



Faculty: Science
Program: B.Sc.
Subject: Data Science
Academic year: 2024-2025

Syllabus for Semester- III and Semester – IV

**Choice Based Credit System Syllabi (as per NEP)
approved by Board of Studies in Data Science to
be brought into effect from June 2024.**

PREAMBLE

Data has become the most important factor in this era of digital transformation. The technological advancements are seen in all walks of life and therefore we are flooded with enormous data. Every business relies on data to deliver better products as well as services. All data are stored in cloud, and so accessed and processed easily. Data analytics has helped in better decision making with sufficient data insights.

Predictive Analysis has played a crucial role in making businesses smarter with improvised strategies. Machine Learning and Artificial Intelligence are used together to optimize business operations and data management. Augmented analytics uses machine learning and natural language processing to automate the process of data analysis. Global data is predicted to grow due to data generated by the Internet of Things(IoT) and cloud computing advancements. These developments have given rise to a new area of study, called Data Science.

Data Science as an area has evolved out of the applications of various tools and techniques in the field of Computer Science, Mathematics and Statistics. There is an increasing demand to capture, analyse the enormous data present in a number of application domains. The data in these applications then needs to be converted into actionable strategies for effective decision making. So, the study of data science has become essential to cater to the growing need of data scientists and data analysts.

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

PROGRAM OUTCOMES AND PROGRAM SPECIFIC OUTCOME

NO.	Details
PO1.	Solving Complex Problems:- Apply the knowledge gained in breaking down complex problems into simple components; and to design processes required for problem solving.
PO2.	Critical Thinking: - Ability to apply the acquired knowledge to identify assumptions and evaluate their accuracy and validity.
PO3.	Reasoning ability and Rational thinking: - Ability to analyse, interpret data and draw logical conclusions; to evaluate ideas rationally.
PO4.	Research Aptitude: - Ability to ask relevant questions to identify and define the problem, applying research tools for analysis and interpretation of data. Understand and comply with research ethics.
PO5.	Effective Communication skill: - Demonstrate the ability to listen and to clearly express ideas verbally. Equip to write reports, make presentations effectively.
PO6.	Information and Digital Literacy: - Equip to use appropriate tools and techniques inclusive of internet and electronic media for acquiring, assessing and analysing data from diverse resources.
PO7.	Social Interactive Skills and team work: - Exhibit networking and social interactive skills; function effectively as an individual and as a member in diverse groups; demonstrate leadership quality useful for employability
PO8.	Self-directed and Lifelong Learning: Ability to explore and gain knowledge in independent and self-reliant ways. Demonstrate ability to adapt and upgrade with the global , social and technological changes.
PO9.	Awareness towards Environment and Sustainable Development: Exhibit awareness and a concern for environmental issues; understand and realize the significance of co-habitation and co-evolution in attaining the needs of sustainable development.
PSO1.	Sound Knowledge: Demonstrate the knowledge of core data science concepts and apply them to develop a user- friendly, scalable, and robust applications
PSO2.	Critical and Rational Thinking: Exhibit higher order skills to adapt to the everchanging technological environment
PSO3.	Logic Building and Programming Skills: The ability to apply logic to problem solving and acquire proficiency in various programming languages.
PSO4.	Data Analysis : Apply quantitative modelling and data analysis techniques to solve real world business problems, Learn tools and techniques for transformation of data and statistical data analysis
PSO5.	Pursue Higher Education: Make students competent to take up advanced degree courses like MSc(Data Science),MCA, MSc(CS), MSc(IT) and MBA etc.

S.Y.B.Sc.(DS) under NEP - Semester III

Course Code	Course Type	Course Title	Credits
SIUDSMJ211	Core Subject (Major)	Data Science with R	3
SIUDSMJ212	Core Subject (Major)	Probability and Distributions	3
SIUDSMN211	Core Subject (Minor)	Database Technology – I	3
SIUDSVS211	Vocational Skill Course (VSC)	Introduction to Kafka	1
SIUDSOE211	Open Elective	Digital Marketing – I	1
SIUHNAE211	Ability Enhancement Course (AEC)	Hindi Bhasha evam Prayojanmulak Hindi	2
SIUDSFP211	Field Project (FP)		2
SIUEXCC211	Co-Curricular (CC)		2
SIUDSMJP211	Core Subject (Major) Practical	Data Science with R – Practical	1
SIUDSMJP212	Core Subject (Major) Practical	Probability and Distributions – Practical	1
SIUDSMNP211	Core Subject (Minor) Practical	Database Technology – I – Practical	1
SIUDSVS211	Vocational Skill Course Practical (VSC)	Introduction to Kafka – Practical	1
SIUDSOE211	Open Elective Practical	Digital Marketing – I Practical	1
TOTAL CREDITS			22

SIUDSMJ211 : Data Science with R

B.Sc. (Data Science)	Semester – III
Course Name: Data Science with R	Course Code: SIUDSMJ211
Credits	3

Course Objective:

- Master the use of the R interactive environment and expanding by installing R packages
- Read Structured Data into R from various sources
- Understand the different data types and data structures in R
- Manipulate strings, dates in R
- Understand basic regular expressions in R
- Understand base R graphics
- Focus on GGplot2 graphics for R and be familiar with trellis (lattice) graphics.

Course Outcomes:

After completion of this course, student will be able to:

- **CO1:** To use R Studio and explore the features for R programming
- **CO2:** To use R functions and graphics with in R programming for solving problems.
- **CO3:** To work with advanced graphics of R, import and use the data and represent the data into tables.
- **CO4:** To manipulating Data Frames and make use of Dates in R application.

Unit	Contents	No. of Lectures
I	<p>Getting started with R: R Software: Obtaining R and RStudio, First R Encounter, Getting started: R as a big calculator, Assignment, Basic operators, Help with functions and features</p> <p>R Interfaces: Using R and RStudio: R Software, Obtaining R and RStudio, The default R interface, RStudio Interface, Example Datasets in R, R Packages, Installing new R libraries, Customizing R Start-up</p> <p>Objects in R: Using ls and rm to managing R Objects, Types of R objects, Attributes of R Objects, Creating and accessing objects, Modifying elements.</p> <p>Reading and writing data to and from R: Importing and reading text files data into RStudio, Importing data using R command read.table(), Importing text files Using scan(), Parsing each line – Readlines, Writing Data table from R, Importing Data from other Software, Reading data from Excel into R, Import/Export from other statistical software, From a Database Connection, Sampling</p>	15

	and Creating simulated data. Introduction to programming and writing Functions in R: Functions, Conditional statements (if, ifelse, switch), Repetitive execution: For and While loops	
II	R Apply Package: The Apply Functions, Functions for parsing text, Programming in R: More advanced, Viewing Code of functions from R packages, Parsing Real Data - World Population Data from Wikipedia, Writing functions: more technical discussion -Scoping, Options for Running memory or CPU intensive jobs in R, Efficient R coding Importing Data – readr: Functions for Reading Data, File Headers, Column Types, String-based Column Type Specification, Function based Column Type Specification Parsing Time and Dates, Space-separated Columns, Functions for Writing Data Introduction to graphics in R: The R function plot(), Customize plot with low-level plotting commands, Default parameters – par, Interacting with graphics, Saving plots, Useful Graphics Resources Advanced Graphics: Advanced plotting using Trellis; ggplots2, Lattice, Examples that Present Panels of Scatterplots using xyplot(), Simple use of xyplot	15
III	Reformatting Tables – tidyr: Tidy Data, Gather and Spread, Complex Column Encodings, Expanding, Crossing, and Completing, Missing Values, Nesting Data Pipelines – magrittr: The Problem with Pipelines, Pipeline Notation, Pipelines and Function Arguments, Function Composition, Other Pipe Operations Working with Strings – stringr: Counting String Patterns, Splitting Strings, Capitalizing Strings, Wrapping, Padding, and Trimming, Detecting Substrings, Extracting Substrings, Transforming Strings Working with Factors – forcats: Creating Factors, Concatenation, Projection, Adding Levels, Reorder Levels Manipulating Data Frames – dplyr: Selecting Columns, Filter, Sorting, Modifying Data Frames, Grouping and Summarizing, Joining Tables. Working with Dates – lubridate: Time Points, Time Zones, Time Intervals	15

Books and Reference

Sr. No.	Title	Author/s	Publisher	Edition	Year
1	R Programming for Data Science	Roger D Peng		1 st	2015
2	Data Science from Scratch	Joel Grus	O'Reilly Media, Inc.	2 nd	2019
3	An Introduction to Statistical	Gareth James, Daniela	Springer Science &	Illustrated	2013

	Learning	Witten, Trevor Hastie, Robert <u>Tibshirani</u>	Business Media, 2013		
4	Practical Statistics for Data Scientists	Peter Bruce, Andrew Bruce	O'Reill y Media, Inc.	3 rd	2018

SIUDSMJP211 : Data Science with R Practical

B.Sc. (Data Science)	Semester – III
Course Name: Data Science with R Practical	Course Code: SIUDSMJP211
Credits	1

List of Practical:

1	Introduction to R Programming Elements
a.	Write an R Program to implement expressions, assignment and decision making
b.	Write an R Program to design and implement loops.
c.	Write a R program to demonstrate the use of essential data structures in R [Hint: Vectors, Matrix, Arrays]
2	Using List, Data Frames and Functions in R
a.	Write an R program to manage data and exhibit operations on it using List data structure
b.	Write an R program to manage data and exhibit operations on it using Data Frames
c.	Write an R program to demonstrate the use of: User-defined functions Built-in numeric function, character functions etc.
3	Implement decision making in R
a.	Write an R program to implement if, if-else, if-else-if ladder, nested if-else, and switch
4	Implementing Strings in R
a.	Write an R program to store and access string in R objects(vectors, matrix, arrays, data frames, and lists)
b.	Write an R program to demonstrate use of various string manipulation functions. [Hint: paste(), print(), noquote(),format(), cat(), toString(), sprintf()]
5	Performing Statistics with R-I
a.	Write an R program to apply built-in statistical functions. [Hint: mean, median, standard deviation and others]
b.	Write an R program to demonstrate Linear and Multiple Regression analysis.
6	Performing Statistics with R-II
a.	Write an R program to implement Normal Distribution. [Hint: dnorm(), pnorm(), qnorm(), rnorm()] Binomial Distribution: [Hint: dbinom(), pbinom(),qbinom(),rbinom()]
7	Data Visualization and Analysis
a.	Write an R program to demonstrate various ways of performing Graphical analysis.
8	Data Interfaces in R
a.	Write an R program to demonstrate data interface with CSV files [Hint: creating data for CSV, analyzing, writing CSV files]

b.	Write an R program to work with spreadsheet (Excel) programs. [Hint: installing, loading, verifying, creating data for xlsx file]
c.	Write an R program to manage data using XML files. [Develop data interface for maintaining Employee Information]
9	Handling Errors in R
a.	Write an R program to demonstrate various error messages in R Programming
b.	Write an R program to implement Error Handling in R [Hint: warning(),stop(),try(), tryCatch(), CallingHandlers()]
10	Handling Dates in R
a.	Write R program to demonstrate dates in R.

SIUDSMJ212 : Probability and Distributions

B.Sc. (Data Science)	Semester – III
Course Name: Probability and Distributions	Course Code: SIUDSMJ212
Credits	3

Course Objective:

- To explore about random variables and implement various distribution functions
- To familiarize with concepts of probability and learn implementation of different types of probabilities.
- Learn and implement the concept of expectation, related theorems and generating functions.
- To know the concept and implementation of discrete distributions including Bernoulli, Binomial and power series distributions.

Course Outcomes:

After completion of this course, student will be able to:

- **CO1:** Organize, manage and present data.
- **CO2:** Analyse statistical data graphically using frequency distributions and cumulative frequency distributions.
- **CO3:** Use the basic probability rules, including additive and multiplicative laws, using the terms independent and mutually exclusive events.
- **CO4:** Derive the probability density function of the transformation of random variables.
- **CO5:** Calculate probabilities and derive the marginal and conditional distributions of bivariate random variables.

Unit	Contents	No. of Lectures
I	<p>Theory of Probability: Introduction, history, different terms, mathematical tools, Axiomatic approach to probability, Mathematical notation, multiplication and conditional probability, Bayes theorem, Geometric probability.</p> <p>Random Variables and Distribution Functions: Random Variable, distribution function, discrete random variable, continuous random variable, joint probability law, transformation of one-dimensional random variable, transformation of two-dimensional random variable</p>	15
II	<p>Mathematical Expectation and Generating Functions: Mathematical expectation, Expectation of a Function of a Random Variable, Addition Theorem of Expectation, Multiplication Theorem of Expectation, Expectation of a Linear Combination of Random Variables, Covariance, Variance of a Linear Combination of Random Variables, Moments of Bivariate Probability Distributions, Conditional Expectation, and Conditional Variance,</p>	15

	Moment Generating Function, Cumulants, Characteristic Function, Chebychev's Inequality, Convergence in- Probability, Weak Law of Large Numbers, Borel Canteli Lemma, Probability Generating Function.	
III	<p>Theoretical Discrete Distributions: Introduction, Bernoulli distribution, Binomial Distribution, Poisson Distribution, Negative Binomial Distribution, Geometric Distribution, Hypergeometric Distribution, Multinomial Distribution, Discrete Uniform Distribution, Power Series Distribution</p> <p>Theoretical Continuous Distributions: Rectangular or Uniform Distribution, Normal Distribution, Gamma Distribution, Beta Distribution of First Kind, Beta Distribution of Second Kind, The Exponential Distribution, Laplace Double Exponential Distribution, Weibul Distribution, Cauchy Distribution, Central Limit Theorem, Compound Distributions, Pearson's Distributions, Variate Transformations, Order Statistics, Truncated Distributions.</p>	15

Books and References

Sr. No.	Title	Author/s	Publisher	Edition	Year
1	Fundamentals of Mathematical Statistics	S.C, Gupta, and V. K. Kapoor	S. Chand and Sons	10 th	2002
2	Applied Statistics and Probability for Engineers	Douglas C. Montgomery and George C. Runger	Wiley	6 th	2014
3	Probability, Statistics, and Stochastic Processes	Peter Olofsson And Mikael Andersson	Wiley	2 nd	2012

SIUDSMJP212 : Probability and Distributions Practical

B.Sc. (Data Science)	Semester – III
Course Name: Probability and Distributions Practical	Course Code: SIUDSMJP212
Credits	1

List of Practical: (Can be done in MS-Excel)

1	Introduction to Probability
a.	Formulate and apply Bayes' Theorem Calculations for problems like the "Two Supplier Example". [Hint: Use Prior Probabilities and Conditional Probabilities to compute Joint and posterior probabilities.]
b.	Design a spreadsheet to demonstrate the association Between Two Variables by Computing the Covariance and Correlation Coefficient. [Hint: Use COVAR and CORREL)]
2	Discover Probability using formulas
a.	Design and spreadsheet experiment to compute the probability using the geometric distribution formula.
b.	Create a spreadsheet application to compute the Conditional Probability. Also determine the probability that a randomly chosen event.
3	Random Variables and Distribution Functions
a.	Create a spreadsheet application to Compute the Expected Value, Variance, and Standard Deviation
b.	Create a spreadsheet application to Compute Binomial Probabilities. [Hint: Use BINOM DIST]
4	Probability Distribution and Law
a.	Create a spreadsheet application to Poisson Probability Distribution. [Hint: Use POISSON]
b.	Create a spreadsheet application to implement joint probability law.
5	Mathematical Expectation and Chebychev's Theorem
a.	Create a spreadsheet application to compute the expectation of a Function of a Random Variable
b.	Create a spreadsheet application to apply Chebychev's Theorem
6	Conditional Expectation and Generating Functions
a.	Create a spreadsheet application to compute Conditional Expectation and Conditional Variance.
b.	Create a spreadsheet application to demonstrate the use of Generating Functions
7	Theoretical Discrete Distributions1
a.	Create a spreadsheet application to demonstrate Bernoulli Distribution.
b.	Create spreadsheet application to use excel function for computing hypergeometric probabilities.
8	Theoretical Discrete Distributions2

a.	Create spreadsheet application to Calculate Binomial Distribution in Excel. [Hint: Use BINOM.DIST]
b.	Create suitable spreadsheet application to work with PowerSeries Distribution.
9	Theoretical Continuous Distributions1
a.	Create spreadsheet application for computing probabilities and zvalues for the standard normal distribution. [Hint: Use NORMSDIST and NORMSINV]
b.	Create spreadsheet application for computing probabilities for the exponential probability distribution. [Hint: Use EXPONDIST]
10	Theoretical Continuous Distributions2
a.	Create spreadsheet application for demonstrating Weibull Distribution to obtain a model for data sets. [Hint: Use WEIBULL.DIST]
b.	Create spreadsheet application for demonstrating Pearson's Distributions.

SIUDSMN211: Database Technology-I

B.Sc. (Data Science)	Semester – III
Course Name: Database Technology-I	Course Code: SIUDSMN211
Credits	3

Course Objective:

- This course concentrates the concept of the DBMS with respect To principles, design and implementation of DBMS. It aims to specify the functional and Data requirements for a typical database application and to understand creation, manipulation And querying of data in databases.
- To understand Organizing, structuring and storing data
- Understand Database as Relational model
- To understand SQL to retrieve data and concept of redundancy
- To specify the functional and data requirements for a typical database application
- To understand creation, manipulation and querying of data in databases

Course Outcomes:

After completion of this course, student will be able to:

- **CO1:** Students should be able to evaluate business information problem and find the requirements of a problem in terms of data.
- **CO2:** Students should be able to draw database design in logical structure and can identify the entities which exist in a system
- **CO3:** Students should be able to construct normalized database and functional dependencies between attributes and relational algebra queries
- **CO4:** Students should be able to design the database schema with the use of appropriate datatypes for storage of data in database.
- **CO5:** Students should be able to create, manipulate, query and backup the databases with features of SQL.

Unit	Contents	No. of Lectures
I	Introduction to Databases and Transactions What is database system, purpose of database system, view of data, relational databases, database architecture, transaction management Data Models The importance of data models, Basic building blocks, Business rules, The evolution of data models, Degrees of data abstraction. Database Design and ER Diagram Database design and ER Model: overview, ER Model, Constraints, ER Diagrams, ERD Issues, weak entity sets, Codd's rules, Relational Schemas	15

	<p>Relational database model: Logical view of data, keys, integrity rules, Relational Database design: features of good relational database design, atomic domain and Normalization (1NF, 2NF, 3NF, BCNF).</p> <p>Relational Algebra</p>	
II	<p>Structured Query Language(SQL):Introduction to SQL, Basic SELECT Queries, SELECT Statement Options, FROM Clause Options, ORDERBY Clause Options, WHERE Clause Options, Aggregate Processing, Subqueries, SQL Functions, Relational Set Operators, Crafting SELECT Queries</p>	15
III	<p>Constraints, Views and SQL Constraints, types of constrains, Integrity constraints, Views: Introduction to views, data independence, security, updates on views, comparison between tables and views SQL: data definition, aggregate function, Null Values, nested sub queries, Joined relations. Triggers</p> <p>Transaction management and Concurrency Control Transaction management: ACID properties, serializability and concurrency control, Lock based concurrency control (2PL, Deadlocks), Time stamping methods, optimistic methods, database recovery management.</p>	15

Books and References

Sr. No.	Title	Author/s	Publisher	Edition	Year
1	Database System and Concepts	A Silberschatz, H Korth, SSudarshan	Mc Graw Hill	5 th	
2	Database Systems	Rob Coronel	Cengae Learning	12 th	
3	Introduction to Database System	C.J.Date	Pearson	1 st	2003

SIUDSMNP211 : Database Technology-I Practical

B.Sc. (Data Science)	Semester – III
Course Name: Database Technology-I Practical	Course Code: SIUDSMNP211
Credits	1

List of Practical:

1	Create ER-diagram for Hospital and library system.
2	Draw E-R diagram and convert entities it to table.
3	Write SQL query for given problem statement
a.	Creating a Database
b.	Viewing all databases
c.	Viewing all Tables in a Database
4	Perform the following:
a.	Create table
b.	Drop table
c.	Truncate
5	Manipulating Data
a.	Using INSERT statement
b.	Using DELETE statement
c.	Using UPDATE statement
6	Creating and Managing Tables
a.	Creating and Managing Tables
b.	Including Constraints
7	Perform following:
a.	Simple Queries with Where Operators
b.	Where with Keywords and Logical Operators
c.	Queries with Aggregate functions (group by and having clause)
8	Retrieving Data from Multiple Table:
a.	Joining Tables (Inner-Joins, Outer-Joins)
b.	Aliases for Table Names
9	Subqueries:
a.	With IN clause

b.	With EXISTS clause														
10	I r E E	<table border="0"> <tr> <td data-bbox="331 188 464 215">1. Pay_scale</td> <td data-bbox="564 188 647 215">Number</td> <td data-bbox="820 188 1038 215">Pay scale of employee</td> </tr> <tr> <td data-bbox="331 224 464 250">2. Basic_pay</td> <td data-bbox="564 224 647 250">Number</td> <td data-bbox="820 224 1023 250">Employee's basic pay</td> </tr> <tr> <td data-bbox="331 259 448 286">3. Net Pay</td> <td data-bbox="564 259 647 286">Number</td> <td data-bbox="820 259 1034 286">Employee's deduction</td> </tr> <tr> <td data-bbox="331 295 464 322">4. Gross Pay</td> <td data-bbox="564 295 647 322">Number</td> <td data-bbox="820 295 1023 322">Employee's gross pay</td> </tr> </table> <ol style="list-style-type: none"> 1. Set emp_Id as primary key 2. Set the validation rule for basic_pay as (>8000) and the validation text as "Basic pay shouldbe greater than 8000" 3. Set the caption value of the field "Emp_Name" as "Employee Name" 4. Set the field "pay_scale" as a required(mandatory) fieldv) 5. Enter 10 records 	1. Pay_scale	Number	Pay scale of employee	2. Basic_pay	Number	Employee's basic pay	3. Net Pay	Number	Employee's deduction	4. Gross Pay	Number	Employee's gross pay	base and where each
1. Pay_scale	Number	Pay scale of employee													
2. Basic_pay	Number	Employee's basic pay													
3. Net Pay	Number	Employee's deduction													
4. Gross Pay	Number	Employee's gross pay													

SIUDSVS211: Introduction to Kafka

B.Sc. (Data Science)	Semester – III
Course Name: Introduction to Kafka	Course Code: SIUDSVS211
Credits	1

Course Objective:

- Kafa Introduction.
- A Deep Dive into Kafka Architecture.
- Understanding Brokers.
- Understanding Producers.
- Understanding Consumers.

Course Outcomes:

After completion of this course, student will be able to:

- **CO1:** To understand Kafka Architecture.
- **CO2:** To Understanding Brokers.
- **CO3:** To Understanding Producers.
- **CO4:** To Understanding Consumers.
-

Unit	Contents	No. of Lectures
I	<p>Introduction: Kafka origin , Benefits, Use Cases, Messaging System.</p> <p>Fundamentals of Kafka: Topics and Partitions, Producers and Consumers, Brokers and Clusters.</p> <p>Kafka CLI : Creating Kafka Topics, Sending data to Kafka , Kafka Console Consumer, Kafka Consumer Group , Kafka Brokers, Topic Replication, and Controller</p> <p>Apache Kafka With Zookeeper: Start ZooKeeper, Single Node-Single Broker Configuration, Start Producer to Send Messages, Start Consumer to Receive Messages, Single Node-Multiple Brokers Configuration</p> <p>Creating a Topic: Start Producer to Send Messages, Start Consumer to Receive Messages</p> <p>Kafka Real Time Example: Creating Twitter Producer, Application</p> <p>Kafka monitoring, Kafka Connect</p>	15

Books and References

Sr. No.	Title	Author/s	Publisher	Edition	Year
1	Kafka : The Definitive Guide	Neha Narkhede, Gwen Shapira , Todd Palino	O'Reilly	1 st	2015
2	Kafka Streams in Action: Real-time apps and microservices with the Kafka Streams API	Bill Bejeck	Manning	1 st	2018
3	Building Data Streaming Applications with Apache Kafka: Design, develop and streamline applications using Apache Kafka, Storm, Heron and Spark	Manish Kumar , Chanchal Singh	Packt	1 st	2017

SIUDSVS211 : Introduction to Kafka - Practical

B.Sc. (Data Science)	Semester – III
Course Name: Introduction to Kafka - Practical	Course Code: SIUDSVS211
Credits	1

List of Practical:

1.	Installation of Kafka.
2.	Start the Zookeeper. Perform Single Node- Single Broker Configuration.
3.	Demonstrate Start Consumer to Receive Messages.
4.	Demonstrate Single Node- Multiple Broker Configuration.
5.	A Practical to creating a topic in zookeeper.
6.	Demonstrate Real time application(Twitter).

SIUDSOE211: Digital Marketing-I

B.Sc. (Data Science)	Semester – III
Course Name: Digital Marketing-I	Course Code: SIUDSOE211
Credits	1

Course Objective:

- To understand the basic Concepts of Digital marketing and the road map for successful Digital marketing strategies
- To know the importance of Social media Platforms importance in Digital Marketing
- To understand the technological importance of SEO

Course Outcomes:

After completion of this course, student will be able to:

- **CO1:** to do digital marketing using various marketing strategies
- **CO2:** to do marketing on social media platform
- **CO3:** understand page ranking algorithm.

Unit	Contents	No. of Lectures
I	<p>Introduction to Digital Marketing: Fundamentals of Digital marketing & Its Significance, Traditional marketing Vs Digital Marketing, Evolution of Digital Marketing, Digital Marketing Landscape, Key Drivers, Digital Consumer & Communities, Gen Y & Netizen’s expectation & influence wrt Digital Marketing.</p> <p>Display adverting, Buying Models, different type of ad tools, Display advertising terminology, types of display ads, different ad formats, Ad placement techniques, Important ad terminology, Programmatic Digital Advertising.</p> <p>Social Media Marketing: Fundamentals of Social Media Marketing& its significance, Necessity of Social media Marketing, Building a Successful strategy: Goal Setting, Implementation.</p> <p>Facebook Marketing: Facebook for Business, Facebook Insight, Different types of Ad formats, Setting up Facebook Advertising Account, Facebook audience & types</p> <p>LinkedIn Marketing: Importance of LinkedIn presence, LinkedIn Strategy, Content Strategy, LinkedIn analysis, Targeting, Ad Campaign</p> <p>Search Engine Optimization: Introduction to SEO, How Search engine works, SEO Phases, History Of SEO, How SEO Works, What is Googlebot (Google Crawler), Types Of SEO technique, Keywords, Keyword Planner tools</p> <p>Off page Optimization- About Off page optimization, Authority & hubs, Backlink, Blog Posts, Press Release, Forums, Unnatural links.</p>	15

Books and References

Sr. No.	Title	Author/s	Publisher	Edition	Year
1	Digital Marketing	Kamat and Kamat-	Himalaya	2 nd	
2	Marketing Strategies for Engaging the Digital Generation	D. Ryan	Kogan Page	5 th	2020
3	Digital Marketing	S.Gupta	McGraw-Hill	3 rd	2022

SIUDSOE211: Digital Marketing-I Practical

B.Sc. (Data Science)	Semester – III
Course Name: Digital Marketing-I Practical	Course Code: SIUDSOE211
Credits	1

List of Practical:

Use Canva / Adobe Spark / Prezi / Vennage

1	Content writing
2	Create flyers
3	Invitation letter
4	Data representation using external data
5	Design business card
6	Create infographics
7	Resume
8	Email marketing
9	Mobile marketing
10	Social media marketing

S.Y.B.Sc.(DS) under NEP - Semester IV

Course Code	Course Type	Course Title	Credits
SIUDSMJ221	Core Subject (Major)	Big Data Analytics through Spark	3
SIUDSMJ222	Core Subject (Major)	Matrix Theory and Linear Algebra	3
SIUDSMN221	Core Subject (Minor)	Database Technology – II	3
SIUDSSE221	Skill Enhancement Course (SEC)	Web Technology – I	1
SIUDSOE221	Open Elective	Digital Marketing – II	1
SIUDSAE221	Ability Enhancement Course (AEC)	Communication Skills in Hindi	2
SIUSEP221	CEP		2
SIUSCC221	Co-Curricular (CC)		2
SIUDSMJP221	Core Subject (Major) Practical	Big Data Analytics through Spark – Practical	1
SIUDSMJP222	Core Subject (Major) Practical	Matrix Theory and Linear Algebra – Practical	1
SIUDSMNP221	Core Subject (Minor) Practical	Database Technology – II – Practical	1
SIUDSSE221	Skill Enhancement Course (SEC)	Web Technology – I – Practical	1
SIUDSOE221	Open Elective Practical	Digital Marketing – II Practical	1
TOTAL CREDITS			22

SIUDSMJ221 : Big Data Analytics through Spark

B.Sc. (Data Science)	Semester – IV
Course Name: Big Data Analytics through Spark	Course Code: SIUDSMJ221
Credits	3

Course Objective:

- Understand the Big Data Platform and its Use cases.
- Provide an overview of Apache Hadoop.
- Provide HDFS Concepts and Interfacing with HDFS.
- Understand Map Reduce Jobs Provide hands on Hadoop Eco System.
- Apply analytics on Structured, Unstructured Data.

Course Outcomes:

After completion of this course, student will be able to:

- **CO1:** Identify Big Data and its Business Implications.
- **CO2:** List the components of Hadoop and Hadoop Eco-System.
- **CO3:** Access and Process Data on Distributed File System.
- **CO4:** Manage Job Execution in Hadoop Environment.
- **CO5:** Develop Big Data Solutions using Hadoop Eco System.
- **CO6:** Analyze Infosphere BigInsights Big Data Recommendations.

Unit	Contents	No. of Lectures
I	<p>Big Data Technology Landscape: Hadoop, HDFS, MapReduce, Hive</p> <p>Apache Spark: Evolution, Features, Spark Built on Hadoop, Components of Spark.</p> <p>Spark architecture: Resilient Distributed Dataset (RDD), Directed Acyclic Graph (DAG).</p>	15
II	<p>Interactive Data Analysis with Spark Shell: REPL Commands, Using the Spark Shell as a Scala Shell, Number Analysis, Log Analysis.</p> <p>Writing a Spark Application: Hello World in Spark, Compiling and Running the Application, Monitoring the Application, Debugging the Application.</p> <p>Introducing Spark Streaming: Spark Streaming Is a Spark Add-on, High-Level Architecture, Data Stream Sources, Receiver, Destinations, Application Programming Interface (API), Streaming Context, Basic Structure of a Spark Streaming Application, Discretized Stream (DStream), Creating a Dstream, Processing a Data Stream, Output Operations, Window Operation.</p>	15

III	Introducing Spark SQL: Integration with Other Spark Libraries, Performance, Applications, ETL(Extract Transform Load), Application Programming Interface (API), Built-in Functions, Aggregate, Collection, Date/Time, Math, String, Window Introducing Machine Learning: Features, Labels, Models, Training Data, Test Data, Machine Learning Applications,	15
------------	--	----

Books and References

Sr. No.	Title	Author/s	Publisher	Edition	Year
1	Spark : The Definitive Guide	Bill Chambers & Mataei Zaharia	O'Reilly <u>Media, Inc.</u>	1 st	2018
2	Big Data Analytics with Spark : A Practitioner's Guide to Using Spark for Large Scale Data Analysis	Mohammed Guller	Apress	1 nd	2015
3	Practical Big Data Analytics Hands-on Techniques to Implement Enterprise Analytics and Machine Learning Using Hadoop, Spark, NoSQL and R	Nataraj Dasgupta	Packt	Illustrated	2018

SIUDSMJP221 : Big Data Analytics through Spark

B.Sc. (Data Science)	Semester – IV
Course Name: Big Data Analytics through Spark Practical	Course Code: SIUDSMJP221
Credits	1

List of Practical:

1.	Installation of Spark. Write a Hello World in Spark.
2.	Implement a program in Pig.
3.	Implement word count/ frequency program using MapReduce.
4.	Configure the Hive and implement the application in Hive.
5.	Implement Spark SQL.
6.	Implement machine learning with Spark or Hadoop.
7.	Implement Spark Streaming.
8.	Demonstrate Spark Shell commands.

SIUDSMJ222 : Matrix Theory and Linear Algebra

B.Sc. (Data Science)	Semester – IV
Course Name: Matrix Theory and Linear Algebra	Course Code: SIUDSMJ222
Credits	3

Course Objective:

- Understand the basic concepts of matrix algebra and its applications.
- Solving computational problems of linear algebra.

Course Outcomes:

After completion of this course, student will be able to:

- **CO1:** Understand basic matrix properties like rank, determinant, inverse, and a special type of matrices
- **CO2:** Introduce Gaussian / Gauss-Jordan elimination methods, LU factorization technique
- **CO3:** Use computational techniques for singular value decomposition (Computational and Algebraic Skills).
- **CO4:** Understand the concepts of vector space and subspaces.
- **CO5:** Understand the use of linear algebra and matrices in several important, modern applications of research and industrial problems involving statistics.

Unit	Contents	No. of Lectures
I	<p>Matrix theory: Algebra of Matrices, Trace and Rank of a Matrix and their properties, Determinants, Inverse, Eigen values and Eigen vectors, symmetric, orthogonal and idempotent matrices and their properties</p> <p>Matrix Factorization: Gauss elimination, row canonical form, diagonal form, triangular form, Gauss-Jordan-LU decomposition, solving systems of linear equations.</p> <p>Decomposition of Matrices: Spectral decomposition, singular value decomposition, Quadratic forms, definiteness and related results with proofs.</p>	15
II	<p>Vector Spaces: Vector Spaces, Subspaces, Basis and dimension of a vector space, linear dependence and linear independence, spanning set.</p> <p>Linear transformation: Linear transformation, kernel, range, Matrix Representation of a linear transformation, rank- nullity theorem, change of basis and similar matrices.</p> <p>Inner product spaces: Inner-product spaces, orthogonal sets and bases, Orthogonal Projection, Gram-Schmidt orthogonalization process, Orthogonal Bases and Gram-Schmidt, The Fast Fourier Transform</p> <p>Determinants: Introduction, Properties of the Determinant, Formulas for the Determinant, Applications of Determinants</p>	15

III	<p>Eigenvalues and Eigenvectors: Introduction, Diagonalization of a Matrix, Difference Equations and Powers A^k, Differential Equations and e^{At}, Complex Matrices, Similarity Transformations</p> <p>Computations with Matrices: Introduction, Matrix Norm and Condition Number, Computation of Eigenvalues, Iterative Methods for $Ax = b$</p> <p>Linear Programming and Game Theory: Linear Inequalities, The Simplex Method, The Dual Problem, Network Models, Game Theory</p>	15
------------	--	----

Books and References

Sr. No.	Title	Author/s	Publisher	Edition	Year
1	Introduction to linear algebra	Gilbert Strang	Wellesley-Cambridge	6 th	2023
2	Linear Algebra and Its Applications	David C. Lay, Steven R. Lay, Judi J. McDonald	Pearson Education	5 th	2023
3	Linear Algebra, Theory and Applications	Kenneth Kuttlet			2014
4	Basics of Matrix Algebra for Statistics with R	Fieller N	Chapman and Hall/CRC		2015

SIUDSMJP222: Matrix Theory and Linear Algebra Practical

B.Sc. (Data Science)	Semester – IV
Course Name: Matrix Theory and Linear Algebra Practical	Course Code: SIUDSMJP222
Credits	1

List of Practical: (use R/python/scilab/matlab)

1	Matrices and Gaussian Elimination.
a.	Multiplication and transpose of matrix.
b.	Inverses of matrix without using any inbuilt package.
c.	Inverses of matrix using any inbuilt package like numpy.
d.	Linear equation with n unknowns using Gauss Elimination
2	Vector
a.	Addition, subtraction, multiplication and division of vector
b.	Dot product & cross product of vector
c.	Visualizing vector Linear Transformations
3	Orthogonal and Orthonormal Vectors
a.	Computes the orthonormal vectors using the GS algorithm
b.	Projections and Least Squares
c.	Fast Fourier Transform
4	Determinant of matrix
a.	Finding determinant of matrix without using any inbuilt package.
b.	Finding determinant of matrix using any in built package.
5	Calculate Eigenvalues
a.	Compute the eigenvalues and right eigenvectors of a given square array
b.	Program to test diagonalizable matrix
6	i. Tests for Positive Definiteness ii. Singular Value Decomposition iii. The Finite Element Method (Only Demonstration)
7	Simplex Method. (Only Demonstration)
8	The Dual Problem. (Only Demonstration)
9	Implementing Network Models (Only Demonstration)
10	Implementing Game Theory (Only Demonstration)

SIUDSMN221 : Database Technology-II

B.Sc. (Data Science)	Semester – IV
Course Name: Database Technology-II	Course Code: SIUDSMN221
Credits	3

Course Objective:

- To understand the basics of PL/SQL.
- To understand control and conditional statement in PL/SQL.
- To understand working of sequences and cursor in PL/SQL.
- To understand concept of stored procedure and functions.
- To understand triggers and packages in PL/SQL.

Course Outcomes:

After completion of this course, student will be able to:

- **CO1:** Understand the basics of PL/SQL.
- **CO2:** Use of the control and conditional statement in PL/SQL.
- **CO3:** Apply sequences and cursor in PL/SQL.
- **CO4:** Know the concept of stored procedure and functions
- **CO5:** Create the triggers and packages in PL/SQL.

Unit	Contents	No. of Lectures
I	Overview of Oracle Architecture: Oracle Physical architecture and Oracle Instance Architecture. The PL/SQL Block, Lexical Units: Identifiers, Delimiters, Literals, Comments, Variables, PL/SQL Types, Expression, Operators and Control Structures, Records, Cursors: Definition of Cursor, Explicit and Implicit Cursors, Cursor for loops, Cursor variables, parametrized cursor	15
II	Sub Program: Procedures, Functions, Subprogram creation, parameter modes, Procedure versus Functions, Database Triggers: Use of Database Triggers, Types of Triggers, Creation and Deletion of Triggers, Error Handling: Declaring Exception, Raising Exception, Handling Exception, Sequences & Pseudo Columns: CURRVAL & NEXTVAL, ROWID, ROWNUM, PL/SQL Transactions: Transaction Control statements: commit, savepoint and rollback	15
III	Locks: Types of Locks and levels of locks. Database Security: Database Security Issues, Security Models, Different threats to databases, Challenges to maintaining database security, Introduction to Other Databases: Introduction to Parallel and distributed Databases, Introduction to Object Based Databases, XML Databases, NoSQL Database, Multimedia Databases, Big Data Databases	15

Books and References

Sr. No.	Title	Author/s	Publisher	Edition	Year
1	Database System Concepts	Silberschatz A., Korth H., Sudarshan S.	McGraw Hill Education	6 th	
2	SQL, PL/SQL the Programming Language of Oracle	Ivan Bayross	BPB Publications	4 th	
3	Oracle PL/SQL Programming	Steven Feuerstein and Bill Pribyl	O'Reilly Publications	6 th	2014
4	NoSQL Distilled A Brief Guide to the Emerging World of Polyglot Persistence	Pramod J. Sadalage, Martin Fowler, Addison-Wesley	Pearson Education, Inc.	1 st	2012
5	Database Systems: Concepts, Design and Application	S.K.Singh	Pearson Education		

SIUDSMNP221: Database Technology-II Practical

B.Sc. (Data Science)	Semester – IV
Course Name: Database Technology-II Practical	Course Code: SIUDSMNP221
Credits	1

List of Practical:

1	PL/SQL Basics
a.	Use of variables.
b.	Write executable statement.
c.	Interacting with Oracle Server.
d.	Create anonymous PL/SQL block
2	Control Structure in PL/SQL
a.	Using while loop
b.	Do loop
c.	For loop
d.	Use of GOTO statement
3	Create conditional statement using PL/SQL
a.	Using if statement
b.	Using if else statement
c.	Using elsif ladder
d.	Using case expression.
4	Creation of Sequence in PL/SQL
5	Create cursor in PL/SQL
a.	Implicit cursor
b.	Explicit
c.	Parameterized cursor
d.	Cursor for loop
6	Creation of Procedures in PL/SQL
7	Functions in PL/SQL
a.	Compute and returns the maximum value
b.	Compute factorial of given number.
8	Creation of Trigger
a.	Create Row level trigger
b.	Create Statement level trigger
c.	Create instead of trigger
9	Handling exceptions
a.	Creation of user defined exception
b.	Creation of system defined exception
10	Design and Develop Application for Student Mark Sheet Processing

SIUDSSE221: Web Technology-I

B.Sc. (Data Science)	Semester – IV
Course Name: Web Technology-I	Course Code: SIUDSSE221
Credits	1

Course Objective:

- Introducing the basic concepts of Internet and web design to learners
- Providing brief knowledge about HTML5 concepts
- Giving insight of the Page layout and navigation with HTML5
- Making students aware about use of Tables, Forms and Media with HTML5
- Teaching data validation using Java Script. Giving knowledge about transmission of data

Course Outcomes:

After completion of this course, student will be able to:

- **CO1:** Understand the meaning of the basic terminologies of web technology and explore, use The HTML5 concepts, Understand the basic requirement of web design. Understand and use the Page layout, Navigation, Tables, Forms and Media features of HTML5.
- **CO2:** Understand and use Cascading Style sheet for beautifying the webpages.
- **CO3:** Understand and use the Java Script for validation of user forms in web pages

Unit	Contents	No. of Lectures
I	<p>HTML5: Introduction, Why HTML5? Formatting text by using tags, using lists and backgrounds, Creating hyperlinks and anchors.</p> <p>HTML5 Page layout and navigation: Creating navigational aid, creating image map, redirecting to another URL, HTML5 semantic tags</p> <p>HTML5 Tables, Forms and Media: Creating tables and formatting tables, creating user forms, Incorporating sound and video.</p> <p>CSS: Style sheets, CSS, formatting text using style sheets, formatting paragraphs using style sheets.</p> <p>Java Script: Introduction, Client-Side JavaScript, Server-Side JavaScript, JavaScript Objects, JavaScript Security Core JavaScript (Properties and Methods of Each) : Array, Boolean, Date, Function, Math, Number, Object, String, regExp Document and its associated objects: document, document object methods, Link, Area, Anchor, Image, Layer</p> <p>Events and Event Handlers : General Information about Events, Defining Event Handlers, event, onAbort, onBlur, onChange, onClick, onDbClick, onDragDrop, onError, onFocus, onKeyDown, onKeyPress, onKeyUp, onLoad, onMouseDown, onMouseMove, onMouseOut, onMouseOver, onMouseUp, onMove, onReset, onResize, onSelect, onSubmit, onUnload</p>	15

Books and References

Sr. No.	Title	Author/s	Publisher	Edition	Year
1	Web Design The Complete Reference	Thomas Powell	Tata McGraw Hill	Fifth Edition	2009
2	HTML5 Step by Step	Faithe Wempen	Microsoft Press	Twelfth Edition	2011
3	JavaScript 2.0: The Complete Reference	Thomas Powell and Fritz Schneider	Tata McGraw Hill	Second Edition	2004

SIUDSSE221 : Web Technology-I Practical

B.Sc. (Data Science)	Semester – IV
Course Name: Web Technology-I Practical	Course Code: SIUDSSE221
Credits	1

List of Practical:

1	Use of Basic Tags:
a.	Design a web page using different text formatting tags.
b.	Demonstrate use of Font tag with its attributes and HTML various color options in web page.
c.	Design a web page with links to different pages and allow navigation between web pages.
2.	Navigation, list and paragraph:
a.	Design a web page to demonstrate text-based navigation bar.
b.	Demonstrate use of lists and backgrounds in web page.
c.	Demonstrate use of paragraph and its associated tags in web page.
3.	Lists, images and semantics:
a.	Demonstrate use of multiple image tag in web page.
b.	Design a web page with Imagemaps.
c.	Design a web page demonstrating use of various semantics tags
4.	Multimedia and User controls:
a.	Design a web page with a form that uses all types of user controls.
b.	Design a web page embedding with multimedia features.
c.	Design a 3 page static website with appropriate tags and attributes.
5.	CSS with list, links and table:
a.	Create and use different style rules with available types of lists.
b.	Create and use different style rules with hyperlinks.
c.	Create and use different style rules with tables.
6.	CSS with font, paragraph and types:
a.	Create and use different style rules with font elements.
b.	Demonstrate the use of inline, internal and external CSS in one web page.
7.	Demonstrate the use of Document object methods
a.	Using java script, demonstrate validating Text Input Fields, Drop-down Lists and Checkboxes.
b.	Using java script, demonstrate validating Radio buttons and Validating Multi-Select Boxes.
8.	Java Script : Handling the events
a.	Using java script, demonstrate the use of onAbort, onBlur, onChange, onClick, onDbClick events.

b	Using java script, demonstrate the use of o onError, onFocus events.
c	Using java script, demonstrate the use of onKeyDown, onKeyPress, onKeyUp, onLoad, onReset, onResize, onSelect, onSubmit, onUnload events.
9.	Demonstrate complete validation of User Registration form using appropriate fields of html and events of java script.

SIUDSOE221 : Digital Marketing-II

B.Sc. (Data Science)	Semester – IV
Course Name: Digital Marketing-II	Course Code: SIUDSOE221
Credits	1

Course Objective:

- How to creates effective Ad Words campaign & Advertising Positioning with respect to the Digital marketing
- How to create website using Word Press
- Web analytics focuses on optimizing an organization's digital ecosystem by collecting, analyzing and enabling the making of data-informed decisions.

Course Outcomes:

After completion of this course, student will be able to:

- **CO1:** to use advertising tools effectively.
- **CO2:** to create website using wordpress
- **CO3:** analyze web and can take effective decision.

Unit	Contents	No. of Lectures
I	<p>Advertising Tools & Its Optimization : Display adverting, different type of ad tools, Display advertising terminology, types of display ads, different ad formats, Ad placement techniques, Important ad terminology, ROI measurement techniques, AdWords & Adsense.</p> <p>Wordpress: What is WordPress, History, Versions, features,.com vs org,Installing WordPress. Dashboard:Post, Pages, Post vs Pages, Comments, Categories, Tags, Link, Appearance, Users, Menus, Media Library Themes and Plugins:Installing themes, Free vs Paid themes, Plugins,Installing Plugins, Live Chat Plugin, Popup Plugins, Search Plugin, Slider Plugin, Web Hosting</p> <p>Web Analytics: Introduction- What's analysis?, Is analysis worth the effort?, Small businesses, Medium and Large scale businesses, Analysis vs intuition, Introduction to web analytic Google Analytics -Getting Started With Google Analytics, How Google Analytics works?, Accounts, profiles, and users navigating Google Analytics, Basic metrics, The main sections of Google Analytics reports Traffic Sources Direct, referring, and search traffic Campaigns AdWords, Adsense</p>	15

Books and References

Sr. No.	Title	Author/s	Publisher	Edition	Year
1	Digital Marketing	Kamat and Kamat-	Himalaya	2 nd	
2	Marketing Strategies for Engaging the Digital Generation	D. Ryan	Kogan Page	5 th	2020
3	Digital Marketing	S.Gupta	McGraw-Hill	3 rd	2022
4.	Practical Web Analytics for User Experience, How Analytics Can Help You Understand Your Users	Michael Beasley	Morgan Kaufmann		2013
5	WordPress for Web Developers: An Introduction for Web Professionals	Stephanie Leary	APress	2 nd	2013

SIUDSOE221: Digital Marketing-II Practical

B.Sc. (Data Science)	Semester – IV
Course Name: Digital Marketing-II Practical	Course Code: SIUDSOE221
Credits	1

List of Practical:

1	Installation of wordpress and create a simple project.
2	Create different pages and provide the connectivity between them.
3	Create different categories and also implement a menu bar in a page.
4	Create a page to implement media libraries
5	Create a wordpress application by implementing different suitable them according to website type.
6	Implement the following plugins.
7	i) search plugin ii) slider plugin iii) live chat plugin.
8	Perform web hosting of the wordpress application.
9	Youtube marketing