



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

Department of Computer Science and Engineering (AIML)

MACHINE LEARNING LAB

II B.TECH-II SEMESTER

MLRS-BT25 REGULATION



A.Y.2026-2027



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

INDEX

S.No	Details	Pg. No
1	Certificate	IV
2	Preface	V
3	Acknowledgement	VI
4	General Instructions	VII
5	Safety Measures	8–9
6	Vision and Mission of the Institute and the Department along with PEOs of the Program	10–12
7	Course Descriptor	—
8	Previous CO Attainment and Target for Present Semester	—
9	Academic Calendar	
10	Lab Timetable	
11	Syllabus Copy	16
12	Virtual Lab Details (If applicable)	18
13	Lab Planner	21
14	Rubrics used to assess learning's in laboratories	22
List of Experiments		
1	Write a Python program to compute Central Tendency Measures: Mean, Median, Mode, and Measures of Dispersion: Variance and Standard Deviation.	25
2	Study Python basic libraries such as Statistics, Math, NumPy, and SciPy.	27
3	Study Python libraries for machine learning applications such as Pandas and Matplotlib.	30
4	Write a Python program to implement Simple Linear Regression.	33
5	Implement Multiple Linear Regression for House Price Prediction using sklearn.	36
6	Implement Decision Tree using sklearn and perform parameter tuning.	39
7	Implement K-Nearest Neighbors (KNN) using sklearn.	42
8	Implement Logistic Regression using sklearn.	45
9	Implement K-Means Clustering.	48
10	Perform performance analysis of classification algorithms on a specific dataset.	51
Open Ended Experiments		
1	Implement Multiple Linear Regression for House Price Prediction using sklearn.	
2	Implement Decision Tree using sklearn and perform parameter tuning.	
3	Implement K-Nearest Neighbors (KNN) using sklearn.	
4	Implement Logistic Regression using sklearn.	
5	Implement K-Means Clustering.	



MARRI LAXMAN REDDY **INSTITUTE OF TECHNOLOGY AND MANAGEMENT**

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

S.No	Details	Pg. No
6	Perform performance analysis of classification algorithms on a specific dataset.	
7	Implement Support Vector Machine (SVM) for sample data.	
8	Write a Python program to implement Hidden Markov Models.	
9	Implement Linear Discriminant Analysis (LDA).	



MARRI LAXMAN REDDY
INSTITUTE OF TECHNOLOGY AND MANAGEMENT
(AN AUTONOMOUS INSTITUTION)
(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)
Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

Department of Computer Science and Engineering
(AIML)

CERTIFICATE

This is to certify that this manual is a Bonafide record of practical work carried out in **MACHINE LEARNING LAB** for the **B.Tech Computer Science Engineering (AI&ML) IV Semester** Programme during the academic year **2026–2027**.

This manual has been prepared by **Mrs. K UDAYA SRI (Assistant Professor)**, Department of Computer Science Engineering (AI&ML), with my own efforts and to the best of our knowledge.

Signature of Lab Faculty

Signature of HOD



MARRI LAXMAN REDDY **INSTITUTE OF TECHNOLOGY AND MANAGEMENT**

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

PREFACE

This Lab Manual entitled “*Machine Learning Laboratory*” is intended for II B.Tech II Semester Computer Science and Engineering (AIML) students of Marri Laxman Reddy Institute of Technology and Management, Dundigal, Hyderabad.

The laboratory provides hands-on experience in fundamental machine learning concepts using Python, including statistical analysis, data preprocessing, and visualization with libraries such as NumPy, SciPy, Pandas, and Matplotlib. Students implement key supervised and unsupervised algorithms such as Linear Regression, Logistic Regression, Decision Trees, KNN, SVM, and K-Means Clustering.

This course enhances student’s analytical, programming, and problem-solving skills, preparing them for real-world applications and advanced studies in machine learning and artificial intelligence.

By,

Mrs.K.Udaya Sri



ACKNOWLEDGEMENT

It was a valuable and enriching experience working on the *Machine Learning Laboratory* manual. I would like to express my sincere gratitude to **Mrs.K.Udaya Sri**, Assistant Professor, Department of Computer Science & Engineering (AI&ML), Marri Laxman Reddy Institute of Technology and Management, for her constant guidance and technical support in preparing this document.

I extend my heartfelt thanks to **Dr. B. Ravi Prasad**, Head of the Department, Computer Science & Engineering (AI&ML), for his encouragement and for providing me the opportunity to prepare this *Machine Learning Laboratory* manual.

I am deeply indebted to **Dr. B. Ravi Prasad**, Dean Academics, Marri Laxman Reddy Institute of Technology and Management, for his continuous support, valuable suggestions, and scholarly guidance throughout this work.

I also express my sincere thanks to **Dr. P. Sridhar**, Director, and **Dr. R. Murali Prasad**, Principal, Marri Laxman Reddy Institute of Technology and Management, for giving me this wonderful opportunity to prepare the *Machine Learning Laboratory* manual.

Finally, I would like to thank all the faculty members of the Department of Computer Science & Engineering for their encouragement and support in successfully completing this work.

By,
Mrs.K.Udaya Sri

Department of Computer Science & Engineering (AI&ML)



GENERAL INSTRUCTIONS

1. Students are instructed to come to the *Machine Learning Laboratory* on time. Latecomers will not be entertained.
2. Students should be punctual. Experiments missed due to late arrival will not be repeated.
3. Students are expected to come prepared with the experiments to be performed during the lab session.
4. Students must display their identity cards before entering the laboratory.
5. Students are strictly prohibited from bringing mobile phones into the laboratory.
6. Any damage or loss of system components such as keyboard or mouse during the lab session will be the responsibility of the student concerned, and a penalty will be imposed.
7. Students should update their records and observation books session-wise. Before leaving the lab, they must get their observation book signed by the faculty.
8. Students should submit their lab records in the next lab session to the concerned faculty for correction and return.
9. Students should not move around unnecessarily during the lab session.
10. In case of any emergency, students must obtain written permission from the concerned faculty member.
11. Faculty members have the authority to suspend any student from the lab session on disciplinary grounds.
12. Students should not copy outputs from others and must record their own results.



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
(AI&ML)

SAFETY MEASURES

To ensure the safe and efficient use of the Computer Science and Engineering (AI&ML) laboratory, all students must strictly adhere to the following safety guidelines:

1. General Conduct

- Maintain silence and discipline during lab sessions.
- Do not bring food, drinks, or chewing gum into the laboratory.
- Use lab resources responsibly and follow all instructions given by the instructor or lab assistant.

2. Electrical Safety

- Do not touch electrical switches, sockets, or plugs with wet hands.
- Avoid overloading power sockets with unauthorized devices.
- Immediately report any loose connections, sparks, or unusual noises from equipment.

3. Computer and Equipment Handling

- Handle all computer systems, keyboards, mice, and peripherals with care.
- Do not attempt to open or tamper with any hardware components.
- Use only the assigned system; do not switch systems without permission.

4. Software and Data Safety

- Use only authorized software installed by the lab administrator.
- Do not attempt to install, uninstall, or modify any software without approval.
- Save your work frequently and maintain backups of important files.



5. Cyber security and Network Usage

- Keep your login credentials confidential.
- Do not attempt to access restricted websites or servers.
- Avoid activities such as hacking, gaming, or using pirated content.

6. Emergency Preparedness

- Be familiar with the location of emergency exits, fire extinguishers, and first aid kits.
- In case of fire, electrical hazards, or any emergency, remain calm and inform the lab instructor immediately.
- Follow the evacuation procedures as instructed.

7. Post-Lab Procedures

- Log out of your session and shut down the system properly after use.
- Leave your workstation clean and organized.
- Return any borrowed materials or equipment to their proper place.

8. Hygiene and Cleanliness

- Wash or sanitize your hands before and after using shared devices.
- Do not write on or place unnecessary items on the workstation.
- Report any spills or cleanliness issues to the lab staff.



MARRI LAXMAN REDDY **INSTITUTE OF TECHNOLOGY AND MANAGEMENT**

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

VISION & MISSION OF THE INSTITUTE

Vision of the Institute:

To be a globally recognized institution that fosters innovation, excellence, and leadership in education, research, and technology development, empowering students to create sustainable solutions for the advancement of society.

Mission of the Institute:

- To foster a transformative learning environment that empowers students to excel in engineering, innovation, and leadership.
- To produce skilled, ethical, and socially responsible engineers who contribute to sustainable technological advancements and address global challenges.
- To shape future leaders through cutting-edge research, industry collaboration, and community engagement.



VISION & MISSION OF THE DEPARTMENT

Department Vision:

To nurture globally competent professionals in Artificial Intelligence and Machine Learning through excellence in education, research, and innovation, committed to developing sustainable and impactful solutions for the betterment of society.

Department Mission:

- To provide a transformative learning environment that equips students with in-depth knowledge and practical skills in Artificial Intelligence and Machine Learning, fostering innovation, leadership, and lifelong learning.
- To advance AI and ML through cutting-edge research, strong industry collaboration, and community engagement, preparing students to address real-world challenges on a global scale.
- To produce competent and ethical AI professionals who contribute to technological progress while addressing societal and environmental challenges with sustainable solutions.
- To foster a research-driven culture by partnering with industry and academia, encouraging entrepreneurship, and engaging in community-centered technology development.



Program Educational Objectives (PEOs)

Professional Competence:

Graduates will possess strong theoretical and practical knowledge in Artificial Intelligence and Machine Learning, enabling them to solve complex real-world problems, pursue higher education, or excel in professional careers.

Innovation and Research Orientation:

Graduates will engage in innovative practices, cutting-edge research, and contribute to the advancement of AI and ML technologies through collaboration with industry and academia.

Leadership and Lifelong Learning:

Graduates will exhibit leadership qualities, effective communication, and teamwork skills, and will continuously upgrade their knowledge to adapt to evolving technological landscapes.

Entrepreneurial and Community Engagement:

Graduates will leverage entrepreneurial skills and a sense of civic responsibility to create AI-driven solutions that benefit local and global communities.

Course Outcomes (Cos)

CO No.	Course Outcome
CO1	To Perform statistical analysis and compute central tendency and dispersion measures using Python Programming.
CO2	To Utilize Python libraries such as NumPy, SciPy, Pandas, and Matplotlib for data manipulation, Visualization, and Analysis.
CO3	To Implement and evaluate supervised machine learning including Linear Regression, Logistic Regression, Decision Trees, and KNN
CO4	To Apply unsupervised learning techniques such as K-Means Clustering to extract patterns from datasets.
CO5	To analyze and compare the performance of machine learning models on real-world datasets to draw meaningful insights.



Program Outcomes (POs)

PO:

PO1.Engineering Knowledge:

Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.

PO 2. Problem Analysis:

Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

PO 3.Design/Development of Solutions:

Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

PO 4. Conduct Investigations of Complex Problems:

Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

PO 5. Modern Tool Usage:

Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

PO 6. The Engineer and Society:

Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.



MARRI LAXMAN REDDY **INSTITUTE OF TECHNOLOGY AND MANAGEMENT**

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

PO 7. Environment and Sustainability:

Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

PO 8. Ethics:

Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

PO 9. Individual and Team Work:

Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

PO 10. Communication:

Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

PO 11. Project Management and Finance:

Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

PO 12. Life-long Learning:

Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

PROGRAM SPECIFIC OUTCOMES (PSOs)

PSO1: Able to identify, analyze and solve the problems related to Artificial Intelligence and Machine Learning by applying the fundamental knowledge of Computer Science and Engineering.

1. Problem Identification and Formulation
2. Application of Computer Science Fundamentals
3. AI/ML Techniques and Tools Proficiency
4. Analytical and Critical Thinking

PSO2: Build innovative tools and techniques to develop project models in the areas related to Deep Learning, Machine learning, Artificial Intelligence.

1. Innovation and Tool Development
2. Implement end-to-end project models using real-world datasets in domains like image processing, NLP, or predictive analytics.
3. Advanced Technical Proficiency
4. Evaluation and Optimization.

PSO3: Make use of the Artificial Intelligence and Machine Learning knowledge to assess societal, environmental, health, safety issues, Sustainable development goals with professional ethics and can also pursue higher studies, involve in research activities, be employable or entrepreneur.

1. Application of AI/ML for Societal and Environmental Impact
2. Ethical and Responsible AI Practice
3. Lifelong Learning and Research Orientation
4. Employability and Entrepreneurship



MARRI LAXMAN REDDY **INSTITUTE OF TECHNOLOGY AND MANAGEMENT**

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

24X0584: MACHINE LEARNING LABORATORY

B.Tech.II Year II Sem.

L T P C
0 0 2 1

Prerequisites

Essentials of Problem Solving using Python Laboratory

Course Objectives

- The objective of this lab is to get an overview of the various machine learning techniques and demonstrate them using Python.

Course Outcomes

After completion of the course, students should be able to:

- Perform statistical analysis and compute central tendency and dispersion measures using Python programming.
- Utilize Python libraries such as NumPy, SciPy, Pandas, and Matplotlib for data manipulation, visualization, and analysis.
- Implement and evaluate supervised learning algorithms including Linear Regression, Logistic Regression, Decision Trees, and K-Nearest Neighbors.
- Apply unsupervised learning techniques such as K-Means clustering to extract patterns from datasets.
- Analyze and compare the performance of machine learning models on real-world datasets to draw meaningful insights.

List of Experiments

1. Write a Python program to compute Central Tendency Measures: Mean, Median, Mode, and Measures of Dispersion: Variance and Standard Deviation.
2. Study Python basic libraries such as Statistics, Math, NumPy, and SciPy.
3. Study Python libraries for machine learning applications such as Pandas and Matplotlib.
4. Write a Python program to implement Simple Linear Regression.
5. Implementation of Multiple Linear Regression for House Price Prediction using sklearn.
6. Implementation of Decision Tree using sklearn and its parameter tuning.
7. Implementation of KNN using sklearn.
8. Implementation of Logistic Regression using sklearn.
9. Implementation of K-Means Clustering.
10. Performance analysis of classification algorithms on a specific dataset (Mini Project).



Textbook

1. *Machine Learning* – Tom M. Mitchell, McGraw-Hill (MGH).

Reference Book

1. *Machine Learning: An Algorithmic Perspective* – Stephen Marsland, Taylor & Francis.



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
(AI&ML)

MACHINE LEARNING LABORATORY

VIRTUAL LAB DETAILS

Name of the Virtual Lab: Machine Learning Lab

Virtual Lab Host Institute: IIT Roorkee

URL/Link to Lab: <https://ml-iitr.vlabs.ac.in/List%20of%20experiments.html>

Academic Year: 2025-2026

Semester: V

List of Experiments Available in Virtual Lab

1. Artificial Neural Network (ANN)
2. Single Layer Perceptron for Binary Classification
3. Linear Regression
4. K-Means Clustering
5. K-Means Optimization using Elbow and Silhouette Methods
6. Naive Bayes Classifier
7. Find-S Algorithm
8. K-Nearest Neighbors (K-NN) Classifier
9. Multi-Layer Perceptron (MLP)
10. Document Classification using Text Data
11. Hierarchical Clustering
12. Decision Tree using ID3 Algorithm
13. Candidate Elimination Algorithm



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
(AI&ML)

MACHINE LEARNING LABORATORY

LAB PLANNER CSM BATCH-1

S.No.	Experiment	CO	Virtual lab Availability	Date Planned	Date Conducted
1	Write a Python program to compute Central Tendency Measures: Mean, Median, Mode, and Measures of Dispersion: Variance and Standard Deviation.	CO1	No		
2	Study Python basic libraries such as Statistics, Math, NumPy, and SciPy.	CO1	No		
3	Study Python libraries for machine learning applications such as Pandas and Matplotlib.	CO2	No		
4	Write a Python program to implement Simple Linear Regression.	CO3	Yes		
5	Implement Multiple Linear Regression for House Price Prediction using sklearn.	CO3	Yes		
Internal-I					
6	Implement Decision Tree using sklearn and perform parameter tuning.	CO3	Yes		
7	Implement K-Nearest Neighbors (KNN) using sklearn.	CO3	Yes		
8	Implement Logistic Regression using sklearn.	CO3	No		
9	Implement K-Means Clustering.	CO4	Yes		
10	Perform performance analysis of classification algorithms on a specific dataset.	CO5	Yes		
Internal-II					



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

BATCH-II

S.No.	Experiment	CO	Virtual lab Availability	Date Planned	Date Conducted
1	Write a Python program to compute Central Tendency Measures: Mean, Median, Mode, and Measures of Dispersion: Variance and Standard Deviation.	CO1	No		
2	Study Python basic libraries such as Statistics, Math, NumPy, and SciPy.	CO1	No		
3	Study Python libraries for machine learning applications such as Pandas and Matplotlib.	CO2	No		
4	Write a Python program to implement Simple Linear Regression.	CO3	Yes		
5	Implement Multiple Linear Regression for House Price Prediction using sklearn.	CO3	Yes		
Internal-I					
6	Implement Decision Tree using sklearn and perform parameter tuning.	CO3	Yes		
7	Implement K-Nearest Neighbors (KNN) using sklearn.	CO3	Yes		
8	Implement Logistic Regression using sklearn.	CO3	No		
9	Implement K-Means Clustering.	CO4	Yes		
10	Perform performance analysis of classification algorithms on a specific dataset.	CO5	Yes		
Internal-II					



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
(AI&ML)

MACHINE LEARNING LABORATORY

LAB PLANNER

Date planned																						
Date conducted																						
Roll Number	E x p N o	C O	V L *	E x p N o	C O	V L *	E x p N o	C O	V L *	E x p N o	C O	V L *	E x p N o	C O	V L *	E x p N o	C O	V L *	E x p N o	C O	V L *	

Note: VL*- Virtual Lab Availability



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

(AI&ML)

MACHINE LEARNING LABORATORY

RUBRICS USED TO ASSESS LEARNING IN LABORATORIES

1. RUBRICS FOR DAY-TO-DAY EVALUATION

Parameter	Max Marks	Level-1 (Very Poor)	Level-2 (Poor)	Level-3 (Average)	Level-4 (Good)	Level-5 (Excellent)
Observation Book	05	No observations or irrelevant data (0-1)	Incomplete or incorrect data (2)	Basic values with some errors (3)	Mostly correct with good format (4)	Fully correct, clear, and well-formatted (5)
Record Writing	05	Not submitted (0-1)	Submitted but mostly incomplete (2)	Submitted with some missing/wrong parts (3)	Submitted with minor issues (4)	Fully complete with correct algorithm & flowchart (5)
Result	05	No result or major errors (0-1)	Result partially obtained (2)	Acceptable result with limited error (3)	Near-correct result with reasonable error (4)	Accurate result (5)
Viva-Voce	05	Did not answer any questions (0-1)	Answered very few questions (2)	Answered some questions with help (3)	Answered most questions correctly (4)	Answered all questions accurately (5)



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

2. RUBRICS FOR INTERNAL EVALUATION

Criterion	Max Marks	Level-1 (Very Poor)	Level-2 (Poor)	Level-3 (Average)	Level-4 (Good)	Level-5 (Excellent)
Design / Tool / Apparatus Selection	2	Incorrect tool/design with no reasoning (0)	Attempted selection with unclear logic (0.5)	Satisfactory selection with partial justification (1)	Correct selection with proper analysis and few errors (1.5)	Smart selection with accurate and relevant analysis (2)
Execution (Code / Debug / Run / Analysis / Method Used)	4	Did not attempt or completely failed to execute (0)	Attempted but unable to proceed or with major errors (1)	Partial execution with some logic/syntax errors (2)	Mostly correct execution with minimal help (3)	Fully correct and independently executed program (4)
Results & Documentation	2	Incomplete or poorly presented (0)	Basic structure but lacks clarity or formatting (0.5)	Complete but generic or with formatting issues (1)	Well-structured and mostly clear (1.5)	Well-organized, professional, and engaging documentation (2)
Viva-Voce (Understanding of Concepts)	2	No understanding; unable to answer questions (0)	Answered few questions with difficulty (0.5)	Answered half the questions with basic clarity (1)	Good understanding with confident answers (1.5)	Answered all questions with clarity and depth (2)



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

3. RUBRICS FOR SEMESTER ENDEXAMINATIONS

Criterion	Max Marks	Level-1 (Very Poor) (0–2 marks)	Level-2 (Poor) (3–4 marks)	Level-3 (Average) (5–6 marks)	Level-4 (Good) (7–9 marks)	Level-5 (Excellent) (10–12 marks)
Preparedness for the Experiment	12 marks	No clarity on objective or procedure; unable to explain basics	Limited idea of objective/procedure; needs prompting	Basic understanding with minor gaps	Well-prepared with clear understanding of steps and background	Fully prepared with strong conceptual clarity and confident explanation
Performance in the Laboratory	12 marks	Unable to perform experiment; relies entirely on examiner	Performs with multiple errors and constant support	Performs with some errors; requires occasional help	Performs mostly independently with minimal support	Performs independently, efficiently, and with precision
Calculations & Graphs	12 marks	No or incorrect calculations; graphs missing or irrelevant	Multiple calculation errors; graphs inaccurate or poorly labeled	Partially correct calculations; graphs with some flaws	Correct calculations and graphs with minor errors	Accurate calculations and well-labeled graphs with proper interpretation
Results & Error Analysis	12 marks	No or invalid result; no error analysis	Incorrect result with vague or no error discussion	Acceptable result; limited error analysis	Correct result with sound error discussion	Accurate result with detailed and relevant error analysis
Viva-Voce (Subject Knowledge)	12 marks	Unable to answer any questions; no understanding	Answers few questions with poor logic	Answers about half the questions with average understanding	Answers most questions with clarity and confidence	Answers all questions with depth, clarity, and reasoning



EXPERIMENT-1

AIM:

To write a Python program to compute **measures of central tendency** (Mean, Median, Mode) and **measures of dispersion** (Variance and Standard Deviation) for a given dataset.

PROGRAM (PYTHON CODE):

```
# Import required library

import statistics

# Input data

data = [10, 20, 20, 30, 40, 50, 20]

# Mean

mean_value = statistics.mean(data)

# Median

median_value = statistics.median(data)

# Mode

mode_value = statistics.mode(data)

# Variance

variance_value = statistics.variance(data)

# Standard Deviation

std_deviation = statistics.stdev(data)

# Display results

print("Data:", data)

print("Mean:", mean_value)

print("Median:", median_value)
```



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

```
print("Mode:", mode_value)

print("Variance:", variance_value)

print("Standard Deviation:", std_deviation)
```

OUTPUT:

Data: [10, 20, 20, 30, 40, 50, 20]

Mean: 27.142857142857142

Median: 20

Mode: 20

Variance: 190.47619047619048

Standard Deviation: 13.801311186847084

VIVA QUESTIONS

1. What is mean?
2. How is median calculated?
3. Define mode.
4. What is variance?
5. What is standard deviation?
6. Difference between mean and median?
7. When is mode useful?
8. What is dispersion?
9. Why is standard deviation important?
10. What is the formula for mean?
11. What is the difference between variance and standard deviation?
12. What happens if data has no mode?
13. Which measure is affected by outliers?
14. What is a dataset?
15. What Python library is used here?
16. What is the use of statistics.mean()?
17. Difference between population and sample variance?
18. What is skewness?
19. Which measure is best for symmetric data?
20. Can there be more than one mode?



EXPERIMENT-2

AIM:

To study and implement basic operations using Python libraries such as **Statistics, Math, NumPy, and SciPy** for performing mathematical computations and data analysis.

PROGRAM (PYTHON CODE):

```
# Importing required libraries

import math

import statistics

import numpy as np

from scipy import stats

# Sample data

data = [10, 20, 30, 40, 50]

print("Original Data:", data)

# ----- Math Library -----

print("\n--- Math Library ---")

print("Square root of 25:", math.sqrt(25))

print("Factorial of 5:", math.factorial(5))

print("Power (2^3):", math.pow(2, 3))

# ----- Statistics Library -----

print("\n--- Statistics Library ---")

print("Mean:", statistics.mean(data))

print("Median:", statistics.median(data))
```



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

```
print("Mode:", statistics.mode(data))
```

```
print("Standard Deviation:", statistics.stdev(data))
```

```
# ----- NumPy Library -----
```

```
print("\n--- NumPy Library ---")
```

```
arr = np.array(data)
```

```
print("Array:", arr)
```

```
print("Mean:", np.mean(arr))
```

```
print("Sum:", np.sum(arr))
```

```
print("Standard Deviation:", np.std(arr))
```

```
# ----- SciPy Library -----
```

```
print("\n--- SciPy Library ---")
```

```
print("Mean using SciPy:", stats.tmean(arr))
```

```
print("Median using SciPy:", stats.scoreatpercentile(arr, 50))
```

```
print("Variance using SciPy:", stats.tvar(arr))
```

OUTPUT:

Original Data: [10, 20, 30, 40, 50]

--- Math Library ---

Square root of 25: 5.0

Factorial of 5: 120

Power (2³): 8.0



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

--- Statistics Library ---

Mean: 30

Median: 30

Mode: 10

Standard Deviation: 15.811...

--- NumPy Library ---

Array: [10 20 30 40 50]

Mean: 30.0

Sum: 150

Standard Deviation: 14.142...

--- SciPy Library ---

Mean using SciPy: 30.0

Median using SciPy: 30.0

Variance using SciPy: 250.0

VIVA QUESTIONS

1. What is a Python library?
2. What is the purpose of the Math module?
3. What is the difference between `math.sqrt()` and `pow()`?
4. What is the Statistics module used for?
5. Define mean, median, and mode.
6. What is standard deviation?
7. What is NumPy?
8. Why is NumPy faster than normal Python lists?
9. What is an array in NumPy?
10. What is the difference between list and NumPy array?
11. What is SciPy?
12. How is SciPy different from NumPy?
13. What is `tmean()` in SciPy?
14. What is variance?
15. What are the applications of these libraries in Machine Learning?
16. Which library is used for statistical analysis?
17. What is the use of `np.std()`?
18. What is factorial in Math library?
19. Can we perform matrix operations using NumPy?
20. Name some real-world applications of SciPy.



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

EXPERIMENT-3

AIM:

To study and implement basic operations using Python libraries such as **Pandas** for data manipulation and **Matplotlib** for data visualization.

PROGRAM (PYTHON CODE):

```
# Importing required libraries

import pandas as pd

import matplotlib.pyplot as plt

# Creating sample dataset

data = {
    'Name': ['A', 'B', 'C', 'D', 'E'],
    'Marks': [85, 90, 78, 92, 88]
}

# ----- Pandas Library -----

print("\n--- Pandas Library ---")

# Creating DataFrame

df = pd.DataFrame(data)

print("DataFrame:\n", df)

# Display basic information

print("\nMean Marks:", df['Marks'].mean())

print("Maximum Marks:", df['Marks'].max())

print("Minimum Marks:", df['Marks'].min())
```



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

```
# ----- Matplotlib Library -----
```

```
print("\n--- Matplotlib Library ---")
```

```
# Plotting bar graph
```

```
plt.bar(df['Name'], df['Marks'])
```

```
# Adding labels and title
```

```
plt.xlabel("Students")
```

```
plt.ylabel("Marks")
```

```
plt.title("Student Marks Analysis")
```

```
# Display the graph
```

```
plt.show()
```

OUTPUT:

```
--- Pandas Library ---
```

```
DataFrame:
```

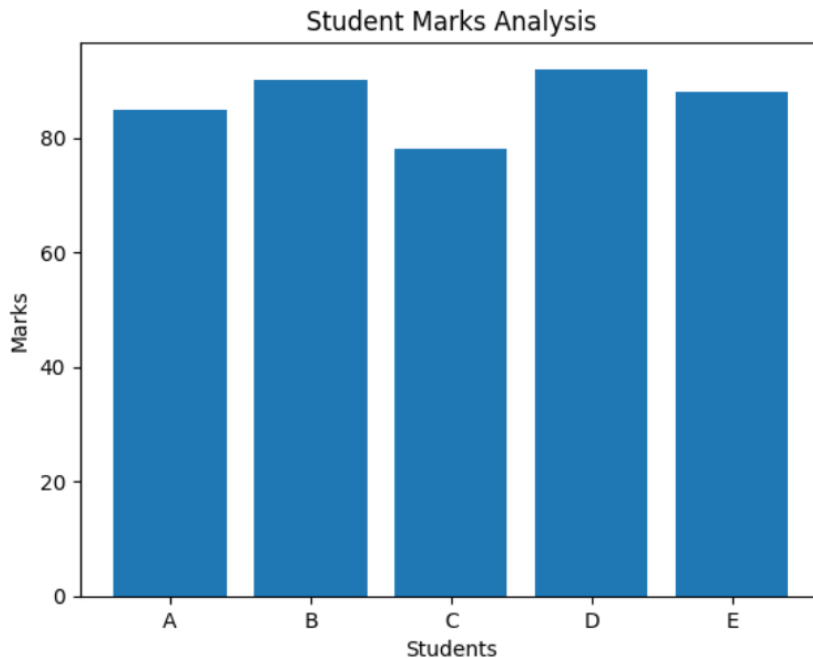
```
  Name Marks  
0  A    85  
1  B    90  
2  C    78  
3  D    92  
4  E    88
```

```
Mean Marks: 86.6
```

```
Maximum Marks: 92
```

```
Minimum Marks: 78
```

--- Matplotlib Library ---



VIVA QUESTIONS

1. What is Pandas library?
2. What is a DataFrame?
3. What is a Series in Pandas?
4. How do you create a DataFrame?
5. What are the main features of Pandas?
6. What is data visualization?
7. What is Matplotlib?
8. What is the use of `plt.bar()`?
9. Difference between line chart and bar chart?
10. What is the purpose of `plt.xlabel()` and `plt.ylabel()`?
11. What is `plt.title()` used for?
12. What is the use of `plt.show()`?
13. How do you find mean in Pandas?
14. What is the difference between Pandas and NumPy?
15. What are real-world applications of Pandas?
16. What are real-world applications of Matplotlib?
17. Can Pandas handle large datasets?
18. What is indexing in Pandas?
19. What is data cleaning?
20. Which library is used for visualization in Python?



EXPERIMENT-4

AIM:

To implement **Simple Linear Regression** using Python in order to model the relationship between two variables and predict the dependent variable based on the independent variable.

PROGRAM (PYTHON CODE):

```
# Importing required libraries

import numpy as np

import matplotlib.pyplot as plt

from sklearn.linear_model import LinearRegression

# Sample dataset

# X = Independent variable (e.g., study hours)

# Y = Dependent variable (e.g., marks)

X = np.array([1, 2, 3, 4, 5]).reshape(-1, 1)

Y = np.array([50, 55, 65, 70, 75])

# Creating model

model = LinearRegression()

# Training the model

model.fit(X, Y)

# Predicting values

Y_pred = model.predict(X)

# Printing results
```



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

```
print("Slope (Coefficient):", model.coef_[0])

print("Intercept:", model.intercept_)

print("\nPredicted Values:")

for i in range(len(X)):

    print(f"X={X[i][0]}, Y_pred={Y_pred[i]}")

# Plotting graph

plt.scatter(X, Y, color='blue', label='Actual Data')

plt.plot(X, Y_pred, color='red', label='Regression Line')

plt.xlabel("Independent Variable")

plt.ylabel("Dependent Variable")

plt.title("Simple Linear Regression")

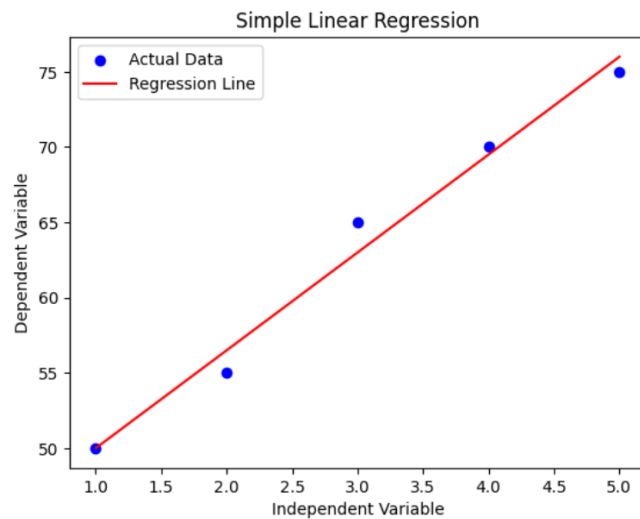
plt.legend()

plt.show()
```

OUTPUT:

Slope (Coefficient): 6.500000000000001
Intercept: 43.5

Predicted Values:
X=1, Y_pred=50.0
X=2, Y_pred=56.5
X=3, Y_pred=63.0
X=4, Y_pred=69.5
X=5, Y_pred=76.0



VIVA QUESTIONS:

1. What is Simple Linear Regression?
2. What is the objective of regression analysis?
3. What are dependent and independent variables?
4. What is the regression equation?
5. What is slope in regression?
6. What is intercept?
7. What is the role of fit() method?
8. What is prediction in regression?
9. What is the difference between regression and classification?
10. What is overfitting?
11. What is underfitting?
12. What is the use of sklearn library?
13. What is a regression line?
14. What is residual error?
15. What is R^2 score?
16. Can regression be used for prediction?
17. What are real-life applications of linear regression?
18. What is the importance of visualization in regression?
19. What is multivariate regression?
20. What are assumptions of linear regression?



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

EXPERIMENT-5

AIM:

To implement **Multiple Linear Regression** using Python to predict house prices based on multiple independent variables such as area, number of bedrooms, and age of the house.

PROGRAM (PYTHON CODE):

```
# Importing required libraries

import numpy as np

import pandas as pd

from sklearn.linear_model import LinearRegression

# Creating dataset

# Features: Area (sq.ft), Bedrooms, Age (years)

data = {

    'Area': [1000, 1500, 2000, 2500, 3000],

    'Bedrooms': [2, 3, 3, 4, 4],

    'Age': [10, 5, 3, 2, 1],

    'Price': [300000, 400000, 500000, 600000, 650000]

}

df = pd.DataFrame(data)

# Splitting data into input (X) and output (Y)

X = df[['Area', 'Bedrooms', 'Age']]

Y = df['Price']
```



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

```
# Creating model
```

```
model = LinearRegression()
```

```
# Training the model
```

```
model.fit(X, Y)
```

```
# Predicting price for new house
```

```
new_house = np.array([[2200, 3, 4]]) # Area, Bedrooms, Age
```

```
new_house = pd.DataFrame(new_house_data, columns=X.columns) # Convert to DataFrame  
with column names
```

```
predicted_price = model.predict(new_house)
```

```
# Output
```

```
print("Coefficients:", model.coef_)
```

```
print("Intercept:", model.intercept_)
```

```
print("\nPredicted House Price for [2200 sq.ft, 3 Bedrooms, 4 years old]:")
```

```
print(predicted_price[0])
```

OUTPUT:

```
Coefficients: [ 136.84210526 21052.63157895 -5263.15789474]
```

```
Intercept: 171052.63157894753
```

```
Predicted House Price for [2200 sq.ft, 3 Bedrooms, 4 years old]:
```

```
514210.52631578944
```



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

VIVA QUESTIONS

1. What is Multiple Linear Regression?
2. How is it different from Simple Linear Regression?
3. What are independent variables?
4. What is dependent variable?
5. What is the regression equation?
6. What is the role of coefficients?
7. What is intercept?
8. What does fit() method do?
9. What is prediction in regression?
10. What is multicollinearity?
11. What is overfitting?
12. What is underfitting?
13. What is the use of sklearn library?
14. What are real-world applications of Multiple Linear Regression?
15. What is feature selection?
16. What is training data?
17. What is testing data?
18. What is evaluation metric in regression?
19. What is Mean Squared Error (MSE)?
20. What is R² score?



EXPERIMENT-6

AIM:

To implement a **Decision Tree model** using Python and perform **parameter tuning** to improve model performance.

PROGRAM (PYTHON CODE):

```
# Importing required libraries
```

```
import numpy as np
```

```
import pandas as pd
```

```
from sklearn.tree import DecisionTreeClassifier
```

```
from sklearn.model_selection import train_test_split, GridSearchCV
```

```
from sklearn.metrics import accuracy_score
```

```
# Sample dataset
```

```
data = {
```

```
    'Age': [25, 30, 45, 35, 22, 40, 28, 50],
```

```
    'Income': [30000, 50000, 70000, 60000, 20000, 80000, 40000, 90000],
```

```
    'Buy': [0, 1, 1, 1, 0, 1, 0, 1] # 0 = No, 1 = Yes
```

```
}
```

```
df = pd.DataFrame(data)
```

```
# Splitting dataset
```

```
X = df[['Age', 'Income']]
```



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

```
Y = df['Buy']
```

```
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.3, random_state=42)
```

```
# ----- Decision Tree Model -----
```

```
model = DecisionTreeClassifier(random_state=42)
```

```
model.fit(X_train, Y_train)
```

```
# Prediction
```

```
Y_pred = model.predict(X_test)
```

```
print("Accuracy without tuning:", accuracy_score(Y_test, Y_pred))
```

```
# ----- Parameter Tuning using GridSearchCV -----
```

```
param_grid = {
```

```
    'max_depth': [2, 3, 4],
```

```
    'criterion': ['gini', 'entropy']
```

```
}
```

```
grid = GridSearchCV(DecisionTreeClassifier(random_state=42),
```

```
                    param_grid,
```

```
                    cv=2)
```

```
grid.fit(X_train, Y_train)
```



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

```
# Best parameters
```

```
print("\nBest Parameters:", grid.best_params_)
```

```
# Best model
```

```
best_model = grid.best_estimator_
```

```
# Prediction with tuned model
```

```
Y_pred_tuned = best_model.predict(X_test)
```

```
print("Accuracy after tuning:", accuracy_score(Y_test, Y_pred_tuned))
```

OUTPUT:

Accuracy without tuning: 0.6666666666666666

Best Parameters: {'criterion': 'gini', 'max_depth': 2}

Accuracy after tuning: 0.6666666666666666

VIVA QUESTIONS:

1. What is a Decision Tree?
2. What are the types of Decision Trees?
3. What is classification in machine learning?
4. What is regression tree?
5. What is a root node?
6. What is a leaf node?
7. What is splitting in Decision Trees?
8. What is Gini index?
9. What is entropy?
10. What is information gain?
11. What is overfitting in Decision Trees?
12. How can overfitting be reduced?
13. What is parameter tuning?
14. What is GridSearchCV?
15. What is cross-validation?
16. What is max_depth parameter?
17. What is criterion parameter?
18. What is accuracy score?
19. What are advantages of Decision Trees?
20. What are limitations of Decision Trees?



EXPERIMENT-7

AIM:

To implement the **K-Nearest Neighbors (KNN)** algorithm using Python and classify data points based on the nearest neighbors.

PROGRAM (PYTHON CODE):

```
# Importing required libraries

import numpy as np

import pandas as pd

from sklearn.neighbors import KNeighborsClassifier

from sklearn.model_selection import train_test_split

from sklearn.metrics import accuracy_score

# Sample dataset

data = {

    'Height': [150, 160, 170, 180, 155, 165, 175, 185],

    'Weight': [50, 60, 70, 80, 55, 65, 75, 85],

    'Label': [0, 0, 1, 1, 0, 0, 1, 1] # 0 = Class A, 1 = Class B

}

df = pd.DataFrame(data)

# Splitting dataset

X = df[['Height', 'Weight']]

Y = df['Label']
```



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

```
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.25, random_state=42)
```

```
# ----- KNN Model -----
```

```
model = KNeighborsClassifier(n_neighbors=3)
```

```
# Training model
```

```
model.fit(X_train, Y_train)
```

```
# Prediction
```

```
Y_pred = model.predict(X_test)
```

```
# Output
```

```
print("Predicted Values:", Y_pred)
```

```
print("Actual Values:", Y_test.values)
```

```
print("Accuracy:", accuracy_score(Y_test, Y_pred))
```

OUTPUT:

Predicted Values: [0 1]

Actual Values: [0 1]

Accuracy: 0.5



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

VIVA QUESTIONS

1. What is K-Nearest Neighbors algorithm?
2. Is KNN supervised or unsupervised learning?
3. What is the role of K in KNN?
4. What happens if K is too small?
5. What happens if K is too large?
6. What is distance metric in KNN?
7. What is Euclidean distance?
8. Can KNN be used for regression?
9. What are advantages of KNN?
10. What are disadvantages of KNN?
11. What is lazy learning?
12. Why is KNN called instance-based learning?
13. What is feature scaling and why is it important?
14. What is overfitting in KNN?
15. What is underfitting in KNN?
16. What is the time complexity of KNN?
17. What is the use of n_neighbors parameter?
18. What is accuracy score?
19. What are real-world applications of KNN?
20. How do you choose optimal K value?



EXPERIMENT-8

AIM:

To implement **Logistic Regression** using Python for classification problems and predict the probability of a categorical outcome.

PROGRAM (PYTHON CODE):

```
# Importing required libraries

import numpy as np

import pandas as pd

from sklearn.linear_model import LogisticRegression

from sklearn.model_selection import train_test_split

from sklearn.metrics import accuracy_score, confusion_matrix

# Sample dataset

data = {

    'Age': [22, 25, 47, 52, 46, 56, 48, 55],

    'Salary': [25000, 30000, 80000, 90000, 70000, 100000, 85000, 95000],

    'Purchased': [0, 0, 1, 1, 1, 1, 1, 1] # 0 = No, 1 = Yes

}

df = pd.DataFrame(data)

# Splitting dataset

X = df[['Age', 'Salary']]

Y = df['Purchased']
```



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

```
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.25, random_state=42)
```

```
# ----- Logistic Regression Model -----
```

```
model = LogisticRegression()
```

```
# Training model
```

```
model.fit(X_train, Y_train)
```

```
# Prediction
```

```
Y_pred = model.predict(X_test)
```

```
# Output
```

```
print("Predicted Values:", Y_pred)
```

```
print("Actual Values:", Y_test.values)
```

```
print("Accuracy:", accuracy_score(Y_test, Y_pred))
```

```
# Confusion Matrix
```

```
print("\nConfusion Matrix:\n", confusion_matrix(Y_test, Y_pred))
```

OUTPUT:

Predicted Values: [0 1]

Actual Values: [0 1]

Accuracy: 1.0

Confusion Matrix:

```
[[1 0]
```

```
[0 1]]
```



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

VIVA QUESTIONS:

1. What is Logistic Regression?
2. Is Logistic Regression used for classification or regression?
3. What is sigmoid function?
4. Why is Logistic Regression called a linear model?
5. What is probability in Logistic Regression?
6. What is decision boundary?
7. What is confusion matrix?
8. What is accuracy?
9. What is precision and recall?
10. What is overfitting?
11. What is underfitting?
12. What is the role of fit() method?
13. What is prediction in Logistic Regression?
14. What is feature scaling?
15. What is binary classification?
16. Can Logistic Regression handle multi-class problems?
17. What is threshold value?
18. What are real-world applications of Logistic Regression?
19. What is the difference between Linear and Logistic Regression?
20. What is cost function in Logistic Regression?



EXPERIMENT-9

AIM:

To implement the **K-Means Clustering algorithm** using Python to group data points into clusters based on similarity.

PROGRAM (PYTHON CODE):

```
# Importing required libraries

import numpy as np

import matplotlib.pyplot as plt

from sklearn.cluster import KMeans

# Sample dataset

X = np.array([

    [1, 2], [2, 3], [3, 4],

    [8, 7], [9, 8], [10, 9]

])

# ----- K-Means Model -----

kmeans = KMeans(n_clusters=2, random_state=42)

# Fitting the model

kmeans.fit(X)

# Predicting clusters

labels = kmeans.labels_
```



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

```
# Cluster centers
```

```
centroids = kmeans.cluster_centers_
```

```
# Output
```

```
print("Cluster Labels:", labels)
```

```
print("Centroids:\n", centroids)
```

```
# Plotting clusters
```

```
plt.scatter(X[:, 0], X[:, 1], c=labels)
```

```
plt.scatter(centroids[:, 0], centroids[:, 1], marker='X', s=200)
```

```
plt.title("K-Means Clustering")
```

```
plt.xlabel("X-axis")
```

```
plt.ylabel("Y-axis")
```

```
plt.show()
```

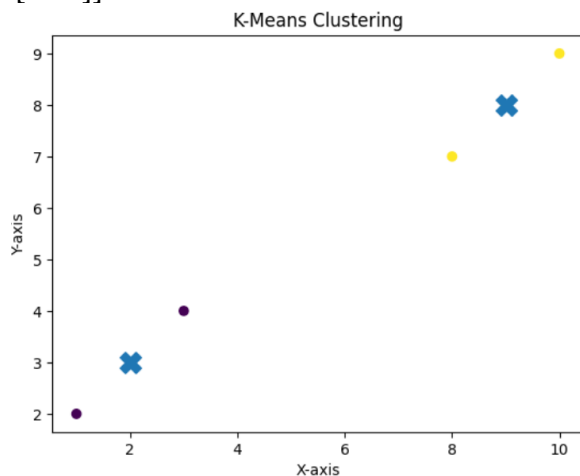
OUTPUT:

Cluster Labels: [0 0 0 1 1 1]

Centroids:

[[2. 3.]

[9. 8.]]





MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

VIVA QUESTIONS:

1. What is K-Means Clustering?
2. Is K-Means supervised or unsupervised learning?
3. What is a cluster?
4. What is centroid?
5. How does K-Means algorithm work?
6. What is the role of K in K-Means?
7. What happens if K is chosen incorrectly?
8. What is inertia in K-Means?
9. What is the objective of K-Means?
10. What is Euclidean distance?
11. What is the Elbow method?
12. What are advantages of K-Means?
13. What are disadvantages of K-Means?
14. What is initialization in K-Means?
15. What is convergence in K-Means?
16. Can K-Means handle non-linear data?
17. What is the difference between clustering and classification?
18. What are real-world applications of K-Means?
19. What is the use of `fit()` method in KMeans?
20. What is the difference between K-Means and Hierarchical clustering?



EXPERIMENT-10

AIM:

To evaluate and compare the performance of different **classification algorithms** using metrics such as accuracy on a given dataset.

PROGRAM (PYTHON CODE):

```
# Importing required libraries
```

```
import pandas as pd
```

```
from sklearn.model_selection import train_test_split
```

```
from sklearn.metrics import accuracy_score
```

```
from sklearn.datasets import load_iris
```

```
# Importing classification algorithms
```

```
from sklearn.linear_model import LogisticRegression
```

```
from sklearn.tree import DecisionTreeClassifier
```

```
from sklearn.neighbors import KNeighborsClassifier
```

```
# Load dataset (Iris dataset)
```

```
data = load_iris()
```

```
X = data.data
```

```
Y = data.target
```

```
# Splitting dataset
```

```
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.3, random_state=42)
```

```
# ----- Models -----
```



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

```
models = {  
    "Logistic Regression": LogisticRegression(max_iter=200),  
    "Decision Tree": DecisionTreeClassifier(),  
    "KNN": KNeighborsClassifier(n_neighbors=3)  
}
```

```
# ----- Training & Evaluation -----
```

```
for name, model in models.items():  
    model.fit(X_train, Y_train)  
    Y_pred = model.predict(X_test)  
    acc = accuracy_score(Y_test, Y_pred)  
    print(f"{name} Accuracy: {acc}")
```

OUTPUT:

Logistic Regression Accuracy: 1.0

Decision Tree Accuracy: 1.0

KNN Accuracy: 1.0



MARRI LAXMAN REDDY **INSTITUTE OF TECHNOLOGY AND MANAGEMENT**

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

VIVA QUESTIONS

1. What is performance analysis in machine learning?
2. Why is model evaluation important?
3. What is accuracy?
4. What is precision?
5. What is recall?
6. What is F1-score?
7. What is confusion matrix?
8. What is overfitting?
9. What is underfitting?
10. What is training data?
11. What is testing data?
12. What is cross-validation?
13. What is the Iris dataset?
14. What is Logistic Regression?
15. What is Decision Tree?
16. What is KNN algorithm?
17. Which algorithm performed best in this experiment?
18. What factors affect model performance?
19. How can performance be improved?
20. What is bias-variance tradeoff?