27-08-2024

Google classroom code: 562biah

classification & Regression

supervised
unsupervised
semi-supervised

04-08-2024

Linear Regression:

Mathematically the relationship can be represented with the help of following equation.

$$y = mx + b$$

Dependent (y) ; Independent (x)

m = slop
b = constant ; $x = 0 \rightarrow (y = b)$

model with multiple features:

$\rightarrow y' = b + w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + w_5x_5$

$\boxed{* \ b = \text{bias}}$

→ Pounds → x_1

→ Engine Displacement → x_2

→ Acceleration → x_3

→ Number of cylinders → x_4

→ Horsepower → x_5

Loss:

Train

Test

Absolute value

$|2-5|$

$= |-3|$

$= 3$

L1 Loss → $\sum |\text{actual value} - \text{predicted value}|$

MAE (Mean absolute ERROR) → $\frac{\sum |\text{actual value} - \text{predicted value}|}{N}$

L2 Loss → $\sum |\text{actual value} - \text{predicted value}|^2$

MSE (Mean squared ERROR) → $\frac{\sum |\ |^2}{N}$

Consider a very simple data set

$x = (3, 6, 9)$
$y = (6.9, 12.1, 16)$

$$h_{\theta_0, \theta_1}(x) = \theta_0 + \theta_1 x$$

[our hypothesises is a function parameterised by θ_0, θ_1].

10-09-2024

Linear Regression

$$y = mx + b$$

$$y = \theta_0 + \theta_1 x_1 + \theta_2 x_2$$

Cost Function:

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2$$

$$\text{Price} = \theta_0 + \theta_1 \times \text{size} + \theta_2 \times \text{Bedroom}$$

$$= (-80000) + 200 \times 1850 + 10000 \times 4$$

$$= -80000 + 370000 + 40000$$

$$= 330000 \text{ dollar}$$

$$\boxed{\begin{aligned} \theta_0 &= -80000 \\ \theta_1 &= 200 \\ \theta_2 &= 10000 \end{aligned}}$$

error

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (\text{Actual MPG} - \text{predicted MPG})^2$$

$$= \frac{1}{5}\left(2^2 + (-1)^2 + (-2)^2 + 1^2 + 1^2\right)$$

$$= \frac{1}{5}(4 + 1 + 4 + 1 + 1)$$

$$= \frac{11}{5}$$

$$= 2.2$$

Gradient Descent:

Gradient descent is a mathematical technique that iteratively find the weights and bias that produce the model with the lowest loss.

Batch Gradient Descent:

Initial Guess:

while Not converged:

$\alpha \rightarrow$ The learning rate.

$\theta \rightarrow$ Parameters (weights) at the n-th iteration.

stochastic Gradient descent:

11-09-2024

If only 1 variable,

$$h_\theta(x) = \theta_0 + \theta_1 x_1$$

Batch + stochastic

Equation: [for Gradient Descent]

Batch Gradient Descent

Update Rule:

$$\begin{cases} \theta_0 = \theta_0 - \alpha\,(h_\theta(x^{(i)} - y^{(i)}) \\ \theta_1 = \theta_1 - \alpha\,((h_\theta(x^{(i)}) - y^{(i)})\,x_1^{(i)}) \end{cases}$$

<u>First iteration</u>

$$\alpha = 0.01$$

$\rightarrow$ $h_\theta(x) = 0.25 x_1 - 100$

<u>Data points:</u>

For $x_1^{(1)} = 2104$ $\quad y_1 = 400$

" $x_1^{(2)} = 1600$ $\quad y_2^{(2)} = 330$

" $x_1^{(3)} = 2400$ $\quad y_3^{(3)} = 369$

Batch GD (one iteration)

$\Rightarrow h_\theta(x^{(1)}) = 0.25 \times 2104 - 100$
$= 426$

$\Rightarrow h_\theta(x^{(2)}) = 0.25 \times 1600 - 100$
$= 300$

$\Rightarrow h_\theta(x^{(3)}) = 0.25 \times 2400 - 100$
$= 500$

Loss for $x_1^{(1)} = 426 - 400 = 26$

Loss for $x_1^{(2)} = 300 - 330 = -30$

Loss for $x_1^{(3)} = 500 - 369 = 131$

update θ_0

$\theta_0 = \theta_0 - \alpha \cdot \frac{1}{3}(26 + (-30) + 131)$

$= -100 - 0.01 \times \frac{1}{3} \times 127$

$= -100.42$

$$\theta_1 = \theta_1 - \alpha \cdot \frac{1}{3}\big((26 \times 2104) + (-30 \times 1600) + (131 \times 2400)\big)$$

$$= 0.25 - 0.01 \times \frac{1}{3} \times \{54704 - 48000 + 314400\}$$

$$= 0.25 - 1070.35$$

$$= -1070.10$$

17-09-2024

Next Tuesday: class Test

overfitting & underfitting

Q: Define linear Regression with a proper example-

$\boxed{\text{Loss, Gradient discent, Model}}$

Linear Regression is a process of mathematical system where multiple function are used in graphical and mathematical.

The equation are

$$y = mx + b$$

Here y are dependent variable and x are Independent variable. b is also know constant.

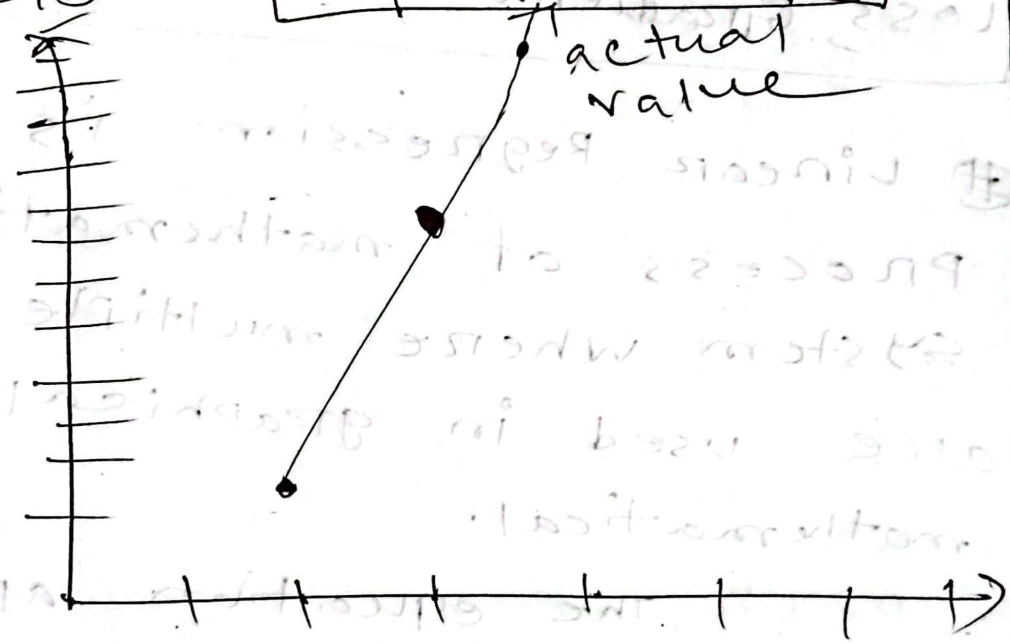
$\therefore y = 5x + 10$ let assume a data set

$m = 5$

$b = 10$

y	x	m	b
7	3	2	1
14	4	3	2
2	2	1	0

$y = mx + b$



$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2$

$x_1 \rightarrow$ bedroom

$x_2 \rightarrow$ size,

$\theta_0 = -10$

the $h_\theta(x)$ ~~20~~ $\theta_1 = 110$

$\theta_2 = 70$

18-09-2024

Exercise

$\boxed{\text{stochastics}}$

$h_\theta(x) = \theta_0 + \theta_1 x_1$

$\theta_0 = \theta_0 - \alpha\left((h_\theta(x^{(i)}) - y^{(i)}\right)$

$\theta_1 = \theta_1 - \alpha\left((h_\theta(x^{(i)}) - y^{(i)}) x_1^{(i)}\right)$

For the first data point:

$x_1^{(i)} = 2104, \quad y_1^{(i)} = 400 \qquad \alpha = 0.01$

$\Rightarrow h_\theta(x^{(1)}) = -100 + 0.25 \times 2104$

$= 426$

Update $\theta_0 = \theta_0 - \alpha\left(h_\theta(x_1^{(i)}) - y_1^{(i)}\right)$

$= -100 - 0.01(426 - 400)$

$= -100.26$

update θ_1

$= 0.25 - 0.01\,(26 \times 2104)$

$= -546.79$

For the second data point:

$h_\theta(x^{(2)}) = -100 + 0.\cancel{25}^{26} \times 1600$

$= \cancel{300}\ 299.74$

update $\theta_0 = -100 - 0.01\,(\underbrace{\cancel{290 - 330}}_{}\ 299.74 - 330)$

$= -100 + 0.3$

$= -99.7$

$\theta_1 = -100 - 0.01\,(\cancel{300} - 30 \times 1600)$

$=$

$h_\theta(x^{(2)})$

$= -100.26 + (-546.79 \times 1600)$

$= -874964.26$

Update

$\theta_0 = -100.26 - 0.01(-874964.26 - 330)$

$= -100.26 - 0.01 \times (-875294.26)$

$= -100.26 + 8752.9426$

$= 8652.6826$

$\theta_1 = \cancel{-100.26}\ -546.79 - 0.01(8752.9426 \times 1600)$

$= \cancel{-140147.341}6$

$= -140593.8716$

$$\theta_0 = \theta_0 - \alpha \cdot \frac{1}{m}(h_\theta(x^{(i)}) - y^{(i)})$$
$$\theta_1 = \theta_1 - \alpha \cdot \frac{1}{m}((h_\theta(x^{(i)}) - y^{(i)}) \cdot x^{(i)})$$

own Preparation for class Test:

Batch GD

size	price
2104	400
1600	330
2400	369

$\theta_0 = -100$

$\theta_1 = 0.25$

$h_\theta(x) = \theta_0 + \theta_1 x$

$= -100 + 0.25x$

$= 0.25x - 100$

Batch GD:

one iteration:

$h_\theta(x^{(1)}) = \cancel{0.25+} -100 + 0.25 \times 2104$

$= 426$

$h_\theta(x^{(2)}) = -100 + 0.25 \times 1600$

$= 300$

$h_\theta(x^{(3)}) = -100 + 0.25 \times 2400$

$= 500$

Error for

$(x^{(1)}) = 426 - 400$
$= 26$

$(x^{(2)}) = 300 - 330$
$= -30$

$(x^{(3)}) = 500 - 369$
$= 131$

Update θ_0:

* $\theta_0 = \theta_0 - \alpha \cdot \frac{1}{3}(26 + (-30) + 131)$

$= -100 - 0.01 \times \frac{1}{3} 127$

$= -100.4233$

* $\theta_1 = \theta_1 - \alpha \cdot \frac{1}{3}((26 \times 2104) + (-30 \times 1600)$
$+ (131 \times 2400))$

$= 0.25 - 0.01 \times \frac{1}{3} \times 321104$

$= -1070.1$

So, After one iteration the updated values are: $\theta_0 = -100.4233$
$\theta_1 = -1070.1$

$x_1 = 2104$
$y_1 = 400$

stochastic GD

For the first data point:

$\Rightarrow h_\theta(x^{(1)}) = 0.25 \times 2104 - 100$
$= 426$.

ERROR is:
$= 426 - 400$
$= 26$

update θ_0:

α = learning Rate

$\theta_0 = \theta_0 - \alpha(26)$
$= -100 - 0.01 \times 26$
$= -100.26$

update θ_1:

$\theta_1 = \theta_1 - \alpha(26 \times 2104)$
$= -546.79$

so, after one iteration the update values are:

$\theta_0 = -100.26$
$\theta_1 = -546.79$

$x_2 = 1600$
$y_2 = 330$

For the second data point:

$x_2 = 1600, \quad y_2 = 330$

$h_\theta(x_2) = \theta_0 + \theta_1 x_2$

$= -100.26 + (-546.79) \times 1600$

$= -874964.26$

Error is $= -874964.26 - 330$

$= -875294.26$

update $\theta_0 = \theta_0 - \alpha(\cancel{-874964.26}\ -875294.26)$

$= -100.26 - 0.01 \times (\cancel{-874964.26}\ -875294.26)$

$= \cancel{8652.6826}$ $= 8652.6826$

update θ_1

$= \theta_1 - \alpha(\cancel{-874964.26}\ -875294.26 \times 1600)$

$= -546.79 - 0.01 \times (\cancel{-1399942816}\ -1400470816)$

$= -546.79 + \cancel{13999428.16}\ 14004708.16$

$= \cancel{13998881.37}$ 14004161.37

$\theta_0 = 8652.6826$

$\theta_1 = \cancel{13998881.37}$ 14004161.37

Update θ_0

$$\cancel{= \theta_0 - \alpha(33597\,329571.7)}$$

For the third data point:

$$x_3 = 2400$$
$$y_3 = 369$$

$$h_\theta(x_3) = \theta_0 + \theta_1 x_3$$
$$= 8652.6826 + 14004161.37 \times 2400$$
$$= 33609995940.7$$

$$\text{Error} = 33609995940.7 - 369$$
$$= 33609995571.7$$

Update:

$$\theta_0 = \theta_0 - \alpha(33609995571.7)$$
$$= 8652.6826 - 0.01(33609995571.7)$$

$$= -336091303$$

Update:

$$\theta_1 = \theta_1 - \alpha(33609995571.7 \times 2400)$$
$$= 14004161.37 - 0.01 \times (33609995571.7 \times 2400)$$
$$= 14004161.37 - 806639893721$$
$$= -806625889560$$

$\theta_0 = -336091303$

$\theta_1 = -806625889560$

25-09-2024

লিখতে হবে

Logistic Regression :- page (3)

x_1	x_2	x_3	y
27	25	29	happy
50	50	60	sad

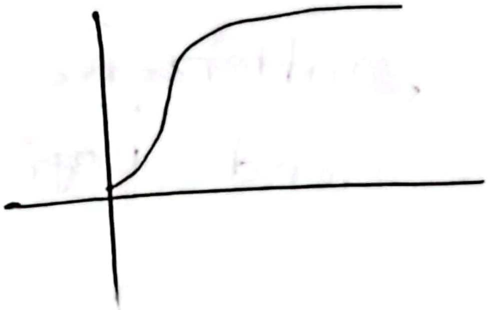
0, 1 →

Probability

Logistic
- → Binomial
- → Multinomial

sigmoid function

$\therefore f(x) = \frac{1}{1+e^{-x}}$



(মান)

$\therefore Y \to (0,1)$

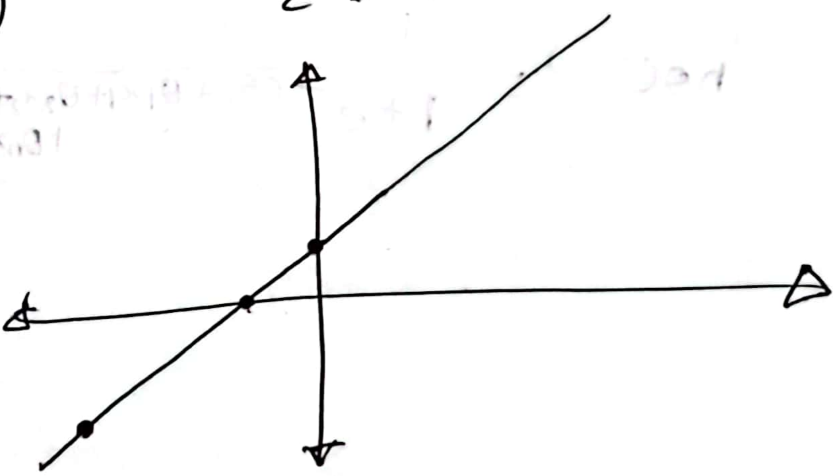
(logistic Regression model:)

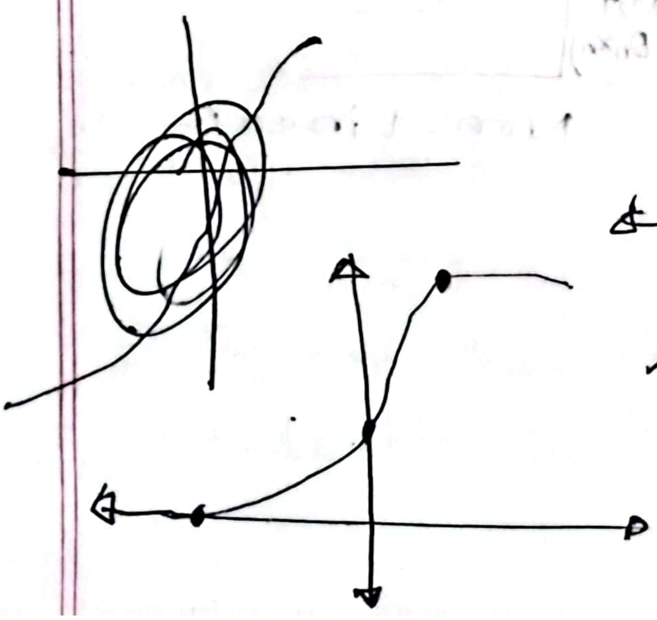
$$\boxed{z = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n}$$

$$\boxed{y' = \frac{1}{1+e^{-z}}}$$

$Y' = 1/(1+e^{-z})$

$z = 2x+5$





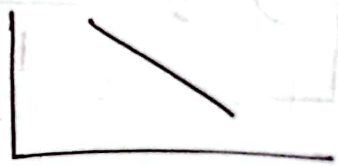
Page-(09) question is Important

Difference Between Linear and Logistic Regression.

01-10-2024

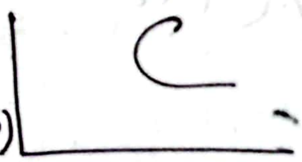
Decision Boundary.

The hypothesis function for logistic regression.



Linear

$$h_\theta(x) = \frac{1}{1+e^{-(\theta_0+\theta_1 x_1+\theta_2 x_2+\cdots+\theta_n x_n)}}$$



Non-Linear

[Decision boundary को - 2 Marks:]

Exercise - 01 (Page - 10) - [2 Marks:]

$\theta_0 = -50, \ \theta_1 = 6, \ \theta_2 = 1$

so the Hypothesis function

$$h_\theta(x) = \frac{1}{1+e^{-(-50+6x_1+1 \cdot x_2)}}$$

$\Rightarrow -(-50+6x_1+x_2) = 0$

$\Rightarrow 50 - 6x_1 - x_2 = 0$

$\Rightarrow 50 = 6x_1 + x_2$

$\Rightarrow x_2 = 6x_1 - 50$

This line represents the decision boundary.

$\Rightarrow$ Above the line ($x_2 > 6x_1 - 50$): the student is predicted to pass (class 1)

$\Rightarrow$ Below the line ($x_2 < 6x_1 - 50$); the student is predicted to fail (class 0)