

SRI VENKATESWARA UNIVERSITY
BCA Data Science Honours– W.E.F. 2025-26

Year	Semester	Course	Title of the Course	No. of Hrs /Week	No. of Credits
III	V	12	Data Mining and Data Wrangling	3	3
			Data Mining and Data Wrangling Lab	2	1
		13	Machine Learning using R	3	3
			Machine Learning using R Lab	2	1
		14	Cloud Computing	3	3
			Cloud Computing Lab	2	1
		15	Data visualization	3	3
			Data visualization Lab	2	1

SRI VENKATESWARA UNIVERSITY::TIRUPATI

BCA Data Science Honours

III Year V Semester

COURSE 12: Data Mining and Data Wrangling

(w.e.f. 2025-26)

Theory

Credits: 3

3 hrs/week

UNIT-I

Data warehouse: What is Data warehouse, Difference between operational database systems and data warehouses, Data warehouse Characteristics, Data warehouse Architecture and its Components. Multi- Dimensional: Tables and Spreadsheets to Data Cubes, Data Modeling, Schema Design, Star and Snow-Flake Schema, Fact Constellation, Fact Table. Data Warehouse Architecture: Three-Tier Data ware house architecture, Meta Data Repository, Types of OLAP servers ROLAP versus MOLAP versus HOLAP

UNIT- II

Introduction: What is data mining, Data Mining Functionalities, Classification of Data Mining systems, Data Mining Task Primitives, Integration of a Data Mining System with a Database or Data Warehouse System, Major issues in Data Mining Data Preprocessing: Need for Preprocessing the Data, Data Cleaning, Data Integration &Transformation, Data Reduction, Discretization and Concept Hierarchy Generation.

UNIT- III

Mining Frequent Patterns Associations: Frequent Pattern Mining, Market basket Analysis, Frequent Itemsets and Association Rules, Apriori Algorithm, Improving Efficiency of Apriori, Generating Rules for Frequent Itemsets, Mining various kinds of Association Rules: Mining Multi-dimensional Association Rules form Relational Data warehouses.

UNIT- IV

Classification and Prediction: What is Classification and Prediction, Preparing Data for Classification and Prediction, Classification techniques: Decision Trees Induction, Tree Pruning, scalability and Decision Tree Induction Rule Based Classification: Using IF-THEN classification, Rule Extraction from Decision Tree. Bayesian Classification: Naïve Bayesian classification, K- Nearest neighbor Classifiers, Case-Based Reasoning and Genetic Algorithm Prediction: Linear Regression, Non-Linear regression and Other Regression-based methods

UNIT- V

Cluster Analysis: Types of Data in Cluster Analysis, A Categorization of Major Clustering Methods, Partitioning Methods, Hierarchical Methods, Density-Based Methods, Grid-Based Methods, Model based Clustering Methods, Outlier Analysis.

TEXT BOOKS:

1. Data Mining- Concepts and Techniques- Jiawei Han, Micheline Kamber, Morgan Kaufmann Publishers, Elsevier, 2 Edition,2006.
2. Data Mining Techniques, Arun K Pujari, 3rd Edition, Universities Press.

REFERENCE BOOKS:

1. Data Mining Introductory and Advanced Topics, by Margaret Dunhum and Sridhar Pearson Edition
2. Data Warehousing Fundament's, Pualraj Ponnaiah, Wiley Student Edition

BCA Data Science Honours
III Year V Semester
COURSE 12: Data Mining and Data Wrangling Lab

Practical

Credits: 1

2hrs/week

LIST OF EXPERIMENTS:

1. Draw the **three-tier architecture** diagram, Label and explain: data sources, staging, warehouse, OLAP tools.
2. Draw.io to sketch Star Schema and Snowflake Schema for a sales database.
3. Perform basic data cleaning tasks on a sample dataset like Remove duplicates, Fill missing values ,Format dates and numbers.
4. Apply preprocessing like null handling and data type conversion.
5. Apriori Algorithm using Online Tools (or WEKA)
6. Simple Classification using Decision Tree in WEKA
7. Naive Bayes Classification in WEKA
8. K-Means Clustering using WEKA
9. Simple Regression using Excel
10. Outlier Detection using Box Plot in Excel

MODEL QUESTION PAPER
SRI VENKATESWARA UNIVERSITY::TIRUPATI
BCA Data Science Honours
III Year V Semester
COURSE 12: Data Mining and Data Wrangling
(w.e.f. 2025-26)

Time : 3 Hrs

Max Marks : 75

SECTION - A

Answer any Five of the following

5 X 3= 15 Marks

1. Short answer question from Unit-1
2. Short answer question from Unit-1
3. Short answer question from Unit-2
4. Short answer question from Unit-2
5. Short answer question from Unit-3
6. Short answer question from Unit-3
7. Short answer question from Unit-4
8. Short answer question from Unit-4
9. Short answer question from Unit-5
10. Short answer question from Unit-5

SECTION - B

Answer any Five of the following

5 X 12= 60 Marks

11. Long answer question from Unit-1
12. Long answer question from Unit-1
13. Long answer question from Unit-2
14. Long answer question from Unit-2
15. Long answer question from Unit-3
16. Long answer question from Unit-3
17. Long answer question from Unit-4
18. Long answer question from Unit-4
19. Long answer question from Unit-5
20. Long answer question from Unit-5

Note: The question paper setter is requested to set question paper based on a model question paper and ensure coverage across all units equally.

SRI VENKATESWARA UNIVERSITY::TIRUPATI
BCA Data Science Honours
III Year V Semester
COURSE 13: Machine Learning using R
(w.e.f. 2025-26)

Theory

Credits: 3

3 hrs/week

I. Learning Outcomes: After completing this course, the student will be able to

1. Understand the basic concepts such as decision trees and neural networks.
2. Ability to formulate machine learning techniques to respective problems.
3. Apply machine learning algorithms to solve problems of moderate complexity.
4. Use and program in the programming language R

UNIT – I

INTRODUCTION: Overview of Machine learning, machine learning types, well posed learning algorithms and issues in machine learning. **CLASSICAL DATA ANALYSIS:** Mean, variance, Regression- linear. **CONCEPT LEARNING:** Introduction, a concept learning task and learning as search, version space and candidate elimination algorithm, inductive bias.

UNIT – II

DECISION TREE LEARNING: - introduction, Representation, basic decision tree learning algorithm, space search in decision tree learning algorithm ,cross validation and over fitting.

NEURAL NETWORK LEARNING: Introduction, Perceptrons - Representational power of Perceptrons, Back propagation algorithm.

UNIT – III

INTRODUCTION TO R: Basic features of R, Benefits of R, data types in R, Installing R, Getting started with the RStudio IDE, Running R, Packages in R, variable names and assignment, operators, Input/output functions.

UNIT-IV

Preview of Some Important R Data Structures: Vectors, Character Strings, Matrices, Lists, Data Frames, and Classes.

Control Structure: Conditional statements, Loops, dates and times, functions, String manipulations.

UNIT-V

Vectors: Scalars, Vectors, Arrays and Matrices: Adding and Deleting Vector Elements, Obtaining the Length of a Vector, Vector Indexing, Common vector operations, Generating vector sequences, Repeating vector constants, using `all()` and `any()`, Vectorized operations.

TEXTBOOKS:

1. Tom Michel, Machine Learning, McGraw Hill, 1997.
2. The Art of R Programming by Norman Matloff, No starch press, SAN FRANCISCO, 2011.

REFERENCE BOOKS:

1. Machine Learning Methods in the Environmental Sciences, Neural Networks, William W Hsieh, Cambridge University Press.
2. An Introduction to R for Beginners by SASHA HAFNER, on AUG-2019.

BCA Data Science Honours
III Year V Semester
COURSE 13: Machine Learning using R

Practical

Credits: 1

2hrs/week

LIST OF EXPERIMENTS:

1. Write an R-Program to take input from the user.
2. Write an R-Program to compute a product of two integer values.
3. Write an R-Program to demonstrate working with operators (Arithmetic, Relational, Logical, Assignment operators).
4. Write an R Program to Check if a Number is Odd or Even
5. Write an R program to check whether the biggest of two values.
6. Write an R Program to check if the given Number is a Prime Number
7. Write an R Program to Find the Factorial of a Number
8. Write an R Program to Find the Fibonacci sequence Using Recursive Function.
9. Write an R Program to create a Vector and to access elements in a Vector.
10. Write an R Program to create a Matrix from a Vector using dim() function.

MODEL QUESTION PAPER
SRI VENKATESWARA UNIVERSITY::TIRUPATI
BCA Data Science Honours
III Year V Semester
COURSE 13: Machine Learning using R
(w.e.f. 2025-26)

Time : 3 Hrs

Max Marks : 75

SECTION - A

Answer any Five of the following

5 X 3= 15 Marks

21. Short answer question from Unit-1
22. Short answer question from Unit-1
23. Short answer question from Unit-2
24. Short answer question from Unit-2
25. Short answer question from Unit-3
26. Short answer question from Unit-3
27. Short answer question from Unit-4
28. Short answer question from Unit-4
29. Short answer question from Unit-5
30. Short answer question from Unit-5

SECTION - B

Answer any Five of the following

5 X 12= 60 Marks

31. Long answer question from Unit-1
32. Long answer question from Unit-1
33. Long answer question from Unit-2
34. Long answer question from Unit-2
35. Long answer question from Unit-3
36. Long answer question from Unit-3
37. Long answer question from Unit-4
38. Long answer question from Unit-4
39. Long answer question from Unit-5
40. Long answer question from Unit-5

Note: The question paper setter is requested to set question paper based on a model question paper and ensure coverage across all units equally.

SRI VENKATESWARA UNIVERSITY::TIRUPATI

BCA Data Science Honours

III Year V Semester

COURSE 14 : CLOUD COMPUTING

(w.e.f. 2025-26)

Theory

Credits: 3

3 hrs/week

Course Objectives:

1. This course introduces the basic principles of cloud computing, Cloud Computing models and Services, Virtualization Techniques, Resource Sharing, Load Balancing and Security issues in Cloud Computing.
2. It will also equip the students to understand major industry players in the public cloud domain.

Course Outcomes: Upon Completion of the course, the students will be able to

1. Understand the basic principles of cloud computing.
2. Analyze different types of cloud services – Cloud Computing Models and Cloud Computing Services.
3. Understand Virtualization techniques
4. Use different Cloud Platforms
5. Learn different File systems and Security issues in Cloud Computing Platforms

Unit-I: Introduction to Cloud Computing, Evaluation and Enabling Technologies, Benefits and Challenges, Cloud Computing Model

CASE STUDY: Identify different types of Public Cloud, Private Cloud, and Hybrid Cloud

Unit-II: Cloud Computing Services, Elements of Cloud Security model, Cloud Security reference model, Examining Cloud Security against traditional computing, Security Policy

CASE STUDY: Identify various real-time examples of Cloud Computing Services and work with them

Unit-III: Virtualization basics, Machine or Server Level Virtualization, Major Server virtualization products and vendors, Types of Virtualization, Advantages and Disadvantages of Virtualization, Virtualization Security Threats

CASE STUDY: Install any Virtualization Software and Perform any task

Unit-IV: Resource Pooling and Sharing, Multi-tenancy, Resource Provisioning, Scaling, Scaling strategies in Cloud, Types of Scaling, Load balancing in Cloud, Categories of Load Balancing, Service Oriented Architecture (SOA), Elements of a service, Benefits of SOA

CASE STUDY: Perform different operations using Google Cloud and Amazon Elastic Compute Cloud (EC2)

Unit-V: Cloud Native File System, Storage types, Relational DBMS in Cloud, Non-Relational DBMS in Cloud, Cloud Security, Treats to cloud security, Infrastructure Security in Cloud, Information Security in Cloud, Identity Management and Access Control

CASE STUDY: Identify and work with Popular cloud storages for developers, Popular General Purpose Cloud Storages

Prescribed Text Books:

1. **Cloud Computing: Sandeep Bhowmik , Cambridge University Press**
2. **CLOUD COMPUTING:Principles and Paradigms by Rajkumar Buyya, James Broberg, Andrzej Goscinski**
3. **The Definite Guide to Cloud Computing by Dan Sullivan**

BCA Data Science Honours
III Year V Semester
COURSE 14 : CLOUD COMPUTING

Practical

Credits: 1

2hrs/week

LIST OF EXPERIMENTS:

Course Outcomes: Upon Completion of the course, the students will be able to

1. Working with Virtual machines.
2. Run simple programs on virtual box.
3. Acquire knowledge in Google App Engine
4. Get practical exposure on CloudSim

Experiments:

1. Install Virtualbox / VMware Workstation with different flavours of linux or windows OS on top of windows10 or 11.
2. Install a C compiler in the virtual machine created using virtual box and execute Simple Programs
3. Install Google App Engine. Create hello world app and other simple web applications using python/java.
4. Use GAE launcher to launch the web applications.
5. Simulate a cloud scenario using CloudSim and run a scheduling algorithm that is not present in CloudSim.
6. Find a procedure to transfer the files from one virtual machine to another virtual machine.
7. Find a procedure to launch virtual machine using trystack (Online Openstack Demo Version)

Web Resources:

1. [Cloud Computing Fundamentals Labs - Cloud Academy](#)

MODEL QUESTION PAPER
SRI VENKATESWARA UNIVERSITY::TIRUPATI
BCA Data Science Honours
III Year V Semester
COURSE 14 : CLOUD COMPUTING
(w.e.f. 2025-26)

Time : 3 Hrs

Max Marks : 75

SECTION - A

Answer any Five of the following

5 X 3= 15 Marks

1. Short answer question from Unit-1
2. Short answer question from Unit-1
3. Short answer question from Unit-2
4. Short answer question from Unit-2
5. Short answer question from Unit-3
6. Short answer question from Unit-3
7. Short answer question from Unit-4
8. Short answer question from Unit-4
9. Short answer question from Unit-5
10. Short answer question from Unit-5

SECTION - B

Answer any Five of the following

5 X 12= 60 Marks

11. Long answer question from Unit-1
12. Long answer question from Unit-1
13. Long answer question from Unit-2
14. Long answer question from Unit-2
15. Long answer question from Unit-3
16. Long answer question from Unit-3
17. Long answer question from Unit-4
18. Long answer question from Unit-4
19. Long answer question from Unit-5
20. Long answer question from Unit-5

Note: The question paper setter is requested to set question paper based on a model question paper and ensure coverage across all units equally.

SRI VENKATESWARA UNIVERSITY::TIRUPATI

BCA Data Science Honours

III Year V Semester

COURSE 15: Data visualization

(w.e.f. 2025-26)

Theory

Credits: 3

3 hrs/week

Aim and objectives of Course:

- To know the importance of data Visualization in the world of DataAnalytics and Prediction
- To know the important libraries in Tableau
- To get equipped with Tableau Tool

Learning outcomes of Course:

- Students should be able to visualize data through seven stages of data analysisprocess
- Should be able to do explanatory and hybrid types of data visualization
- Should be able to understand various stages of visualizing data

UNIT I:

Creating Visual Analytics with tableau desktop, connecting to your data-How to Connect to your data, What are generated Values? Knowing when to use a direct connection, Joining tables with tableau, blending different data sources in a single worksheet.

UNIT II:

Building your first Visualization- How Me works- Chart types, Text Tables, Maps, bar chart, Line charts, Area Fill charts and Pie charts, scatter plot, Bullet graph, Gantt charts, Sorting data in tableau, Enhancing Views with filters, sets groups and hierarchies.

UNIT III:

Creating calculations to enhance your data- What is aggregation, what arecalculated values and table calculations, Using the calculation dialog box to create,Building formulas using table calculations, Using table calculation functions

UNIT IV:

Using maps to improve insights-Create a Standard Map View, Plotting your ownlocations on a map, Replace Tableau's standard maps, Shaping data to enable Point-to-Point mapping.

UNIT V:

Developing an Adhoc analysis environment- generating new data with forecasts, providing self evidence adhoc analysis with parameters, Editing views in tableau Server.

TEXT BOOKS:

1. Tableau your data-Daniel G. Murray and the Inter works BI team, Wiley Publications
2. Tableau Data Visualizaton Cookbook, AshutoshNandeshwar, PACKT publishing.
3. Storytelling with Data: A Data Visualization Guide for BusinessProfessionals by Cole NussbaumerKnafllic (2014)
4. ggplot2: Elegant Graphics for Data Analysis by Hadley Wickham (2009)

BCA Data Science Honours
III Year V Semester
COURSE 15 : Data visualization

Practical

Credits: 1

2hrs/week

LIST OF EXPERIMENTS:

1. Connect to data Sources
2. Create Univariate Charts
3. Create Bivariate and Multivariate charts
4. Create Maps
5. Calculate user-defined fields
6. Create a workbook data extract
7. Save a workbook on a Tableau server and web
8. Export images, data.



MODEL QUESTION PAPER
SRI VENKATESWARA UNIVERSITY::TIRUPATI
BCA Data Science Honours
III Year V Semester
COURSE 15: Data visualization
(w.e.f. 2025-26)

Time : 3 Hrs

Max Marks : 75

SECTION - A

Answer any Five of the following

5 X 3= 15 Marks

41. Short answer question from Unit-1
42. Short answer question from Unit-1
43. Short answer question from Unit-2
44. Short answer question from Unit-2
45. Short answer question from Unit-3
46. Short answer question from Unit-3
47. Short answer question from Unit-4
48. Short answer question from Unit-4
49. Short answer question from Unit-5
50. Short answer question from Unit-5

SECTION - B

Answer any Five of the following

5 X 12= 60 Marks

51. Long answer question from Unit-1
52. Long answer question from Unit-1
53. Long answer question from Unit-2
54. Long answer question from Unit-2
55. Long answer question from Unit-3
56. Long answer question from Unit-3
57. Long answer question from Unit-4
58. Long answer question from Unit-4
59. Long answer question from Unit-5
60. Long answer question from Unit-5

Note: The question paper setter is requested to set question paper based on a model question paper and ensure coverage across all units equally.

