Lesson 12

Machine Learning – Logistic Regression

Kush Kulshrestha

Introduction

- One of the most simple algorithm used for 2 class classification problem.
- Logistic Regression can be used for various classification problems such as spam detection. Diabetes prediction, if a given customer will purchase a particular product or will they churn another competitor, whether the user will click on a given advertisement link or not.
- Just like Linear Regression uses optimize Least Square function to calculate the coefficients, Logistic Regression uses Log Likelihood function to find the value of the best fit parameters.

$$LogarithmicLoss = \frac{-1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij} * \log(p_{ij})$$

- Logistic Regression is the basic algorithm used to define a Neural Network.

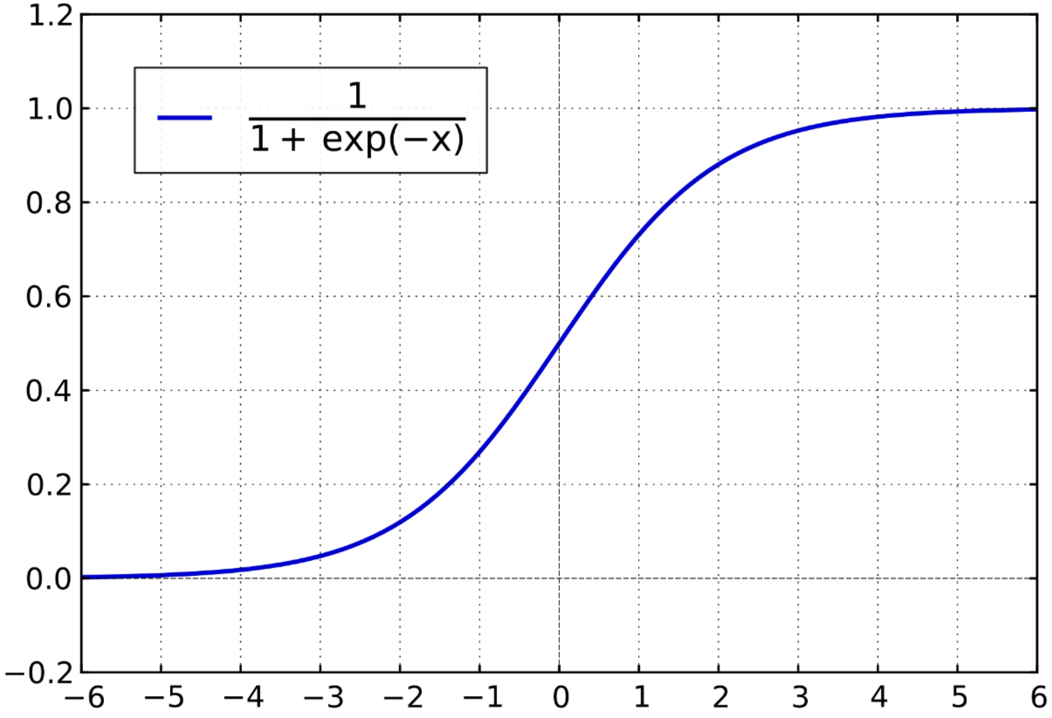
Sigmoid Function Refresher

The logistic function or a Sigmoid function is a function which takes any real value between zero and one. It is defined as

$$y = 1 / (1 + e^{-x})$$

where y is the scaled value and x is the input value.

The logistic function transforms the original range of [-∞,∞] to [0,1] and also has a linear part on the transform.



Probability, Odds and Log odds

Probability of an event = (No of Success / Total no of cases)

Odds = (One outcome / all other outcomes)

Examples:

- For rolling a dice and getting 1 –

Probability = (1/6)
Odds = (1/5)

- Rolling a dice and getting an even number –

Probability = (3/6) = 0.5
Odds = (3/3) = 1

- Dice roll less than 5 –

Probability = (4/6) = (2/3)
Odds = (4/2) = 2

Conversion:

$$odds = \frac{probability}{1 - probability}$$

$$probability = \frac{odds}{1 + odds}$$

What is Logistic Regression

Linear regression is continuous response is modelled as a linear combination of the features.

$$y = \beta_0 + \beta_1 x$$

In Logistic regression, log-odds of a categorical response being "true" (1) is modelled as a linear combination of the features. This is called the **logit function**.

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

The equation can be rearranged into the **logistic function**:

$$\mathbf{p} = \frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}}$$

Conclusion –

- Logistic regression outputs the **probabilities of a specific class.**
- Those probabilities can be converted into **class predictions.**
- The **logistic function** takes on an "s" shape and is bounded by 0 and 1.

Interpreting Logistic Regression Coefficients

Final equation of Logistic regression is:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

Interpreting beta:

‘1’ unit increase in ‘x' is associated with a B1 unit increase in the log-odds of event for which p is probability of.

Log odds are: $\log\left(\frac{p}{1-p}\right)$

Hence, positive coefficients increase the log-odds of the response (and thus increase the probability), and negative coefficients decrease the log-odds of the response (and thus decrease the probability).