R. A. Podar College of Commerce and Economics
(Autonomous)
Matunga, Mumbai-19

Syllabus
For Under Graduate Programme

T. Y. B.Sc. (Data Science and Analytics)
Semester V & VI

CHOICE BASED CREDIT AND GRADING SYSTEM
(CBCGS) With effect from the academic year 2024-25

College Website: www.rapodar.ac.in
Bachelor of Science (Data Science and Analytics) Programme
Course Structure
T.Y.B.S.C (Data Science and Analytics)
(To be implemented from Academic Year - 2024-25)

x

Course Code	Semester V	Credits
	Core Courses (CC)	
60501	Communication in Data Science	**02**
60502	Project Dissertation	**02**
	Ability Enhancement Courses (AEC)	
	Ability Enhancement Compulsory Courses (AECC)	
60503	Design Thinking	**02**
60504	Design Thinking Practical	**02**
	Skill Enhancement Courses (SEC)	
60505	Apache Spark and Scala	**02**
60506	Apache Spark and Scala Practical	**02**
	Discipline Specific Elective (DSE) Courses	
60507	Machine Learning	**02**
60508	Machine Learning Practical	**02**
60509	Introduction to Cloud Computing	**02**
60510	Introduction to Cloud Computing Practical	**02**
Total Credits		**20**

x

Semester V

x

Core Courses (CC)

B. Sc. (Data Science and Analytics)	Semester – V
Course Name: Communication in Data Science	Course Code: 60501
Periods per week (1 Period is 1 hour)	3
Credits	2

		Hours	Marks
Evaluation System	Theory Examination	2	60
	Internal	--	40

Course Objectives:
The main objective of this Course is to
● Develop active listening skills and understand the importance of empathy in communication.
● Enhance speaking skills for confident and clear communication.
● Gain knowledge about different types of interviews and improve interview performance.
● Understand intrapersonal and interpersonal communication dynamics and conflict management in the workplace
● Learn effective strategies for delivering negative news and crisis communication

Unit	Details	Lectures
I	**Active Listening:** Meaning and Art of Listening, Importance of Listening and Empathy in Communication, Reasons for Poor Listening, Poor Listening Habits, Traits of a Good Listener, Listening Modes and Types, Barriers to Effective Listening, Listening for General Content and Specific Information **Effective Speaking:** Basic Sounds of English, Word Stress, Sentence Stress, Intonation, Achieving Confidence, Clarity, and Fluency, Vocal Cues **Interviews:** Objectives of Interviews, Types of Interviews, Job Interviews, Résumés, Media Interviews, Press Conference.	12
II	**Intrapersonal and Interpersonal Business Communication:** Intrapersonal Communication, Self-Concept and Dimensions of Self, Interpersonal Needs, Social Penetration Theory, Rituals of Conversation and Interviews, Conflict in the Work Environment **Negative News and Crisis Communication:** Delivering a Negative News Message, Eliciting Negative News, Crisis Communication Plan, Press Conferences	12
III	**Sampling Design:** Census and Sample Survey, Implications of a Sample **Intercultural and International Business Communication:** Intercultural Communication: How to Understand Intercultural Communication, Common Cultural Characteristics, Divergent Cultural Characteristics, **International Communication:** The Global Marketplace, Styles of Management, The International Assignment	12

I V	**Group Communication:** Use of Body Language in Group, Discussions, Group Discussions, Organizational GD, GD as Part of Selection Process, Meetings, Conferences **Group Communication, Teamwork, and Leadership:** Group Life Cycles and Member Roles, Group Problem Solving, Business and Professional Meetings, Teamwork and Leadership**.**	**12**
V	**Data Interpretation:** Tabulation, Bar Graphs, Pie Chart, Line Graphs **Logical Reasoning:** Argument forms, structure of categorical propositions, Mood and Figure, Formal and Informal fallacies, Uses of language, Connotations and denotations of terms, Classical square of opposition, deductive and inductive reasoning, Analogies, Venn diagram	**12**

Course Outcome:

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

● Apply active listening techniques and overcome barriers to become a better listener.
● Demonstrate improved speaking skills with clarity, confidence, and fluency.
● Utilize interview techniques to enhance job interview performance and create impactful résumés.
● Apply interpersonal communication skills to build effective relationships and manage conflicts in professional settings.
● Effectively deliver negative news messages, develop crisis communication plans, and handle press conferences in challenging situations.

Books and References:

● Business Communication for Success, University of Minnesota, University of Minnesota, 2015

● Technical Communication: Principles and Practice, Meenakshi Raman, Oxford University Press, 3rd Edition, 2015.

Evaluation scheme
I Continuous Internal Assessment (C.A.)– 40 Marks
1) Assessment 1 (20 Marks)
2) Class Test (20 Marks)
II Semester End Examination (SEE) - 60 Marks
QUESTION PAPER PATTERN OF (SEE)
Maximum Marks: 60 Marks
Time: 2 Hours
Note: 1) Attempt all Questions
2) Any three sub questions to be attempted. All sub questions carry equal marks

x

Question No	Particulars	Marks
Q -1 (Unit I)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 Marks
Q -2 (Unit II)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 Marks
Q -3 (Unit III)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 Marks
Q4 (Unit IV)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 marks
Q5 (Unit IV)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 marks

x

Core Courses (CC)

B. Sc. (Data Science and Analytics)		Semester – V	
Course Name: Project Dissertation		Course Code: 60502	
Periods per week (1 Period is 1 hour)		2	
Credits		2	
		Hours	Marks
Evaluation System	Practical Examination	$2^{1/2}$	50
	Internal	--	-

Course Objectives:
- To provide students with an opportunity to apply the knowledge and skills acquired throughout the B.Sc Data Science program in a real-world project.
- To enable students to independently plan, design, and execute a research project in the field of data science.
- To develop students' critical thinking and problem-solving abilities through the identification and analysis of complex data science problems.
- To enhance students' research and project management skills, including data collection, data analysis, and result interpretation.
- To foster effective communication and presentation skills by requiring students to document and present their research findings in a comprehensive project dissertation.

The project report should contain the following:

Table of Contents

- What is a Data Science Project Report?

- 6 Fundamental Steps to Create a Data Science Report
    - Define the Data Science Project Topic and Problem Statement
    - Explain How You Intend to Address the Problem
    - Describe the Dataset and its Attributes
    - Outline the Design of Your Data Science Project
    - Conduct an In-depth Project Analysis
    - Wrap Up Your Project
- How to Write a Detailed Project Report on Data Science?
    - Defining the Data Science Problem Statement
    - Discuss Your Approach to Address the Problem Statement
    - Explain the Dataset and its Attributes
    - Define the Project Structure/Layout
    - Analyze the Steps in Detail
    - Summarize the Project Outcomes and Add References

- 10 Main Components of a Data Science Project Report
  - Project Title/Topic o Table of Contents
  - Abstract or Project Summary
  - Introduction of Dataset Description
  - Methods and Algorithms
  - Project Analysis
  - Final Results
  - Conclusion and Future Scope
  - References
  - Best Practices for A Data Science Project Report
  - Define the Objective of The Project Report
  - Focus More on the Outcomes, Not the Report
  - Develop a Thorough Project Strategy in Advance
  - Document Everything- Data, Design, Algorithm

Course Outcome:

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

- Students will demonstrate the ability to identify and formulate a research problem in the field of data science and define clear objectives and research questions.
- Students will independently plan and execute a research project, including selecting appropriate research methodologies and data collection techniques.
- Students will analyze and interpret data using advanced data analysis techniques, demonstrating proficiency in statistical analysis, machine learning, or other relevant methodologies.
- Students will present their research findings in a well-structured, comprehensive project dissertation, adhering to academic writing standards and providing appropriate references.
- Students will effectively communicate their research findings through oral Presentations, demonstrating clear and concise delivery of information and responding to questions and feedback from peers and evaluators.

Evaluation scheme

Project Dissertation: 50 marks
Certified copy of the Project is essential to appear for the practical examination.

1.	**Project Dissertation Submission**	20
2.	**Presentation**	20
3.	**Viva Voce**	10

x

Ability Enhancement Courses (AEC)
Ability Enhancement Compulsory Courses (AECC)

B. Sc. (Data Science and Analytics)		Semester – V	
Course Name: Design Thinking		Course Code: 60503	
Periods per week (1 Period is 1 hour)		3	
Credits		2	
		Hours	Marks
Evaluation System	Theory Examination	2	60
	Internal	--	40

Course Objectives:

   The main objective of this Course is to
   - To Understand the Design thinking process.
   - To help in developing empathy and define problem
   - To generate creative ideas
   - To develop prototype for product
   - Learners will able to test and take feedback for improvement
   - Learners will able to collaborate team work
   - Learners will able to create centered design
   - Learners will able to take risk and learned from failure\
   - Learners will able to apply design thinking for real world problems
   - Learn effective strategies for delivering negative news and crisis communication

Unit	Details	Lectures
I	Introduction to Design Thinking: Basics of Design thinking: Definition and Principles of Design Thinking, Historical context and evolution, Need of Empathy Design Thinking Process: Five stages of design thinking: Empathize, Define, Ideate, Prototype, Test.	12
II	Empathize: User Research, Conducting Surveys and data analysis for getting User Insights, Visualizing user needs, emotions, and behaviors through Empathy mapping, Developing User Persona Define: Problem definition and reframing, Creating user-centric problem statements, Diverse and creative problem perspectives	12
III	Design Thinking Ideation Techniques: Cultivating creativity through rapid prototyping and collaboration. Prototyping: Prototypes to test and refine ideas, Importance of prototyping in design, Rapid prototyping techniques,User Feedback and Iteration, Data-Driven Prototyping	12
IV	Testing: Methodologies for Testing, Need for Testing, User Testing and Feedback Refine and Implement Solutions: Finalizing solutions based on user testing and data analysis.	12
V	Ethical Considerations in Design Thinking: Ensuring inclusivity and responsible data usage, How to Build a Design Thinking Culture Emerging Trends in Design Thinking: New techniques and applications of design thinking.	12

Course Outcomes:
Upon the successful completion of this course, Learners will be able to:
● Develop Design Thinking Mindset
● Develop solutions which are user-centric and feasible
● Gain skills in collaborations and ideation
● improve ability to analyze data and synthesize information

Books and References:
● Change by Design: How Design Thinking Transforms Organizations and Inspires Innovation" by Tim Brown
● "The Design of Everyday Things" by Don Norman
● Sprint: How to Solve Big Problems and Test New Ideas in Just Five Days" by Jake Knapp, John Zeratsky, and Braden Kowitz
● Design Thinking for Strategic Innovation: What They Can't Teach You at Business or Design School" by Idris Mootee
● Design Thinking: Integrating Innovation, Customer Experience, and Brand Value" by Thomas Lockwood and Edgar Papke

x

<center>**Evaluation scheme**</center>

I Continuous Internal Assessment (C.A.)– 40 Marks
 1) Assessment 1 (20 Marks)
2) Class Test (20 Marks)
II Semester End Examination (SEE) - 60 Marks
<center>QUESTION PAPER PATTERN OF (SEE)</center>

Maximum Marks: 60 Marks

Time: 2 Hours

Note: 1) Attempt all Questions

   2) Any three sub questions to be attempted. All sub questions carry equal marks

Question No	Particulars	Marks
Q -1 (Unit I)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 Marks
Q -2 (Unit II)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 Marks
Q -3 (Unit III)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 Marks
Q4 (Unit IV)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 marks
Q5 (Unit IV)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 marks

Ability Enhancement Courses (AEC)
Ability Enhancement Compulsory Courses (AECC)

B. Sc. (Data Science and Analytics)	Semester – V	
Course Name: Design Thinking Practical	Course Code: 60504	
Periods per week (1 Period is 1 hour)	2	
Credits	2	
	Hours	Marks
Evaluation System	Practical Examination	$2^{1/2}$	50
	Internal	--	--

Practical Based on Application Case Studies real world problem

Practical Examination : 50 marks
Certified copy of the Presentation is essential to appear for the practical examination.

1.	Presentation	30
2.	Viva Voce	20

Skill Enhancement Courses (SEC)

B. Sc. (Data Science and Analytics)		Semester – V	
Course Name: Apache and Scala		Course Code: 60505	
Periods per week (1 Period is 1 hour)		3	
Credits		2	
		Hours	Marks
Evaluation System	Theory Examination	2	60
	Internal	--	40

Course Objectives:
The main objective of this Course is to
- To manipulating big data distributed over a cluster using functional concepts
- To widespread industrial uses of functional ideas.
- To evidence by the popularity of MapReduce and Hadoop, and most recently Apache Spark, a fast, in-memory distributed collections framework written in Scala.

Unit	Details	Lectures
I	Introduction : What is Scala , Setup and configuration of Scala, Developing and running basic Scala Programs, Scala operations, Functions and procedures in Scala, Different Scala APIs for common operations, Loops and collections- Array, Map, Lists, Tuples, Pattern matching for advanced operations, Eclipse with Scala, Batch vs Real-time vs in-Memory, MapReduce and its limitations, Apache Storm and its limitations, Need for a general purpose solution - Apache Spark	12
II	What is Apache Spark?, Components of Spark architecture, Apache Spark design principles , Spark features and characteristics , Apache Spark ecosystem components and their insights , Setting up the Spark Environment, Installing and configuring prerequisites, Installing Apache Spark in local mode, Working with Spark in local mode, Troubleshooting encountered problems in Spark.	12
III	Installing Spark in standalone mode, Installing Spark in YARN mode , Installing & configuring Spark on a real multi-node cluster, Playing with Spark in cluster mode , Best practices for Spark deployment ,Playing with the Spark shell	12
IV	Executing Scala and Java statements in the shell, Understanding the Spark context and driver, Reading data from the local filesystem Integrating Spark with HDFS , Caching the data in memory for further use, Distributed persistence , Testing and troubleshooting	12

| V | What is an RDD in Spark, How do RDDs make Spark a feature-rich framework, Transformations in Apache Spark RDDs , Spark RDD action and persistence, Spark Lazy Operations - Transformation and Caching , Fault tolerance in Spark, Loading data and creating RDD in Spark , Persist RDD in memory or disk , Pair operations and key-value in Spark, Spark integration with Hadoop , Apache Spark practicals. | **12** |

Course Outcomes:
Upon the successful completion of this course, Learners will be able to:
● Read data from persistent storage and load it into Apache Spark
● Manipulate data with Spark and Scala
● Express algorithms for data analysis in a functional style
● Recognize how to avoid shuffles and recomputation in Spark

Books and References:
● Learning Spark by Matei Zaharia, Patrick Wendell, Andy Konwinski, Holden Karau.
● Advanced Analytics with Spark by Sandy Ryza, Uri Laserson, Sean Owen and Josh Wills.
● Mastering Apache Spark by Mike Frampton.
● Programming Scala by Dean Wampler, Alex Payne.

Evaluation scheme
I Continuous Internal Assessment (C.A.)– 40 Marks
1) Assessment 1 (20 Marks)
2) Class Test (20 Marks)
II Semester End Examination (SEE) - 60 Marks
QUESTION PAPER PATTERN OF (SEE)
Maximum Marks: 60 Marks
Time: 2 Hours
Note: 1) Attempt all Questions
     2) Any three sub questions to be attempted. All sub questions carry equal marks

Question No	Particulars	Marks
Q -1 (Unit I)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 Marks
Q -2 (Unit II)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 Marks
Q -3 (Unit III)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 Marks

Q4 (Unit IV)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 marks
Q5 (Unit IV)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 marks

Skill Enhancement Courses (SEC)

B. Sc. (Data Science and Analytics)		Semester – V	
Course Name: Apache and Scala Practical		Course Code: 60506	
Periods per week (1 Period is 1 hour)		2	
Credits		2	
		Hours	Marks
Evaluation System	Practical Examination	2 $^{1/2}$	50
	Internal	--	--

List of Practical

1	Setup and configure Scala
2	Setup and configure Apache Spark
3	Developing and Running basic Scala Program
4	Perform basic mathematical operations in scala
5	Perform basic operation on collection and loops
6	Create Functions and procedures in scala
7	Implement a basic Word Count program using Spark RDDs
8	Implement an ETL (Extract, Transform, Load) pipeline using Spark

x

Practical Exam: 50 marks

Certified copy journal is essential to appear for the practical examination.

1.	Practical Work	30 Marks
2.	Journal	10 Marks
3.	Viva	10 Marks

Discipline Specific Elective (DSE) Courses

B. Sc. (Data Science and Analytics)		Semester – V	
Course Name: Machine Learning		Course Code: 60507	
Periods per week (1 Period is 1 hour)		3	
Credits		2	
		Hours	Marks
Evaluation System	Theory Examination	2	60
	Internal	--	40

Course Objectives:
● To introduce students to the fundamental concepts and principles of Machine Learning.
● To familiarize students with various types of machine learning algorithms and their applications.
● To provide hands-on experience in implementing and evaluating machine learning models.
● To develop critical thinking and problem-solving skills in the context of machine learning.
● To enable students to apply machine learning techniques to real-world problems and datasets.
● To understand the limitations and challenges of machine learning and develop strategies to address them.

Unit	Details	Lectures
I	**Introduction to Machine Learning :** Machine Learning(ML), Need for Machine Learning, ML from Knowledge-driven to Data Driven, Applications of Machine Learning, Problems suitable for Machine Learning, Advantages, Disadvantages and Challenges of Machine Learning, Challenges of ML. General architecture of ML systems, **Underlying Concepts in Machine Learning:** Inductive Learning, Generalization, Bias and Variance, Overfitting and Underfitting, Parametric and Non Parametric algorithms **Types of Machine Learning:** Supervised and Unsupervised Learning, Workflow, Semi supervised Learning, Reinforced Learning, Deep Learning, introduction to Neural Networks	12
II	**Introduction to Classification Algorithms:** Concept of Classification, Binary classification, Multi-Class Classification, Multi Label Classification **K-Nearest Neighbor Method**: need and Working of KNN, Computing Distance, Pros and Cons of KNN, **Decision Tree based Algorithm**: Terminologies and assumptions, Working of Decision Trees, ID3 Algorithm, Attribute selection Methods(Entropy, Gini Impurity, Information Gain) **ANN – Artificial Neural Network**	12
III	**Support Vector Machines:** Work in of SVM, SVM Concepts Support Vectors, Hard Margin, soft Margin, Kernels, Advantages and Disadvantages of SVM **Probabilistic Learning**: Introduction to Bayes Learning, Interpretation of Bayes Rule, Benefits and shortfalls of Bayesian Learning, Naïve Bayes Classifier, Characteristics of Naïve **Bayes Regression Methods:** Linear Regression Models, Logistic Regression	12
IV	**Performance Evaluation:** Classification Metrics-Accuracy, Sensitivity, Precision, F1 Score, ROC/AUC Curve, Cross Validation **Unsupervised Learning**: Concept of unsupervised Learning, Importance and Challenges of unsupervised Learning, Clustering and its applications **Hierarchical Clustering:** Introduction, Types of Hierarchical Clustering, Issues with Hierarchical Clustering,	12
V	**Partition algorithm:** K-means Clustering, steps of K-means Clustering, Issues, Strength and Weakness of K-means clustering. Curse of Dimensionality. **Dimensionality Reduction:** Criteria for Reduction, Feature Reduction and Selection, Principal Component Analysis (PCA) **Association rule Mining:** Basic Concepts, Market Basket Analysis, and Apriori algorithm.	12

Course Outcomes:
Upon the successful completion of this course, students will be able to:
● Understand the foundational concepts and principles of Machine Learning
● Apply supervised and unsupervised learning techniques, including classification algorithms and clustering algorithms
● Evaluate the performance of Machine Learning models using classification metrics,
● ROC/AUC curve analysis, and cross-validation techniques.
● Implement regression models (such as linear regression and logistic regression) and understand their applications in predictive analysis.
● Utilize dimensionality reduction techniques (Like PCA) for feature reduction and selection, and apply association rule mining algorithms (such as the Apriori algorithm) for discovering meaningful patterns in datasets.

Books and References:
● Machine Learning: Concepts, Techniques and Applications, T V Geetha S, Sendhil kumar CRC Press, Taylor and Francis, 1 st Edition, 2023
● Machine Learning for Decision Sciences with Case Studies in Python, S. Sumathi, Suresh V. Rajappa, CRC Press, Taylor and Francis, 1 st Edition, 2022
● Introduction to Machine Learning with Python, Andreas C. Müller, Sarah Guido O'Reilly Media, Inc. 1 st Edition, 2016
● Machine Learning for Beginners, Harsh Bhasin, BPB, 1st Edition, 2020
● Machine Learning S Sridhar, Oxford University, Press, 1 st Edition, 2021
● Machine Learning Ruchi Doshi, Kamal Kant Hiran, BPB, 1 St Edition, 2021

Evaluation scheme
I Continuous Internal Assessment (C.A.)– 40 Marks
1) Assessment 1 (20 Marks)
2) Class Test (20 Marks)
II Semester End Examination (SEE) - 60 Marks
QUESTION PAPER PATTERN OF (SEE)
Maximum Marks: 60 Marks
Time: 2 Hours
Note: 1) Attempt all Questions
      2) Any three sub questions to be attempted. All sub questions carry equal marks

x

Question No	Particulars	Marks
Q -1 (Unit I)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 Marks
Q -2 (Unit II)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 Marks
Q -3 (Unit III)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 Marks
Q4 (Unit IV)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 marks
Q5 (Unit IV)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 marks

Discipline Specific Elective (DSE) Courses

B. Sc. (Data Science and Analytics)		Semester – V	
Course Name: Machine Learning Practical		Course Code: 60508	
Periods per week (1 Period is 1 hour)		2	
Credits		2	
		Hours	Marks
Evaluation System	Practical Examination	$2^{1/2}$	50
	Internal	--	--

x

	List of Practical
1	Implementing a K-Nearest Neighbour algorithm (e.g. to classify handwritten digits)
2	Building a decision tree model using the ID3 algorithm (e.g. to predict whether a customer will churn or not)
3	Developing a Support Vector Machine model (e.g. to classify email messages as spam or not spam)
4	Building a Naïve Bayes classifier (e.g. to classify movie reviews as positive or negative sentiment)
5	Implementing linear regression (e.g. to predict housing prices based on features)
6	Using logistic regression (e.g. to predict whether a credit card transaction is fraud)
7	Evaluating a classification model using metrics such as accuracy, precision, recall, and F1 score
8	Applying hierarchical clustering (e.g. to group customer segments based on their purchasing behaviour)
9	Implementing the K-means clustering algorithm (e.g. to identify distinct clusters in a customer demographic dataset).
10	Utilizing Principal Component Analysis (PCA) for dimensionality reduction to improve the efficiency and interpretability of a model.

Practical Exam: 50 marks

Certified copy journal is essential to appear for the practical examination.

1.	**Practical Work**	**30 Marks**
2.	**Journal**	**10 Marks**
3.	**Viva**	**10 Marks**

x

Discipline Specific Elective (DSE) Courses

B. Sc. (Data Science and Analytics)		Semester – V	
Course Name: Introduction to Cloud Computing		Course Code: 60509	
Periods per week (1 Period is 1 hour)		3	
Credits		2	
		Hours	Marks
Evaluation System	Theory Examination	2	60
	Internal	--	40

Course Objectives:
The main objective of this Course is to
● Introduce cloud computing technology through its various aspects,
● To understand virtualization technology and its functionality in cloud computing,
● To get familiar with Mobile Cloud Computing Concept, to
● To be aware of different service and privacy aspects of cloud computing.

Unit	Details	Lectures
I	Definitions, Components, Characteristics / Properties / Features, Cloud Computing Vs On Premise System, Evolutions of Cloud Computing, CC Opportunities and Challenges, Advantages and Disadvantages of Cloud Comp, Benefits and Risks in Cloud Computing Cloud Computing Architecture, Service Oriented architecture (SOA) and Cloud Computing Reference Architecture by IBM.	12
II	Introduction, Types of Virtualization, I/O Virtualization, Storage virtualization, Network Virtualization, Memory Virtualization, Virtualization and Cloud Computing.	12
III	Mobile Cloud and Cloud Security. Mobile Computing, Architecture, Need, Characteristics, Advantages and Disadvantages, Mobile cloud Applications, Issues or CSF for Mobile Cloud Computing.	12
IV	Cloud Security Fundamentals, Cloud Security Challenges, Cloud Security and Privacy, Vulnerability Assessment tools for Cloud computing, Virtualization Security, Management, Security and Privacy in Mobile Cloud, Virtualization Security and Challenges, Cloud Computing and Data Security Risk. Cloud Computing and Identity, The Cloud, Digital Identity, and Data Security,	12
V	Azure Web Services: Compute, Database, Networking, Storage Management.	12

Course Outcomes:
Upon the successful completion of this course, Learners will be able to:
● Understand the foundational concepts and principles of Machine Learning
● Apply supervised and unsupervised learning techniques, including classification algorithms and clustering algorithms
● Evaluate the performance of Machine Learning models using classification metrics, ROC/AUC curve analysis, and cross-validation techniques.
● Implement regression models (such as linear regression and logistic regression) and understand their applications in predictive analysis.
● Utilize dimensionality reduction techniques (Like PCA) for feature reduction and selection, and apply association rule mining algorithms (such as the Apriori algorithm) for discovering meaningful patterns in datasets.

Books and References:

● "Distributed and Cloud Computing" From Parallel Processing to the Internet of Things By Kai Hwang, Geoffrey C. Fox, Morgen Kaufmann Publication.
● Fundamentals of Computers and information system, Harendra nath Tiwari, Dr. Hemchand Jain, International Boo House Pvt. Ltd.

x

- Cloud Computing, Implementation, Management, and Security, Jhon W. Rittinghouse, James F. Ransome, CRC Press.
- Cloud Security and Privacy, Tim Mather, Subra Kumarswamy, Shahad Latif, O'REILLY, Shroff Publication and Distribution.
- Cloud Computing A Practical Approach, Anthony T. Velte, Toby J. Velte, Robert Elsenpeterm, McGraw Hill Education (India)
- Enterprise Cloud Computing, Technology, Architecture, Applications, Gautam Shroff, Cambridge University Press.
- Cloud Computing Bible. Barrie Sosinsky, Wiley Publication
- Mastering Cloud Computing, Rajkumar Buyya, Chritian Vecchiola, S Thamarai, McGraw Hill Education (India)

Evaluation scheme

I Continuous Internal Assessment (C.A.)– 40 Marks

1) Assessment 1 (20 Marks)

2) Class Test (20 Marks)

II Semester End Examination (SEE) - 60 Marks

QUESTION PAPER PATTERN OF (SEE)

Maximum Marks: 60 Marks

Time: 2 Hours

Note: 1) Attempt all Questions

2) Any three sub questions to be attempted. All sub questions carry equal marks

Question No	Particulars	Marks
Q -1 (Unit I)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 Marks
Q -2 (Unit II)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 Marks
Q -3 (Unit III)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 Marks
Q4 (Unit IV)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 marks
Q5 (Unit IV)	**Answer the following questions (Any 3)** A. Theory/Concept Based question B. Theory/Concept Based question C. Theory/Concept Based question D. Theory/Concept Based question	12 marks

x

Discipline Specific Elective (DSE) Courses

B. Sc. (Data Science and Analytics)		Semester – V	
Course Name: Introduction to Cloud Computing Practical		Course Code: 60510	
Periods per week (1 Period is 1 hour)		2	
Credits		2	
		Hours	Marks
Evaluation System	Practical Examination	$2^{1/2}$	50
	Internal	--	--

List of Practical	
1	Creating Azure Cloud Instances
2	Deploy a virtual machine on Azure instance
3	Connect to VM and explore file systems
4	Configure basic security
5	Create Cloud Database
6	Implement Azure Storage
7	Experiment with Azure ML

x

Practical Exam: 50 marks

Certified copy journal is essential to appear for the practical examination.

1.	**Practical Work**	**30 Marks**
2.	**Journal**	**10 Marks**
3.	**Viva**	**10 Marks**