Regression deals with a predictive modelling technique. It shows the relation between independent also known as predictor variable and dependent or the target variable.

Logistic regression comes into use when the output value is of categorical format such as yes/no, 1 or 0, true or false , high or low.

As in linear regression we can get output values over a large range but for categorical output it should be between 1 and 0 hence the linear regression curve is to be clipped between 1 and 0 which is performed using logistic regression .so to clip values between 0 or 1, we use a sigmoid function (1/1 + exp(-(h(x))).

Concept of threshold:

This is used to decide whether the output value (between 0 to 1) rounds off to give the output as 0(high) or 1(low).

The output values between threshold value (usually 0.5) and 1 is rounded to and the value below threshold (usually less than 0.5)is rounded to 0.

Concept of log likelihood:

Using the concept of linear regression:

h(x) =β1X1+β2X2+β3X3+……….βnXn

putting the above equation in sigmoid function:

Y = (1/1 + exp(-(h(x))).

Now, by taking inspiration from the Bernoulli’s traits, we find the log likelihood function and differentiating it to find the gradient ascent update equation. Hence the likelihood is defined as:

Log likelihood becomes:

Differentiating it,

The gradient ascent update equation becomes:

Comparison of linear and logistic regression:

|  |  |
| --- | --- |
| Linear Regression | Logistic Regression |
| Continuous variables | Categorical variables |
| Solve regression problem | Solve classification problem |
| Straight line graph | Uses sigmoid curve |

Application of logistic regression:

1. It is used for weather forecast such as whether it is cloudy or not or it will rain or not.
2. Multiclass classification like dividing organisms into various sub-groups like vertebrates, amphibians etc.
3. Medical uses like diagnosis of various disease.