SRI VENKATESWARA UNIVERSITY
BCA DEGREE COURSE IN DATA SCIENCE
III- SEMESTER

(Syllabus under CBCS w.e.f. 2021-22)

Semester	Course Code	Course Title	Hours/Week	Hours	Credits
III	C7	Probability and Statistics	4	60	3

Course Objective

1. The main objective of this course is to provide students with the foundations of probabilistic and statistical analysis mostly used in varied applications in engineering and science like disease modelling, climate prediction and computer networks etc.

Course Outcomes

After successful completion of this course, the student will be able to;

1. Able to know the concepts of the set theory and operations in sets.
2. Knowledge to conceptualize the probabilities of events including frequent and axiomatic approach. Simultaneously, they will learn the notion of conditional probability.
3. Knowledge related to concept of discrete and continuous random variables and their probability distributions including expectation and moments,
4. Knowledge related to concept of random variable, Probability mass function and probability density function.
5. Knowledge related to concept of Mathematical expectation.

UNIT-I

Set theory for Probability: Definition of Set- Types of Sets-Union of Sets-Intersection of Sets-Venn diagrams- Operations on Sets-Complement of Set-Distributive Laws-De'Morgan's Laws

UNIT -II

Introduction to Probability: Basic concepts of Probability, random experiment, trial, outcome, sample space, event, mutually exclusive event, equally likely events, favourable events, classical, statistical and axiomatic definitions of probability.

UNIT-III

Probability theorems: Conditional probability, Independent events Addition and multiplication theorems of probability for 2 events (Statement and proof), Addition and multiplication theorems of probability for n events statements only, Bayees theorem Statement and its applications.

UNIT-IV

Random variable: Definition of Random variable, discrete and continuous random variables, functions of random variable, probability mass function, probability density function, distribution function and its properties, For a given probability mass function calculation of mean and variance, For a given probability density function calculation of mean and variance.

UNIT-V

Mathematical expectation : Mathematical Expectation of random variable and function of random variable, Moments and covariance using mathematical expectation, with examples. Addition and multiplication theorem on expectation. Definitions of M.G.F, P.G.F,C.F and their properties, and applications.

Note:

1. Concentration on numerical problems only.

2. Proofs of theorems and Derivations of expressions are omitted.

Text Book:

1. Statistical Methods by S.P. Gupta.

2. Fundamentals of Mathematical statistics - S.C. Gupta & V.K. Kapoor.

Reference Books:

1. Samba Vyatka - Telugu Academy.

2. Fundamentals of statistics - Goon, Gupta and Das Gupta.

SRI VENKATESWARA UNIVERSITY
BCA DEGREE COURSE IN DATA SCIENCE
III- SEMESTER

(Syllabus under CBCS w.e.f. 2021-22)

Semester	Course Code	Course Title	Hours/Week	Hours	Credits
III	C7-P	Probability and Statistics Lab	2	30	1

1. Operation on sets.
2. Distributive Laws-De'Morgan's Laws.
3. Basic concepts of probability.
4. Conditional Probability.
5. Addition and Multiplication theorem problems.
6. Payees theorem applications.
7. Random variable
8. Mathematical Expectation.

SRI VENKATESWARA UNIVERSITY
BCA DEGREE COURSE IN DATA SCIENCE
III- SEMESTER

(Syllabus under CBCS w.e.f. 2021-22)

PROBABILITY AND STATISTICS

(Statistical tables and Electronic Calculators are allowed)

MODEL QUESTION PAPER

TIME: 3 HOURS **MAX.MARKS:75**

SECTION-A

ANSWER ANY FIVE QUESTIONS. EACH QUESTION CARRIES 5 MARKS 5 X 5 = 25

1. If $A=\{5,6,7,8,9\}, B=\{2,4,6,8,10\}$, find $(i)\, A\cup B (ii)\, A\cap B$ and examine they are equal or not

2. If $A=\{3,4,5,6,7\}, B=\{5,6,7,8,9\}$, find $(i)\, A-B (ii)\, B-A$ and examine they are equal or not.

3. Define sample space and random experiment.

4. Define Equally likely events and Exhaustive events.

5. If A and B are events such that $P(A)=\frac{3}{4}$ *and* $P(B)=\frac{5}{8}$ then show that $\frac{3}{8}\le P(A\cap B)\le\frac{5}{8}$

6. State Addition theorem for "n" events in probability.

7. Define Random variable.

8. A random variable X has the following probability function. Find 'K' value :

$X=x$	-2	-1	0	1	2	3
$P(X=x)$	0.1	K	0.2	2k	0.3	K

9. Define Mathematical Expectation.

10. Write Addition and multiplication theorems on mathematical expectations.

SECTTION-B

ANSWER ANY FIVE QUESTIONS.EACH QUESTION CARRIES 10 MARKS 5X10=50

11. If $\mu=\{1,3,5,7,9,11\}, A=\{3,5,7,11\}, B=\{1,5,7,11\}$ then find

$(i)A\cup\mu(ii)B\cup\mu(iii)A\cap\mu(iv)B\cap\mu(v)A^1(vi)B^1$

12. If $A=\{1,2,3,5,8\}, B=\{5,6,7,8\}C=\{1,3,5,7\}$ prove that $A\cup(B\cup C)=(A\cup B)\cup C$.

13. Define the following (i) trial (ii) sample space (iii) event (iv) out come.

14. Write classical, statistical and axiomatic definitions of probability.

15. State and prove addition theorem for two events.

16. State Bayees theorem and its applications.

17. Explain properties of distribution function.

18. A random variable X has the following probability function.

$X=x$	0	1	2	3	4	5	6	7
$P(X=x)$	0	K	$2K$	$2K$	$3K$	K^2	$2K^2$	$7K^2+K$

Find $(i)K$ (ii) Mean (iii) Variance

19.

x :	**2**	**3**	**6**
P(x) :	Y_2	Y_4	Y_4

Find (i) E(x) (ii) V(x).

20. Define M.G.F and write its applications.

<u>Instruction to Paper Setter:</u>

Paper Setter must select <u>TWO</u> Short Questions and <u>TWO</u> Essay Questions from Each Unit

SRI VENKATESWARA UNIVERSITY
BCA DEGREE COURSE IN DATA SCIENCE
III- SEMESTER

(Syllabus under CBCS w.e.f. 2021-22)

Semester	Course Code	Course Title	Hours/Week	Hours	Credits
III	C8	Data Structures	4	60	4

Course Objectives

To introduce the fundamental concept of data structures and to emphasize the importance of data structures in developing and implementing efficient algorithms.

Course Learning Outcomes:
Upon successful completion of the course, a student will be able to:

1. Understand how arrays, records, linked structures, stacks, queues, trees, and graphs are represented in memory and used by algorithms.
2. Learn the benefits of dynamic and static data structures implementations
3. Design and develop various programs using data structures
4. Demonstrate different methods for traversing trees
5. Develop ability to implement different Sorting and Search methods
6. Understand the concept of recursion and how it can be implemented using a stack .

UNIT – I

INTRODUCTION TO DATA STRUCTURE: Definition, Data Types, Abstract Data Types (ADT), classification of data structure - primitive & non-primitive data structures, Linear and Non-linear data structures
ARRAYS: Definition, one dimensional array, two dimensional arrays, Applications, pointers.
LINKED LIST: Definition, linked list ADT, single linked list, double linked list, circular linked list, comparison of linked list with Arrays.

UNIT – II

STACKS: Definition, Stack as an ADT & Operations on stack, Applications of stack, Representation of stack.
QUEUES: Definition, Queue as an ADT & Operations on Queue, Application of Queues, Representation of Queues, Various Queue Structures: circular Queue, DEQueue.

UNIT – III

TREES: Definition, Basic Tree Terminology. **Binary Tree** – Definition, Properties of Binary Trees, Types of Binary Trees, Representation of Binary Tree, Binary Tree Traversals. **Binary Search Tree (BST)** – Definition, Operations on a Binary Search Tree, Examples of BST.

UNIT - IV

GRAPHS: Definition, Basic Graph Terminology, Representation of Graphs, Graph Traversal – DFS and BFS. Topological sort, Shortest Path problem, Minimum Spanning Tree.

UNIT – V

SORTING: Definition, Sorting methods - Bubble Sort, Selection Sort, Quick Sort, Insertion Sort, Merge Sort.

SEARCHING: Definition, searching methods - Linear or Sequential Search, Binary Search.

Text Books:

1. "Classic Data Structures ", by DEBASIS SAMANTHA 2nd EDITION, PHI publications , 2009
2. "Data Structures and Algorithms", by NARASIMHA KARUMANCHI , CAREERMONK Publications , 2017

Reference Books:

1. Data structures by Lipchitz, McGraw Hill Education
2. Fundamentals of Data Structures in C by Shane Horowitz, University Press
3. Data Structures And Algorithms by Alfred V Ahoy and John E Hop croft and Jeffrey D Pullman, Pearson Education
4. "Data Structures through C", Yashavant Kanetkar, BPB Publications

SRI VENKATESWARA UNIVERSITY
BCA DEGREE COURSE IN DATA SCIENCE
III- SEMESTER

(Syllabus under CBCS w.e.f. 2021-22)

Semester	Course Code	Course Title	Hours/Week	Hours	Credits
III	C8-P	Data Structures Lab	2	30	1

1. Program to generate Fibonacci series using recursion
2. Program for implementation of stack using arrays.
3. Program for implementation stack using linked list.
4. Program for implementation queue using array.
5. Program for implementation queue using linked list.
6. Program for implementation of circular queue.
7. Program for bubble sorting
8. Program for selection sorting.
9. Program for insertion sorting.
10. Program for merge sorting.
11. Program for linear searching.
12. Program for binary searching.
13. Program for Binary search tree operations.
14. Program to implement Graph traversal using DFS
15. Program to implement Graph traversal using BFS

SRI VENKATESWARA UNIVERSITY
BCA DEGREE COURSE IN DATA SCIENCE
III- SEMESTER

(Syllabus under CBCS w.e.f. 2021-22)

DATA STRUCTURES

MODEL QUESTION PAPER

Time: 3 Hours Max. Marks : 75

SECTION-A

Answer any FIVE of the following Questions: **(5 x 5= 25 Marks)**

1. Explain about Abstract Data Types.
2. Write about one dimensional arrays
3. What are the applications of stack? Explain
4. Write the differences between stack and queue
5. Write about fully and complete binary Tree
6. Write about tree terminologies
7. What is Graph? Write the representation of adjacency matrix
8. Briefly explain shortest path problem
9. What is sorting? Write an algorithm for Bubble sort
10. What is searching? Explain about sequential search.

SECTION – B

Answer any FIVE of the following Questions **(5 × 10 =50 Marks)**

11. Write the classification of Data structures with diagram
12. Explain about primitive data structures.
13. What are the operations performed on stacks
14. Write the algorithm for Queue insert and Queue Delete
15. What is Binary tree? Write the Tree traversal techniques in Binary tree.
16. What is Binary search tree? Give an example
17. Write about Minimum Spanning tree?
18. Explain BFS with an example
19. What is insertion sort? Explain the procedure with an Example.
20. Explain about Binary search.

Instruction to Paper Setter:

Paper Setter must select TWO Short Questions and TWO Essay Questions from Each Unit

SRI VENKATESWARA UNIVERSITY
BCA DEGREE COURSE IN DATA SCIENCE
III- SEMESTER

(Syllabus under CBCS w.e.f. 2021-22)

Semester	Course Code	Course Title	Hours/Week	Hours	Credits
III	C9	Introduction to Data Science	4	60	4

Course Objectives

1. Provide you with the knowledge and expertise to become a proficient data scientist.
2. Demonstrate an understanding of statistics and machine learning concepts that are vital for data science;
3. Produce R code to statistically analyze a dataset.
4. Critically evaluate data visualizations based on their design and use for communicating stories from data

Course Outcomes

After completing this course, the student will be able to:

1. Explain how data is collected, managed and stored for data science.
2. Understand the key concepts in data science, including their real-world applications and the toolkit used by data scientists.
3. Implement data collection and management scripts using database.

UNIT-I

Introduction to Data Science: Definition, benefits and uses of data science and big data.
Facets of Data: Structured data, unstructured data, natural language, machine generated data, network data, audio, images and video streaming data.
Data science process: overview of data science process, defining the goal, Retrieving data, data preparation, data exploration, build the models, cleaning and transforming data, presentation and automation.

UNIT-II

DATA: Definition, characteristics of data, classification of digital data.
The Data Science Fundamentals: Distributed file system, data integration frame work, machine learning framework, system deployment, security.
Data Mining: definition, languages for data science, collection data –hunting, logging, scraping, cleaning data –error vs. artifacts, data compatibility, dealing with missing values, outlier detection.

UNIT-III

BIG DATA: Definition, Evolution of big data and its importance, four V's in big data, Drivers for Big data, Big data analytics, Big data applications, designing data architecture, Big data Vs Little data

UNIT-IV

Machine Learning: Definition, Applications of machine learning in data science, Types of Machine Learning (Degree) - supervised learning, semi supervised learning, un-supervised learning, Linear regression, Decision Tree classifier – constructing decision Tree, Bayes - Naive Bayes

UNIT-V

Data Visualization: Definition, importance of data visualization in data science, Exploratory Data analysis -confronting new data set, visualization tools, developing a visualization aesthetic – maximizing data link ratio, proper scaling and labeling, effective use of color and shading, the power of repetition.
Chart Types: Tabular data, dot and line plots, scatter plots, bar plots and pie charts.

Text Books:

1. Introducing Data Science by Davy Coleen , Arno D.B.Meysman and Mohamed Ali, Published by Manning
2. Steven S.Skiena, The Data Science Design Manual, Published by Springer. Nature.

Reference Books:

3. Cathy O'Neil and Rachel Schutt. Doing Data Science, Straight Talk from The Frontline.O'Reilly.
4. Jure Leskovek, AnandRajaraman and Jeffrey D.Ullman. Mining of Massive Datasets. v2.1, Cambridge University Press.

SRI VENKATESWARA UNIVERSITY
BCA DEGREE COURSE IN DATA SCIENCE
III- SEMESTER

(Syllabus under CBCS w.e.f. 2021-22)

Semester	Course Code	Course Title	Hours/Week	Hours	Credits
III	C9-P	Introduction to Data Science Lab	2	30	1

1. Write and implement algorithm for Decision Tree classification.
2. Write and implement algorithm for Navie Bayes.
3. Write and implement algorithm for Back Propagation
4. Write Steps to Installation of R language
5. Write a program in R to draw scatter chart for data visualization.
6. Write a program in R to draw pie chart for data visualization.
7. Write program in R to import the data from outside of the sources
8. Write steps to install Python language
9. Write a program in Python to find the mean of list of numbers
10. Write a program in Python to find the variance of numbers

SRI VENKATESWARA UNIVERSITY
BCA DEGREE COURSE IN DATA SCIENCE
III- SEMESTER

(Syllabus under CBCS w.e.f. 2021-22)

INTRODUCTION TO DATA SCIENCE

MODEL QUESTION PAPER

Time: 3 Hours Max. Marks : 75

SECTION-A

Answer any FIVE of the following Questions: **(5 x 5= 25 Marks)**

1. What is Data science? Write benefits of data science
2. Contrast between structured and unstructured data
3. Write the characteristics of data
4. Explain machine learning framework
5. Define Big Data. Write its importance
6. Write various drivers for Big data
7. Explain the supervised machine learning
8. Briefly explain decision tree classifier
9. What is Visualization? Write its importance in data science
10. Define Chart. Explain types of charts.

SECTION – B

Answer any FIVE of the following Questions **(5 × 10 =50 Marks)**

11. Write about various facets of data
12. Explain the process of data science
13. Write and explain distributed file system in data science
14. Define Data Mining. Explain the different stages of data mining
15. Explain characteristics of big data
16. Write about Big data applications
17. Explain the applications of machine learning in data science
18. Write about different types of machine learning
19. Elaborate various tools of visualization
20. Explain the process of developing a visualization aesthetic

Instruction to Paper Setter:

Paper Setter must select TWO Short Questions and TWO Essay Questions from Each Unit