. As per the paper read (1) above, The Principal, Mahatma Gandhi College, Iritty submitted the scheme & syllabus of M.Sc. Data Science programme to be offered in the Mahatma Gandhi College, Iritty.
2. As per the paper read (2) above, the same was forwarded to the Dean, Faculty of Technology, for remarks. The Dean approved the syllabus vide Paper (3) above.
3. The Vice Chancellor, after considering the recommendation of the Dean and in exercise of the powers of the Academic Council conferred under section 11 (1), chapter III of the Kannur University Act, 1996, read with all other enabling provisions, approved the Scheme & Syllabus of M.Sc. Data Science programme and implemented with effect from the academic Year 2026-2027. The Vice Chancellor has further ordered to report the matter before the Academic Council for ratification.
4. Scheme & Syllabus are appended to this university order and uploaded on the University website(www.kannuruniv.ac.in).
5. Orders are issued accordingly.

Sd/-

Jisha K P
Assistant Registrar II
For REGISTRAR

To: The Principal, Mahatma Gandhi College, Iritty
The Dean, Faculty of Technology

Copy To: PS to VC / P A to Registrar

P A to CE

To web Manager (to publish in the website)



Forwarded / By Order

[Signature]
SECTION OFFICER





KANNUR UNIVERSITY

SYLLABI FOR M.Sc. DATA SCIENCE

Under the Choice Based Credit and Semester System (CBCSS)

2026 Admission Onwards



PROGRAMME OUTCOME

PO1	Advanced Analytical and Critical Thinking Apply advanced analytical reasoning, statistical thinking, and computational methods to identify, analyze, and solve complex real-world data-driven problems.
PO2	Data-Driven Decision Making Develop the ability to extract meaningful insights from structured and unstructured data using statistical, mathematical, and machine learning techniques for informed decision-making.
PO3	Technical Competency and Innovation Demonstrate proficiency in modern data science tools, programming languages, algorithms, artificial intelligence, and emerging technologies for innovative solution development.
PO4	Research, Communication and Collaboration Effectively communicate technical concepts, research findings, and analytical results through scientific reports, visualizations, presentations, and collaborative interdisciplinary teamwork.
PO5	Ethical, Responsible and Sustainable Data Practices Apply ethical principles, data privacy regulations, cybersecurity awareness, and sustainable computing practices while handling data and deploying intelligent systems.
PO6	Leadership, Lifelong Learning and Global Perspective Exhibit leadership qualities, professional responsibility, adaptability, and continuous learning to address evolving global challenges in academia, industry, and society.



PROGRAMMESPECIFICOUTCOMES

PSO1	Apply advanced statistical, mathematical, and computational techniques for acquiring, processing, analysing, and interpreting large-scale datasets.
PSO2	Design and develop intelligent data-driven solutions using machine learning, deep learning, artificial intelligence, and predictive analytics techniques.
PSO3	Utilize modern programming languages, big data platforms, cloud technologies, and visualization tools for efficient data management and scalable analytics.
PSO4	Develop and deploy data science models and analytical frameworks to solve interdisciplinary problems in domains such as healthcare, finance, business, environment, education, and governance.
PSO5	Conduct independent and collaborative research by integrating domain knowledge, scientific methodologies, and emerging technologies in data science.
PSO6	Apply ethical standards, data security principles, and responsible AI practices while handling sensitive data and developing intelligent systems for societal benefit.
PSO7	Demonstrate professional competency, entrepreneurship, innovation, and lifelong learning skills to adapt to rapidly evolving trends in data science and artificial intelligence.



STRUCTURE OF THE PROGRAMME

The Programme of instruction for M.Sc. Data Science shall consist of Lecture Courses, Practical/Laboratory Courses, Tutorial Sessions, Seminars, Viva-Voce, Mini Project, Major Project/Dissertation, and Internship/Industrial or Institutional Visit.

1. Lecture Courses:

Courses involving classroom lectures delivered by faculty members to provide advanced theoretical knowledge in areas such as statistics, machine learning, artificial intelligence, big data analytics, data visualization, cloud computing, and related domains of Data Science.

2. Tutorial Courses:

Courses involving guided problem-solving sessions, case studies, discussions, coding exercises, and analytical assignments related to Data Science concepts under the supervision of qualified faculty members.

3. Practical / Laboratory Work:

Courses requiring students to perform laboratory experiments, programming exercises, data analysis tasks, model development, simulations, and implementation of algorithms using modern software tools, programming languages, and data science platforms under faculty supervision.

4. Elective Courses:

Three elective courses are in the IVth semester. One from Statistics and Two from Computer Science. Students can select the courses from the basket of courses proposed in the syllabus.

5. Seminar:

A course requiring students to present and discuss recent developments, research papers, case studies, emerging technologies, and applications in Data Science and allied fields. The seminar aims to enhance presentation skills, technical communication, critical thinking, and research aptitude.

6. Comprehensive Viva-Voce:

This is an essential component of assessment intended to evaluate the student's conceptual understanding, analytical ability, technical knowledge, research orientation, and application



skills in Data Science. The viva-voce also provides an opportunity for academic interaction and expert evaluation.

7. Mini Project:

Students shall undertake a mini project individually or in groups during the Programme under the guidance of a faculty member. The mini project shall focus on applying Data Science techniques and tools to solve a specific real-world problem, thereby strengthening practical knowledge, teamwork, innovation, and problem-solving skills.

8. Internship / Industrial or Institutional Visit:

Students may undergo internship training, industrial training, field visits, or institutional visits in industries, research organizations, startups, IT firms, analytics companies, laboratories, Higher Education Institutions (HEIs), or research centres relevant to Data Science and Artificial Intelligence. Such exposure shall help students understand industrial practices, research methodologies, professional ethics, and current technological trends.

9. Major Project / Dissertation:

Every student shall carry out a Major Project/Dissertation under the guidance of a faculty member during the final phase of the Programme. The project shall involve independent research, development, implementation, or analytical work in a specialized area of Data Science. Students shall submit a dissertation/report and successfully defend the work through project presentation and viva-voce examination for the award of credits. The project may be carried out in collaboration with industries, research institutions, or academic organizations wherever feasible.



COURSE STRUCTURE

COURSE CODE	NAME OF THE COURSE	HOURS / WEEK	CREDITS	TOTAL MARKS
SEMESTER I				
KU1DSMAT01	Mathematical foundations for Data Science	2	2	50
KU1DSSTA01	Statistical Methods – I	3	3	70
KU1DSSTA02	An Introduction to Statistical Data Analysis Using R	2	2	50
KU1DSCSC01	Advanced DBMS	3	3	70
KU1DSCSC02	Python for Data Science	3	3	70
KU1DSSTA02-LAB	Statistical Data Analysis Using R	4	2	40
KU1DSCSC01-LAB	Advanced DBMS	4	2	40
KU1DSCSC02-LAB	Python for Data Science	4	2	40
TOTAL		25	19	450
SEMESTER II				
KU2DSMAT01	Mathematical Methods for Data Science	2	2	50
KU2DSSTA01	Statistical Methods - II	3	3	70
KU2DSSTA02	Multivariate Analysis	2	2	50
KU2DSCSC01	Design and Analysis of Algorithms	3	3	70
KU2DSCSC02	Data Visualization	3	3	70
KU2DSSTA02-LAB	Advanced Statistical Data Analysis using R/Python	4	2	40
KU2DSCSC01-LAB	Design and Analysis of Algorithms	4	2	40
KU2DSCSC02-LAB	Data Visualization	4	2	40
KU2DSCSC06	Mini Project	-	2	50
TOTAL		25	21	480
SEMESTER III				



KU3DSSTA01	Bayesian Statistics, Time Series Analysis and Forecasting Techniques	3	3	70
KU3DSCSC01	Machine learning Techniques	2	2	50
KU3DSCSC02	Natural Language Processing	3	3	70
KU3DSCSC03	Cloud Computing and Data Engineering	3	3	70
KU3DSSTA01-LAB	LAB-Bayesian Statistics, Time Series Analysis and Forecasting Techniques	4	2	40
KU3DSCSC01-LAB	Machine learning Techniques for big data	3	2	40
KU3DSCSC02-LAB	Natural Language Processing	4	2	40
KU3DSCSC03-LAB	Cloud Computing and Data Engineering	3	2	40
TOTAL		25	19	420
SEMESTER IV				
KU4DSSTA01	Elective I (Statistics)	3	3	70
KU4DSCSC-ELE01	Elective II (Computer Science)	3	3	70
KU4DSCSC-ELE02	Elective III (Computer Science)	3	3	70
KU4DSSTA02-LAB	Statistics (Elective)	4	2	40
KU4DSCSC03-LAB	Computer Science I	3	2	40
KU4DSCSC04-LAB	Computer Science II	3	2	40
KU4DSCSC05	Project	6	6	100
TOTAL		25	21	430

Baskets of Electives	
Elective I (Statistics)	1. SURVIVAL ANALYSIS AND RISK ANALYTICS 2. FINANCIAL STATISTICS AND RISK ANALYTICS 3. BIOSTATISTICS AND RISK ANALYTICS 4. STOCHASTIC PROCESSES 5. OPTIMIZATION TECHNIQUES FOR DATA SCIENCE
Elective II (Computer Science)	1. DEEP LEARNING TECHNIQUES 2. GENERATIVE AI AND LARGE LANGUAGE MODELS 3. SCALABLE DATA SYSTEMS



	4. DATA VISUALIZATION AND VISUAL ANALYTICS 5. BUSINESS INTELLIGENCE AND DECISION SUPPORT SYSTEMS
Elective III (Computer Science)	1. ETHICAL AI AND RESPONSIBLE COMPUTING 2. SOFT COMPUTING TECHNIQUES 3. ADVANCED COMPUTER VISION 4. SOCIAL MEDIA ANALYTICS 5. INTERNET OF THINGS



SYLLABUS

SEMESTER I

KU1DSMAT01

Course Title: MATHEMATICAL FOUNDATIONS FOR DATA SCIENCE

Course Structure

- Lecture Hours: 2 Hours/Week
- Credits: 2

Course Objectives

The course aims to:

1. Develop a strong mathematical foundation required for data science and machine learning.
2. Introduce matrix algebra and vector spaces for data representation and analysis.
3. Explain eigenvalue methods and matrix decompositions used in dimensionality reduction.
4. Develop computational skills in differential and multivariable calculus for optimization problems.

Course Outcomes (COs)

CO	Course Outcome	Bloom's Knowledge Level
CO1	Perform matrix operations and solve systems of linear equations for data analysis.	Apply (K3)
CO2	Explain vector spaces, linear transformations, eigenvalues, eigenvectors and Singular Value Decomposition.	Understand (K2)
CO3	Apply differential calculus to analyze functions and optimization problems.	Apply (K3)
CO4	Compute partial derivatives, gradients, Jacobians and Hessians of multivariable functions.	Apply (K3)
CO5	Apply mathematical concepts using computational tools to solve elementary data science problems.	Apply (K3)

Knowledge Levels

- K1 – Remember
- K2 – Understand
- K3 – Apply
- K4 – Analyze
- K5 – Evaluate
- K6 – Create

Module I – Matrix Algebra (7 Hours)

- Introduction to Linear Algebra
- Scalars, Vectors and Matrices
- Types of Matrices
- Matrix Operations
- Transpose and Symmetric Matrices
- Determinants



- Inverse of a Matrix
- Rank of a Matrix
- Trace of a Matrix
- Elementary Row Operations
- Systems of Linear Equations
- Gaussian Elimination

Module II – Vector Spaces and Matrix Decomposition (8 Hours)

Vector Spaces

- Vector Spaces
- Subspaces
- Span of Vectors
- Linear Independence and Dependence
- Basis and Dimension
- Orthogonality and Orthogonal Projections
- Inner Product and Norms
- Linear Transformations

Eigen Analysis

- Characteristic Equation
- Eigenvalues and Eigenvectors
- Diagonalization
- Positive Definite Matrices
- Spectral Decomposition (Introduction)

Module III – Differential Calculus for Data Science (7 Hours)

- Functions and Graphs
- Limits and Continuity
- Differentiability
- Rules of Differentiation
- Higher Order Derivatives
- Taylor Series
- Maxima and Minima of Single Variable Functions

Module IV – Multivariable Calculus (8 Hours)

- Functions of Several Variables
- Partial Derivatives
- Total Derivatives
- Chain Rule
- Higher Order Partial Derivatives
- Jacobian Matrix
- Hessian Matrix

Reference Books

1. Strang, G. (2016). *Introduction to Linear Algebra* (5th ed.). Wellesley-Cambridge Press.
2. Strang, G. (2019). *Linear Algebra and Learning from Data*. Wellesley-Cambridge Press.
3. Lay, D. C., Lay, S. R., & McDonald, J. J. (2021). *Linear Algebra and Its Applications* (6th ed.). Pearson.
4. Anton, H., & Rorres, C. (2013). *Elementary Linear Algebra* (11th ed.). Wiley.
5. Stewart, J. (2020). *Calculus: Early Transcendentals* (9th ed.). Cengage Learning.

Additional References

6. Marsden, J. E., & Tromba, A. J. (2012). *Vector Calculus* (6th ed.). W. H. Freeman.
7. Kreyszig, E. (2011). *Advanced Engineering Mathematics* (10th ed.). Wiley.
8. Boyd, S., & Vandenberghe, L. (2018). *Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares*. Cambridge University Press.
9. Deisenroth, M. P., Faisal, A. A., & Ong, C. S. (2020). *Mathematics for Machine Learning*. Cambridge University Press.



10. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. (Relevant chapters on Linear Algebra and Calculus)

KU1DSSTA01

Course Title: STATISTICAL METHODS – I

COURSE STRUCTURE

- Lecture Hours: 3 Hours/Week
- Credits: 3

COURSE OBJECTIVES

The main objectives of this course are to:

1. Introduce the fundamental concepts of descriptive and inferential statistics.
2. Develop knowledge of probability theory and random variables for data analysis.
3. Understand various probability distributions and their applications in data science.
4. Apply correlation and regression techniques for analysing relationships among variables.
5. Build statistical foundations required for machine learning and predictive analytics.

EXPECTED COURSE OUTCOMES

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

CO No.	Course Outcome	Knowledge Level
CO1	Explain the concepts of descriptive statistics and data summarization techniques	K1/K2
CO2	Apply measures of central tendency and dispersion to analyze datasets	K3
CO3	Understand probability concepts and random variables in statistical modelling	K2/K3
CO4	Analyse probability distributions and their real-life applications	K4
CO5	Apply correlation and regression methods for predictive analysis and interpretation	K3/K4/K5

Knowledge Levels

- K1 – Remember
- K2 – Understand
- K3 – Apply
- K4 – Analyze
- K5 – Evaluate
- K6 – Create

Module I – Descriptive Statistics and Data Summarization (12 Hours)

- Introduction to Statistics
- Collection and Classification of Data
- Tabular and Graphical Representation of Data
- Frequency Distribution
- Histogram, Bar Diagram, Pie Chart, Ogive
- Descriptive Statistics and Data Summarization Techniques- Bar plot, Q-Q plots, Stem and Leaf plots.
- Arithmetic Mean
- Median and Mode
- Geometric Mean and Harmonic Mean
- Partition Values: Quartiles, Deciles and Percentiles
- Measures of Dispersion
- Range
- Quartile Deviation



- Mean Deviation
- Variance and Standard Deviation
- Coefficient of Variation
- Skewness and Kurtosis

Module II – Probability Theory and Random Variables (10 Hours)

- Basic Concepts of Probability
- Sample Space and Events
- Addition and Multiplication Theorems
- Conditional Probability
- Independent Events
- Random Variables
- Discrete and Continuous Random Variables
- Probability Mass Function and Density Function
- Mathematical Expectation and Variance

Module III – Probability Distributions, Correlation and Regression (12 Hours)

Probability Distributions

- Binomial Distribution
- Poisson Distribution
- Exponential Distribution
- Gamma Distribution
- Beta Distribution
- Normal Distribution
- Standard Normal Distribution
- Simple Properties and Applications

Module IV – Correlation and Regression (11 hours)

Correlation

- Scatter Diagram
- Karl Pearson's Correlation Coefficient
- Spearman's Rank Correlation
- Interpretation of Correlation

Regression

- Simple Linear Regression
- Regression Lines
- Regression Coefficients
- Prediction and Estimation
- Applications in Data Science

Multiple and Partial Correlation-An introduction with three variables.

Reference Books

1. Fundamentals of Mathematical Statistics – S. C. Gupta and V. K. Kapoor, Sultan Chand Publications.
2. Introduction to Probability and Statistics – William Mendenhall, Robert J. Beaver and Barbara M. Beaver, Cengage Learning.
3. Applied Statistics and Probability for Engineers – Douglas C. Montgomery and George C. Runger, Wiley Publications.
4. Business Statistics – S. P. Gupta and M. P. Gupta, Sultan Chand & Sons.
5. Statistics for Data Science – Packt Publishing.
6. Practical Statistics for Data Scientists – Peter Bruce, Andrew Bruce and Peter Gedeck, O'Reilly Publications.

KU1DSSTA02

Course Title: AN INTRODUCTION TO STATISTICAL DATA ANALYSIS USING R



Credits: 2

Contact Hours: 2 Hours/Week

Course Type: Core/Foundation Course

Course Description

This course introduces students to statistical data analysis using the R programming language. It covers data management, exploratory data analysis, statistical methods, visualization, and basic predictive modelling. The course emphasizes hands-on learning using real-world datasets and equips students with essential analytical skills required for data science applications.

Course Objectives

1. To introduce the fundamentals of R programming for statistical computing.
2. To develop skills in data manipulation, visualization, and exploratory data analysis.
3. To apply descriptive and inferential statistical techniques using R.
4. To perform basic predictive modeling and interpret statistical results.

Course Outcomes

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

- Use R and RStudio effectively for data analysis.
- Import, clean, and manipulate datasets.
- Generate statistical summaries and visualizations.
- Conduct hypothesis tests and interpret results.
- Develop and evaluate simple regression models.
- Communicate analytical findings through reports and visualizations.

Module I: Fundamentals of R Programming and Data Handling (8 Hours)

Topics

- Introduction to R and RStudio
- R Environment and Workspace
- Variables, Data Types, and Operators
- Vectors, Matrices, Lists, and Data Frames
- Control Structures and User-Defined Functions

Module II: Exploratory Data Analysis and Data Visualization (7 Hours)

Topics

- Importing and Exporting Data
- Data Cleaning and Preprocessing
- Handling Missing Values and Outliers
- Introduction to Exploratory Data Analysis (EDA)
- Data Summarization Techniques
- Measures of Central Tendency and Dispersion
- Frequency Distributions

Module III: Probability and Statistical Inference Using R (7 Hours)

Topics

- Data Aggregation and Grouping
- Graphical Representation of Data
- Histograms, Bar Charts, Pie Charts
- Box Plots, Scatter Plots, Line Charts
- Introduction to ggplot2
- Basic Probability Concepts
- Conditional Probability and Bayes' Theorem
- Random Variables and Probability Distributions
 - Binomial Distribution
 - Poisson Distribution
 - Normal Distribution etc.



Module IV: Correlation, Regression, and Reporting (8 Hours)

Topics

- Correlation Analysis
 - Pearson Correlation
 - Spearman Correlation
- Simple Linear Regression
- Model Evaluation and Diagnostics
- Residual Analysis
- Case Studies and Applications in Data Science

Suggested Text Books

1. Purohit, Gore, Deshmukh, Statistics Using R
2. Peter Dalgaard, *Introductory Statistics with R*.
3. Norman Matloff, *The Art of R Programming*.

References

1. Gareth James et al., *An Introduction to Statistical Learning with Applications in R*.
2. Peter Bruce & Andrew Bruce, *Practical Statistics for Data Scientists*.

KU1DSCSC01

Course Title: ADVANCED DATABASE MANAGEMENT SYSTEMS

Course Structure

- Lecture Hours: 3 Hours/Week
- Credits: 3

Course Objectives

The main objectives of this course are to:

1. Introduce the concepts and architecture of modern database management systems.
2. Develop understanding of relational database design and advanced normalization techniques.
3. Explain transaction processing and concurrency control mechanisms in DBMS.
4. Familiarize students with distributed databases, graph databases and modern database technologies for data science.
5. Provide practical knowledge of SQL and NoSQL databases for large-scale data management and analytics.

Expected Course Outcomes

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

CO No.	Course Outcome	Knowledge Level
CO1	Explain database concepts, architectures and data models used in modern applications	K1/K2
CO2	Design efficient relational databases using advanced normalization techniques	K3/K4
CO3	Apply transaction processing and concurrency control techniques for reliable database systems	K3/K4
CO4	Analyze distributed, graph and NoSQL database systems for data science applications	K4/K5
CO5	Develop SQL and NoSQL database solutions for large-scale analytics and network-based applications	K5/K6

Knowledge Levels

- K1 – Remember



- K2 – Understand
- K3 – Apply
- K4 – Analyze
- K5 – Evaluate
- K6 – Create

Module I – Fundamentals of Database Management Systems and Database Design (15 Hours)

- Introduction to DBMS
- File System vs DBMS
- Database Users and DBA
- Database Architecture
- Data Models
- Database Design Process

Advanced Normalization Techniques

- First Normal Form (1NF)
- Second Normal Form (2NF)
- Third Normal Form (3NF)
- Boyce-Codd Normal Form (BCNF)

Module II – SQL(15 Hours)

- Data Definition Language (DDL)
- Data Manipulation Language (DML)
- Data Control Language (DCL)
- Queries using SELECT
- Joins and Subqueries
- Aggregate Functions
- Views and Indexes
- Stored Procedures and Triggers
- Transaction Concepts
- ACID Properties
- Database Security and Authorization

Module III – Advanced Database Architecture and Distributed DBMS (15 Hours)

Advanced Database Architecture

- Client-Server Architecture
- Parallel Database Systems
- Data Warehousing Architecture
- Cloud Database Systems

Distributed DBMS

- Concepts of Distributed Databases
- Fragmentation and Replication
- Distributed Query Processing
- Distributed Transactions
- Distributed Concurrency Control
- CAP Theorem
- Fault Tolerance and Recovery

Database Technologies for Data Science

- Data Warehousing
- Data Lakes
- ETL Processes
- Real-time Data Processing
- Database Scalability for Big Data Analytics

Module IV – NoSQL Databases and Graph Databases (15 Hours)

NoSQL Databases

- Need for NoSQL Systems
- SQL vs NoSQL Databases



- Key-Value Databases
- Document Databases
- Column-oriented Databases
- Graph Databases
- Graph Data Model
- Nodes, Edges and Relationships
- Graph Query Languages
- Neo4j Introduction
- Applications in Recommendation Systems and Fraud Detection
- Modern Database Technologies

Reference Books

1. Database System Concepts, Abraham Silberschatz, Henry F. Korth, S. Sudarshan McGraw-Hill
2. Fundamentals of Database Systems, Ramez Elmasri, Shamkant B. Navathe, Pearson
3. Database Management Systems, Raghu Ramakrishnan, Johannes Gehrke, McGraw-Hill
4. Distributed Database Systems, M. Tamer Özsu, Patrick Valduriez, Springer
5. NoSQL Distilled, Pramod J. Sadalage, Martin Fowler, Addison-Wesley
6. Graph Databases. Ian Robinson, Jim Webber, Emil Eifrem, O'Reilly Media
7. Mining of Massive Datasets, Jure Leskovec, Anand Rajaraman, Jeffrey Ullman, Cambridge University Press
8. MongoDB: The Definitive Guide, Kristina Chodorow, Shannon Bradshaw, O'Reilly Media

KU1DSCSC02

Course Title: PYTHON FOR DATA SCIENCE

Course Structure

- Lecture Hours: 3 Hours/Week
- Total Credits: 3

Course Objectives

The main objectives of this course are to:

1. Introduce the fundamentals of Python programming for data science applications.
2. Develop problem-solving skills using Python programming constructs and data structures.
3. Familiarize students with object-oriented programming concepts and file handling techniques.
4. Enable students to perform scientific computing and data manipulation using Python libraries.
5. Provide practical knowledge for developing data processing applications using Python.

Expected Course Outcomes

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

CO No.	Course Outcome	Knowledge Level
CO1	Explain Python programming concepts and syntax	K1/K2
CO2	Develop Python programs using control structures, functions and data structures	K3
CO3	Apply object-oriented programming concepts and file handling techniques	K3/K4
CO4	Perform scientific computing and data manipulation using Python libraries	K4/K5
CO5	Develop Python-based applications for data processing and analysis	K5/K6

Knowledge Levels

- K1 – Remember



- K2 – Understand
- K3 – Apply
- K4 – Analyze
- K5 – Evaluate
- K6 – Create

Module I – Fundamentals of Python Programming (15 Hours)

- Introduction to Python
- Features and Applications of Python
- Python Installation and IDEs
- Variables, Data Types, Operators and Expressions
- Input and Output Functions
- Conditional Statements
- Looping Statements
- Functions in Python
- Modules and Packages

Module II – Data Structures and Object-Oriented Programming (15 Hours)

Data Structures

- Lists
- Tuples
- Sets
- Dictionaries
- String Manipulation
- List Comprehensions

Object-Oriented Programming

- Classes and Objects
- Constructors
- Inheritance
- Polymorphism
- Encapsulation

Module III – Exception Handling and File Processing (15 Hours)

- Exception Handling
- Types of Errors and Exceptions
- Try, Except and Finally Blocks
- User-defined Exceptions
- File Handling
- Reading and Writing Text Files
- CSV File Processing
- Introduction to Regular Expressions
- Python Standard Libraries

Module IV – Scientific Computing and Data Manipulation (15 Hours)

NumPy

- Introduction to NumPy
- Arrays and Array Operations
- Matrix Operations
- Broadcasting
- Mathematical Functions

Pandas

- Series and Data Frames
- Importing and Exporting Data
- Data Cleaning
- Data Transformation
- Data Aggregation and Grouping
- Handling Missing Values



Reference Books

1. Python for Data Analysis, Wes McKinney, O'Reilly Media
2. Learning Python, Mark Lutz, O'Reilly Media
3. Python Crash Course, Eric Matthes, No Starch Press
4. Automate the Boring Stuff with Python, Al Sweigart, No Starch Press
5. Think Python, Allen B. Downey, O'Reilly Media
6. Programming Python, Mark Lutz, O'Reilly Media
7. Core Python Programming, R. Nageswara Rao, Dreamtech Press

KU1DSSTA02-LAB

SEMESTER – I: LAB- STATISTICAL DATA ANALYSIS USING R COURSE STRUCTURE

- Lab: 4 Hours/Week
- Credits: 2

Practical Exercises:

Students shall perform practical exercises using Rin

1. Data Import and Cleaning
2. Descriptive statistics
3. Graphical representation of data
4. Probability calculations
5. Monte Carlo Simulations
6. Monte Carlo Integration
7. Distribution fitting
8. Correlation analysis
9. Regression modelling
10. Mini Data Analysis Project

KU1DSCSC01-LAB

Course Title: LAB-I: ADVANCED DBMS

COURSE STRUCTURE

- Lab Hours: 4 Hours/Week
- Credits: 2

Students shall perform the practical exercises in Lab:

- ER Diagram Design
- Database Normalization
- SQL Queries and Stored Procedures
- Distributed Database Operations
- Operations on NoSQL
- Graph Database Queries using Neo4j

KU1DSCSC02-LAB



Course Title: LAB-II: PYTHON FOR DATA SCIENCE

COURSE STRUCTURE

- Lab Hours: 4 Hours/Week
- Credits: 3

Students shall perform the practical exercises in Lab:

- Basic Python programming
- Control structures and functions
- Data structures and OOP concepts
- Exception handling
- File processing
- NumPy array operations
- Data analysis using Pandas
- CSV data handling

SEMESTER II

KU2DSMAT01

Course Title: MATHEMATICAL METHODS FOR DATA SCIENCE

Credits: 2

Instruction Hours: 30 Hours

Course Objectives

The course aims to:

1. Develop mathematical reasoning and analytical skills essential for data science.
2. Introduce vector calculus and numerical methods used in scientific computing and machine learning.



3. Build a foundation in discrete mathematics and graph theory for network analysis and algorithm design.
4. Familiarize students with mathematical tools used in optimization, probabilistic modelling, and artificial intelligence.
5. Apply mathematical methods using computational tools to solve real-world data science problems.

Course Outcomes (COs)

CO	Course Outcome	Bloom's Knowledge Level
CO1	Explain the concepts of sequences, series and multivariable functions relevant to mathematical modelling.	Understand (K2)
CO2	Apply vector calculus concepts such as gradient, divergence, curl and multiple integrals in solving data science problems.	Apply (K3)
CO3	Apply numerical methods to solve nonlinear equations, perform numerical differentiation and integration, and analyze computational errors.	Apply (K3)
CO4	Apply concepts of logic, combinatorics and graph theory to mathematical modelling and network analysis.	Apply (K3)
CO5	Use Python/R libraries to implement numerical and graph-theoretic algorithms for data science applications.	Apply (K3)

Knowledge Levels

- K1 – Remember
- K2 – Understand
- K3 – Apply
- K4 – Analyze
- K5 – Evaluate
- K6 – Create

Module I – Sequences, Series and Multivariable Functions (7 Hours)

- Sequences and Series
- Convergence of Sequences
- Infinite Series
- Power Series
- Real-valued Functions of Several Variables
- Domain and Range
- Limits and Continuity of Multivariable Functions
- Differentiability
- Taylor's Theorem for Several Variables (Introduction)

Module II – Vector Calculus (8 Hours)

- Scalar and Vector Fields
- Gradient
- Divergence
- Curl
- Line Integrals (Introduction)
- Multiple Integrals (Double and Triple Integrals – Basic Concepts)
- Change of Variables
- Tangent Plane (Introduction)

Module III – Numerical Methods for Data Science (7Hours)

- Sources of Numerical Errors
- Floating Point Arithmetic
- Error Analysis
- Solution of Nonlinear Equations
 - Bisection Method



- Newton–Raphson Method
- Numerical Differentiation
- Numerical Integration
 - Trapezoidal Rule
 - Simpson's Rule
- Numerical Solution of Linear Systems (Introduction)
- Numerical Differentiation (Introduction)

Module IV – Discrete Mathematical Foundations (8 Hours)

- Sets and Relations (Review)
- Functions and Mappings
- Mathematical Logic
- Proof Techniques (Direct, Contradiction, Induction)
- Counting Principles
- Basic Combinatorics
- Recurrence Relations (Introduction)
- Graphs and Networks
- Trees
- Introduction to Graph Matrices (Adjacency and Incidence Matrices)

Reference Books

1. Kreyszig, E. (2011). *Advanced Engineering Mathematics* (10th ed.). Wiley.
2. Burden, R. L., & Faires, J. D. (2015). *Numerical Analysis* (10th ed.). Cengage Learning.
3. Rosen, K. H. (2019). *Discrete Mathematics and Its Applications* (8th ed.). McGraw-Hill.
4. Anton, H., Bivens, I., & Davis, S. (2013). *Calculus* (10th ed.). Wiley.
5. Deisenroth, M. P., Faisal, A. A., & Ong, C. S. (2020). *Mathematics for Machine Learning*. Cambridge University Press.

Additional References

6. Marsden, J. E., & Tromba, A. J. (2012). *Vector Calculus* (6th ed.). W. H. Freeman.
7. Gilbert Strang (2019). *Linear Algebra and Learning from Data*. Wellesley-Cambridge Press.
8. Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2022). *Introduction to Algorithms* (4th ed.). MIT Press. (Relevant chapters on graph algorithms)
9. Bondy, J. A., & Murty, U. S. R. (2008). *Graph Theory*. Springer.
10. Sauer, T. (2022). *Numerical Analysis* (4th ed.). Pearson.

KU2DSSTA01

Course Title: STATISTICAL METHODS – II

Course Structure

- Lecture Hours: 3 Hours/Week
- Credits: 3

Course Objectives

The main objectives of this course are to:

1. Introduce the concepts and methods of statistical sampling and estimation.
2. Develop understanding of statistical inference and hypothesis testing procedures.
3. Provide knowledge of analysis of variance techniques for comparing multiple populations.
4. Familiarize students with non-parametric statistical methods for distribution-free analysis.
5. Explain the principles and applications of design of experiments in data science and research.

Expected Course Outcomes

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

CO No. Course Outcome

- | | |
|-----|--|
| CO1 | Explain the principles and methods of statistical sampling |
| CO2 | Apply estimation techniques and construct confidence intervals |

Knowledge Level

- | |
|-------|
| K1/K2 |
| K3 |



CO No.	Course Outcome	Knowledge Level
CO3	Perform hypothesis testing using appropriate statistical tests	K3/K4
CO4	Analyze experimental data using ANOVA and non-parametric methods	K4/K5
CO5	Design and interpret statistical experiments for real-world applications	K5/K6

Knowledge Levels

- K1 – Remember
- K2 – Understand
- K3 – Apply
- K4 – Analyze
- K5 – Evaluate
- K6 – Create

Module I – Sampling Theory and Estimation (10 Hours)

Sampling Theory

- Introduction to Sampling
- Census and Sample Methods
- Sampling and Non-sampling Errors
- Probability and Non-probability Sampling
- Simple Random Sampling
- Stratified Sampling
- Systematic Sampling

Estimation

- Statistical Estimation
- Point and Interval Estimation
- Properties of Good Estimators
- Confidence Intervals for Mean and Proportion

Module II – Hypothesis Testing (12 Hours)

- Statistical Hypothesis
- Null and Alternative Hypotheses
- Level of Significance
- Type I and Type II Errors
- Critical Region and p-value
- One-tailed and Two-tailed Tests
- Tests based on Normal Distribution
- *t*-Test
- Chi-square Test
- F-Test
- Applications of Hypothesis Testing in Data Science
- A brief introduction to Stochastic Process and Markov Chain

Module III – Analysis of Variance and Non-parametric Tests (10 Hours)

Analysis of Variance (ANOVA)

- Introduction to ANOVA
- One-way ANOVA
- Two-way ANOVA
- Assumptions of ANOVA

Non-parametric Tests

- Introduction to Non-parametric Methods
- Sign Test
- Wilcoxon Signed Rank Test
- Mann–Whitney U Test
- Kruskal–Wallis Test



- Chi-square Test for Independence

Module IV – Statistical Simulation Methods (13 Hours)

- Acceptance–Rejection Method
- Importance Sampling (Basic Concepts)
- Applications in Bayesian Computation
- Numerical Optimization for Maximum Likelihood Estimation
- Newton–Raphson Method
- Simulation-based Evaluation of Estimators
- Bootstrap Estimation of Standard Errors
- Introduction to the EM Algorithm
- Missing Data Problems
- E-step and M-step

Reference Books

1. Fundamentals of Applied Statistics – Sultan Chand Publications.
2. Statistical Methods – Sultan Chand & Sons.
3. Introduction to the Theory of Statistics – McGraw-Hill Publications.
4. Design and Analysis of Experiments – Wiley Publications.
5. Nonparametric Statistical Methods – Wiley Publications.
6. Practical Statistics for Data Scientists – O'Reilly Publications.
7. Simulation – Sheldon M. Ross.
8. Monte Carlo Statistical Methods – Christian P. Robert & George Casella.
9. The EM Algorithm and Extensions – Geoffrey J. McLachlan & Thiriyambakam Krishnan.

KU2DSSTA02

Course Title: MULTIVARIATE ANALYSIS

Credits: 2

Contact Hours: 2 Hours/Week

Course Description

This course introduces the theory and applications of multivariate statistical methods used for analyzing datasets involving multiple variables simultaneously. The course focuses on data reduction, classification, clustering, and interpretation of complex datasets with applications in data science, business analytics, finance, healthcare, and social sciences.

Course Objectives

1. To understand the structure and characteristics of multivariate data.
2. To apply dimensionality reduction techniques for high-dimensional datasets.
3. To perform classification and clustering using multivariate statistical methods.
4. To interpret multivariate analytical results for decision-making and predictive analytics.

Course Outcomes

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

- Analyze and summarize multivariate datasets.
- Apply Principal Component Analysis and Factor Analysis for dimensionality reduction.
- Use Discriminant Analysis for classification problems.
- Perform Cluster Analysis for data segmentation and pattern discovery.
- Interpret and communicate results from multivariate analyses.

Module I – Introduction to Multivariate Analysis (7 Hours)

- Introduction to Multivariate Data
- Types and Structure of Multivariate Data
- Mean Vector and Covariance Matrix



- Correlation Matrix
- Essentials of Multivariate Normal Distribution
 - Definition and Properties
 - Marginal and Conditional Distributions
 - Linear Combinations of Normal Variables
 - Mahalanobis Distance

Module II – Principal Component Analysis (8 Hours)

- Introduction to Principal Component Analysis
- Eigenvalues and Eigenvectors
- Principal Components
- Proportion of Variance Explained
- Interpretation of Principal Components
- Applications of PCA in Data Science

Module III – Factor Analysis and Cluster Analysis (7 Hours)

Factor Analysis

- Introduction and Objectives
- Common Factor Model
- Factor Extraction
- Interpretation of Factor Loadings

Cluster Analysis

- Measures of Similarity and Distance
- Hierarchical Clustering
- K-means Clustering

Module IV – Discriminant Analysis and Applications (8 Hours)

- Introduction to Discriminant Analysis
- Linear Discriminant Analysis (LDA)
- Classification and Model Evaluation
- Introduction to MANOVA
- Applications of Multivariate Analysis in Data Science

Suggested Text Books

1. Johnson, R. A. and Wichern, D. W., *Applied Multivariate Statistical Analysis*, Pearson.
2. Rencher, A. C. and Christensen, W. F., *Methods of Multivariate Analysis*, Wiley.
3. Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E., *Multivariate Data Analysis*, Pearson.

Reference Books

1. Everitt, B. and Hothorn, T., *An Introduction to Applied Multivariate Analysis with R*, Springer.
2. Härdle, W. and Simar, L., *Applied Multivariate Statistical Analysis*, Springer.

KU2DSCSC01

Course Title: DESIGN AND ANALYSIS OF ALGORITHMS

Course Structure

- Lecture Hours: 3 Hours/Week
- Credits: 3

Course Objectives

The main objectives of this course are to:

1. Introduce fundamental concepts and techniques for algorithm design and analysis.



2. Develop problem-solving skills using efficient data structures and algorithms.
3. Familiarize students with algorithmic paradigms used in computational and data science applications.
4. Analyze the time and space complexity of algorithms.
5. Enable students to apply optimization and algorithmic techniques in data science problems.

Expected Course Outcomes

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

CO No.	Course Outcome	Knowledge Level
CO1	Explain the principles of algorithm design and complexity analysis	K1/K2
CO2	Apply suitable data structures and algorithmic techniques to solve computational problems	K3
CO3	Analyze the efficiency of algorithms using asymptotic notations	K3/K4
CO4	Design algorithms using greedy, dynamic programming and graph techniques	K4/K5
CO5	Develop optimized algorithmic solutions for data science and real-world applications	K5/K6

Knowledge Levels

- K1 – Remember
- K2 – Understand
- K3 – Apply
- K4 – Analyze
- K5 – Evaluate
- K6 – Create

Module I – Introduction to Algorithms and Complexity Analysis (15 Hours)

- Introduction to Algorithms
- Characteristics of Algorithms
- Performance Analysis
- Time and Space Complexity
- Asymptotic Notations
- Best Case, Worst Case and Average Case Analysis
- Recurrence Relations
- Solving Recurrences
- Algorithmic Problem-Solving Techniques

Module II – Fundamental Data Structures and Algorithms (15 Hours)

Data Structures

- Arrays
- Linked Lists
- Stacks and Queues
- Trees and Binary Search Trees
- Heaps and Priority Queues
- Hashing Basics

Searching and Sorting

- Linear Search
- Binary Search
- Bubble Sort
- Selection Sort
- Merge Sort
- Quick Sort



- Heap Sort
- Complexity Analysis of Sorting Algorithms

Module III – Algorithm Design Techniques (15 Hours)

Divide and Conquer

- Concept and Applications
- Merge Sort
- Quick Sort
- Binary Search

Greedy Algorithms

- Greedy Strategy
- Activity Selection Problem
- Huffman Coding
- Minimum Spanning Tree

Dynamic Programming

- Principles of Dynamic Programming
- Fibonacci Sequence
- Knapsack Problem
- Longest Common Subsequence

Module IV – Graph Algorithms and Applications in Data Science (15 Hours)

Graph Algorithms

- Graph Representations
- Breadth First Search (BFS)
- Depth First Search (DFS)
- Shortest Path Algorithms
- Dijkstra's Algorithm
- Minimum Spanning Trees
- Topological Sorting

Applications

- Randomized Algorithms
- Approximation Algorithms
- Algorithms in Machine Learning and Data Science
- Computational Complexity and NP-Completeness (Introduction only)

Reference Books

1. Introduction to Algorithms, Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein, MIT Press
2. Algorithm Design, Jon Kleinberg, Éva Tardos, Pearson Publications
3. Data Structures and Algorithms Made Easy, Narasimha Karumanchi, CareerMonk Publications
4. The Algorithm Design Manual, Steven S. Skiena, Springer Publications
5. Algorithms, Robert Sedgewick, Kevin Wayne, Addison-Wesley Publications
6. Foundations of Algorithms, Richard Neapolitan, Kumarss Naimipour, Jones & Bartlett Learning
7. Problem Solving with Algorithms and Data Structures Using Python, Bradley N. Miller, David L. Ranum. Franklin, Beedle & Associates

KU2DSCSC02

Course Title: DATA VISUALIZATION

Course Structure



- Lecture Hours: 3 Hours/Week
- Credits: 3

Course Objectives

The main objectives of this course are to:

1. Introduce the principles and techniques of effective data visualization.
2. Develop skills in creating visual representations using Python visualization libraries.
3. Familiarize students with modern visualization tools such as Tableau and Power BI.
4. Explain dashboard design principles for business intelligence and analytics.
5. Enable students to communicate insights effectively through visual storytelling.

Expected Course Outcomes

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

CO No.	Course Outcome	Knowledge Level
CO1	Explain the principles and importance of data visualization in analytics	K1/K2
CO2	Create visualizations using Matplotlib and Seaborn libraries	K3
CO3	Design interactive dashboards using Tableau and Power BI	K3/K4
CO4	Analyze and present data insights using visualization techniques	K4/K5
CO5	Develop effective visual storytelling solutions for decision making	K5/K6

Knowledge Levels

- K1 – Remember
- K2 – Understand
- K3 – Apply
- K4 – Analyze
- K5 – Evaluate
- K6 – Create

Module I – Principles of Data Visualization (12 Hours)

- Introduction to Data Visualization
- Importance of Visualization in Data Science
- Types of Data Visualization
- Principles of Effective Visualization
- Visual Perception and Human Cognition
- Choosing Appropriate Charts and Graphs
- Data Cleaning and Preparation for Visualization
- Ethical Issues in Data Visualization

Module II – Visualization using Matplotlib and Seaborn (12 Hours)

Matplotlib

- Introduction to Matplotlib
- Line Charts
- Bar Charts
- Histograms
- Scatter Plots
- Pie Charts
- Customization of Graphs
- Multiple Plots and Subplots

Seaborn

- Introduction to Seaborn
- Statistical Visualization
- Distribution Plots
- Box Plots and Violin Plots



- Heatmaps
- Pair Plots
- Regression Plots

Module III – Tableau, Power BI and Dashboard Design (12 Hours)

Tableau

- Introduction to Tableau
- Data Import and Data Preparation
- Charts and Interactive Visualizations
- Filters and Parameters
- Tableau Dashboards

Power BI

- Introduction to Power BI
- Data Modelling
- Interactive Reports
- DAX Basics
- Dashboard Creation

Dashboard Design

- Principles of Dashboard Design
- KPI Visualization
- Interactive Dashboards
- Business Intelligence Dashboards
- Real-time Dashboards

Module IV – Visual Storytelling and Advanced Visualization Techniques (12 Hours)

- Introduction to Visual Storytelling
- Storytelling with Data
- Data-driven Decision Making
- Infographics and Interactive Reports
- Visualization for Business Analytics
- Visualization for Machine Learning Results
- Geospatial and Network Visualization
- Case Studies in Data Visualization

Reference Books

1. Storytelling with Data, Cole Nussbaumer Knaflic, Wiley Publications
2. Python Data Science Handbook, Jake VanderPlas, O'Reilly Media
3. Fundamentals of Data Visualization, Claus O. Wilke, O'Reilly Media
4. Practical Tableau, Ryan Sleeper, O'Reilly Media
5. Microsoft Power BI Cookbook, Brett Powell, Packt Publications
6. Interactive Data Visualization for the Web, Scott Murray, O'Reilly Media
7. Data Visualization: A Practical Introduction, Kieran Healy, Princeton University Press

KU2DSSTA02-LAB

Course Title: LAB – ADVANCED STATISTICAL DATA ANALYSIS USING R/PYTHON

COURSE STRUCTURE

- Lab: 4 Hours/Week
- Credits: 2

Problems relating to the following topics using R/Python programming shall form the basis for setting the question paper.

- Sampling techniques
- Confidence interval estimation



- Hypothesis testing
- ANOVA analysis
- Non-parametric tests
- Confidence interval estimation
- Hypothesis testing
- Acceptance–Rejection algorithm implementation
- Efficiency comparison of different simulation methods
- Implementation of the EM algorithm
- Comparison of EM estimates with direct MLE
- Exploratory Analysis of Multivariate Data
- Assessment of Multivariate Normality
- Principal Component Analysis (PCA) and Factor Analysis
- Cluster Analysis
- Discriminant Analysis and MANOVA

KU2DSCSC01- LAB

Course Title: LAB-I: DESIGN AND ANALYSIS OF ALGORITHMS

COURSE STRUCTURE

- Lab Hours: 4 Hours/Week
- Credits: 2

Students shall perform the practical exercises in Lab:

- Complexity analysis of algorithms
- Searching and sorting implementations
- Stack and queue applications
- Tree and graph traversal algorithms
- Greedy algorithms
- Dynamic programming problems
- Graph-based data science applications

KU2DSCSC02- LAB

Course Title: LAB-II: DATA VISUALIZATION

COURSE STRUCTURE

- Lab Hours: 4 Hours/Week
- Credits: 2

Students shall perform the practical exercises in Lab:

- Creating charts and graphs using Matplotlib
- Statistical plots using Seaborn
- Interactive dashboards in Tableau
- Business reports in Power BI
- KPI dashboard design
- Visual storytelling projects



Course Title: Mini Project
Programme: M.Sc. Data Science

Credits: 2

COURSE STRUCTURE

The Mini Project is intended to provide students with practical exposure to real-world problems in Data Science and to develop analytical, programming, research, and presentation skills through independent or group-based project work.

**No separate instructional hours are allotted per week for the Mini Project. Students shall carry out the work independently under the guidance of a faculty supervisor.*

Project Supervision

1. Each project shall be carried out under the guidance of a faculty member of the Department/College.
2. The guide shall:
 - Monitor the progress of the work,
 - Provide technical guidance,
 - Evaluate the quality and originality of the project,
 - Ensure timely submission of the report.

Project Proposal

Students shall submit a brief project proposal and this shall be approved by the Department before commencement of the work.

Project Report

At the completion of the project, students shall submit a duly certified project report containing:

- Title Page
- Certificate
- Declaration
- Acknowledgement
- Abstract
- Introduction
- Literature Review
- Methodology
- Implementation/Analysis
- Results and Discussion
- Conclusion and Future Scope
- References

The report shall follow the format prescribed by the University.

Evaluation Scheme

The Mini Project shall be evaluated based on:



Component	Marks
Problem Identification and Objectives	5
Methodology and Implementation	10
Innovation and Technical Content	5
Report Quality and Documentation	5
Presentation and Demonstration	10
Viva-Voce	15
Total	50 Marks

SEMESTER III

KU3DSSTA01

Course Title:

BAYESIAN STATISTICS, TIME SERIES ANALYSIS AND FORECASTING TECHNIQUES

Credits: 3

Contact Hours: 3 Hours/Week

Course Type: Core/Elective

Course Description



This course provides a comprehensive treatment of Bayesian statistical methods and time series analysis techniques used in modern data science. The course focuses on Bayesian inference, probabilistic modelling, forecasting methodologies, and advanced time series models. Students will develop skills in handling uncertainty, analysing temporal data, constructing forecasting models, and applying statistical methods to real-world problems in business analytics, finance, healthcare, economics, and artificial intelligence.

Course Objectives

1. To develop a strong understanding of Bayesian statistical principles and inference.
2. To apply Bayesian methods for estimation, prediction, and decision-making.
3. To understand the structure and characteristics of time series data.
4. To develop forecasting models for temporal data.
5. To evaluate forecasting accuracy and model performance.
6. To apply Bayesian and time series methods in practical data science applications.

Course Outcomes

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

- Apply Bayesian inference for statistical analysis and predictive modeling.
- Construct and interpret prior and posterior probability models.
- Analyse and decompose time series data.
- Develop forecasting models using classical and modern techniques.
- Evaluate forecasting performance using appropriate statistical measures.
- Apply Bayesian and forecasting methods to real-world data science problems

Module I – Foundations of Bayesian Statistics (11 Hours)

Introduction to Bayesian Statistics

- Historical Development of Bayesian Statistics
- Classical versus Bayesian Approaches
- Probability as a Measure of Uncertainty
- Applications in Data Science

Bayesian Probability Models

- Bayes' Theorem
- Likelihood Functions

Prior, Likelihood and Posterior Distributions

- Prior Distributions
- Informative and Non-informative Priors
- Likelihood Functions
- Posterior Distributions
- Conjugate Priors
- Bayesian Updating

Bayesian Estimation

- Bayesian Point Estimation
- Bayesian Interval Estimation
- Credible Intervals
- Comparison with Confidence Intervals

Module II – Bayesian Inference and Computational Methods (11 Hours)

Bayesian Inference

- Bayesian Hypothesis Testing
- Bayesian Model Comparison
- Bayes Factors

Computational Bayesian Methods

- Introduction to Monte Carlo Methods
- Markov Chain Monte Carlo (MCMC)
- Gibbs Sampling



- Metropolis–Hastings Algorithm

Bayesian Decision Theory

- Loss Functions
- Bayesian Risk
- Decision Making under Uncertainty

Bayesian Applications

- Bayesian Regression
- Bayesian Classification
- Applications in Predictive Analytics

Module III – Fundamentals of Time Series Analysis (11 Hours)

Introduction to Time Series

- Characteristics and Components of Time Series
- Time Series Visualization
- Applications in Business and Data Science

Time Series Decomposition

- Additive and Multiplicative Models
- Trend Estimation
- Seasonal Adjustment

Stationarity and Dependence

- Stationary and Non-stationary Processes
- Differencing
- White Noise
- Autocorrelation Function (ACF)
- Partial Autocorrelation Function (PACF)

Exploratory Time Series Analysis

- Correlograms
- Seasonal Plots
- Identification of Temporal Patterns

Module IV – Time Series Forecasting (12 Hours)

Forecasting Techniques

- Principles of Forecasting
- Forecast Accuracy Measures
- Moving Average Forecasting
- Exponential Smoothing
- Holt's Linear Trend Method
- Holt–Winters Seasonal Method

Suggested Text Books

1. Bolstad, W. M. and Curran, J. M. – *Introduction to Bayesian Statistics*
2. Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., & Rubin, D., *Bayesian Data Analysis*, CRC Press.
3. Arora, S. and Bansal, S., *New Mathematical Statistics*, Satya Prakashan.
4. Mukhopadhyay, P., *Mathematical Statistics*, Books and Allied (P) Ltd., Kolkata.
5. Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B., *Bayesian Data Analysis*, 3rd Edition, CRC Press.
6. Montgomery, D. C., Jennings, C. L., and Kulahci, M., *Introduction to Time Series Analysis and Forecasting*, 2nd Edition, Wiley.
7. Hyndman, R. J., & Athanasopoulos, G., *Forecasting: Principles and Practice*.
8. Hoff, P. D., *A First Course in Bayesian Statistical Methods*.
9. Shumway, R. H., & Stoffer, D. S., *Time Series Analysis and Its Applications*.

Reference Books



1. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M., *Time Series Analysis: Forecasting and Control*.
2. Kruschke, J. K., *Doing Bayesian Data Analysis*.
3. West, M., & Harrison, J., *Bayesian Forecasting and Dynamic Models*.
4. Barber, D., *Bayesian Reasoning and Machine Learning*.
5. Ross, S. M. *Introduction to Probability and Statistics for Engineers and Scientists*, Academic Press.

KU3DSCSC01

Course Title: MACHINE LEARNING TECHNIQUES

Course Structure

- Lecture Hours: 2 Hours/Week
- Credits: 2

Course Objectives

The main objectives of this course are to:

1. Introduce the fundamental concepts and techniques of machine learning.
2. Develop understanding of supervised and unsupervised learning methods.
3. Enable students to build regression, classification and clustering models.
4. Familiarize students with model evaluation techniques and performance metrics.
5. Explain bias-variance trade-off and its role in machine learning model optimization.

Expected Course Outcomes

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

CO No.	Course Outcome	Knowledge Level
CO1	Explain the concepts and workflow of machine learning systems	K1/K2
CO2	Apply supervised learning techniques for predictive modelling	K3
CO3	Implement unsupervised learning methods for pattern discovery	K3/K4
CO4	Evaluate machine learning models using appropriate performance measures	K4/K5
CO5	Analyze bias-variance tradeoff and optimize machine learning models	K5/K6

Knowledge Levels

- K1 – Remember
- K2 – Understand
- K3 – Apply
- K4 – Analyze
- K5 – Evaluate
- K6 – Create

Module I – Introduction to Machine Learning (15 Hours)

- Introduction to Machine Learning
- Applications of Machine Learning
- Types of Machine Learning
- Machine Learning Workflow
- Data Preprocessing
- Feature Selection and Feature Engineering



Module II – Supervised Learning & Unsupervised Learning (15 Hours)

Supervised Learning

- Introduction to Supervised Learning
- Training and Testing Data
- Linear Regression
- Multiple Regression
- Logistic Regression
- Overfitting and Underfitting

Unsupervised Learning

- Introduction to Unsupervised Learning
- Dimensionality Reduction Basics
- Principal Component Analysis (PCA)

Module III – Classification Techniques (15 Hours)

- Introduction to Classification
- k-Nearest Neighbour (k-NN)
- Decision Trees
- Naive Bayes Classifier
- Support Vector Machines (SVM)
- Ensemble Learning Basics
- Random Forest
- Classification Performance Metrics
- Confusion Matrix
- Precision, Recall and F1-Score
- ROC Curve and AUC

Module IV – Clustering and Model Evaluation (15 Hours)

Clustering

- k-Means Clustering
- Hierarchical Clustering
- Density-based Clustering
- Cluster Validation Techniques
- Applications of Clustering in Data Science

Model Evaluation

- Training, Validation and Test Sets
- Cross Validation Techniques
- Evaluation Metrics for Regression
- Mean Squared Error (MSE)
- Root Mean Squared Error (RMSE)
- R-squared Measure

Reference Books

1. Introduction to Machine Learning with Python, Andreas C. Müller, Sarah Guido, O'Reilly Media
2. Hands-On Machine Learning with Scikit-Learn, Keras and TensorFlow, Aurélien Géron, O'Reilly Media
3. Pattern Recognition and Machine Learning, Christopher M. Bishop, Springer Publications
4. Machine Learning, Tom M. Mitchell, McGraw-Hill Publications
5. The Elements of Statistical Learning, Trevor Hastie, Robert Tibshirani, Jerome Friedman, Springer Publications
6. Data Mining: Concepts and Techniques, Jiawei Han, Micheline Kamber, Jian Pei, Morgan Kaufmann Publications
7. Machine Learning for Absolute Beginners, Oliver Theobald, Scatterplot Press



Course Title: NATURAL LANGUAGE PROCESSING**Course Structure**

- Lecture Hours: 3 Hours/Week
- Credits: 3

Course Objectives

The main objectives of this course are to:

1. Introduce the fundamental concepts and techniques of Natural Language Processing (NLP).
2. Develop skills in text preprocessing, representation, and linguistic analysis.
3. Familiarize students with machine learning and deep learning approaches for NLP tasks.
4. Explain advanced NLP techniques for text analytics and language understanding.
5. Enable students to build NLP applications for real-world data science problems.

Expected Course Outcomes

CO No.	Course Outcome	Knowledge Level
CO1	Explain the concepts, components, and applications of Natural Language Processing	K1/K2
CO2	Apply text preprocessing and feature extraction techniques for NLP tasks	K3
CO3	Develop NLP models using machine learning and deep learning frameworks	K3/K4
CO4	Analyze and interpret textual data using NLP techniques	K4/K5
CO5	Design intelligent NLP solutions for real-world applications	K5/K6

Knowledge Levels

- K1 – Remember
- K2 – Understand
- K3 – Apply
- K4 – Analyze
- K5 – Evaluate
- K6 – Create

Module I – Introduction to Natural Language Processing (15 Hours)

- Introduction to Natural Language Processing
- Applications of NLP
- NLP Pipeline
- Challenges in NLP
- Text Preprocessing Techniques
- Tokenization
- Stop Word Removal
- Stemming and Lemmatization
- Part-of-Speech Tagging
- Named Entity Recognition
- Syntactic and Semantic Analysis
- Corpus and Linguistic Resources



Module II – Text Representation (12 Hours)

- Bag of Words Model
- N-Gram Models
- TF-IDF Representation
- Word Embeddings
- Word2Vec and GloVe
- Feature Extraction Techniques

Module III - Machine Learning for Natural Language Processing (15 Hours)

- Text Classification
- Sentiment Analysis
- Spam Detection
- Topic Modelling
- Naïve Bayes and Support Vector Machines for NLP
- Evaluation Metrics for NLP Models

Module III – Deep Learning for Natural Language Processing (18 Hours)

Deep Learning Techniques

- Introduction to Deep Learning in NLP
- Neural Networks for NLP
- Recurrent Neural Networks (RNN)
- Long Short-Term Memory (LSTM)
- Gated Recurrent Unit (GRU)
- Sequence-to-Sequence Models

Transformer Models

- Attention Mechanism
- Transformer Architecture
- BERT and GPT Models
- Transfer Learning in NLP
- Text Generation Techniques

Reference Books

1. Speech and Language Processing, Daniel Jurafsky and James H. Martin, Pearson
2. Natural Language Processing with Python, Steven Bird, Ewan Klein, and Edward Loper, O'Reilly Media
3. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, Aurélien Géron, O'Reilly Media
4. Deep Learning for Natural Language Processing, Palash Goyal, Sumit Pandey, and Karan Jain, Apress
5. Practical Natural Language Processing, Sowmya Vajjala et al., O'Reilly Media
6. Foundations of Statistical Natural Language Processing, Christopher D. Manning and Hinrich Schütze, MIT Press
7. Natural Language Processing in Action, Hobson Lane, Cole Howard, and Hannes Hapke, Manning Publications

KU3DSCSC03

Course Title: CLOUD COMPUTING AND DATA ENGINEERING

Course Structure

- Lecture Hours: 3 Hours/Week
- Credits: 3

Course Objectives

The main objectives of this course are to:



1. Introduce the concepts, architecture, and service models of cloud computing.
2. Understand of data engineering principles and big data processing techniques.
3. Familiarize students with cloud platforms and distributed data processing frameworks.
4. Explain data pipelines, storage systems, and scalable data architectures.
5. Enable students to design cloud-based data engineering solutions for analytics applications.

Expected Course Outcomes

CO No.	Course Outcome	Knowledge Level
CO1	Explain the concepts, architectures, and services of cloud computing and data engineering	K1/K2
CO2	Apply cloud platforms and distributed computing techniques for data processing	K3
CO3	Develop data pipelines and scalable storage solutions using cloud technologies	K3/K4
CO4	Analyze big data processing frameworks and cloud-based analytics systems	K4/K5
CO5	Design cloud-enabled data engineering solutions for real-world applications	K5/K6

Knowledge Levels

- K1 – Remember
- K2 – Understand
- K3 – Apply
- K4 – Analyze
- K5 – Evaluate
- K6 – Create

Module I – Introduction to Cloud Computing and Data Engineering (15 Hours)

- Introduction to Cloud Computing
- Characteristics and evolution of Cloud Computing
- Cloud Service Models: IaaS, PaaS, SaaS
- Cloud Deployment Models
- Virtualization and Containerization
- Introduction to Data Engineering
- Data Engineering Lifecycle
- Big Data Concepts and Characteristics
- Applications of Cloud Computing and Data Engineering

Module II – Cloud Platforms and Data Storage (12 Hours)

Cloud Platforms

- Introduction to Amazon Web Services (AWS)
- Introduction to Microsoft Azure
- Introduction to Google Cloud Platform (GCP)
- Cloud Storage Services
- Cloud Databases
- Compute and Networking Services

Data Storage and Management

- Structured and Unstructured Data
- Data Security and Privacy in Cloud

Module III – Big Data Processing (15 Hours)

Big Data Processing Frameworks

- Introduction to Hadoop Ecosystem
- Hadoop Distributed File System (HDFS)
- MapReduce Programming Model
- Introduction to Apache Spark
- Batch and Stream Processing



- Data Transformation Techniques

Module IV - Pipeline Engineering and Data Engineering Applications (18 Hours)

Data Pipelines

- ETL and ELT Processes
- Data Integration Techniques
- Workflow Orchestration
- Data Pipeline Automation
- Real-time Data Processing
- Monitoring and Logging of Data Pipelines

Data Engineering Applications

- Cloud-native Applications
- Microservices Architecture
- Serverless Computing
- Data Engineering for Business Analytics
- Ethical and Security Issues in Cloud Computing

Reference Books

1. Cloud Computing: Concepts, Technology & Architecture, Thomas Erl, Pearson
2. Designing Data-Intensive Applications, Martin Kleppmann, O'Reilly Media
3. Data Engineering with Python, Paul Crickard, Packt Publications
4. Hadoop: The Definitive Guide, Tom White, O'Reilly Media
5. Learning Spark, Jules S. Damji et al., O'Reilly Media
6. AWS Certified Cloud Practitioner Study Guide, Ben Piper and David Clinton, Sybex
7. Fundamentals of Data Engineering, Joe Reis and Matt Housley, O'Reilly Media

KU3DSSTA01-LAB

SEMESTER – III: LAB –

BAYESIAN STATISTICS, TIME SERIES ANALYSIS AND FORECASTING TECHNIQUES

COURSE STRUCTURE

- Lab: 4 Hours/Week
- Credits: 2

Practical Components

1. Computation of prior, likelihood, and posterior distributions; Bayesian point estimation and credible intervals.
2. Bayesian hypothesis testing, Bayes factors, and Bayesian model comparison.
3. Implementation of Monte Carlo simulation, Gibbs Sampling, and the Metropolis-Hastings algorithm for Bayesian inference.
4. Bayesian regression and Bayesian classification using real-world datasets.
5. Visualization and decomposition of time series data into trend, seasonal, and irregular components.
6. Assessment of stationarity using differencing, Autocorrelation Function (ACF), and Partial Autocorrelation Function (PACF).
7. Forecasting using moving averages, exponential smoothing, Holt's linear trend, and Holt-Winters seasonal methods.

KU3DSCSC01-LAB

Course Title:



LAB-I: MACHINE LEARNING TECHNIQUES FOR BIG DATA

Course Structure

- Lab Hours: 3 Hours/Week
- Credits: 2

Students shall perform the practical exercises in Lab:

1. Data preprocessing
2. Regression models
3. Classification models
4. Clustering algorithms
5. PCA implementation
6. Model evaluation metrics

KU3DSCSC02-LAB

Course Title: LAB-II: NATURAL LANGUAGE PROCESSING

Course Structure

- Lab Hours: 4 Hours/Week
 - Credits: 2
1. Text preprocessing using Python
 2. Tokenization and stop word removal
 3. Stemming and lemmatization
 4. Part-of-Speech tagging
 5. Sentiment analysis on sample datasets
 6. Visualization of word frequencies and word clouds
 7. Spam detection using machine learning techniques
 8. Sentiment analysis using real-world datasets
 9. Topic modelling using LDA
 10. Word embedding implementation using Word2Vec

KU3DSCSC03-LAB

Course Title: LAB-III: CLOUD COMPUTING AND DATA ENGINEERING

Course Structure

- Lab Hours: 3 Hours/Week
 - Credits: 2
1. Exploring AWS/Azure/GCP cloud services
 2. Working with cloud storage services
 3. Database creation using cloud databases
 4. Uploading and managing datasets in cloud storage
 5. Basic operations using Hadoop Distributed File System (HDFS)
 6. Monitoring cloud resources and usage
 7. Working with NoSQL databases
 8. Creating dashboards using cloud analytics tools



SEMESTER IV

ELECTIVE-I (Statistics)

KU4DSSTA01

Title of the Course: In Semester IV – To choose an elective course from the basket of Statistics



courses:

KU4DSSTA01-A

Course I: SURVIVAL ANALYSIS AND RISK ANALYTICS

Credits: 3

Instruction Hours: 3 Hours per week

Course Description

This course introduces the principles and techniques of survival analysis and risk analytics, focusing on the analysis of time-to-event data and risk prediction. Students will learn methods for estimating survival probabilities, modeling hazard rates, analyzing censored data, and assessing risks in various domains such as healthcare, finance, insurance, marketing, and industrial systems. The course emphasizes both theoretical foundations and practical applications using statistical software.

Course Objectives

The course aims to:

1. Introduce the concepts of survival and time-to-event data analysis.
2. Develop an understanding of survival functions, hazard functions, and censoring mechanisms.
3. Apply statistical methods for estimating and comparing survival experiences.
4. Build regression models for survival data and risk prediction.
5. Explore applications of survival analysis in healthcare, finance, business, and industrial analytics.

Course Outcomes

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

- Understand and analyze survival and time-to-event data.
- Estimate and interpret survival probabilities and hazard rates.
- Apply non-parametric and semi-parametric survival models.
- Compare survival experiences across groups.
- Develop survival-based risk prediction models.
- Utilize survival analytics techniques in real-world decision-making problems.

Module I – Fundamentals of Survival Analysis (11 Hours)

Introduction to Survival Analysis

- Nature and Characteristics of Survival Data
- Time-to-Event and Failure-Time Data
- Terminology in Survival Analysis
 - Survival Time
 - Event Time
 - Failure Time
 - Risk Set
- Reliability and Lifetime Concepts
- Applications of Survival Analysis

Probability Models for Survival Data

- Survival and Lifetime Distributions
- Weibull Distribution
- Reliability Function
- Mean and Median Survival Time

Module II – Survival Functions and Estimation (11 Hours)

Survival and Hazard Functions

- Survival Function
- Probability Density Function
- Cumulative Distribution Function
- Hazard Function
- Cumulative Hazard Function



- Relationships among Survival, Density, and Hazard Functions

Censoring and Survival Estimation

- Types of Censoring
- Right, Left, and Interval Censoring
- Kaplan–Meier Estimator
- Survival Curves and Their Interpretation
- Life-Table (Actuarial) Method

Module III – Survival Models and Statistical Inference (11 Hours)

Comparing Survival Distributions

- Log-rank Test
- Wilcoxon Test
- Interpretation of Test Results

Regression Models for Survival Data

- Introduction to Survival Regression
- Proportional Hazards Concept
- Cox Proportional Hazards Model
- Hazard Ratios and Their Interpretation
- Partial Likelihood Estimation

Model Assessment

- Goodness-of-Fit
- Checking the Proportional Hazards Assumption
- Interpretation of Survival Models

Module IV – Risk Analytics and Applications (12 Hours)

Fundamentals of Risk Analytics

- Risk and Uncertainty
- Risk Measurement and Assessment
- Time-dependent Risk Prediction
- Event-Risk Modelling

Survival-Based Predictive Analytics

- Risk Scoring Models
- Survival-Based Prediction
- Decision-Making under Uncertainty

Reference Books

1. Kleinbaum, D. G., & Klein, M. *Survival Analysis: A Self-Learning Text*. Springer.
2. Lawless, J. F. *Statistical Models and Methods for Lifetime Data*, 2nd Edition, Wiley, 2003.
3. Hosmer, D. W., Lemeshow, S., & May, S. *Applied Survival Analysis*. Wiley.
4. Lee, E. T., & Wang, J. W. *Statistical Methods for Survival Data Analysis*. Wiley.
1. Collett, D. *Modelling Survival Data in Medical Research*. CRC Press.
2. Klein, J. P., & Moeschberger, M. L. *Survival Analysis: Techniques for Censored and Truncated Data*. Springer.
3. Allison, P. D. *Survival Analysis Using SAS: A Practical Guide*. SAS Institute.
4. Kuhn, M., & Johnson, K. *Applied Predictive Modeling*. Springer.

KU4DSSTA01-A-LAB

SEMESTER – IV: LAB - SURVIVAL ANALYSIS AND RISK ANALYTICS

COURSE STRUCTURE

- Lab: 4 Hours/Week
- Credits: 2

Practical Component



1. Import, preprocess, and summarize survival data; create time-to-event datasets and perform exploratory analysis.
2. Estimate and interpret survival functions using the Kaplan–Meier estimator and construct survival curves.
3. Analyze censored survival data using the Kaplan–Meier and Life-Table methods and compare survival estimates.
4. Compare survival distributions using the Log-rank test and Wilcoxon test.
5. Fit and interpret the Cox Proportional Hazards model; estimate hazard ratios and assess model adequacy.
6. Apply survival analysis techniques for risk prediction using real-world datasets from healthcare, finance, insurance, or customer churn analytics. Credit risk analysis using time-to-event data.
7. Mini-project on survival analysis and risk analytics.

KU4DSSTA01-B

Course II:
Financial Statistics and Risk Analytics

Credits: 3

Instruction Hours: 3 Hours per week

Major Topics

- Financial Data Analysis
- Portfolio Theory
- Risk Measurement
- Value at Risk (VaR)
- Time Series in Finance
- Credit Risk Modelling

KU4DSSTA01-C

Course III:
Biostatistics and Health Analytics

Credits: 3

Major Topics

- Clinical Trial Design
- Epidemiological Statistics
- Survival Analysis
- Bioinformatics Basics
- Public Health Data Analysis
- Statistical Methods in Medicine

KU4DSSTA01-D

Course IV:
Stochastic Processes

Credits: 3

Major Topics

- Markov Chains
- Poisson Processes
- Birth-Death Processes
- Queueing Models
- Brownian Motion
- Applications in AI and Finance

KU4DSSTA01-E



Course V:
Optimization Techniques for Data Science
Credits: 3
Major Topics

- Linear Programming
- Non-linear Optimization
- Convex Optimization
- Gradient Methods
- Lagrange Multipliers
- Optimization in Machine Learning

ELECTIVE-II **(Computer Science)**

KU4DSCSC-ELE01-A

Course Title: DEEP LEARNING TECHNIQUES

Course Structure

- Lecture Hours: 2 Hours/Week
- Credits: 2

Course Objectives

The main objectives of this course are to:

1. Introduce the fundamentals and architectures of deep learning models.
2. Develop practical skills in implementing deep learning algorithms using Python frameworks.
3. Explain training, optimization, and evaluation techniques for neural networks.
4. Familiarize students with advanced deep learning models for computer vision and sequence data.
5. Enable students to apply deep learning techniques for real-world data science applications.

Expected Course Outcomes

CO No.	Course Outcome	Knowledge Level
CO1	Explain the concepts, architectures, and applications of deep learning	K1/K2
CO2	Develop and train neural network models using deep learning frameworks	K3
CO3	Apply optimization and regularization techniques to improve model performance	K3/K4
CO4	Analyze deep learning models for image, text, and sequence data	K4/K5
CO5	Design deep learning solutions for real-world analytics and intelligent systems	K5/K6

Knowledge Levels

- K1 – Remember
- K2 – Understand
- K3 – Apply
- K4 – Analyze
- K5 – Evaluate
- K6 – Create

Module I – Introduction to Deep Learning and Neural Networks (12 Hours)

- Introduction to Artificial Intelligence, Machine Learning, and Deep Learning



- Biological Neuron and Artificial Neuron
- Perceptron Model
- Activation Functions
- Feedforward Neural Networks
- Multilayer Perceptron (MLP)
- Forward and Backpropagation
- Loss Functions
- Gradient Descent and Optimization Basics
- Introduction to Deep Learning Frameworks
- Applications of Deep Learning

Module II – Deep Learning using TensorFlow and Keras (12 Hours)

TensorFlow and Keras Basics

- Introduction to TensorFlow
- Introduction to Keras
- Building Neural Network Models
- Data Preprocessing for Deep Learning
- Model Compilation and Training
- Evaluation Metrics

Training and Optimization

- Batch Processing and Epochs
- Hyperparameter Tuning
- Regularization Techniques
- Dropout and Batch Normalization
- Model Saving and Loading

Practical Deep Learning Models

- Image Classification using Neural Networks
- Regression using Deep Learning
- Binary and Multiclass Classification

Module III – Convolutional and Recurrent Neural Networks (12 Hours)

Convolutional Neural Networks (CNN)

- Introduction to CNN
- Convolution and Pooling Operations
- CNN Architectures
- Image Feature Extraction
- Image Classification Applications
- Transfer Learning Basics

Recurrent Neural Networks (RNN)

- Sequential Data and Time Series
- Recurrent Neural Networks
- Vanishing Gradient Problem
- Long Short-Term Memory (LSTM)
- Gated Recurrent Unit (GRU)
- Applications in Text and Sequence Analysis

Module IV – Advanced Deep Learning Techniques and Applications (12 Hours)

- Introduction to Autoencoders
- Generative Adversarial Networks (GANs)
- Deep Learning for Natural Language Processing
- Deep Learning for Computer Vision
- Recommendation Systems
- Deep Reinforcement Learning Basics
- Explainable AI in Deep Learning
- Ethical Issues in Deep Learning
- Case Studies and Real-world Applications



Reference Books

1. Deep Learning, Ian Goodfellow, Yoshua Bengio and Aaron Courville, MIT Press
2. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, Aurélien Géron, O'Reilly Media
3. Neural Networks and Deep Learning, Charu C. Aggarwal, Springer
4. Deep Learning with Python, François Chollet, Manning Publications
5. Python Deep Learning, Ivan Vasilev, Packt Publications
6. Practical Deep Learning for Cloud, Mobile, and Edge, Anirudh Koul, Siddha Ganju, and Meher Kasam, O'Reilly Media
7. Fundamentals of Deep Learning, Nikhil Buduma and Nicholas Locascio, O'Reilly Media

KU4DSCSC-ELE01-A-LAB

Course Title: LAB- DEEP LEARNING TECHNIQUES

Course Structure

- Lab Hours: 3 Hours/Week
 - Credits: 2
1. Installation and setup of TensorFlow and Keras
 2. Implementing basic perceptron model
 3. Working with activation functions
 4. Creating simple feedforward neural networks
 5. Binary classification using deep learning
 6. Multiclass classification using neural networks
 7. Visualizing training and validation accuracy
 8. Handwritten digit recognition using MNIST dataset
 9. Text classification using deep learning
 10. Sentiment analysis using LSTM networks

KU4DSCSC-ELE01-B

Course II: Course Title: GENERATIVE AI AND LARGE LANGUAGE MODELS

Course Structure

- Lecture Hours: 3 Hours/Week
- Credits: 3

Course Objectives

The main objectives of this course are to:

1. Introduce the concepts and architectures of Generative AI and Large Language Models (LLMs).
2. Develop understanding of transformer-based deep learning models.
3. Familiarize students with text generation and prompt engineering techniques.
4. Explain applications and deployment of generative AI systems.
5. Enable students to build responsible AI-powered applications using LLMs.

Expected Course Outcomes

CO No.	Course Outcome	Knowledge Level
CO1	Explain the concepts and applications of Generative AI and LLMs	K1/K2
CO2	Apply transformer models and prompt engineering techniques	K3
CO3	Develop applications using pre-trained generative AI models	K3/K4



CO No.	Course Outcome	Knowledge Level
CO4	Analyze performance and ethical challenges of LLMs	K4/K5
CO5	Design AI-powered intelligent systems using generative models	K5/K6

Knowledge Levels

- K1 – Remember
- K2 – Understand
- K3 – Apply
- K4 – Analyze
- K5 – Evaluate
- K6 – Create

Module I – Introduction to Generative AI and LLMs (12 Hours)

- Introduction and applications of Generative AI
- Evolution of Language Models
- Neural Language Models
- Introduction to Transformers
- Attention Mechanism
- Encoder and Decoder Architecture
- Large Language Models
- Foundation Models
- AI-powered Conversational Systems
- Challenges and Ethical issues in Generative AI

Module II – Transformer Models (12 Hours)

- Transformer Architecture
- BERT Model
- GPT Models
- T5 and Other Transformer Variants
- Transfer Learning in LLMs
- Fine-tuning Pre-trained Models

Module III – Prompt Engineering (12 Hours)

- Prompt Design Techniques
- Zero-shot and Few-shot Learning
- Chain-of-Thought Prompting
- Text Generation Techniques
- Evaluation of Generated Content
- Hallucination and Bias in LLMs

Module IV – Generative AI Applications and Tools (12 Hours)

Generative AI Applications

- Text Generation
- Question Answering Systems
- Chatbots and Virtual Assistants
- Summarization and Translation
- Code Generation
- Image and Multimodal AI Basics

Tools and Frameworks

- Hugging Face Transformers
- Open-source LLM Platforms
- APIs for Generative AI
- Model Deployment Basics
- Retrieval Augmented Generation (RAG)



Reference Books

1. Hands-On Large Language Models, Jay Alammar and Maarten Grootendorst, O'Reilly Media
2. Deep Learning, Ian Goodfellow, Yoshua Bengio and Aaron Courville, MIT Press
3. Natural Language Processing with Transformers, Lewis Tunstall, Leandro von Werra, and Thomas Wolf, O'Reilly Media
4. Generative AI on AWS, Chris Fregly and Antje Barth, O'Reilly Media
5. Deep Learning with Python, François Chollet, Manning Publications
6. Transformers for Natural Language Processing, Denis Rothman, Packt Publications
7. Building LLM Powered Applications, Valentina Alto, O'Reilly Media

KU4DSCSC-ELE01-B-LAB

Course Title: LAB-GENERATIVE AI AND LARGE LANGUAGE MODELS

Course Structure

- Lab Hours: 3 Hours/Week
 - Credits: 2
1. Load and use a pre-trained Transformer model to generate text from a given prompt.
 2. Experiment with different prompt engineering techniques such as Zero-shot and Few-shot prompting, and compare the outputs.
 3. Generate text using a GPT-based model for different prompts and observe the quality of the generated content.
 4. Perform text summarization using a pre-trained Transformer model and compare the summary with the original text.
 5. Develop a simple Question Answering application using a pre-trained language model.
 6. Create a basic AI chatbot using a pre-trained conversational model and test it with different user queries.

KU4DSCSC-ELE01-C

Course III: Course Title: SCALABLE DATA SYSTEMS

Course Structure

- Lecture Hours: 3 Hours/Week
- Credits: 3

Course Objectives

The main objectives of this course are to:

1. Introduce scalable architectures and distributed data systems.
2. Develop understanding of large-scale data storage and processing techniques.
3. Familiarize students with distributed computing frameworks and cloud-based systems.
4. Explain performance optimization and fault tolerance mechanisms.
5. Enable students to design scalable systems for big data applications.

Expected Course Outcomes

CO No.	Course Outcome	Knowledge Level
CO1	Explain concepts and architectures of scalable data systems	K1/K2
CO2	Apply distributed storage and processing techniques	K3
CO3	Develop scalable data solutions using distributed frameworks	K3/K4
CO4	Analyse performance and reliability of scalable systems	K4/K5
CO5	Design large-scale data processing systems for real-world applications	K5/K6

Knowledge Levels



- K1 – Remember
- K2 – Understand
- K3 – Apply
- K4 – Analyse
- K5 – Evaluate
- K6 – Create

Module I – Introduction to Scalable Data Systems (12 Hours)

- Introduction to Distributed Systems
- Scalability Concepts
- Data-intensive Applications
- CAP Theorem
- Distributed Architectures
- Fault Tolerance
- Replication and Partitioning
- Consistency Models
- Distributed File Systems
- Applications of Scalable Systems

Module II – Distributed Storage and Processing (12 Hours)

Distributed Storage

- Relational and NoSQL Databases
- Key-value Stores
- Column-oriented Databases
- Graph Databases
- Data Warehousing
- Data Lakes

Distributed Processing

- Hadoop Ecosystem
- HDFS Architecture
- MapReduce Programming Model
- Apache Spark
- Batch and Stream Processing
- Data Transformation Techniques

Module III – Scalable Architectures and Cloud Systems (12 Hours)

Scalable Architectures

- Microservices Architecture
- Event-driven Systems
- Load Balancing Techniques
- High Availability Systems

Cloud Data Systems

- Cloud Storage Services
- Distributed Cloud Databases
- Serverless Computing
- Real-time Analytics Systems
- Monitoring and Logging
- Security in Distributed Systems

Module IV – Advanced Scalable Systems and Applications (12 Hours)

- Data Pipeline Engineering
- Stream Processing Frameworks
- Scalable Machine Learning Systems
- Performance Optimization Techniques
- Data Governance and Reliability
- Scalability Testing
- Ethical Issues in Large-scale Data Systems



- Industrial Applications of Scalable Systems

Reference Books

1. Designing Data-Intensive Applications, Martin Kleppmann, O'Reilly Media
2. Hadoop: The Definitive Guide, Tom White, O'Reilly Media
3. Learning Spark, Jules S. Damji et al., O'Reilly Media
4. Distributed Systems: Concepts and Design, George Coulouris et al., Pearson
5. Cloud Computing: Concepts, Technology & Architecture, Thomas Erl, Pearson
6. Data Engineering with Python, Paul Crickard, Packt Publications
7. Streaming Systems, Tyler Akidau et al., O'Reilly Media

KU4DSCSC-ELE01-C-LAB

Course Title: SCALABLE DATA SYSTEMS

Course Structure

- Lab Hours: 3 Hours/Week
 - Credits: 2
1. Implement simple MapReduce programs (e.g., Word Count) and observe the execution process.
 2. Perform basic data processing using Apache Spark to read, filter, and analyze a dataset.
 3. Study cloud storage services and upload, retrieve, and manage files using a cloud platform.
 4. Develop a simple microservices-based application and demonstrate communication between services using REST APIs.
 5. Implement a basic data pipeline to collect, process, and store data from a given source.

KU4DSCSC-ELE01-D

Course IV: Course Title: DATA VISUALIZATION AND VISUAL ANALYTICS

Course Structure

- Lecture Hours: 3 Hours/Week
- Credits: 3

Course Objectives

The main objectives of this course are to:

1. Introduce principles and techniques of effective data visualization.
2. Develop skills in creating visual representations using modern visualization tools.
3. Familiarize students with dashboard design and visual analytics techniques.
4. Explain storytelling approaches for analytical insights.
5. Enable students to communicate data-driven decisions effectively.

Expected Course Outcomes

CO No.	Course Outcome	Knowledge Level
CO1	Explain the principles and importance of data visualization and visual analytics	K1/K2
CO2	Develop visualizations using Python libraries and BI tools	K3
CO3	Design dashboards and interactive visual analytics systems	K3/K4
CO4	Analyse and interpret data insights using visualization techniques	K4/K5
CO5	Create visual storytelling solutions for decision-making applications	K5/K6

Knowledge Levels



- K1 – Remember
- K2 – Understand
- K3 – Apply
- K4 – Analyze
- K5 – Evaluate
- K6 – Create

Module I – Principles of Data Visualization (12 Hours)

- Introduction to Data Visualization
- Importance of Visual Analytics
- Types of Data Visualization
- Principles of Effective Visualization
- Visual Perception and Cognition
- Choosing Appropriate Charts and Graphs
- Data Preparation for Visualization
- Ethical Issues in Visualization
- Storytelling with Data
- Exploratory Data Analysis
- Interactive Visualization Concepts
- Applications of Visual Analytics

Module II – Visualization Tools and Techniques (12 Hours)

Python Visualization

- Introduction to Matplotlib
- Seaborn Visualization
- Plotly Basics
- Statistical Visualization
- Heatmaps and Pair Plots
- Interactive Charts

BI and Dashboard Tools

- Tableau Basics
- Power BI Basics
- Dashboard Components
- KPI Visualization
- Filters and Parameters
- Dashboard Publishing

Module III – Dashboard Design (12 Hours)

- Principles of Dashboard Design
- User-centric Dashboard Development
- Business Intelligence Dashboards
- Real-time Dashboards
- Mobile and Web Dashboards
- Visualization Best Practices

Module IV – Visual Analytics (12 Hours)

- Analytical Reporting
- Data Exploration Techniques
- Trend and Pattern Analysis
- Comparative Analytics
- Geospatial Visualization
- Network Visualization Basics
- Explainable Visualization Techniques

Applications

- Infographics and Interactive Reports
- Visualization for Business Analytics
- Ethical and Accessibility Issues



Reference Books

1. Storytelling with Data, Cole Nussbaumer Knafl, Wiley Publications
2. Fundamentals of Data Visualization, Claus O. Wilke, O'Reilly Media
3. Python Data Science Handbook, Jake VanderPlas, O'Reilly Media
4. Practical Tableau, Ryan Sleeper, O'Reilly Media
5. Microsoft Power BI Cookbook, Brett Powell, Packt Publications
6. Interactive Data Visualization for the Web, Scott Murray, O'Reilly Media
7. Data Visualization: A Practical Introduction, Kieran Healy, Princeton University Press

KU4DSCSC-ELE01-D-LAB

Course Title: DATA VISUALIZATION AND VISUAL ANALYTICS

Course Structure

- L Hours: 3 Hours/Week
 - Credits: 2
1. Create basic charts using Matplotlib such as line charts, bar charts, pie charts, and histograms for a given dataset.
 2. Visualize data using Seaborn by creating scatter plots, box plots, count plots, and pair plots.
 3. Create interactive charts using Plotly and compare them with static charts.
 4. Design a simple dashboard using Tableau or Power BI to display key performance indicators (KPIs) from a sample dataset.
 5. Apply filters and interactive controls in Tableau or Power BI to improve dashboard usability.
 6. Create a business dashboard that includes multiple visualizations such as charts, tables, and KPI cards.

KU4DSCSC-ELE01-E

Course V:

Course Title: BUSINESS INTELLIGENCE AND DECISION SUPPORT SYSTEMS

Course Structure

- Lecture Hours: 3 Hours/Week
- Credits: 3

Course Objectives

The main objectives of this course are to:

1. Introduce the concepts and components of Business Intelligence (BI) systems.
2. Develop understanding of decision support systems and analytical reporting.
3. Familiarize students with data warehousing and dashboard technologies.
4. Explain data-driven decision-making techniques in organizations.
5. Enable students to design business intelligence solutions for enterprise applications.

Expected Course Outcomes

CO No.	Course Outcome	Knowledge Level
CO1	Explain concepts and applications of business intelligence and decision support systems	K1/K2
CO2	Apply BI tools and analytical techniques for business reporting	K3
CO3	Develop dashboards and data warehousing solutions	K3/K4
CO4	Analyse business data for strategic decision-making	K4/K5



CO No.	Course Outcome	Knowledge Level
CO5	Design enterprise-level BI and decision support applications	K5/K6

Knowledge Levels

- K1 – Remember
- K2 – Understand
- K3 – Apply
- K4 – Analyse
- K5 – Evaluate
- K6 – Create

Module I – Introduction to Business Intelligence and DSS (12 Hours)

- Introduction to Business Intelligence
- Components of BI Systems
- Decision Support Systems
- Types of Decision Support Systems
- Business Analytics Concepts
- Data-driven Decision Making
- Enterprise Information Systems
- Data Sources for BI
- Organizational Decision Processes
- Applications of BI
- Challenges in BI Systems
- Ethical Issues in Business Analytics

Module II – Data Warehousing and BI Tools (12 Hours)

Data Warehousing

- Data Warehouse Architecture
- ETL Processes
- Data Integration Techniques
- OLAP Concepts
- Data Marts
- Metadata Management

BI Tools

- Tableau for Business Analytics
- Power BI for Reporting
- KPI Dashboards
- Interactive Reports
- Data Visualization Techniques
- Dashboard Publishing and Sharing

Module III – Analytical Techniques and Decision Support (12 Hours)

Analytical Techniques

- Descriptive Analytics
- Predictive Analytics
- Prescriptive Analytics
- Statistical Analysis for BI
- Trend and Forecasting Analysis
- Performance Measurement Techniques

Decision Support Systems

- Model-driven DSS
- Data-driven DSS
- Knowledge-driven DSS
- Executive Information Systems
- Group Decision Support Systems



- Real-time Decision Support

Module IV – Advanced BI Applications and Enterprise Analytics (12 Hours)

- Business Performance Management
- Customer Relationship Analytics
- Financial and Marketing Analytics
- Supply Chain Analytics
- Big Data and BI
- Cloud-based BI Systems
- Mobile Business Intelligence
- Ethical and Security Issues in BI

Reference Books

1. Business Intelligence: Data Mining and Optimization for Decision Making, Carlo Vercellis, Wiley Publications
2. Fundamentals of Business Intelligence, Wilfried Grossmann and Stefanie Rinderle-Ma, Springer
3. Data Warehousing Fundamentals, Paulraj Ponniah, Wiley Publications
4. Microsoft Power BI Cookbook, Brett Powell, Packt Publications
5. Successful Business Intelligence, Cindi Howson, McGraw-Hill
6. Decision Support Systems and Intelligent Systems, Efraim Turban et al., Pearson
7. Business Intelligence Guidebook, Rick Sherman, Morgan Kaufmann

KU4DSCSC-ELE01-E-LAB

Course Title: BUSINESS INTELLIGENCE AND DECISION SUPPORT SYSTEMS

Course Structure

- Lab Hours: 3 Hours/Week
 - Credits: 2
1. Design a simple data warehouse schema (Star or Snowflake Schema) for a business scenario such as sales or inventory management.
 2. Perform a basic ETL process by extracting, cleaning, transforming, and loading data into a database or spreadsheet.
 3. Create OLAP-style reports by summarizing and analyzing data using pivot tables or a BI tool.
 4. Create basic visualizations using Tableau or Power BI to represent sales, customer, or financial data.
 5. Develop a simple KPI dashboard in Tableau or Power BI using charts, cards, and filters.
 6. Generate an interactive business report with filtering and sorting features using a BI tool.
 7. Perform descriptive analytics on a given business dataset and interpret the results using charts and summary statistics.
 8. Analyze business trends and prepare a simple forecasting report using spreadsheet tools or a BI platform.

ELECTIVE III (Computer Science)

KU4DSCSC-ELE02-A

Course VI: Course Title: ETHICAL AI AND RESPONSIBLE COMPUTING

Course Structure

- Lecture Hours: 3 Hours/Week



- Credits: 3

Course Objectives

The main objectives of this course are to:

1. Introduce ethical principles and responsible practices in Artificial Intelligence.
2. Develop understanding of fairness, accountability, transparency, and privacy in AI systems.
3. Familiarize students with ethical challenges in data-driven technologies.
4. Explain governance frameworks and regulations related to AI.
5. Enable students to design socially responsible and trustworthy AI solutions.

Expected Course Outcomes

CO No.	Course Outcome	Knowledge Level
CO1	Explain ethical principles and responsible computing practices in AI systems	K1/K2
CO2	Apply fairness, transparency, and privacy techniques in AI applications	K3
CO3	Analyse ethical and societal impacts of AI technologies	K3/K4
CO4	Evaluate governance, accountability, and regulatory frameworks for AI	K4/K5
CO5	Design responsible and trustworthy AI solutions for real-world applications	K5/K6

Knowledge Levels

- K1 – Remember
- K2 – Understand
- K3 – Apply
- K4 – Analyse
- K5 – Evaluate
- K6 – Create

Module I – Introduction to Ethical AI and Responsible Computing (12 Hours)

- Introduction to Ethics in Computing
- Responsible AI Concepts
- Fairness and Bias in AI
- Transparency and Explainability
- Accountability in AI Systems
- Privacy and Data Protection
- AI and Society
- Ethical Decision-making Models
- Human-centered AI
- Ethical Challenges in Data Science
- AI Misuse and Risks
- Case Studies in Ethical AI

Module II – Fairness, Privacy and Explainable AI (12 Hours)

Fairness and Bias

- Sources of Bias in AI
- Bias Detection Techniques
- Fairness Metrics
- Bias Mitigation Techniques
- Inclusive AI Systems
- Ethical Dataset Preparation

Privacy and Explainability

- Data Privacy Concepts
- Privacy-preserving Techniques
- Differential Privacy Basics



- Explainable AI Concepts
- Model Interpretability
- Transparent AI Systems

Module III – AI Governance and Responsible Technologies (12 Hours)

AI Governance

- AI Regulations and Policies
- AI Governance Frameworks
- Ethical Standards in AI
- Responsible Innovation
- Risk Assessment in AI Systems
- Compliance and Auditing

Responsible Technologies

- Responsible Machine Learning
- Ethical Issues in Generative AI
- Cybersecurity and AI
- Sustainable Computing
- Green AI Concepts
- Social Impacts of Automation

Module IV – Applications and Case Studies in Responsible AI (12 Hours)

- Responsible AI in Healthcare
- Ethical AI in Finance and Business
- AI in Education and Governance
- Trustworthy Autonomous Systems
- AI for Social Good
- Human-AI Collaboration
- Ethical Issues in Emerging Technologies

Reference Books

1. Ethics of Artificial Intelligence and Robotics, Vincent C. Müller, Routledge
2. Artificial Intelligence: A Guide for Thinking Humans, Melanie Mitchell, Farrar, Straus and Giroux
3. Responsible Artificial Intelligence, Virginia Dignum, Springer
4. Fairness and Machine Learning, Solon Barocas, Moritz Hardt, and Arvind Narayanan
5. Human Compatible, Stuart Russell, Viking Publications
6. Ethical Machines, Reid Blackman, Harvard Business Review Press
7. The Ethical Algorithm, Michael Kearns and Aaron Roth, Oxford University Press

KU4DSCSC-ELE02-A-LAB

Course Title: ETHICAL AI AND RESPONSIBLE COMPUTING

Course Structure

- Lab Hours: 3 Hours/Week
 - Credits: 2
1. Study the principles of Ethical AI and Responsible Computing.
 2. Identify bias in a given AI dataset.
 3. Compare fair and biased AI predictions.
 4. Perform basic data anonymization on a dataset.
 5. Explore an Explainable AI (XAI) tool.
 6. Analyze privacy issues in an AI application.
 7. Study AI ethics guidelines of a popular organization.
 8. Prepare a report on Responsible AI case studies.
 9. Evaluate fairness in a simple machine learning model.
 10. Present an Ethical AI solution for a real-world application.



KU4DSCSC-ELE02-B

Course VII: Course Title: SOFT COMPUTING TECHNIQUES

Course Structure

- Lecture Hours: 3 Hours/Week
- Credits: 3

Module I – Introduction to Soft Computing (12 Hours)

- Introduction to Soft Computing
- Characteristics
- Applications
- Artificial Neural Networks
- Fuzzy Logic
- Genetic Algorithms
- Evolutionary Computing
- Hybrid Systems
- Learning Systems

Module II – Neural Networks (12 Hours)

- Perceptron
- Feed Forward Networks
- Back propagation
- Activation Functions
- Training Process
- Performance Evaluation

Module III - Fuzzy Systems (12 Hours)

- Fuzzy Sets
- Membership Functions
- Fuzzy Rules
- Fuzzy Inference
- Defuzzification
- Applications

Module IV – Evolutionary Computing (12 Hours)

- Genetic Algorithms
- Chromosomes
- Selection
- Crossover
- Mutation
- Fitness Function
- Particle Swarm Optimization
- Ant Colony Optimization
- Simulated Annealing
- Optimization Problems
- Hybrid Optimization
- Applications
- Future Trends

KU4DSCSC-ELE02-B- LAB

Course Title: SOFT COMPUTING TECHNIQUES

Course Structure

- Lecture Hours: 3 Hours/Week
- Credits: 3

1. Implement a simple Perceptron model.
2. Demonstrate activation functions.



3. Implement a simple Artificial Neural Network.
4. Create a basic fuzzy inference system.
5. Perform fuzzy decision making.
6. Simulate a simple Genetic Algorithm.
7. Compare Neural Networks and Fuzzy Logic.
8. Implement a small Soft Computing application.

KU4DSCSC-ELE02-C

Course Title: ADVANCED COMPUTER VISION

Course Structure

- Lecture Hours: 3 Hours/Week
- Credits: 3

Module I – Introduction to Computer Vision (12 Hours)

- Computer Vision Fundamentals
- Digital Images
- Image Representation
- Image Enhancement
- Image Filtering
- Edge Detection
- Feature Detection
- Image Segmentation
- Object Detection
- Image Classification
- Applications
- Challenges

Module II – Deep Learning for Vision (12 Hours)

CNN Models

- CNN Basics
- Convolution Layers
- Pooling
- Activation Functions
- Transfer Learning
- Image Classification

Module III – Detection Models (12 Hours)

- YOLO
- R-CNN
- SSD
- Object Tracking
- Image Annotation
- Model Evaluation

Module IV – Advanced Vision Applications (12 Hours)

- Face Detection
- Face Recognition
- OCR
- Pose Estimation
- Video Analytics
- Medical Imaging
- Autonomous Vehicles
- Surveillance
- Gesture Recognition



- Augmented Reality
- Explainable Vision
- Ethics
- Industrial Applications
- Future Trends

KU4DSCSC-ELE02-C-LAB

Course Title: ADVANCED COMPUTER VISION

Course Structure

- Lab Hours: 3 Hours/Week
- Credits: 2

Simple Lab Questions

1. Install OpenCV and load an image.
2. Perform image resizing and cropping.
3. Convert images into grayscale.
4. Detect edges using OpenCV.
5. Detect faces using Haar Cascade.
6. Perform image classification using a pre-trained CNN.
7. Detect objects using YOLO.
8. Apply image augmentation techniques.
9. Compare different image filtering methods.
10. Develop a simple computer vision application.

KU4DSCSC-ELE02-D

Course Title: SOCIAL MEDIA ANALYTICS

Course Structure

- Lecture Hours: 3 Hours/Week
- Credits: 3

Module I – Introduction to Social Media Analytics (12 Hours)

- Social Media Platforms
- Social Media Data
- Digital Communication
- Social Networks
- Social Media Metrics
- Data Collection
- User Engagement
- Content Analytics
- Business Applications

Module II – Social Media Data Collection (12 Hours)

- APIs
- Web Scraping Basics
- Data Cleaning
- Text Preprocessing
- Sentiment Analysis
- Keyword Analysis

Module III – Social Data Visualization

- Trend Analysis
- Word Cloud
- Hashtag Analysis
- Dashboard Basics



- Reporting
- Interpretation

Module IV – Applications (12 Hours)

- Marketing Analytics
- Customer Analytics
- Public Opinion Analysis
- Crisis Management
- Political Analytics
- Healthcare Analytics
- Topic Modeling
- Fake News Detection
- Social Listening

KU4DSCSC-ELE02-D-LAB

Course Title: SOCIAL MEDIA ANALYTICS

Course Structure

- Lecture Hours: 3 Hours/Week
- Credits: 2

Simple Lab Questions

1. Study different social media platforms and their analytics.
2. Collect sample social media data.
3. Perform basic text preprocessing.
4. Generate a word cloud.
5. Perform sentiment analysis.
6. Visualize social media trends.
7. Analyze hashtags using sample data.
8. Create a dashboard for social media metrics.
9. Identify influencers using simple metrics.
10. Prepare an analytical report from social media data.

KU4DSCSC-ELE02-E

Course Title: INTERNET OF THINGS

Course Structure

- Lecture Hours: 3 Hours/Week
- Credits: 3

Module I – Introduction to Internet of Things (12 Hours)

- Introduction to IoT
- Characteristics of IoT
- IoT Architecture
- Sensors
- Actuators
- Embedded Systems
- Communication Protocols
- IoT Devices
- IoT Applications

Module II – IoT Platforms and Communication (12 Hours)

IoT Hardware

- Arduino
- ESP32/NodeMCU
- Raspberry Pi
- Sensor Interfacing



- Actuator Control
- Device Programming

IoT Communication

- Wi-Fi
- Bluetooth
- MQTT
- HTTP
- Cloud Platforms
- Data Collection

Module III – IoT Data Processing (12 Hours)

- IoT Data Acquisition
- Cloud Storage
- Data Visualization
- Real-time Monitoring
- Dashboard Development
- Edge Computing
- IoT Analytics
- Automation
- Alerts
- Device Management
- Security
- Performance

Module IV – IoT Applications (12 Hours)

- Smart Home
- Smart Agriculture
- Smart Healthcare
- Smart Cities
- Industrial IoT
- Environmental Monitoring
- Wearable Devices
- Energy Management
- AI and IoT

KU4DSCSC-ELE02-E-LAB

Course Title: INTERNET OF THINGS

Course Structure

- Lecture Hours: 3 Hours/Week
 - Credits: 2
1. Interface an LED with Arduino or NodeMCU.
 2. Read temperature and humidity data using a sensor.
 3. Interface an ultrasonic sensor with a microcontroller.
 4. Display sensor readings on the serial monitor.
 5. Develop a simple IoT monitoring dashboard.
 6. Control an LED remotely using Wi-Fi.
 7. Create a basic smart home automation prototype.
 8. Develop a simple IoT application for environmental monitoring.



KU4DSCSC05- PROJECT

Course Structure

- Project Hours: 6 Hours/Week
- Credits: 6

Project Description:

The project aims to solve a real-world problem using data science techniques by applying data collection, preprocessing, exploratory data analysis, machine learning, visualization, and model evaluation. Students are expected to develop an end-to-end data-driven solution using appropriate tools and frameworks, interpret the results, and present meaningful insights through reports and visual dashboards.

