

Cross Validation

Linear Regression

Print slope , intercept & line

Linear regression

In linear regression, the equation that describes the factor-response relationships is $Y=mX+C$ where Y and X are vectors that describe the response variable and the factor variable respectively.

Multiple regression

In a multiple regression case, we're interested in the impact of not only one, but many different factors, on the response variable.

The equation can be written as $Y = m_1X_1 + m_2X_2 + C$ where m_1 and m_2 are the coefficients of factors X_1 and X_2 respectively.

Polynomial Regression

For polynomial regression, we might use higher powers of X to describe Y , as described in

$$Y = m_1 X + m_2 X^2 + C$$

where m_1 and m_2 are coefficients of the first and second powers of the factor.

Logistic regression

Logistic regression is a statistical method for analysing a dataset in which there are one or more independent variables that determine an outcome.

The outcome is measured with a variable in which there are only two possible outcomes. It is used to predict a binary outcome (1 / 0, Yes / No, True / False) given a set of independent variables.

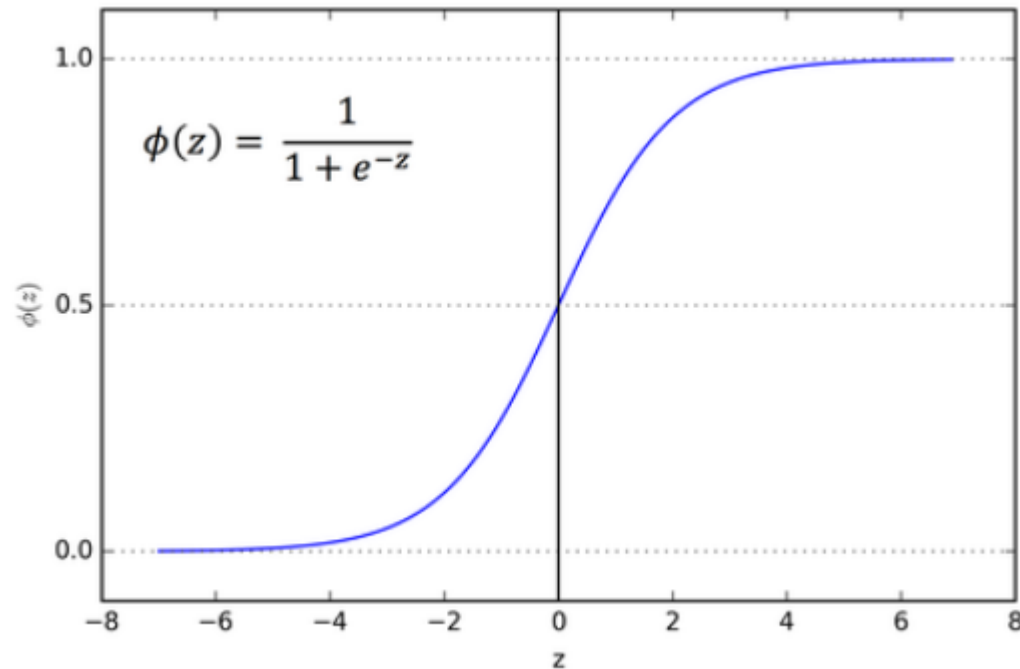
Applications

- Spam Detection : Predicting if an email is Spam or not
- Credit Card Fraud : Predicting if a given credit card transaction is fraud or not
- Health : Predicting if a given mass of tissue is benign or malignant
- Marketing : Predicting if a given user will buy an insurance product or not
- Banking : Predicting if a customer will default on a loan

sigmoid/logistic function

In logistic regression, the goal is to determine a mathematical equation that can be used to predict the probability of event

The sigmoid/logistic function is given by the following equation.



Precision, Recall

$$\text{precision} = \frac{tp}{tp + fp},$$

$$\text{recall} = \frac{tp}{tp + fn},$$