AC: 21/02/2026
Item No. : 3.4

SIES (Nerul) College of Arts, Science & Commerce (Autonomous)
RISE WITH EDUCATION

SIES (Nerul) College of Arts, Science and Commerce (Autonomous)
Syllabus for Approval
B.Sc. (Data Science)

Sr. No.	Heading	Particulars
1	Title of the course	T.Y.B.Sc.(Data Science)
2	Passing Marks	40%
3	Semesters	V and VI
4	Level	UG
5	Pattern	3-4 years & 6-8 semesters Choice Based Grading System
6	Status	New
7	To be implemented from	From Academic year 2026-27 in a progressive manner

Date: 21/02/2026

Signature:

Dr. Koel Roychoudhury

AC Chairperson

Dr. Nutan Sawant

Course Coordinator

SIES (NERUL) COLLEGE OF ARTS, SCIENCE AND COMMERCE (AUTONOMOUS) NERUL NAVI MUMBAI 400 706.

SIES (Nerul) College of Arts, Science and Commerce
(Autonomous)

B.Sc. Data Science

Semester V and VI

(To be implemented from the Academic Year 2026-27)

SIES (Nerul) College of Arts, Science and Commerce
(Autonomous)
(Affiliated to University of Mumbai)
RE-ACCREDITED GRADE "A" BY NAAC (3rd CYCLE)

BOARD OF STUDIES SYLLABUS FOR
B.Sc. (Data Science)

(WITH EFFECT FROM THE ACADEMIC YEAR 2024-2025)

OBJECTIVES OF THE PROGRAMME

- To strengthen the fundamentals of data science and to develop relevant programming abilities.
- To develop the proficiency with statistical analysis of data.
- To develop and exhibit expertise in data management.
- To demonstrate skill and apply tools and techniques for transformation of data and statistical data analysis for multidisciplinary approach.
- To enhance employability skills and provide scope for higher education and research in the field of data science.
- It provides learners with an understanding of the fundamentals and core concepts of data science, which are essential for the industry.

PREAMBLE

Data is the most important asset in this era of digital revolution. The technological innovations are seen in all walks of life and therefore we are flooded with massive data. Every business relies on data to deliver better products as well as services. The study of data science has become essential to meet the growing demand for data scientists and data analysts.

The application of numerous tools and techniques in the fields of computer science, mathematics, and statistics gave rise to the field of data science. The need to gather and evaluate the massive amounts of data found in various application domains is growing.

This course focuses on educating the students about the fundamentals of computer science, applied mathematics, and applied statistics with respect to the data science applications.

Programme Outcomes

PO1: To utilize your understanding of computers to examine practical applications.

PO2: Capacity to deal with the ever-changing technical environment, recognizes and analyzes large amounts of data, and modify and adapt in order to support the expansion of the IT industry.

PO3: Create models and programs to address the different domain-specific issues.

PO4: Understand the many forms and types of data to be able to solve the intricate difficulties in the field of data science.

	SEM V		
Course Code	Course Type	Course Title	Credit
U26DS5MJ01	Major	Data Engineering Fundamentals	3
U26DS5MJP01		Data Engineering Fundamentals Practical	1
U26DS5MJ02	Major	Foundations of Natural Language Processing	3
U26DS5MJP02		Foundations of Natural Language Processing Practical	1
U26DS5MJ03	Major	Project Management	2
U26DS5E01	Elective	Data Science for Health and Social Care	4
U26DS5EP01		Data Science for Health and Social Care Practical	
U26DS5E02		Financial Analytics	
U26DS5EP02		Financial Analytics Practical	
U26DS5MI01	Minor	Regression Analysis	3
U26DS5MIP01		Regression Analysis Practical	1
U26DS5VSC01	VSC	Introduction to Data Mining	1
U26DS5VSCP01		Introduction to Data Mining Practical	1
U26DS5FP01	FP	Field Project	2
	Total		22

SEM VI			
Course Code	Course Type	Course Title	Credit
U26DS6MJ01	Major	Big Data Technologies	3
U26DS6MJP01		Big Data Technologies Practical	1
U26DS6MJ02	Major	Advanced Natural Language Processing	3
U26DS6MJP02		Advanced Natural Language Processing Practical	1
U26DS6MJ03	Major	Data Security Principles and Practices	2
U26DS6E01	Elective	Text, Web and Social Media Analytics	
U26DS6EP01		Text, Web and Social Media Analytics Practical	
U26DS6E02		Business Analytics	
U26DS6EP02		Business Analytics Practical	4
U26DS6MI01	Minor	Quantitative Techniques for Decision Making	3
U26DS6MIP01		Quantitative Techniques for Decision Making Practical	1
U26DS6OJT01	OJT	On Job Training	4
	Total		22

Semester - V

Major (3+1)

Data Engineering Fundamentals

COURSE CODE: U26DS5MJ01 **COURSE CREDIT: 03**

1 credit - 15 lectures **1 lecture is 60 minutes**

Course Objectives:

This course aims to:

1. Introduce basics of Data Engineering and data architecture.
2. Learn data sources, storage, and ingestion methods.
3. Explain data processing, serving, and security concepts.

Course Outcome:

By the end of this course, students will be able to:

1. Explain Data Engineering and the data lifecycle.
2. Identify data sources and storage and apply basic ingestion.
3. Perform basic data queries and understand data security.

Unit	Topics	Lectures
I	**Introduction to Data Engineering:** Meaning and importance of Data Engineering. Role of Data Engineer in Data Science, Difference between Data Engineer, Data Analyst and Data Scientist, Skills and activities of a Data Engineer, Data Engineers inside an organization **Data Engineering Lifecycle:** Data Engineering Lifecycle. Phases of the lifecycle: Data Generation, Storage, Ingestion, Processing, Serving. Major ideas across the lifecycle **Data Generation and Source Systems:** Data creation mechanisms. Types of data sources: Databases, Files, APIs, Logs and IoT, Practical examples of source systems, Roles involved in managing source systems	15

II	**Basics of Data Architecture:** Data Architecture concepts, Principles of good data architecture, Centralized and Distributed systems, Types of data architecture: Data Warehouse, Data Lake and Lakehouse, Stakeholders involved in designing data architecture	15
	Data Storage Concepts: Need for data storage, Storage systems: Relational databases and NoSQL databases, Data Warehouse concepts: Fact tables and Dimension tables, Data Lake; Trends in data storage, Storage abstractions: Tables, Files and Partitions **Data Ingestion:** Data ingestion concepts, Batch and Stream ingestion, Ingestion methods: File-based, API-based and Database import, Overview of ETL tools and connectors, Roles involved in ingestion pipelines	
III	**Queries, Modelling and Transformation:** Data modelling basics, Data transformation concepts and examples, Role of data engineers in data preparation **Serving Data for Analytics and Machine Learning:** Data serving concepts,Data for analytics and reporting, Data for machine learning models,Introduction to Reverse ETL, Support provided by data engineers to ML teams **Security and Privacy in Data Engineering:** Data security fundamentals, Privacy issues in data handling. People, processes and technology involved in data protection	15

References:

- Fundamentals of Data Engineering,Joe Reis and Matt Housley,O'Reilly Media,1st Edition,2022
- Learning Spark: Lightning-Fast Data Analytics,Jules S. Damji, Brooke Wenig O'Reilly Media,2nd Edition ,2020
- Kafka: The Definitive Guide: Real-Time Data and Stream Processing at Scale Neha Narkhede, Gwen Shapira & Todd Palino,O'Reilly Media, 1st edition, 2017
- Data Pipelines Pocket Reference,James Densmore,O'Reilly Media,1st Edition,2021
- Data Engineering with Python,Paul Crickard,Packt Publishing,1st edition,2020

Data Engineering Fundamentals Practical

COURSE CODE: U26DS5MJP01 **COURSE CREDIT: 01**

1 credit - 15 lectures **1 lecture is 120 minutes**

Course Objectives:

This course aims to:

- Develop hands-on skills in basic data engineering tasks.

Course Outcome:

By the end of this course, students will be able to:

- Build simple data pipelines using real data.

Sr. No.	Practical
1	**Study of Data Engineering and its Lifecycle** Explain the meaning and importance of Data Engineering. Study the role of a Data Engineer in an organization and describe the phases of the Data Engineering lifecycle with a neat diagram.
2	**Design of Data Architecture** Design a simple data architecture for an e-commerce or healthcare system showing Data Sources, Storage, Processing, and Serving layers. Compare Data Warehouse, Data Lake, and Lakehouse.
3	**Identification of Data Sources** Identify and classify various data sources such as databases, files, APIs, logs, and IoT. Give real-world examples for each.
4	**Implementation of Relational Data Storage** Create a relational database and tables using MySQL/PostgreSQL for a given dataset. Insert, update, and retrieve data using SQL queries.
5	**Implementation of NoSQL Data Storage** Store and retrieve semi-structured data using MongoDB. Compare it with relational databases.
6	**Study of Data Storage Formats** Store the same dataset in CSV, JSON, and Parquet formats. Compare their size, readability, and access speed.

7	**Batch Data Ingestion Using Files** Perform batch ingestion of data from CSV files into a database using Python or SQL and validate the loaded data.
8	**API-Based Data Ingestion** Use a public API to extract data and store it into a database or file system using Python.
9	**Data Cleaning and Transformation** Perform data cleaning and transformation on raw data using SQL/Python. Handle missing values, duplicates, and incorrect entries.
10	**Data Modelling Using Fact and Dimension Tables** Design a simple data model using fact and dimension tables for a sales or student dataset.
11	**SQL Queries for Analytics and Reporting** Write SQL queries to generate reports such as total sales, top products, or student performance summaries.
12	**Serving Data for Analytics** Connect the prepared dataset to Power BI/Excel and create a basic dashboard for visualization and reporting.
13	**Mini Project: End-to-End Data Engineering Pipeline** Design and implement a mini data pipeline that includes data ingestion, storage, transformation, and serving for analytics.

Major (3+1)
Foundations of Natural Language Processing

COURSE CODE: U26DS5MJ02 **COURSE CREDIT: 03**

1 credit - 15 lectures **1 lecture is 60 minutes**

Course Objectives:

This course aims to:

1. Introduce and explain the fundamental concepts, evolution, characteristics, and challenges of Natural Language Processing, enabling students to understand and analyze the structure and behavior of natural language.
2. Develop the ability to apply core NLP tasks, processing techniques, and linguistic resources for transforming and analyzing textual data.
3. Enable students to apply and evaluate NLP techniques and tools for solving real-world data science problems involving text and speech.

Course Outcome:

By the end of this course, students will be able to:

1. Explain the basic principles, evolution, and linguistic foundations of NLP and analyze language-related challenges such as ambiguity and context in natural language data.
2. Apply text preprocessing methods and core NLP tasks such as tagging, parsing, and named entity recognition, and analyze the role of linguistic resources in NLP systems.
3. Apply NLP techniques to real-world applications such as sentiment analysis, text classification, and machine translation, and evaluate ethical issues, challenges, and future trends in NLP systems.

Unit	Topics	Lectures
I	**Foundations of Natural Language Processing:** Introduction to Natural Language Processing (NLP), Role of NLP in Artificial Intelligence and Data Science, Characteristics of Natural Language.	
	Brief History and Evolution of NLP: Rule-based NLP, Statistical NLP, Machine Learning–based NLP, Neural and Transformer-based NLP (overview). **Challenges and Ambiguity in NLP:** Handling Ambiguity, Polysemy, Language Variations, Morphological, Syntactic & Semantic Challenges. Lexical ambiguity, Syntactic ambiguity, Semantic ambiguity, Types of Ambiguities with Examples. **NLP Pipeline and Processing Steps:** Text collection, preprocessing, tokenization, normalization, morphological analysis, syntactic parsing, semantic analysis, discourse and pragmatic processing, feature extraction, model building, and evaluation.	15
II	**Core NLP Tasks** **Text Preprocessing & Tokenization:** Text Cleaning & Normalization, Tokenization: Word & Sentence Segmentation, Stopword Removal, Lemmatization & Stemming. **Part-of-Speech (POS) Tagging:** Definition & Importance of POS Tagging, POS Tagging Methods: Rule-Based, Hidden Markov Model. **Named Entity Recognition (NER):** Rule-Based vs. Machine Learning-based NER, NER using SpaCy. **Word Sense Disambiguation (WSD):** Lexical Ambiguity & Polysemy, WSD Methods: WordNet-based, Machine Learning-based. **Evaluation Metrics for NLP Tasks**: Precision, Recall, F1 Score, BLEU Score for Machine Translation.	15
III	**NLP Applications, Tools, and Industry Use Cases** NLP Applications in Data Science, Text Classification, Sentiment Analysis and Opinion Mining, Grammar and Spell Checking Systems. **Use cases:** question answering, social media text analytics, information extraction, conceptual overview on Natural Language Generation Applications, Machine Translation, Speech-to-Text (STT) Systems, Text-to-Speech (TTS) Systems applications. **Ethics & Challenges in NLP:** Bias in NLP Models, Fake News Detection &	15

	Misinformation, Fairness & Explainability in NLP Models.	

Books & References:

1. Jurafsky & Martin – Speech and Language Processing, Pearson Education India, 2nd 2013
2. Natural Language Processing with Python (NLTK Book)
3. A Practical Introduction to Natural Language Processing (Sanjeev Kulkarni & Christopher Ré)
4. video Link : https://nptel.ac.in/courses/106105158

Foundations of Natural Language Processing Practical

COURSE CODE: U26DS5MJP02 **COURSE CREDIT: 01**

1 credit - 15 lectures **1 lecture is 120 minutes**

Course Objectives:

This course aims to:

- Introduce and enable students to apply fundamental Natural Language Processing techniques using Python for pre-processing, analysing, and building simple text-based applications relevant to data science.

Course Outcome:

By the end of this course, students will be able to:

- Apply basic NLP tools and techniques to pre-process text data, perform core NLP tasks such as tagging, sentiment analysis, and text classification, and develop simple NLP applications.

Sr. No.	Practical
1	Install Python and understand the NLTK toolkit.
2	To implement and familiarize with the APIs provided by the NLTK python.
3	Text Collection and Basic Exploration: Load text from a file or CSV and display word count, sentence count, and sample text.
4	Text Cleaning and Normalization: Remove punctuation, convert text to lowercase, and remove special characters.
5	Stopword Removal: Identify and remove stopwords using NLTK.
6	Implement text preprocessing techniques stemming and lemmatization.

7	Perform POS tagging using NLTK or spaCy and analyze tagged output.
8	Construct and observe Bag of Words representations of text data and understand word frequency-based features.
9	Apply and identify parts of speech such as nouns, verbs, and adjectives in a given text.
10	Implement Named Entity Recognition (NER) and label named entities such as person names, locations, and organizations from text data using spaCy
11	Implement simple text categorization techniques using predefined classes.
12	Implement a program to classify the sentiment of a given text as positive, negative, or neutral.
13	Design and implement NLP-based applications.

Major (2)
Project Management

COURSE CODE:U26DS5MJ03 **COURSE CREDIT: 02**

1 credit - 15 lectures **1 lecture is 60 minutes**

Course Objectives:

This course aims to:

1. Equip students with foundational project management knowledge and practical planning skills required to effectively select, organize, and plan data science projects by integrating data feasibility, analytics project life cycle, and business value considerations.
2. Enable students to plan, schedule, and monitor data science projects effectively, incorporating resource allocation, risk management, quality control, and performance evaluation using project management tools and statistical techniques.

Course Outcome:

By the end of this course, students will be able to:

1. Apply project management concepts and planning techniques to select, organize, and plan data science projects considering data feasibility, business value, and analytics ROI.
2. Analyse and apply project scheduling, resource management, and risk mitigation techniques to ensure timely, quality-driven execution of data science projects while addressing data, model, and ethical risks.

Unit	Topics	Lectures
I	**Introduction to Project Management:** What is a project? Evolution of project management, the need and relevance of project management, Characteristics of projects, Characteristics of project management, Project life cycle. Data Science Project Life Cycle (CRISP-DM), software projects vs data science projects. **Types of Data Science projects:** Predictive analytics, Descriptive analytics, Prescriptive analytics, Role of project management in analytics-driven organizations. **Project Selection and Appraisal:** Feasibility analysis, Brainstorming and concept evolution, Project selection and evaluation, Selection criteria and models, SWOT	15

	analysis, Cashflow analysis, application to analytics ROI (Payback period, and Net present value) **Project Organization and Planning:** Project manager, Cross-functional team, Matrix organization, Advantages and disadvantages of project organizations, Selection of project organization, Concepts and applications of Work Breakdown Structure (WBS)-Network analysis for time management using CPM, PERT.	
II	**Project Scheduling and Resource Management:** Software Effort estimation techniques, AON and AOA representation, Three time estimates, Timescale version of network, Early start and late start schedules, Resource allocation, Resource loading and levelling, Multi-project scheduling and resource allocation, Crashing a project. **Managing Project Risks:** Identification, Assessment, and Mitigation. Data-related risks: Missing data, Bias and imbalance, Data leakage, Model risks: Overfitting, Poor generalization, Ethical and legal risks: Privacy, Fairness **Earned value concept in project control:** Calculation of Schedule and Cost Variances, Quality management through statistical tools, and Cause and Effect Analysis	15

References:

1. Software Project Management, sixth edition, Rajib Mall.
2. Project Management for Business and Technology: Principles and Practice, John M. Nicholas, Pearson Prentice Hall, New Delhi, 2005.
3. A Guide to the Project Management Body of Knowledge (PMBOKGuide) 5th Edition, PMI.
4. Project Management: A Managerial Approach, Meredith, J. R. and Mantel Jr., S. J., John Wiley, New York. 2004
5. Quantitative Methods in Project Management, Good pasture, J. C., J Ross Publishing, Boca Raton, Florida, USA.2003

Elective (3+1)
Data Science for Health and Social Care

COURSE CODE: U26DS5E01 **COURSE CREDIT: 03**

1 credit - 15 lectures **1 lecture is 60 minutes**

Course Objectives:

This course aims to:

1. To introduce healthcare data, EHR systems, and healthcare analytics.
2. To introduce medical images, signals, and clinical text analysis.
3. To introduce clinical prediction models, decision support systems, and healthcare visualization.

Course Outcome:

By the end of this course, students will be able to:

1. Students will understand healthcare data and apply analytics in healthcare applications.
2. Students will analyze biomedical data and clinical text for healthcare use.
3. Students will understand healthcare prediction, CDSS, and interpret healthcare dashboards.

Unit	Topics	Lectures
I	**Introduction to Data Science in Healthcare:** Role of Data Science in modern healthcare; Evolution of healthcare analytics; Difference between clinical data, operational data, and public health data; Opportunities and limitations of healthcare data analytics.Healthcare Data Sources. **Electronic Health Records (EHR):** Definition and purpose of EHR; Components of EHR: Patient demographics, Medical history, Diagnosis and treatment, Laboratory reports, Imaging records,Benefits and Barriers of EHR,Challenges in EHR Data Analytics **Applications of Data Analytics in Healthcare and Practical Healthcare Systems:** Disease prediction; Hospital resource management; Personalized treatment; Epidemiology and public health surveillance; Healthcare quality improvement	15

II	**Biomedical Image Analysis:** Introduction to medical imaging; Imaging modalities: X-ray, CT, MRI, Ultrasound (conceptual); Image preprocessing; Image segmentation; Object detection; Feature extraction; Applications in diagnosis. **Biomedical Signal Analysis:** Types of biomedical signals: ECG, EEG, EMG; Signal visualization; Basic feature extraction; Applications in health monitoring. **NLP for Clinical Text:** Nature of clinical text data; Challenges in clinical text processing; Information extraction from clinical notes; Rule-based approaches; Machine learning-based approaches; Diagnosis support and symptom detection	15
III	**Clinical Prediction Models:** Introduction to predictive analytics in healthcare; Types of clinical prediction models; Applications in diagnosis and risk assessment; Limitations and challenges **Clinical Decision Support Systems (CDSS):** Definition and objectives; Components of CDSS; Role of data science in CDSS; Examples of CDSS in hospitals **Healthcare Data Visualization:** Importance of visualization in healthcare; Types of healthcare dashboards; Tools for healthcare visualization; Best practices	15

References:

- Chandan K. Reddy and Charu C Aggarwal, "Healthcare data analytics", Taylor & Francis, 2015.
- 2 Hui Yang and Eva K. Lee, "Healthcare Analytics: From Data to Knowledge to Healthcare Improvement, Wiley,2016.
- 1 Madsen, L. B. (2015). Data-driven healthcare: how analytics and BI are transforming the industry. Wiley India Private Limited
- 2 Strome, T. L., & Liefer, A. (2013). Healthcare analytics for quality and performance improvement. Hoboken,NJ, USA: Wiley

Data Science for Health and Social Care Practical

COURSE CODE:U26DS5EP01 **COURSE CREDIT: 01**

1 credit - 15 lectures **1 lecture is 120 minutes**

Course Objectives:

This course aims to:

- To provide hands-on experience in analyzing healthcare data using data analytics, machine learning, image processing, signal processing, NLP, and visualization techniques.

Course Outcome:

By the end of this course, students will be able to:

- Apply data science techniques to solve real-world healthcare problems including prediction, diagnosis support, and healthcare decision-making.

Sr. No.	Practical
1	Load a healthcare dataset and perform exploratory data analysis to identify: (a) types of attributes (clinical, operational, public health), (b) missing values and duplicates, and (c) summary statistics.
2	Given an Electronic Health Record dataset, perform data cleaning by handling missing values, removing duplicates, normalizing numerical fields, and encoding categorical attributes.
3	Analyze patient demographic data to determine age group-wise distribution, gender-wise disease occurrence, and department-wise patient count with suitable visualizations.
4	Build a machine learning model to predict disease occurrence using patient health indicators and evaluate using accuracy and confusion matrix.
5	Analyze hospital admission data to compute bed occupancy rate, average length of stay, and peak admission periods using visual representation.
6	Design an interactive healthcare dashboard showing patient distribution by department, disease

	trends, and recovery/mortality rates.
7	Perform preprocessing on a medical image (X-ray/MRI): grayscale conversion, resizing, smoothing, and edge detection.
8	Apply image segmentation techniques to extract the region of interest from a medical image and display the segmented output.
9	Load and plot a biomedical signal (ECG/EEG), identify peaks, and analyze basic signal patterns.
10	Extract time-domain features such as mean, variance, and heart rate from a biomedical signal and interpret their significance.
11	Perform NLP preprocessing on clinical notes including tokenization, stop-word removal, keyword extraction, and word frequency analysis.
12	Develop a text classification model to categorize clinical notes into disease classes using TF-IDF and a suitable classifier.
13	Design a simple CDSS that accepts patient vitals, predicts health risk, and suggests basic clinical actions based on prediction.

Elective (3+1)
Financial Analytics

COURSE CODE: U26DS5E02 **COURSE CREDIT: 03**

1 credit - 15 lectures **1 lecture is 60 minutes**

Course Objectives:

This course aims to:

1. Introduce students to the basic concepts, scope, and analytics of finance.
2. Explain fundamental risk and return concepts for financial decision-making.
3. Provide an understanding of simple predictive techniques in finance.

Course Outcome:

By the end of this course, students will be able to:

1. Explain basic financial concepts, markets, and key financial terms.
2. Handle, preprocess, and analyze financial data using descriptive analytics techniques.
3. Evaluate risk and return using simple statistical measures

Unit	Topics	No of Lectures
I	**Introduction to Finance:** Meaning and scope of finance, Role of finance in business and economy, Types of financial markets, Key financial terms: price, return, risk **Introduction to Financial Analytics:** What is financial analytics?, Role of analytics in finance, Types of financial analytics, Overview of financial data	15
II	**Financial Data & Basic Data Handling:** Types of financial data, Sources of financial data, Understanding stock price data, Data cleaning and preparation **Descriptive Analytics in Finance:** Measures of central tendency and dispersion, Returns calculation, Data visualization techniques, Trend analysis and interpretation	15

III	**Introduction to Risk & Return Analysis:** Concept of risk and return, Simple risk measures (variance, volatility), Risk–return relationship, Basic diversification concept **Basic Predictive Analytics & Case Studies:** Introduction to prediction in finance, Simple regression and moving averages, Basic forecasting concepts, Simple financial analytics case studies	**15**

References:

1. Mohanty, Pitabas, Financial Analytics, Wiley India.
2. Hilpisch, Yves, Python for Finance, O'Reilly Media.
3. Bodie, Z., Kane, A., Marcus, A. J., Investments, McGraw-Hill Education.
4. Tsay, Ruey S., Analysis of Financial Time Series, Wiley.

Financial Analytics Practical

COURSE CODE: U26DS5EP02 **COURSE CREDIT: 01**

1 Credit – 15 lectures **1 lecture is 120 minutes**

Course Objectives:

This course aims to:

- To familiarize learners with financial data handling, analysis, visualization, risk–return measurement, and basic forecasting techniques using real-world financial datasets.

Course Outcome:

By the end of this course, students will be able to:

- Learners will be able to clean and analyze financial data, compute returns and risk, visualize trends, assess asset relationships, and apply basic predictive and forecasting methods to financial problems.

Sr. No.	List of Practical's
1.	**Familiarization and Loading of Financial Data** Understand financial datasets, identify stock attributes, and load data from CSV files.
2.	**Financial Data Cleaning and Preparation** Handle missing values, sort and filter data
3.	**Exploratory Data Analysis of Financial Data** Examine data distribution, summary statistics, and detect anomalies.
4.	**Calculation of Daily and Monthly Returns** Compute daily and monthly returns.
5.	**Descriptive Statistical Analysis of Financial Data** Calculate mean, median, variance, standard deviation, and interpret volatility.
6.	**Visualization of Stock Prices and Returns** Plot line charts for stock prices, bar charts for returns, and compare multiple assets.

7.	**Trend Analysis Using Moving Averages** Identify price trends, calculate moving averages, and compare actual prices with trends.
8.	**Introduction to Correlation and Covariance** Calculate correlation and covariance between two or more financial assets.
9.	**Basic Risk and Return Analysis** Measure risk using variance and volatility, analyze risk–return relationship.
10.	**Introduction to Predictive Analytics Using Regression** Apply simple linear regression for financial data and interpret prediction results.
11.	**Basic Forecasting Using Time-Based Techniques** Forecast stock prices using moving averages and compare actual versus predicted values.
12.	**Mini Case Study on Financial Analytics** Analyze a selected stock or financial instrument using data cleaning, visualization, risk assessment, and present findings.
13.	**Mini Project**

Minor (3+1)
Regression Analysis

COURSE CODE: U26DS5MI01 **COURSE CREDIT: 03**

1 credit - 15 lectures **1 lecture is 60 minutes**

Course Objectives:

This course aims to:

1. Understand and apply simple and multiple regression models for data analysis
2. Formulate, estimate, and interpret regression models and results for informed decision-making
3. Understand and apply multiple regression using OLS, matrix form, and coefficient testing.

Course Outcome:

By the end of this course, students will be able to:

1. Construct, estimate, and interpret simple and multiple linear regression models, including hypothesis testing of regression coefficients
2. Apply model-building techniques such as transformations, polynomial regression, interaction effects, and dummy variables
3. Able to build, estimate, and interpret multiple regression models and test coefficients.

Unit	Topics	No of Lectures
I	**Introduction**: Meaning and objectives of regression analysis Historic origin of term regression, regression versus causation and correlation, Difference between correlation and regression, Applications of regression in business, economic, social sciences, and decision-making. Introduction to Regression and model building, Data collection, Uses of regression, Role of computer. Types of regression: simple vs multiple.	15
II	**Simple Regression Models with One Independent Variable** - Assumptions, Regression Model and Regression equation. Estimated regression equation, least square method, Properties of least square estimates, coefficient of determination testing of significance, using the estimated regression equation for estimation and prediction residual analysis, validating model assumptions.	15

III	**Multiple Regression Analysis-** Adding a regressor to a simple linear regression model, The Multiple Regression Model Multiple Regression Model with Two Independent Variables (First-Order) ordinary least squares, Determining the Multiple Regression Equation, A Multiple Regression Model. Ordinary least square estimators, simple regression in matrix form, coefficient of determination, Hypothesis concerning one coefficient.	**15**

Text Books:

1. Sanford Weisberg, Applied Linear Regression 4th edition, Wiley publication.
2. Gujarati, D. N., & Porter, D. C. *Basic Econometrics* (5th ed.). McGraw-Hill Education.
3. Montgomery, D. C., Peck, E. A., & Vining, G. G. Introduction to Linear Regression Analysis (5th ed.). Wiley.

References:

1. Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. Applied Linear Statistical Models (5th ed.). McGraw-Hill Education.
2. Wooldridge, J. M. Introductory Econometrics: A Modern Approach (6th ed.). Cengage Learning.
3. Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. Applied Logistic Regression (3rd ed.). Wiley.

Regression Analysis Practical

COURSE CODE: U26DS5MIP01 **COURSE CREDIT: 01**

~~1 credit - 15 lectures~~ ~~1 lecture is 120 minutes~~

Course Objectives:

This course aims to:

- Enable students to apply and analyze regression techniques using real datasets, including model estimation, hypothesis testing, diagnostics, and interpretation of results for data-driven decision-making.

Course Outcome:

By the end of this course, students will be able to:

- Construct, evaluate, and interpret linear and logistic regression models, perform diagnostic analysis, and assess model suitability using statistical measures and tests.

Sr. No.	Practical
1	Visualize Scatter Plot and Correlation Analysis
2	Implement Simple Linear Regression Model
3	Estimation and Interpretation of Regression Coefficients
4	Hypothesis Testing of Regression Coefficients
5	Implement Standard Error of Estimation and Prediction
6	Implement Multiple Regression with Two Independent Variables
7	Implement Residual Analysis and Model Assumption Checking
8	Implement Matrix Formulation and OLS Estimation
9	Implement Adding Second Regressor
10	Implement Two-Variable First-Order Model

11	Implement Multiple Coefficient Tests
12	Implement Coefficient of Determination in Multiple
13	Implement Full Model Building

VSC(credit 2)
Introduction to Data Mining

COURSE CODE:U26DS5VSC01 **COURSE CREDIT: 01**

1 Credit – 15 lectures **1 lecture is 60 minutes**

Course Objectives:

This course aims to:

- Provide a comprehensive understanding of data mining, including fundamental concepts, pattern discovery, classification, decision-making models, and clustering techniques for data analysis.

Course Outcome:

By the end of this course, students will be able to:

- Apply data mining concepts to real-world datasets, including pattern discovery, association rules, classification, and clustering techniques for analysis and decision-making.

Unit	Topics	No of Lectures
I	**Introduction to Data Mining:** Definition of Data Mining, Data Mining vs Data Analysis vs Machine Learning, Applications of Data Mining, Knowledge Discovery in Databases (KDD) Process, Data Mining Tasks and Functionalities **Association Rule Mining:** Frequent Pattern Mining, Apriori Algorithm, Association Rules, Measures: Support, Confidence, Lift, Applications of Association Rule Mining **Classification:** Classification Basic Concepts, Decision Tree Induction, Bayes classification methods, Lazy learners **Cluster analysis:** What is cluster analysis?, Partitioning methods, Hierarchical methods, Density-based methods	15

References:

1. **Han, J., Kamber, M., & Pei, J.**, Data Mining: Concepts and Techniques, 3rd Edition, Morgan Kaufmann.
2. **Tan, P.-N., Steinbach, M.**, & Kumar, V., Introduction to Data Mining, Pearson Education.

Introduction to Data Mining Practical

COURSE CODE: U26DS5VSCP01 **COURSE CREDIT: 01**

1 Credit – 15 lectures **1 lecture is 120 minutes**

Course Objectives:

This course aims to:

- Introduce students to fundamentals of data mining and its practical applications and to implement classification, clustering, association, text and spatial mining algorithms using tools.

Course Outcome:

By the end of this course, students will be able to:

- Pre-process data, apply data mining algorithms, and analyze results.

Sr. No.	List of Practical's
1.	Create a Table with the help of a Data Mining Tool.
2.	Apply pre-processing to the training data-set.
3.	Show implementation of classification rule.
4.	Show implementation of Naïve Bayes algorithm.
5.	Show implementation of Decision Tree.
6.	Show implementation of Time Series Algorithm.
7.	Show implementation of Clustering Algorithm.
8.	Show implementation of k-nearest neighbor.
9.	Show implementation of Apriori Algorithm.
10.	Show implementation of Association Algorithm.
11.	Study of Text Mining.

12.	Study of Spatial mining.
13.	Study of Multimedia Mining.

Field Project (Credit 2)

COURSE CODE: U26DS5FP01 **COURSE CREDIT: 02**

~~1 credit - 15 lectures~~ ~~1 lecture is 60 minutes~~

Course Objectives:

- Provide learners with essential skills and knowledge to develop a project
- Offer hands-on experience in data collection, analysis and report making

Course Outcome:

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

- Understand research Methodology
- Data analysis

The primary objective of incorporating field project work is to foster research analysis and cultivate a scientific temperament, challenging the learner's potential by encouraging curiosity, critical inquiry, and the ability to interpret various aspects of the study.
It is expected that the guiding teacher will provide mentoring sessions, equipping learners with the necessary knowledge and skills for formulating, preparing, and evaluating the project effectively.

The project work should be structured around a research-based methodology relevant to the chosen area of study, ensuring a systematic and analytical approach.

General Guidelines for Field Project-

- The learners' will have identify and approach companies / organizations /Filed area on their own wherein they want to pursue their field project according to their area of interest.
- The project topic may be undertaken in any area of Major/Minor Elective Courses.
- The learner shall decide the topic and title which should be specific, clear and with definite scope in consultation with the faculty-guide concerned.
- The college reserves the right to approve or disapprove the topic/ company / organization. Hence, students must take prior approval of the same from the College before proceeding on field project.
- The learners are required to provide details of the topic and organization.
- The students will conduct the project under the supervision of a mentor.

- After completion, the students will prepare and submit the field project report to the college.
- Duration of the Field Project – The total duration of the Field Project will be of 30 hours since it is of 2 Credits.

Particulars	**No of Hours**
Interactive Sessions with Faculty/Guide	6 hours
Sessions from Industry Experts	3 hours
Project Exibition	4 hours
Field Visits (minimum one visit)	7 hours
Field Project Report	10 hours
Total	30 Hours

FIELD PROJECT FORMAT

1st page (Main Page)

Title of the problem of the Field Project

A Field Project Submitted to
University of Mumbai for partial completion of the degree of Bachelor in BSc (Data Science)
Under the Faculty of Science

By

Name of the Learner

Under the Guidance of

Name of the Guiding Teacher

Name and address of the College

Month and Year

2nd Page

This page to be repeated on 2nd page (i.e. inside after main page)

On separate page

Index

Chapter No. 1
(sub point 1.1, 1.1.1, And so on)
Title of the Chapter Page No.

Chapter No. 2 Title of the Chapter

Chapter No. 3 Title of the Chapter

Chapter No. 4 Title of the Chapter

Chapter No. 5 Title of the Chapter

List of tables, if any, with page numbers. List of Graphs, if any, with page numbers. List of Appendix, if any, with page numbers. Abbreviations used:

Structure to be followed to maintain the uniformity in formulation and presentation of Field Project Work (Model Structure of the Field Project Work)

1. Introduction

- Topic Overview
- Statement of the Problem
- Need for the Study
- Research Objectives
- Scope of the Study

2. Review of Literature

- Summary of relevant past research
- Identification of research gaps

3. Research Methodology

- Hypothesis (if applicable)
- Research Design
- Data Collection Methods
- Data Pre-processing & Cleaning
- Analytical Tools & Techniques

4. Data Analysis and Interpretation

- Presentation of findings
- Statistical or qualitative analysis
- Interpretation in relation to objectives

5. Conclusion & Recommendations

- Summary of key findings
- Limitations of the study
- Practical implications & suggestions for future research

6. References

- Cited sources in appropriate format (APA, MLA, etc.)

7. Appendices

- Supplementary materials (raw data, questionnaires, additional figures, etc.)

Evaluation Rubric:

Criteria	Marks
Participation and Data Collection	10
Data Analysis and Interpretation	10
Innovation of topic and social relevance	10
Report and Presentation	20

The Field project will be evaluated equally by an internal as well as an external Faculty.

Semester - VI

Major (3+1)

Big Data Technologies

COURSE CODE: U26DS6MJ01 **COURSE CREDIT: 03**

1 credit - 15 lectures **1 lecture is 60 minutes**

Course Objectives:

This course aims to:

1. To introduce Big Data concepts and Hadoop architecture.
2. To study NoSQL databases and their role in Big Data solutions.
3. To learn data stream processing and similarity-based analysis.

Course Outcome:

By the end of this course, students will be able to:

1. Students will understand Big Data and Hadoop components for large-scale data processing.
2. Students will be able to choose suitable NoSQL models for Big Data applications.
3. Students will apply clustering and distance measures to analyze large datasets.

Unit	Topics	Lectures
I	**Introduction to Big Data:** Introduction to Big Data, Data characteristics (5Vs),Types of Big Data, Traditional vs Big Data business approach, Case studies of Big Data solutions (Netflix, Amazon, Healthcare, Banking) **Hadoop Framework:** Concept of Hadoop ,Core components: HDFS, YARN, MapReduce, Hadoop, Ecosystem: Hive, Pig, HBase, Sqoop, Flume, Spark (overview) **Hadoop Distributed File System (HDFS):** Distributed file systems, Physical organization of compute nodes, Large-scale file system organization, Data replication and fault tolerance	15

II	**MapReduce Basics:** Map and Reduce tasks, grouping by key, Combiners, Map Reduce execution flow, coping with node failures, Simple Map Reduce applications: Word Count, Log Analysis, Hadoop limitations. **Introduction to NoSQL:** What is NoSQL, NoSQL business drivers, CAP theorem (basic idea), NoSQL Data Models: Key-value stores, Document stores, Column-family stores, Graph stores, Use cases of each, Variations of NoSQL architectures **NoSQL for Big Data Solutions:** Types of Big Data problems, Shared-nothing architecture, Data distribution models, Master-slave vs Peer-to-peer, Selecting suitable NoSQL systems, Case study (MongoDB / Cassandra / Neo4j)	15
III	**Mining Data Streams (Conceptual):** Stream data model, Stream sources and queries, Issues in stream processing, Sampling data in streams, Filtering streams using Bloom filters (conceptual),Count distinct problem (overview only),Frequent items in data streams, Applications in IoT, social media, finance. **Similarity & Distance Measures:** Euclidean distance, Jaccard similarity, Cosine similarity, Edit distance, Hamming distance, Applications in text, images, recommender systems **Clustering in Big Data:** Need for clustering-Means (overview),Hierarchical clustering (overview),Density-based clustering (overview),CURE algorithm (only concept & use case),Clustering for large datasets.	15

References:

1. Anand Rajaraman and Jeff Ullman —Mining of Massive Datasets, Cambridge University Press,
2. Alex Holmes -Hadoop in Practice, Manning Press, Dreamtech Press.
3. Dan Mcary and Ann Kelly —Making Sense of NoSQL|| – A guide for managers and the rest of us, Manning Press.
4. Bill Franks, -Taming The Big Data Tidal Wave: Finding Opportunities In Huge Data Streams With Advanced Analytics, Wiley,
5. Chuck Lam, -Hadoop in Action, Dreamtech Press

Big Data Technologies Practical

COURSE CODE: U26DS6MJP01 **COURSE CREDIT: 01**

1 credit - 15 lectures **1 lecture is 120 minutes**

Course Objectives:

This course aims to:

- To provide hands-on experience in Big Data technologies, enabling students to work with Hadoop, NoSQL databases, and Big Data analytics techniques.

Course Outcome:

By the end of this course, students will be able to:

- Implement Big Data solutions using Hadoop and NoSQL tools and apply analytical techniques on large datasets.

Sr. No	Practical
1	Explain the characteristics of Big Data (5Vs) with examples. Analyse any one Big Data case study from Netflix, Amazon, Healthcare, or Banking.
2	Study the Hadoop ecosystem and explain the role of HDFS, YARN, Hive, Pig, HBase, Sqoop, Flume, and Spark.
3	Demonstrate HDFS operations: create directory, upload file, display contents, delete file, and list directory contents.
4	Explain the data replication mechanism in HDFS and demonstrate how Hadoop ensures fault tolerance.
5	Write and execute a MapReduce program to count the frequency of words in a given text file.
6	Analyse a server log file using MapReduce to generate: requests per IP, error count, or frequently accessed URLs.

7	List and explain Hadoop limitations and compare Hadoop with traditional data processing systems.
8	Perform insert, update, delete, and retrieve operations using MongoDB/Cassandra/Neo4j.
9	Design examples for key-value, document, column-family, and graph data models with suitable use cases.
10	Study and report a real-world application using MongoDB/Cassandra/Neo4j for Big Data.
11	Explain data streams and demonstrate any one sampling technique on streaming or simulated data.
12	Simulate a Bloom Filter and analyse false positive behaviour.
13	Compute similarity measures and apply K-Means clustering on a dataset.

Major (3+1)

Advanced Natural Language Processing

COURSE CODE:U26DS6MJ02 **COURSE CREDIT: 03**

1 credit - 15 lectures **1 lecture is 60 minutes**

Course Objectives:

This course aims to:

1. Introduce students to the fundamentals of natural language parsing, grammar concepts, and syntactic analysis used in Natural Language Processing systems.
2. Develop an understanding of different parsing approaches, algorithms, and tools for analyzing sentence structure in real-world NLP applications.
3. Familiarize students with NLP use cases.

Course Outcome:

By the end of this course, students will be able to:

1. Explain basic NLP parsing concepts, grammar formalisms, and syntactic structures used in Natural Language Processing.
2. Apply appropriate parsing techniques and NLP tools to analyze sentence structure and extract grammatical relationships from text data.
3. Analyze the semantic meaning of text using word similarity, WordNet resources and word sense disambiguation.

Unit	Topics	
I	**NLP Foundations & Processing Pipeline:** Introduction to NLP and role of parsing in Data Science, NLP pipeline overview (tokenization to parsing). Phases of NLP: Tokenize, Parse, Analyze meaning, Resolve ambiguity. Levels of NLP Abstraction: Lexical, Syntactic, Semantic, Pragmatic. Linguistic Levels (overview only): Define Linguistic, Phonetics and Phonology, Morphology, Syntax, Semantics, Sentiment, and Pragmatics. NLP Computing Approaches: Rule-based approaches, Statistical approaches, Machine learning approaches. Introduction to core NLP tasks: POS tagging, NER, WSD	**15**

II	**Parsing & Text Representation:** Role of parsing in NLP pipelines, Top-down parsing vs Bottom-up parsing, Constituency vs dependency grammar, Context Free Grammar (CFG) and Parse trees, Sentence ambiguity and need for probabilities, Probabilistic CFG (PCFG), Probabilistic parsing in NLP tools, Parsing using Python NLP libraries (spaCy, NLTK), Applications of parsing in NLP pipelines.	15
III	**Text Mining, Modern NLP & Applications:** Text representation: Bag of Words, TF-IDF. Word embedding: Word2Vec, GloVe, Fast Text (introductory), Curse of dimensionality in text data, Dimensionality reduction: PCA (conceptual). Text clustering: K-Means, Hierarchical clustering. Association rule mining for text data (conceptual Apriori). Transformer models: BERT, GPT, T5 (overview only). Pretrained models and fine-tuning (conceptual). NLP applications: QA, IE, sentiment analysis, text understanding	15

Books and References:

1. Handbook of Natural Language Processing, Indurkhya, N., & Damerau, F. J., CRC Press
2. Taylor and Francis Group, 2nd , 2010
3. Speech and Language Processing, Martin, J. H., & Jurafsky,D.Pearson Education India, 2nd, 2013
4. Foundations of Statistical Natural Language Processing, Manning, Christopher and Heinrich, Schutze, MIT Press, 1st, 1997
5. Natural Language Processing With Python, Steven Bird, Edward Loper, O'Reilly Media, 2nd, 2016
6. Video Links : https://nptel.ac.in/courses/106105158

Advanced Natural Language Processing Practical

COURSE CODE:U26DS6MJP02 **COURSE CREDIT: 01**

1 credit - 15 lectures **1 lecture is 120 minutes**

Course Objectives:

This course aims to:

- To enable students to apply basic Natural Language Processing techniques and tools to parse text, analyze sentence structure, and extract semantic meaning from textual data using Python-based NLP libraries.

Course Outcome:

By the end of this course, students will be able to:

- Implement simple NLP parsing and semantic analysis tasks using standard NLP tools to process and understand real-world text data.

Sr. No	Practical
1	Tokenization (Word & Sentence):Apply word and sentence tokenization using NLTK / spaCy.
2	Building a Complete Text Preprocessing Pipeline.
3	Analyze sentence structure by performing noun phrase chunking using POS tags.
4	Construct and visualize a parse tree for a given sentence using Context Free Grammar.
5	Examine syntactic ambiguity in sentences by generating multiple parse structures.
6	Differentiate between top-down and bottom-up parsing strategies through manual tracing of parsing steps.
7	Text Vectorization Techniques Comparison

8	Sentiment Analysis for Opinion Mining
9	Perform dependency parsing on text and analyze grammatical relationships between words.
10	Analyze dependency relations by extracting subject, object, and root elements from sentences.
11	Explore WordNet for identifying synonyms, antonyms, and semantic relationships among words.
12	Compute semantic similarity between word pairs using WordNet-based measures.
13	Implement a Mini NLP Application.

Major (2)

Data Security Principles and Practices

COURSE CODE: U26DS6MJ03 **COURSE CREDIT: 02**

1 credit - 15 lectures **1 lecture is 60 minutes**

Course Objectives:

This course aims to:

1. To equip students with fundamental knowledge of data security concepts and security threats.
2. To provide students with an understanding of cryptography concepts, techniques, algorithms, and security tools.

Course Outcome:

By the end of this course, students will be able to:

1. Understand the foundational theory and basic principles of information security, including cryptographic techniques and categories of security attacks.
2. Analyze and apply cryptographic algorithms and security tools to protect information systems against various threats.

Unit	Topics	Lectures
I	**Security Concepts:** Introduction, The need for security, Security approaches, Principles of security, Types of Security attacks, Introduction Security, Attacks, Computer Crime, Security Services. Security Mechanisms, Cyber Crimes, Information Technology ACT. **Cryptography Concepts and Techniques:** Introduction, Plain text and Cipher text, Substitution Techniques, Transposition Techniques, Encryption and Decryption, Confusion, Diffusion.	15
II	**Symmetric and Asymmetric Key Cryptography:** Symmetric Key Ciphers: Block Cipher Principles and Algorithms DES, AES, Block Cipher Operation, Stream Ciphers. Asymmetric Key Ciphers: Principles of Public Key Cryptosystems, RSA Algorithm, Diffie-Hellman Key Exchange. **Cryptographic Tools:** Hash function, Digital Signatures, Digital Certificates.	15

Reference Books:

1. Cryptography and Nctwork Sccurity: Atul Kahatc, McGraw-Hill, 2nd Edition
2. Cryptography and Network Security: C K Shyamala, N Harini, Dr T R Padmanabhan, Wiley India, 1st Edition.
3. Cryptography and Network Security: Forouzan Mukhopadhyay, McGraw-Hill, 2"d Edition
4. Information Security, Principles and Practice: Mark Stamp, Wiley India. Forouzan Mukhopadhyay, McGraw-Hill, , 2nd Edition
5. Information Security, Principles and Practice: Mark Stamp, Wiley India

Elective (3+1)

Text, Web and Social Media Analytics

COURSE CODE:U26DS6E01 **COURSE CREDIT: 03**

1 credit - 15 lectures **1 lecture is 60 minutes**

Course Objectives:

This course aims to:

1. Introduce basic concepts and techniques of text mining and text representation.
2. Develop understanding of text clustering, classification, and sentiment analysis.
3. Familiarize students with web and social media data mining concepts.

Course Outcome:

By the end of this course, students will be able to:

1. Preprocess and represent text data and apply basic information extraction methods.
2. Apply clustering, classification, and sentiment analysis on text datasets.
3. Analyze web and social media data and interpret user behavior and recommendations.

Unit	Topics	Lectures
I	**Introduction to Text Mining:** What is Text Mining, Applications in Data Science, Text Mining vs Data Mining, Basic Text Mining Process, Challenges in Text Mining **Text Pre-processing & Representation:**Tokenization,Stop-word Removal, Stemming and Lemmatization,N-grams,Bag of Words (BoW),TF-IDF Representation **Basic Information Extraction:** Named Entity Recognition (concept only),Keyword Extraction, Simple Pattern-based Extraction, Introduction to Relation Extraction (conceptual)	15

II	**Text Clustering:** Feature Selection for Text, Similarity Measures (Cosine, Jaccard),Distance-based Clustering-Means for Text, Hierarchical Clustering, Word and Document Clustering (basic idea) **Text Classification:** Feature Engineering for Text, Supervised Text Classification: Naïve Bayes, Decision Trees ,k-NN,Rule-based vs Machine Learning Classifiers. **Opinion Mining & Sentiment Analysis:** Introduction to Opinion Mining, Document-level Sentiment Classification, Lexicon-based Sentiment Analysis, Supervised vs Unsupervised Sentiment Methods, Applications	15
III	**Web Mining Basics:** Types: Web Content, Structure, Usage Mining, Search Engines Overview, Web Crawling (basic idea),Indexing and Ranking (conceptual),Introduction to SEO and Web Spam **Web Usage Mining:** Data Collection, Clickstream Data, Session Analysis, User Behaviour Patterns, Simple Association & Sequential Patterns **Social Media Mining:** Introduction to Social Media Mining, Challenges in Social Media Datatypes of Social Network Graphs, Influence and Homophily,Behaviour Analytics, Introduction to Recommendation Systems, Evaluating Recommendations	15

References:

- Daniel Jurafsky and James H. Martin, "Speech and Language Processing," 3rd edition, 2020
- 2 Charu. C. Aggarwal, Cheng Xiang Zhai, Mining Text Data, Springer Science and Business Media, 2012.
- 3 BingLiu, "Web Data Mining-Exploring Hyperlinks, Contents, and Usage Data", Springer, Second Edition, 2011.
- 4 Reza Zafarani, Mohammad Ali Abbasiand Huan Liu, "Social Media Mining- An Introduction", Cambridge University Press, 2014

Text, Web and Social Media Analytics Practical

COURSE CODE: U26DS6EP01 **COURSE CREDIT: 01**

1 credit - 15 lectures **1 lecture is 120 minutes**

Course Objectives:

This course aims to:

- Apply data mining techniques for extracting, analysing, and modelling information from text, web, and social media data to discover meaningful patterns and insights.

Course Outcome:

By the end of this course, students will be able to:

- Preprocess unstructured data, implement mining algorithms, and interpret results for tasks such as clustering, classification, recommendation, sentiment analysis, and user behavior analysis.

Sr. No	Practical
1	**Text Mining Introduction** Perform basic text mining on a given document collection and identify applications of text mining in data science.
2	**Text Pre-processing** Implement tokenization, stop-word removal, stemming and lemmatization on a given text dataset
3	**Text Representation** Convert a text corpus into numerical form using Bag of Words (BoW) and TF-IDF models.
4	**Keyword Extraction & NER** Extract keywords and identify named entities from a given document using suitable NLP techniques.
5	**Similarity Measures for Text** Compute Cosine similarity and Jaccard similarity between pairs of text documents.
6	**Text Clustering using K-Means** Perform document clustering using the K-Means algorithm and analyze the formed clusters.
7	**Hierarchical Text Clustering** Apply hierarchical clustering on text data and visualize the cluster dendrogram.

8	**Text Classification using Naïve Bayes** Build a Naïve Bayes classifier to categorize text documents into predefined classes.
9	**Text Classification using k-NN or Decision Tree** Implement a text classification model using k-NN or Decision Tree and compare its performance with Naïve Bayes.
10	**Sentiment Analysis (Lexicon-based)** Perform lexicon-based sentiment analysis on a set of reviews and classify them as positive, negative, or neutral.
11	**Sentiment Classification (ML-based)** Build a machine learning model for document-level sentiment classification using supervised learning.
12	**Web Usage Mining – User Behavior Analysis** Analyze clickstream or web log data to identify user navigation patterns and frequent access paths.
13	**Social Media Mining & Recommendation** Analyze social media data to study influence and behavior patterns and design a basic recommendation system.

Elective (3+1)

Business Analytics

COURSE CODE:U26DS6E02 **COURSE CREDIT: 03**

1 credit - 15 lectures **1 lecture is 60 minutes**

Course Objectives

This course aims to:

1. To develop an understanding of business analytics concepts and their role in data-driven managerial decision-making through statistical and probabilistic analysis.
2. To analyze business data using statistical techniques, probability, sampling, confidence intervals, and hypothesis testing to address uncertainty in real-world business problems.
3. To apply analytical and business intelligence tools to build and interpret regression, time-series, optimization, and simulation models for forecasting and decision support.

Course Outcomes

By the end of this course, students will be able to:

1. Explain and interpret fundamental concepts of business analytics, probability, and statistical inference in a business decision-making context.
2. Analyze and visualize business data, applying hypothesis testing and confidence interval estimation to draw statistically valid conclusions.
3. Develop and apply regression, time-series forecasting, optimization, and simulation models using business intelligence and Excel-based tools to support managerial decision-making under uncertainty.

Unit	Topics	Lectures
I	Introduction to business analytics, describing the distribution of a variable, finding relationships among variables Business Intelligence Tools for Data Analysis, Probability and Probability Distributions, Decision making under uncertainty.	15

II	Sampling and Sampling Distributions, Confidence Interval Estimation, Hypothesis Testing Regression Analysis: Estimating Relationships, Regression Analysis: Statistical Inference, Time Series Analysis and forecasting	15
III	Optimization modelling: Introduction to Optimization Modelling, A Two-Variable Product Mix Model, Sensitivity Analysis. Simulation modelling: Introduction to Simulation Modelling, Probability Distributions for Input Variables, Simulation and the Flaw of Averages, Simulation with Built-in Excel Tools, Simulation with @RISK, The Effects of Input Distributions on Results.	15

Textbooks and References:

1. Business Analytics Data Analysis and decision making,S.Christian Albright and Wayne L. Winston Cengage, 7th 2020
2. The Applied Business Analytics Casebook: Applications in Supply Chain Management, Operations Management, and Operations Research Matthew Drake O'Reilly 3rd, 2013
3. Applied Business Analytics: Integrating Business Process, Big Data, and Advanced Analytics Nathaniel Lin, FT Press Analytics,1st, 2015

Business Analytics Practical

COURSE CODE: U26DS6EP02 **COURSE CREDIT: 01**

1 credit - 15 lectures **1 lecture is 120 minutes**

Course Objectives:

This course aims to:

- Introduce students to fundamental business analytics concepts and tools, enabling them to collect, explore, visualize, and analyze business data to support data-driven decision-making.

Course Outcomes:

By the end of this course, students will be able to:

- Apply exploratory data analysis, descriptive statistics, correlation and regression techniques, and business intelligence tools to analyze real-world business data and generate actionable insights.

Sr. No.	List of Practicals
1	Introduction to Business Analytics: a. Collect data from a real-life business scenario and perform exploratory data analysis (EDA) to gain insights into the dataset. b. Analyze customer data to identify trends and patterns that can be used for business decision-making.
2	Describing the Distribution of a Variable: a. Obtain a dataset and calculate descriptive statistics (mean, median, mode, variance, etc.) for a specific variable of interest. b. Create visualizations (histograms, box plots) to depict the distribution of a variable and analyze its characteristics

3	Finding Relationships Among Variables: a. Use a dataset with multiple variables and perform correlation analysis to determine the strength and direction of relationships between pairs of variables. b. Apply regression analysis to identify the relationship between an independent variable (e.g., advertising expenditure) and a dependent variable (e.g., sales revenue).
4	Business Intelligence Tools for Data Analysis: Utilize a business intelligence tool (e.g., Tableau, Power BI) to extract insights from a dataset and create interactive visualizations for effective data analysis.
5	Probability and Probability Distributions: a. Simulate a probability experiment (e.g., rolling dice) using programming and calculate the probabilities of different outcomes. b. Generate random numbers from various probability distributions (normal, uniform, exponential) and analyze their properties.
6	Decision Making under Uncertainty: a. Develop a decision tree model to make business decisions considering uncertainties and associated probabilities at each decision point. b. Apply the concept of expected value to evaluate different decision alternatives and select the optimal one
7	Sampling and Sampling Distributions: a. Conduct a survey and collect data from a sample population, ensuring proper sampling techniques are employed. b. Use the Central Limit Theorem to analyze the sampling distribution of a sample mean and estimate population parameters.
8	Confidence Interval Estimation: Calculate confidence intervals for population means or proportions using sample data and interpret the results in a business context.
9	Apply bootstrapping techniques to estimate confidence intervals for nonparametric statistics.
10	Hypothesis Testing: a. Formulate null and alternative hypotheses related to a business problem, conduct a hypothesis test using appropriate statistical tests, and interpret the results

11	Hypothesis Testing: Perform A/B testing on a website or marketing campaign to evaluate the effectiveness of different strategies and make data-driven decisions
12	Regression Analysis and Time Series Analysis:
	Develop a regression model to predict future sales based on historical data, assess model performance, and interpret the significance of predictor variables
13	Apply time series analysis techniques (e.g., ARIMA, exponential smoothing) to forecast future demand for a product or service, and evaluate the accuracy of the forecasts

Minor (3+1)

Quantitative Techniques for Decision Making

COURSE CODE: U26DS6MI01 **COURSE CREDIT: 03**

1 credit - 15 lectures **1 lecture is 60 minutes**

Course Objectives:

This course aims to:

1. Introduce Operations Research as a quantitative tool for decision making.
2. Develop understanding of Linear Programming and its solution methods.
3. Apply optimization models to real-life resource allocation problems.

Course Outcomes

By the end of this course, students will be able to:

1. Understand and formulate basic Operations Research models.
2. Solve Linear Programming problems using appropriate techniques.
3. Apply optimization methods like Goal Programming, Transportation, and Assignment problems

Unit	Topics	Lectures
I	**Exploring Operations Research** – A Quantitative Approach to Decision Making, Definitions, Features, OR Approach to Problem Solving, Models and Modelling in Operations Research, Advantages of Model Building and Operations Research Study, Applications of Operations Research, Computer Software for Operations Research **Formulation of Linear Programming**: Linear Programming: Applications and Model Formulation, Structure of Linear Programming Model, General Structure of an LP Model, Assumptions of an LP Model, Advantages & Limitations Linear Programming, Application Areas, General Mathematical Model of Linear Programming Problem, Examples of LP Model Formulation	15

II	**Linear Programming Problems: Graphical Method:** Maximization & Minimization Type Problems. (Max. Z & Min. Z) ,Two Decision Variables and Maximum Three Constraints Problem ,Constraints can be "less than or equal to", "greater than or equal to" or a combination of both the types i.e. mixed constraints. Concepts: Feasible Region of Solution, Unbounded Solution, Redundant Constraint, Infeasible Solution, Alternative Optima. **Job Sequencing Problem:** Processing Maximum 9 Jobs through Two Machines only. Processing Maximum 6 Jobs through Three Machines only. Calculations of Idle Time, Elapsed Time etc.	15
III	**Transportation Problem:** Introduction, Mathematical Model of Transportation Problem, General Mathematical Model of Transportation Problem, The Transportation Algorithm, Methods for Finding Initial Solution **Assignment Problem:** Introduction, Mathematical Models of Assignment Problem, Solution Methods of Assignment Problem, Hungarian Method for Solving Assignment Problem	15

Textbooks and References:

1. Operations Research: Theory and Applications, J K Sharma, Trinity Press, 6th Edition , 2017
2. Introduction to Operations Research, Frederick S. Hillier, Gerald J. Lieberman, McGraw Hill Education; 11th edition, 2021
3. Operations Research, P K Gupta, S. Chand Publications, 7th Edition, 2018
4. Operations Research, P. Rama Murthy, New Age Publication, 2nd Edition

Quantitative Techniques for Decision Making Practical

COURSE CODE: U26DS6MIP01 **COURSE CREDIT: 01**

1 credit - 15 lectures **1 lecture is 120 minutes**

Course Objectives:

This course aims to:

- To introduce optimization techniques for effective decision making.

Course Outcomes:

By the end of this course, students will be able to:

- Students will be able to solve practical problems using OR models.

Sr. No.	List of Practicals
1	Use Excel Solver / LINGO / Python PuLP to solve a basic optimization problem and identify decision variables, objective function, and constraints.
2	From a given real-life case, formulate a Linear Programming model by defining decision variables, objective function, and constraints.
3	Formulate LP models for two application areas (business/logistics) and write the general mathematical structure.
4	Solve a two-variable LP maximization problem using the Graphical Method and identify feasible region and optimal solution.
5	Solve a two-variable LP minimization problem using the Graphical Method with mixed constraints and comment on infeasible/unbounded cases.
6	Solve an LP problem using the Simplex Method manually and verify the solution using software.
7	Perform sensitivity analysis by changing objective function coefficients and constraint values and study the effect on the optimal solution.
8	Solve a Job Sequencing problem for 2 machines (up to 9 jobs) and compute elapsed time and idle time.

9	Solve a Job Sequencing problem for 3 machines (up to 6 jobs) and calculate total processing time and machine idle time.
10	Solve a Transportation problem using North-West Corner and Least Cost methods to find an initial basic feasible solution.
11	Optimize a Transportation problem using MODI or Stepping Stone method and compare total cost with initial solution.
12	Solve an Assignment problem using the Hungarian Method and interpret the optimal assignments.
13	Solve unbalanced/restricted Assignment problems and develop a small OR mini-project based on LP/Transportation/Assignment model.

On Job Training (Credit 4)

COURSE CODE:U26DS6OJT01 **COURSE CREDIT: 04**

1 credit - 15 lectures **1 lecture is 60 minutes**

Course Objectives:

1. To provide a practical environment where learners can enhance their practical knowledge.
2. To provide the learners the opportunity to integrate theory and practice in his/her professional education

Course Outcome:

Upon successful completion of the On-Job Training (OJT), students will be able to:

1. Apply Theoretical Knowledge to Practical Work – Utilize academic concepts in real-world workplace scenarios to enhance professional competence.
2. Develop Industry-Specific and Technical Skills – Gain hands-on experience, problem-solving abilities, and proficiency in tools and technologies relevant to the field.
3. Enhance Communication and Teamwork Abilities – Demonstrate effective verbal, written, and interpersonal communication while collaborating in a professional environment.
4. Adapt to Workplace Culture and Ethics – Understand organizational behavior, professional responsibility, and ethical standards in the industry.
5. Prepare for Career Growth and Entrepreneurship – Gain exposure to industry trends, job opportunities, and entrepreneurial insights for future career prospects.

Introduction:

- On Job training (OJT) is an important component of the UG program that provides students with a unique opportunity to bridge the gap between theoretical knowledge gained in the classroom and practical application in a real-world environment.
- On Job training (OJT) aims to equip students with both technical and non-technical skills that are essential for success in the industry.

By participating in OJT, students are able to apply the concepts and technologies learned with implementation during their coursework to real-world scenarios. They gain hands-on experience, problem-solving skills, and a deeper understanding of how the industry operates.
From an organizational perspective, hosting OJT programs allows companies to gain insights into the curriculum and content of UG Program. They can provide valuable feedback on the relevance of the coursework and industry requirements, enabling academic institutions to continually improve the program's alignment with industry needs. This collaboration between academia and industry fosters a mutually beneficial relationship,

ensuring that graduates are well-prepared for the job market.

90+30 hours

On-the-Job Training (OJT) for B.Sc. CS, IT, DS, and AI students is structured as 90 + 30 hours to ensure complete professional development. The 90 hours focus on core technical and industry-specific training where students apply their academic knowledge in real work environments. The additional 30 hours are dedicated to holistic skill development, helping students grow beyond technical skills. This includes communication skills, teamwork, time management, ethics, problem-solving, and professional behavior. Together, this structure supports both technical competence and overall personality development, making students industry-ready.

120 hours

The OJT program consists of a total of 120 hours that are entirely completed within the organization. During this period, students focus on gaining continuous hands-on industry experience by working on assigned tasks, projects, and real-time problem-solving, allowing them to develop practical skills and understand professional work environments.

Enhancing practical Skills through OJT:

- The On the Job Training (OJT) program spans 4-6 weeks, requiring a minimum of 120 hours of physical presence at the organization.
- OJT covers any subject within the syllabus, allowing students to align their experience with their academic interests.
- OJT bridges the gap between theoretical knowledge and practical application, preparing students for successful careers in their respective area of interest.

Interning organization: Students have the flexibility to pursue their OJT in various types of organizations, including but not limited to:

- Hardware/Manufacturing Firms: Learn about hardware design, manufacturing processes, and quality assurance.
- Civic Departments: Engage with local civic departments such as ward offices, post offices, police stations, or panchayats to understand their functioning and contribute to their activities.
- Research Centre's/University Departments/Colleges: Contribute as research assistants or in similar roles for research projects or initiatives, fostering collaboration between academia and industry.
- Small-Scale Industries/Service Providers: Explore opportunities in diverse sectors such as banking, clinics, NGOs, and professional institutions like CA firms or law firms
- Industries/Finance and Insurance sector/Management and Media Development Firms: Gain practical experience and to understand their functioning and contribute to their activities.
- Any other organization where technologies can be implemented.

OJT mentors: To enhance the learning experience and ensure the quality of the program,

each student participating in the OJT will be assigned two mentors: a faculty mentor from the institution and an industry mentor from the organization where the student is interning.

Industry Mentor Role: The industry mentor plays a crucial role in guiding the student during the internship. They ensure that the internee fulfills the requirements of the organization and successfully meets the demands of the assigned project. Through their expertise and experience, industry mentors provide valuable insights into real-world practices and industry expectations.

Faculty Mentor Role: The faculty mentor serves as the overall coordinator of the OJT program. They oversee the entire internship process and evaluate the quality of the OJT in a consistent manner across all students. The faculty mentor ensures that the OJT aligns with the program's objectives and provides valuable learning opportunities. They also facilitate communication between the institution, industry mentor, and student to ensure a fruitful OJT experience. By having both an industry mentor and a faculty mentor, students benefit from a comprehensive guidance system that combines industry expertise and academic support

Guidelines for Submission of documentation for OJT

The student will make two documents as part of the OJT:

- Diary (Log Book): This ensures that the student updates daily activity, which could be accessed by both the mentors. Weekly entry can be of 3-4 sentences giving a very brief account of the learning/activities/interaction taken place.
- OJT report: A student is expected to make a report based on the OJT he or she has done in an organization. It should contain the following:

  OJT Report should contain the following:

  1. **Certificate**: A certificate from the organization where the OJT was done.

  2. **Title**: A suitable title giving the idea about what work the student has performed during the OJT.

  3.**Abstract**

  This should be one/two short paragraphs (100-150 words total), summarizing the project work. It is important that this is not just a restatement of the original project outline. A suggested flow is background, project aims and main achievements. From the abstract, a reader should be able to ascertain if the project is of interest to them and, it should present results of which they may wish to know more details.

  4.**Acknowledgements**

  This should express student's gratitude to those who have helped in the preparation of the project. Table of Contents: The table of contents gives the readers a view of the detailed structure of the report. The students would need to provide section and subsection headings with associated pages.

  Chapter 1: Introduction

  The introduction has several parts as given below:

  Background: A description of the background and context of the project and its relation to work already done in the area. Summarise existing work in the area

concerned with the project work.

Objectives: Concise statement of the aims and objectives of the project. Define exactly what is going to be done in the project; the objectives should be about 30 /40 words.

Purpose, Scope and Applicability: The description of Purpose, Scope, and Applicability are given below:

• Purpose: Description of the topic of the project that answers questions on why this project is being done. How the project could improve the system its significance and theoretical framework.

• Scope: A brief overview of the methodology, assumptions and limitations. The students should answer the question: What are the main issues being covered in the project? What are the main functions of the project?

• Applicability: The student should explain the direct and indirect applications of their work. Briefly discuss how this project will serve the computer world and people.

Achievements: Explain what knowledge the student achieved after the completion of the work. What contributions has the project made to the chosen area? Goals achieved - describes the degree to which the findings support the original objectives laid out by the project. The goals may be partially or fully achieved, or exceeded.

Organisation of Report: Summarising the remaining chapters of the project report, in effect, giving the reader an overview of what is to come in the project report.

Chapter 2: Survey of Technologies

In this chapter Survey of Technologies should demonstrate the students awareness and understanding of Available Technologies related to the topic of the project. The student should give the details of all the related technologies that are necessary to complete the project. The should describe the technologies available in the chosen area and present a comparative study of all those

Available Technologies. Explain why the student selected the one technology for the completion of the objectives of the project.

Chapter 3: Requirements and Analysis

Problem Definition: Define the problem on which the students are working in the project. Provide details of the overall problem and then divide the problem into sub-problems. Define each sub-problem clearly.

Requirements Specification: In this phase the student should define the requirements of the system, independent of how these requirements will be accomplished. The Requirements Specification describes the things in the system and the actions that can be done on these things. Identify the operation and problems of the existing system.

Planning and Scheduling: Planning and scheduling is a complicated part of software development. Planning, for our purposes, can be thought of as determining all the small tasks that must be carried out in order to accomplish the goal. Planning also takes into account rules, known as constraints, which control when certain tasks can or cannot happen. Scheduling can be thought of as determining whether adequate resources are available to carry out the plan. The student should show the Gantt chart

and Program Evaluation Review Technique (PERT).

Software and Hardware Requirements: Define the details of all the software and hardware needed for the development and implementation of the project.

• Hardware Requirement: In this section, the equipment, graphics card, numeric co-processor, mouse, disk capacity, RAM capacity etc. necessary to run the software must be noted.

• Software Requirements: In this section, the operating system, the compiler, testing tools, linker, and the libraries etc. necessary to compile, link and install the software must be listed.

Preliminary Product Description: Identify the requirements and objectives of the new system. Define the functions and operation of the application/system the students are developing as project. Conceptual Models: The student should understand the problem domain and produce a model of the system, which describes operations that can be performed on the system, and the allowable sequences of those operations. Conceptual Models could consist of complete Data Flow Diagrams,ER diagrams, Object-oriented diagrams, System Flowcharts etc.

Chapter 4: System Design

Describes desired features and operations in detail, including screen layouts, business rules, process diagrams, pseudocode and other documentation.

Basic Modules: The students should follow the divide and conquer theory, so divide the overall problem into more manageable parts and develop each part or module separately. When all modules are ready, the student should integrate all the modules into one system. In this phase, the student should briefly describe all the modules and the functionality of these modules.

Data Design: Data design will consist of how data is organised, managed and manipulated.

• Schema Design: Define the structure and explanation of schemas used in the project.

• Data Integrity and Constraints: Define and explain all the validity checks and constraints provided to maintain data integrity.

Procedural Design: Procedural design is a systematic way for developing algorithms or procedurals.

• Logic Diagrams: Define the systematic flow of procedure that improves its comprehension and helps the programmer during implementation. e.g., Control Flow Chart, Process Diagrams etc.

• Data Structures: Create and define the data structure used in procedures.

• Algorithms Design: With proper explanations of input data, output data, logic of processes, design and explain the working of algorithms.

User Interface Design: Define user, task, environment analysis and how to map those requirements in order to develop a "User Interface". Describe the external and internal components and the architecture of user interface. Show some rough pictorial views of the user interface and its components.

Security Issues: Discuss Real-time considerations and Security issues related to the project and explain how the student intends avoiding those security problems. What

are the security policy plans and architecture?
Test Cases Design: Define test cases, which will provide easy detection of errors and mistakes within a minimum period of time and with the least effort. Explain the different conditions in which the students wish to ensure the correct working of the project.

Chapter 5: Implementation and Testing

Implementation Approaches: Define the plan of implementation, and the standards the students have used in the implementation.
Coding Details and Code Efficiency: Students do not need to include full source code, instead, include only the important codes (algorithms, applets code, forms code etc). The program code should contain comments needed for explaining the work a piece of code does. Comments may be needed to explain why it does it, or why it does a particular way.
The student can explain the function of the code with a shot of the output screen of that program code.
• Code Efficiency: The student should explain how the code is efficient and how the students have handled code optimisation.
Testing Approach: Testing should be according to the scheme presented in the system design chapter and should follow some suitable model – e.g., category partition, state machine-based. Both functional testing and user-acceptance testing are appropriate. Explain the approach of testing.
• Unit Testing: Unit testing deals with testing a unit or module as a whole. This would test the interaction of many functions but, do confine the test within one module.
• Integrated Testing: Brings all the modules together into a special testing environment, then checks for errors, bugs and interoperability. It deals with tests for the entire application. Application limits and features are tested here.
Modifications and Improvements: Once the students finish the testing they are bound to be faced with bugs, errors and they will need to modify your source code to improve the system. Define what modification are implemented in the system and how it improved the system.

Chapter 6: Results and Discussion

Test Reports: Explain the test results and reports based on the test cases, which should show that the project is capable of facing any problematic situation and that it works fine in different conditions. Take the different sample inputs and show the outputs.
User Documentation: Define the working of the software; explain its different functions, components with screenshots. The user document should provide all the details of the product in such a way that any user reading the manual, is able to understand the working and functionality of the document.

Chapter 7: Conclusions

Conclusion: The conclusions can be summarised in a fairly short chapter (2 or 3 pages). This chapter brings together many of the points that would have been made in

the other chapters.

Limitations of the System: Explain the limitations encountered during the testing of the project that the students were not able to modify. List the criticisms accepted during the demonstrations of the project.

Future Scope of the Project describes two things: firstly, new areas of investigation prompted by developments in this project, and secondly, parts of the current work that was not completed due to time constraints and/or problems encountered.

Interaction between mentors: To ensure the smooth conduct of the OJT a meet-up involving the intern, industry mentor, and the faculty mentor will be scheduled as a review. The meeting ensures the synergy between all stakeholders of the OJT.

OJT workload for the faculty: Every student is provided with a faculty member as a mentor. So, a faculty mentor will have a few students under him/her. A faculty mentor is the overall in charge of the OJT of the student. He/she constantly monitors the progress of the OJT by regularly overseeing the diary, interacting with the industry mentor, and guiding on the report writing etc.

Appendix-II

(Proforma for the certificate for internship in official letter head)

This is to certify that Mr./Ms..... from..... College has worked as an intern towards the partial fulfilment of _____degree in the academic year ____and has not been submitted for any other examination and does not form part of any other course undergone by the candidate.

The particulars of internship are given below:

Internship starting date: _____

Internship ending date: _____

Actual number of days worked: _____

Tentative number of hours worked: _____ Hours

Broad area of work: _____

A small description of work done by the intern during the period:

Signature:

(Seal of the organization)

Designation:

Contact details:

Email:

Evaluation Pattern of On Job Training

90+30 Hours On Job Training (OJT)

Continuous Internal Evaluation (CIE)	25 Marks
External Theory Exam (SEE)	25 Marks
Assessment (Annexure A)	50 Marks
Total	100 Marks

Continuous Internal Evaluation (CIE)-25 marks

Class Test	10 Marks
Class Behaviour and Class Participation	05 Marks
Class Attendance	05 Marks
Attending Training and Workshop	05 Marks
Total	25 Marks

120 Hours On Job Training (OJT)

External Assessment (Annexure B)	50 Marks
Assessment (Annexure A)	50 Marks
Total	100 Marks

Revised Scheme of Examination

Faculty of Science

(Undergraduate Programme)

SCHEME OF EXAMINATION (for 100 marks 3 credits)

The scheme of examination shall be divided into two parts:

- **Internal assessment 40% i.e. 40 marks**
- **Semester end examination 60% i.e. 60 marks**

(A) Internal Assessment 40 marks

Description	Marks
An internal test of 20 marks Q.1 a. Multiple choice Questions - 05 Marks b. True/False - 05 Marks Q.2. Attempt 2 questions out of 3 questions (5 marks each) - 10 Marks **OR** Online MCQ Test	20
Presentation/Case Studies/Assignments	15
Attendance and Class Participation	5
Total	40

B) Semester End examination 60 marks

PAPER PATTERN

Duration: 2 hours	
Total Marks: 60	
Q.1 15 marks OR 15 marks (7 and 8 marks)-Unit 1	15
Q.2 15 marks OR 15 marks (7 and 8 marks)-Unit 2	15
Q.3 15 marks OR 15 marks (7 and 8 marks)-Unit 3	15
Q.4 15 marks-attempt any 3 out of 6 (from Unit 1, Unit 2, Unit 3)	15
Total	60

Passing criteria: Minimum 40% in Internal (16 out of 40) and 40% (24 out of 60) in semester-end examination.

SCHEME OF EXAMINATION (for 50 marks, 2 credits)

The scheme of examination shall be divided into two parts:

- Internal assessment 40% i.e.20 marks
- Semester end examination 60% i.e.30 marks

(A) Internal Assessment 20 marks (Theory)

Description	**Marks**
An internal test of 10 marks Q.1 Multiple choice Questions/True or False - 05 Marks Q.2. Attempt 1 questions out of 2 questions (5 marks each) - 05 Marks **OR** Online MCQ Test	10
Presentation/Case Studies/Assignments	05
Attendance and Class Participation	05
Total	20

OR

(A) Internal Assessment 20 marks (Practical)

Description	**Marks**
Practical Question	10
Journal	05
Viva	05
Total	20

B) Semester End examination 30 marks

PAPER PATTERN

Duration: 1 hour	
Total Marks:30	
Q.1 15 marks OR 15 marks (7 and 8 marks)-Unit 1	15
Q.2 15 marks OR 15 marks (7 and 8 marks)-Unit 2	15
Total	30

OR

PAPER PATTERN (1 credit Theory)

Duration: 1 hour	
Total Marks:30	
Q.1 15 marks OR 15 marks (7 and 8 marks)-Unit 1	15
Q.2 15 marks OR 15 marks (7 and 8 marks)-Unit 1	15
Total	30

Passing criteria: Minimum 40% in Internal (08 out of 20) and 40% (12 out of 30) in semester-end examination.

SCHEME OF PRACTICAL EXAMINATION

(for 50 marks, 1 credit)

Description	Marks
Practical Question 1	20
Practical Question 2	20
Viva	05
Journal	05
Total	50

Passing Standards

The learners to pass a course shall have to obtain a minimum of 40% marks in each head of passing, consisting of Internal Assessment and Semester End Examination. The learners shall obtain a minimum of 40% marks (i.e. 16 out of 40 or 8 Out of 20) in the Internal Assessment and 40% marks in the Semester End Examination (i.e. 24 Out of 60 or 12 Out of 30) separately, to pass the course and a minimum of Grade D, wherever applicable, to pass a particular semester. A learner will be said to have passed the course if the learner passes the Internal Assessment and Semester End Examination together.