|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **Hope Foundation’s**  **Finolex Academy of Management and Technology, Ratnagiri** | | | | | |
| **Department of Computer Science and Engineering (AIML)** | | | | | |
| Subject name: Machine Learning Lab | | | | | | Subject Code: CSL604 | |
| Class | | TE CSE | | | Semester –VI (CBCGS) | Academic year: 2024-25 | |
| Name of Student | |  | | | | **QUIZ Score :** | |
| Roll No | |  | | Experiment No. | | 02 | |
| Title: **Implementation of Logistic Regression.** | | | | | | | |
|  | | | | | | | |
| **1. Lab objectives applicable:**  **LOB1:** To introduce platforms such as Anaconda, COLAB suitable to Machine Learning.  **LOB2:** To implement various Regression techniques. | | | | | | | |
| **2. Lab outcomes applicable:**  **LO1:** Implement various machine learning models. | | | | | | | |
| **3. Learning Objectives:**   1. To predict future continuous value based on past data. | | | | | | | |
| **4. Practical applications of the assignment/experiment:**  .To predict the probability of binary classes | | | | | | | |
| **5. Prerequisites**:   1. Python language | | | | | | | |
| **6. Minimum Hardware Requirements**:-  I series processor, RAM 4GB,  **7. Software Requirements:-**  Colab or Visual Studio or Jupyter notebook (Anaconda) | | | | | | | |
| **8. Quiz Questions :**  [**https://docs.google.com/forms/d/e/1FAIpQLSdRVHVFyYcXUSTrEByBKd5cJkaboF2dWYjEAcmUPVL690ZXwg/viewform?usp=sf\_link**](https://docs.google.com/forms/d/e/1FAIpQLSdRVHVFyYcXUSTrEByBKd5cJkaboF2dWYjEAcmUPVL690ZXwg/viewform?usp=sf_link) | | | | | | | |
| **9. Experiment/Assignment Evaluation:** | | | | | | | |
| **Sr. No.** | **Parameters** | | | | | **Marks obtained** | **Out of** |
| **1** | Technical Understanding (Assessment may be done based on Q & A **or** any other relevant method.) Teacher should mention the other method used - | | | | |  | 6 |
| **2** | Lab Performance | | | | |  | 2 |
| **3** | Punctuality | | | | |  | 2 |
| **Date of performance (DOP)** | | |  | | **Total marks obtained** |  | **10** |

**Signature of Faculty**

**10. Theory:**

Logistic regression is a statistical method used for binary classification problems, where the target variable has two possible outcomes (e.g., success/failure, yes/no, 1/0). It predicts the probability of an event occurring by modeling the relationship between one or more independent variables and the dependent variable using a logistic function (sigmoid curve).

**Key Features:**

1. **Output:** Provides probabilities (values between 0 and 1).
2. **Log-Odds:** The model predicts the log-odds of the dependent variable, which can be converted to probabilities.
3. **Function:** The logistic function ensures the output is bounded between 0 and 1:

**Assumptions**:

* The dependent variable is binary.
* Independent variables have a linear relationship with the log-odds.
* Observations are independent of each other.

From Experiment we get for x=150cm we get y=66.2kg. We apply logistic regression formula for this given value y and predict its binary value.

We define the sigmoid function as:

We know regression equation is given as:

Substituting value of y from experiment 1 we get,

∴For y=66.2 it is classified as class 1.

**11. Installation Steps / Performance Steps and Results :**

import numpy as np

import matplotlib.pyplot as plot

from sklearn.datasets import load\_digits

# Load Digits dataset

digits = load\_digits()

# X is a collection of arrays of 64 (8X8) pixel intensity values for different handwritten digits

X = digits.data

# y is the corresponding value of digits

y = digits.target

def plot\_digit(x):

    'Function to plot heatmaps for pixel intensity arrays'

    plot.imshow(x.reshape(8,8))     # x.reshape(cols,rows)

plot\_digit(X[104])

from sklearn.model\_selection import train\_test\_split

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=1)

from sklearn.linear\_model import LogisticRegression

from sklearn.metrics import accuracy\_score

lrm = LogisticRegression(max\_iter=5000)

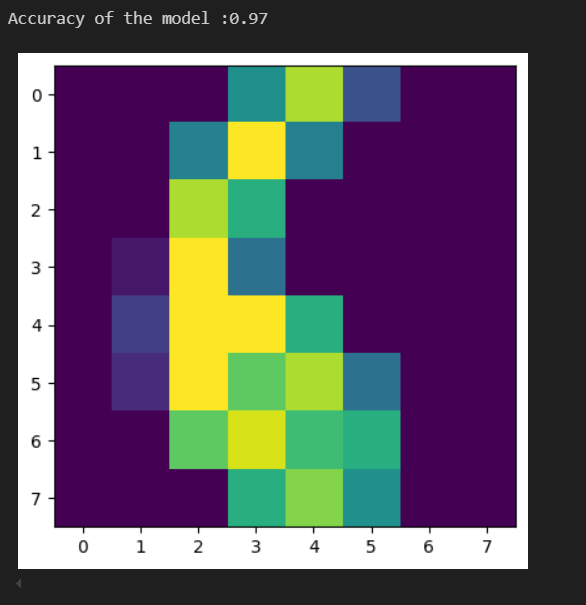
lrm.fit(X\_train, y\_train)

y\_predict = lrm.predict(X\_test)

accuracy = accuracy\_score(y\_predict, y\_test)

print(f"Accuracy of the model :{accuracy:.2f}")

**OUTPUT:**

****

**12. Learning Outcomes Achieved**

1. Students are able toimplement regression to predict future continuos value.

**13. Conclusion:**

**1. Applications of the Studied Technique in Industry**

Linear regression is widely used in various industries due to its simplicity and interpretability. Some notable applications include:

* Finance: Predicting stock prices, evaluating risk factors, or modeling consumer credit risk.
* Marketing: Estimating the impact of advertising spend on sales or customer acquisition metrics.

1. **Engineering Relevance**

Linear regression holds significant relevance in engineering disciplines:

* **Data-Driven Design:** Engineers use regression to model relationships between design parameters and performance metrics.
* **Process Optimization:** Helps in identifying key variables that influence efficiency and quality in manufacturing or chemical processes.

**3. Skills Developed**

**Programming Proficiency:** Gaining experience in Python libraries like NumPy, pandas, scikit-learn, and matplotlib.

**Critical Thinking:** Evaluating the limitations and assumptions of linear regression, such as linearity independence, and normality of residuals.

**14. References**:

1. Nathalie Japkowicz & Mohak Shah, ―Evaluating Learning Algorithms: A Classification Perspective‖, Cambridge.
2. Marc Peter Deisenroth, Aldo Faisal, Cheng Soon Ong, ―Mathematics for machine learning‖
3. Samir Roy and Chakraborty, ―Introduction to soft computing‖, Pearson Edition.
4. Ethem Alpaydın, ―Introduction to Machine Learning‖, MIT Press McGraw-Hill Higher

Education

1. Peter Flach, ―Machine Learning‖, Cambridge University Press
2. Tom M. Mitchell, ―Machine Learning‖, McGraw Hill
3. Kevin P. Murphy, ―Machine Learning ― A Probabilistic Perspective‖, MIT Press