

Assignment No: 3.

Page No.

Date

Title: Assignment based on Logistic Regression for classification using python. Assess the performance of the model using evaluation metrics.

Software Requirements:

python

VS code

Theory

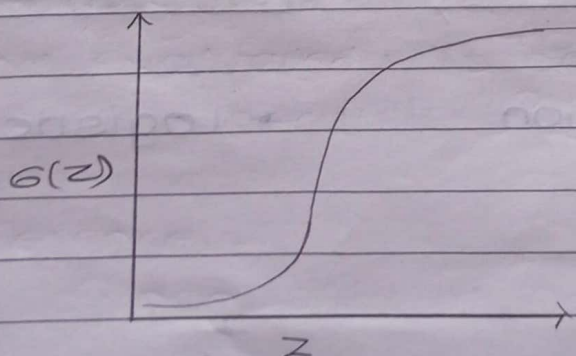
1. Introduction to Logistic Regression

Logistic Regression is a supervised machine learning algorithm used for classification problems. It predicts the probability of a data point belonging to a particular category using the sigmoid function.

2. Logistic Function

The logistic (sigmoid) function is

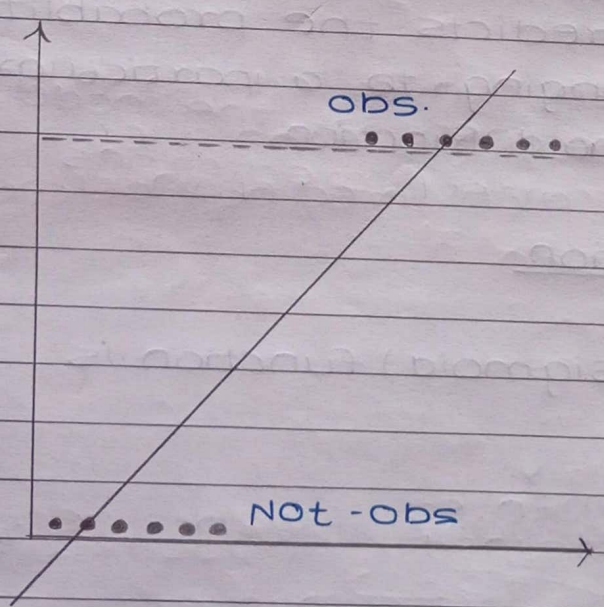
$$\sigma(z) = \frac{1}{1 + e^{-z}}$$



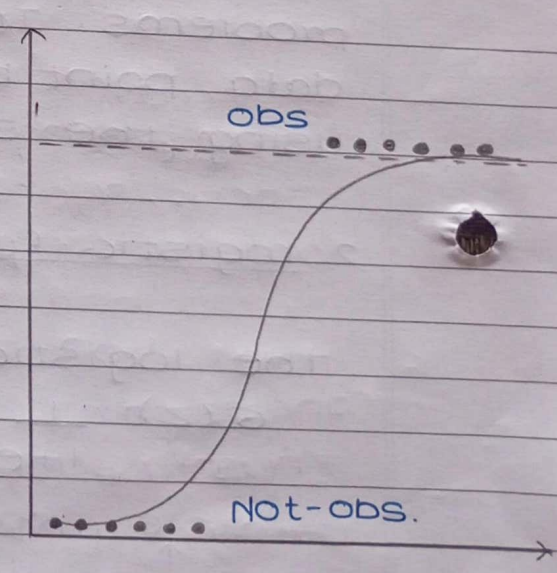
• Logistic Regression (sigmoid function).

3. difference between Target Variable in Logistic and Linear Regression.

Aspect	Linear Regression	Logistic Regression
Target Variable	Continuous value	Categorical value
Output range	Any real number	0 to 1
Use case	prediction of numbers	classification tasks.



• Linear Regression



• Logistic Regression

4. Sigmoid Function in Logistic Regression.

The sigmoid function outputs probability values that can be threshold.

e.g. ≥ 0.5 class 1
 else class 0.

Formula :

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

5. categorical value variable in Database

categorical variable represent categories or groups. (e.g. gender = male / female)

These are often converted to numeric form using encoding techniques (Label Encoding, one Hot encoding) before applying logistic Regression.

6. Evaluation metrics:

1. Confusion matrix

Table showing correct & incorrect prediction

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

2. Precision:-

Correct positive predictions out of all predicted positives

$$\text{precision} = \frac{TP}{TP + FP}$$

3. Recall (sensitivity):

correct positive predictions out of all actual positives.

$$\text{Recall} = \frac{TP}{TP + FN}$$

4. F1-Score

Harmonic mean of precision and recall

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

5. Accuracy :

overall correctness of the model.

$$\text{Accuracy} = \frac{\text{Correct predictions}}{\text{Total prediction}}$$

conclusion:

Thus, we implemented Logistic Regression using python, trained a classification model and evaluated its performance using accuracy, precision, recall, F1-score and confusion matrix.

Reference:

python official Documentation.