

Lab Report on
Logistic Regression



Submitted to
Department of Computer Science and Engineering
Nepal Engineering College

in Partial Fulfillment of the
Requirements for the Degree of B.E. in Computer

Submitted By
Dipson Thapa (020-313)

Date: 28/07/2024

Logistic Regression

Introduction:

Logistic Regression is a fundamental supervised learning algorithm used for binary classification problems. It models the probability that a given input point belongs to a particular class. By using the logistic function (also known as the sigmoid function), logistic regression transforms its linear regression output into a probability value, which is then used to classify the input. This algorithm is widely utilized in various fields such as medical diagnosis, spam detection, and credit scoring due to its simplicity and effectiveness in handling binary classification tasks.

Dataset:

The dataset used for this Logistic Regression analysis is the "Titanic: Machine Learning from Disaster" dataset, which contains information about the passengers of the Titanic and whether they survived the disaster. You can access it here.

Link: <https://www.kaggle.com/datasets/nimapourmoradi/raisin-binary-classification>

Notebook name: Logistic_Regression_020313.ipynb

Libraries Used:

1. Pandas:

- Utilized for data manipulation and analysis. It offers data structures like DataFrames that are ideal for handling tabular data.

2. Pathlib:

- A standard library for handling filesystem paths in a more readable and efficient way.

3. numpy:

- A fundamental package for numerical computations in Python, providing support for arrays and matrices along with a collection of mathematical functions to operate on these data structures.