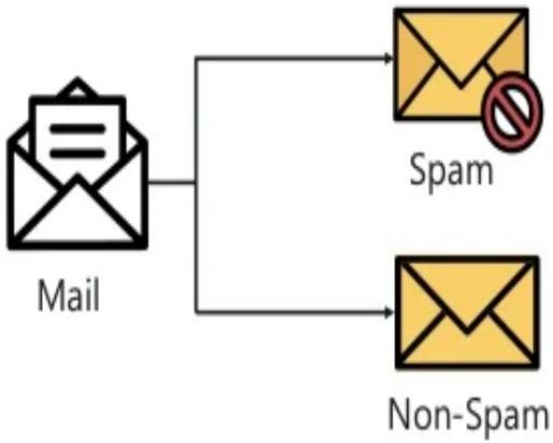
Regression Vs Classification

Classification

Classification is the task of predicting a discrete class label

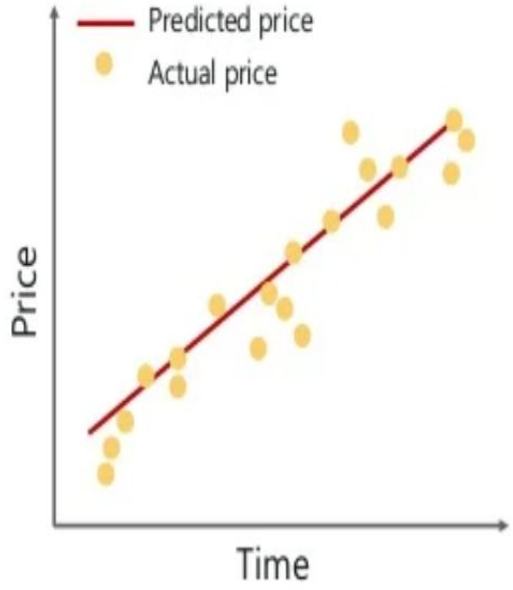
- In a classification problem data is classified into one of two or more classes
- A classification problem with two classes is called binary, more than two classes is called a multi-class classification



Regression

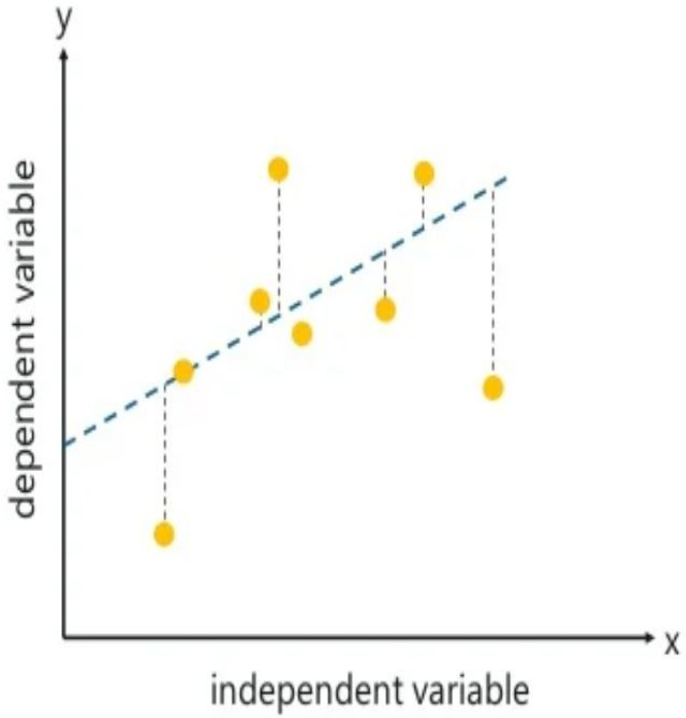
Regression is the task of predicting a continuous quantity

- A regression problem requires the prediction of a quantity
- A regression problem with multiple input variables is called a multivariate regression problem



What Is Linear Regression?

Linear Regression is a method to predict dependent variable (Y) based on values of independent variables (X). It can be used for the cases where we want to predict some continuous quantity.



- *Dependent variable (Y):*

The response variable who's value needs to be predicted.

- *Independent variable (X):*

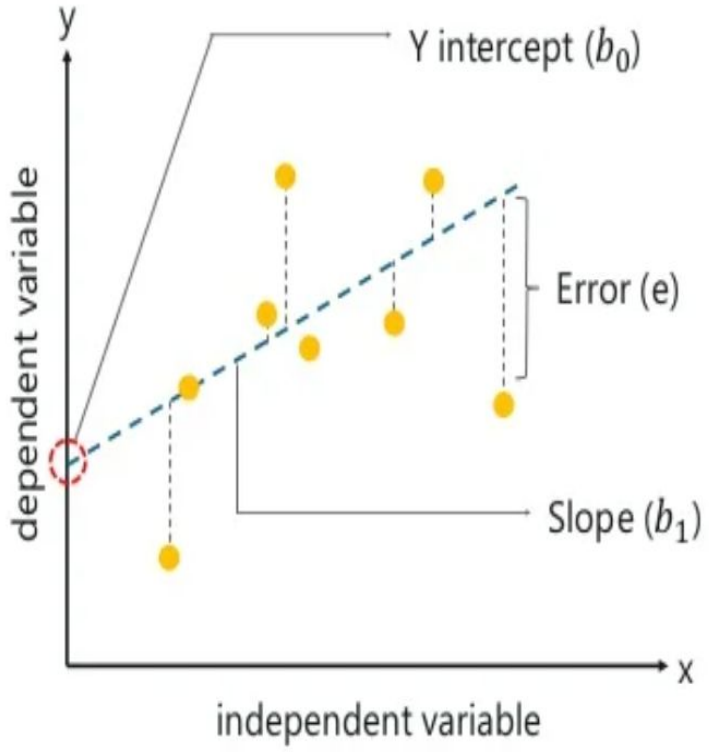
The predictor variable used to predict the response variable.

The following equation is used to represent a linear regression model:

$$Y = b_0 + b_1 x + e$$

What Is Linear Regression?

Linear Regression is a method to predict dependent variable (Y) based on values of independent variables (X). It can be used for the cases where we want to predict some continuous quantity.

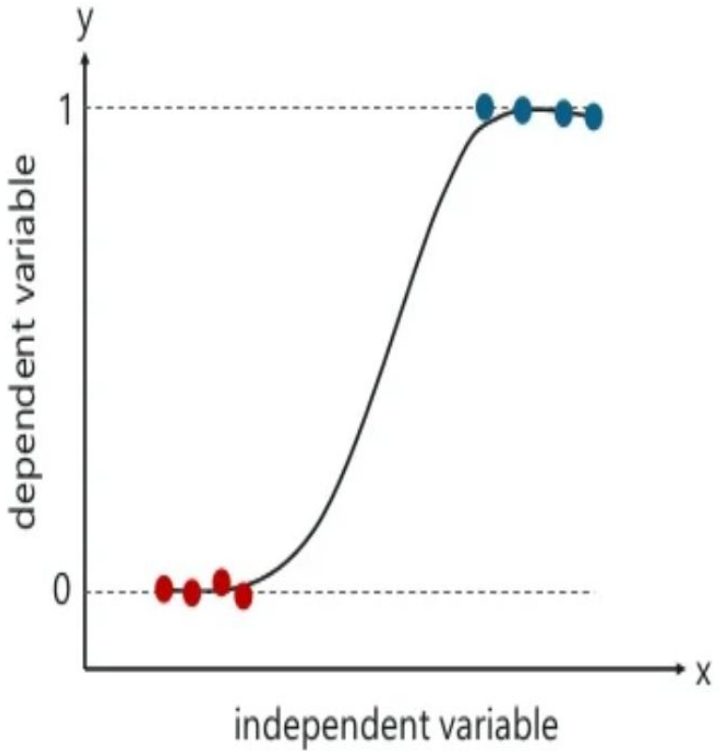


$$Y = b_0 + b_1 x + e$$

- Error
- independent variable
- Slope
- Y intercept
- dependent variable

What Is Logistic Regression?

Logistic Regression is a method used to predict a dependent variable, given a set of independent variables, such that the dependent variable is categorical.



- *Dependent variable (Y):*

The response binary variable holding values like 0 or 1, Yes or No, A, B or C

- *Independent variable (X):*

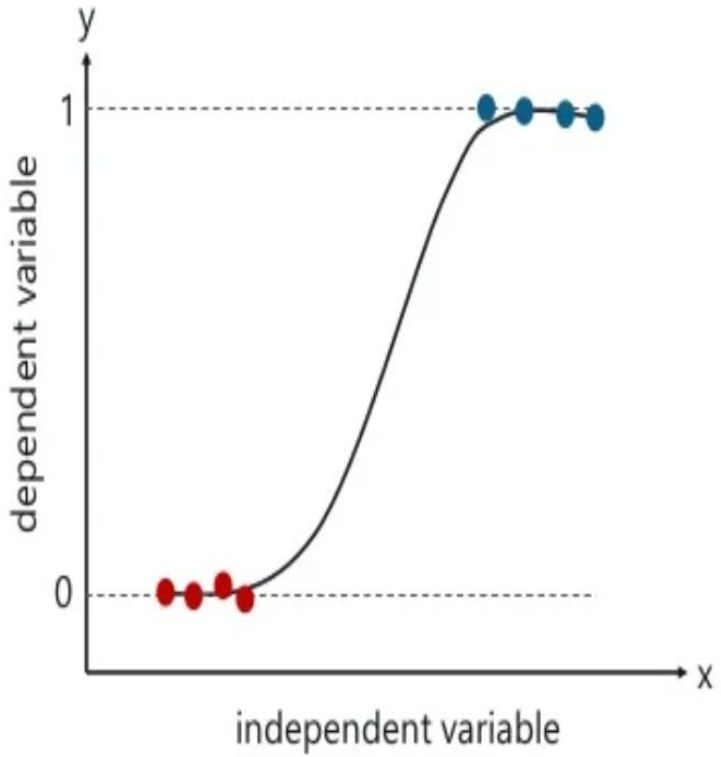
The predictor variable used to predict the response variable.

The following equation is used to represent a linear regression model:

$$\log\left(\frac{Y}{1-Y}\right) = C + B1X1 + B2X2 + \ldots.$$

What Is Logistic Regression?

Logistic Regression is a method used to predict a dependent variable, given a set of independent variables, such that the dependent variable is categorical.



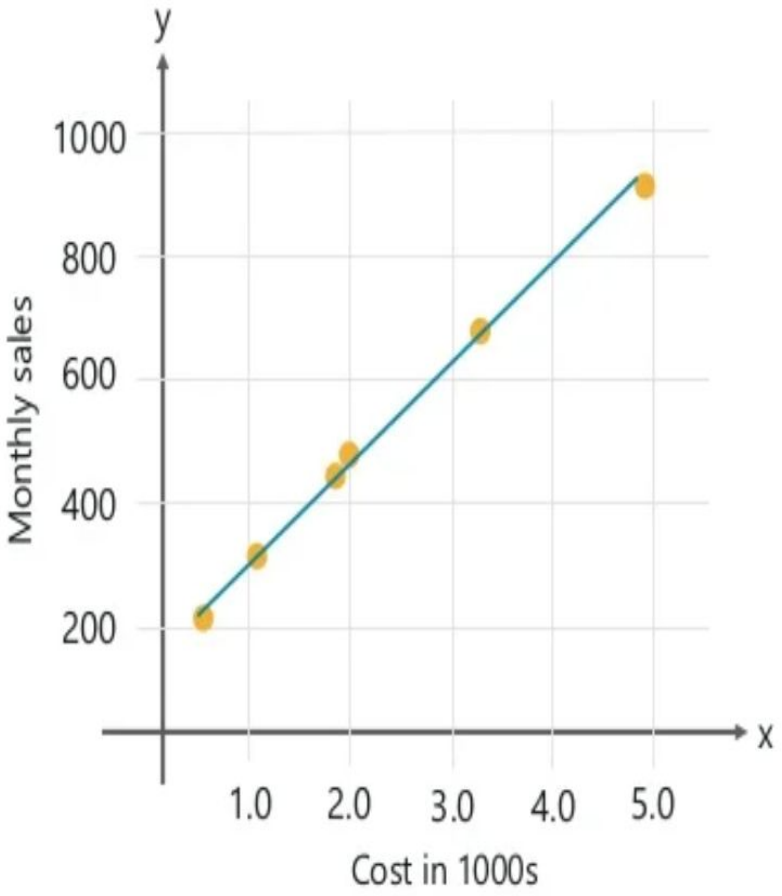
$$\log\left(\frac{Y}{1-Y}\right) = C + B1X1 + B2X2 + ...$$

- Y is the probability of an event to happen which you are trying to predict
- x1, x2 are the independent variables which determine the occurrence of an event i.e. Y
- C is the constant term which will be the probability of the event happening when no other factors are considered

Linear Regression Use Case

To forecast monthly sales by studying the relationship between the monthly e-commerce sales and the online advertising costs.

Monthly sales	Advertising cost In 1000s
200	0.5
900	5
450	1.9
680	3.2
490	2.0
300	1.0



Logistic Regression Use Case

To predict if a student will get admitted to a school based on his CGPA.

Admission	CGPA
0	4.2
0	5.1
0	5.5
1	8.2
1	9.0
1	9.1