Linear Vs Logistic Regression

Steps of Logistic Regression
In logistic regression, we decide a probability threshold. If the probability of a particular element is higher than the probability threshold then we classify that element in one group or vice versa.

Step 1
To calculate the binary separation, first, we determine the best-fitted line by following the Linear Regression steps.

Step 2
The regression line we get from Linear Regression is highly susceptible to outliers. Thus it will not do a good job in classifying two classes.

Thus, the predicted value gets converted into probability by feeding it to the sigmoid function.

Linear Vs Logistic Regression

Step 3
Finally, the output value of the sigmoid function gets converted into 0 or 1(discreet values) based on the threshold value. We usually set the threshold value as 0.5. In this way, we get the binary classification.

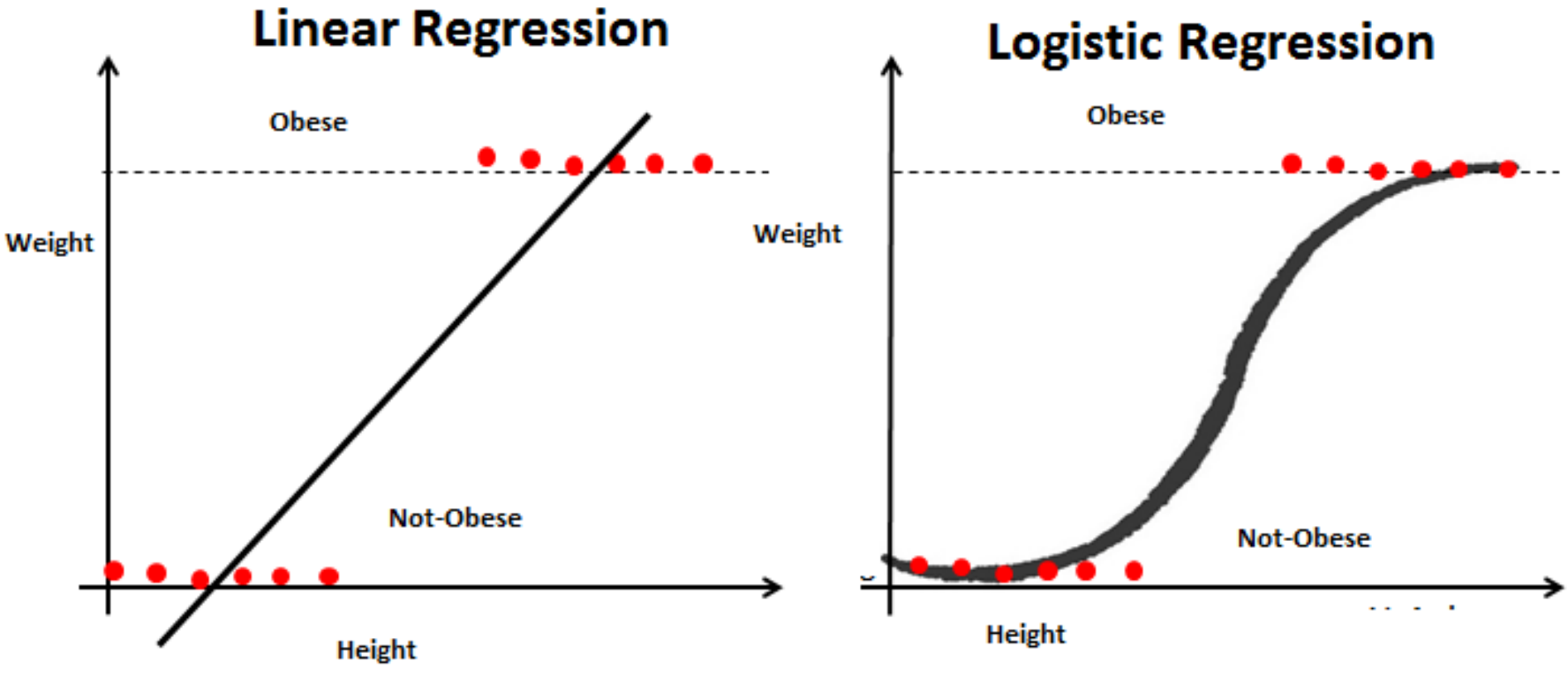
Example
Let us consider a problem where we are given a dataset containing Height and Weight for a group of people. Our task is to predict the Weight for new entries in the Height column.

So we can figure out that this is a regression problem where we will build a Linear Regression model. We will train the model with provided Height and Weight values. Once the model is trained we can predict Weight for a given unknown Height value.

Linear Vs Logistic Regression

Now suppose we have an additional field Obesity and we have to classify whether a person is obese or not depending on their provided height and weight. This is clearly a classification problem where we have to segregate the dataset into two classes (Obese and Not-Obese).
So, for the new problem, we can again follow the Linear Regression steps and build a regression line. This time, the line will be based on two parameters Height and Weight and the regression line will fit between two discreet sets of values. As this regression line is highly susceptible to outliers, it will not do a good job in classifying two classes.
To get a better classification, we will feed the output values from the regression line to the sigmoid function. The sigmoid function returns the probability for each output value from the regression line. Now based on a predefined threshold value, we can easily classify the output into two classes Obese or Not-Obese.

Linear Vs Logistic Regression



The Similarities between Linear Regression and Logistic Regression

- Linear Regression and Logistic Regression both are supervised Machine Learning algorithms.
- Linear Regression and Logistic Regression, both the models are parametric regression i.e. both the models use linear equations for predictions

The Differences between Linear Regression and Logistic Regression

- Linear Regression is used to handle regression problems whereas Logistic regression is used to handle the classification problems.
- Linear regression provides a continuous output but Logistic regression provides discreet output.
- The purpose of Linear Regression is to find the best-fitted line while Logistic regression is one step ahead and fitting the line values to the sigmoid curve.
- The method for calculating loss function in linear regression is the mean squared error whereas for logistic regression it is maximum likelihood estimation.