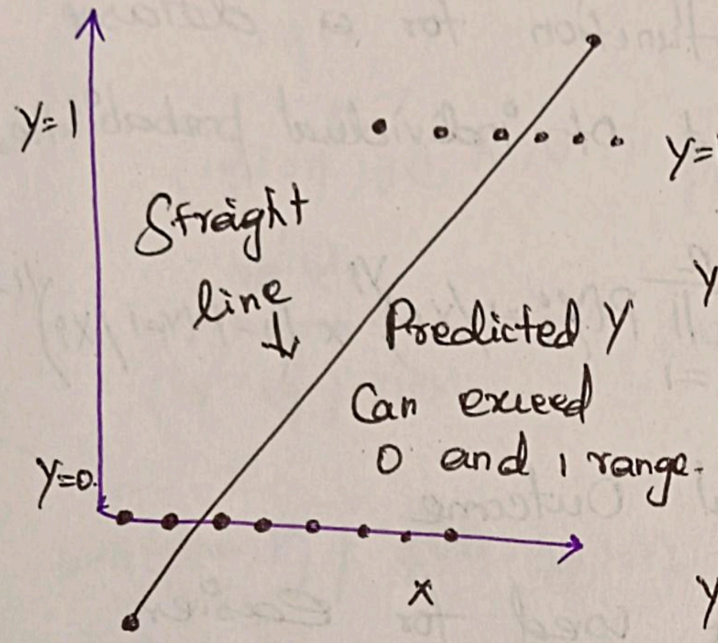
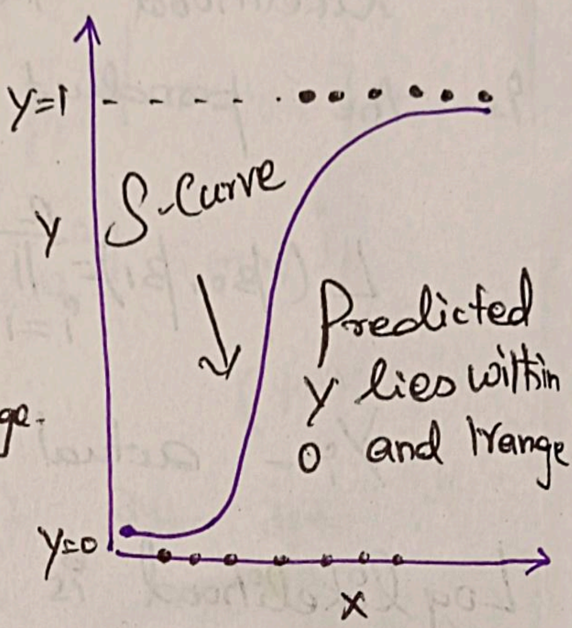
Logistic Regression



Linear Regression



Logistic Regression.

Logistic regression is suitable for binary classification problem.

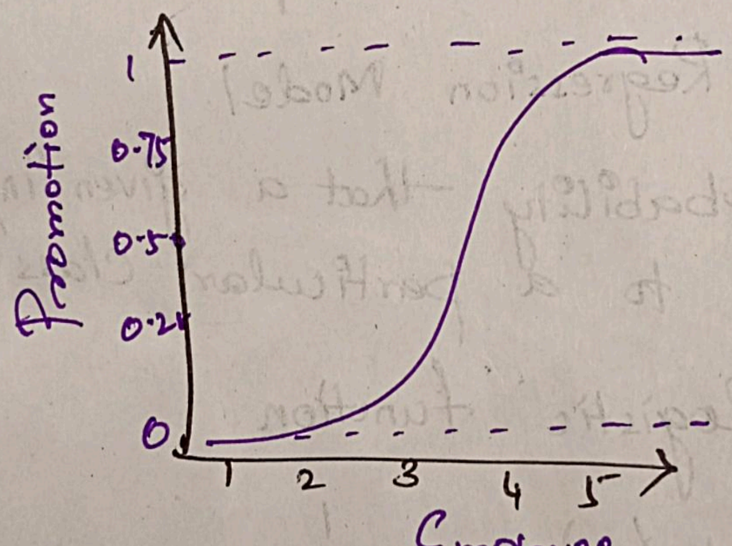
Example:

How logistic regression work?

Example: Organization wants to know whether an employee would get a promotion or not based on their performance?

linear graph is not Suitable.

Person get promotion or not
zero or one
Convert it into S curve



Threshold = 0.5

Odds of Success

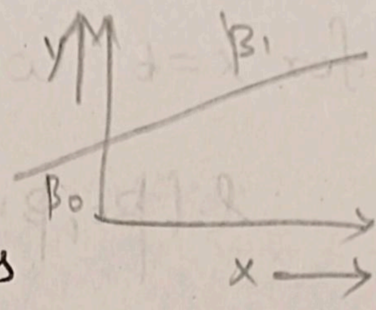
$$\text{Odds}(\theta) = \frac{\text{Probability of an event happening}}{\text{Probability of an event not happening}}$$

logit

$$\text{Odds}(\theta) = \frac{P}{1-P}$$

Value of odds range from 0 to ∞

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



Predict the odds of Success

take log on odds

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

Exponentiating both sides.

$$e^{\ln \frac{P(x)}{1-P(x)}} = e^{\beta_0 + \beta_1 x}$$

$$\frac{P(x)}{1-P(x)} = e^{\beta_0 + \beta_1 x}$$

Let
Assume $Y = e^{\beta_0 + \beta_1 x}$

Then $$\frac{P(x)}{1-P(x)} = Y$$

$$P(x) = Y(1 - P(x))$$

$$P(x) = Y - Y(P(x))$$

$$P(x) + YP(x) = Y$$

$$P(x)(1+y) = y$$

$$P(x) = \frac{y}{1+y}$$

$$P(x) = \frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}}$$

Simplify divide $e^{\beta_0+\beta_1 x}$.

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

The equation of the sigmoid function is

$$\boxed{P(x) = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}}}$$

Let

$$Pr(y=1/x) = P(x)$$

Parameter Estimation

Problem: Predict whether a student passes (1) or fails (0) based on the number of study hours.

Dataset:

Hours Studied (x)	Pass/Fail (y)
2	0
4	0
6	1
8	1
10	1

Logistic Regression Model

model probability that a given input x belongs to a particular class

logistic function.

$$P(y=1/x) = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}}$$

$P(y=1/x)$ - probability that the student passes $(y=1)$

β_0 - intercept β_1 - coefficient of x.

We use Maximum Likelihood function (MLE) to estimate the coefficients.

likelihood function for a dataset is the product of individual probabilities:

$$L(\beta_0, \beta_1) = \prod_{i=1}^{n} P(Y_i = 1/X_i)^{Y_i} \times (1 - P(Y_i = 1/X_i))^{(1-Y_i)}$$

Y_i - actual Outcome

Log likelihood is used for easier Computation:

$$l(\beta_0, \beta_1) - \sum_{i=1} [Y_i \log(P(Y_i - 1) | X_i) + (1 - Y_i) \log(1 - P(Y_i = 1 / X_i))]$$

Initialize Coefficients

Lets Start by

$$\beta_0 = 0, \quad \beta_1 = 0.5$$

Calculate the predicted probability

then

Calculate predicted probabilities

for $X=2$

$$P(y=1/X=2) = \frac{1}{1+e^{-(0+0.5\times2)}} = \frac{1}{1+e^{-1}} = 0.7311$$

for $x=4$

$$P(y=1/x=4) = \frac{1}{1+e^{-(0+0.5\times4)}} =$$

for $x=6$

$P(y=1/x=6)$

for $x=8$

$P(y=1/x=8)$

for $x=10$

$P(y=1/x=10)$

Compute log likelihood.

actual $y = [0, 0, 1, 1, 1]$

for $x=2$ where $y=0$

$$l(\beta_0, \beta_1) = \log(1-0.7311)$$
$$= \log(0.2689) = -1.3133$$

for $x = 2$ where $y = 0$

$$l(\beta_0, \beta_1) = \log(1 - 0.8808)$$

$$= -2.1269$$

for $x = 6$ where $y = 1$

$$l(\beta_0, \beta_1) = \log(0.9526)$$

$$= -0.0480$$

update the coefficient / Gradient Descent)

$$\beta_0^{new} = \beta_0^{old} + \alpha \times \frac{\partial l(\beta_0, \beta_1)}{\partial \beta_0}$$

$$\beta_1^{new} = \beta_1^{old} + \alpha \times \frac{\partial l(\beta_0, \beta_1)}{\partial \beta_1}$$

α - learning rate

Partial derivatives