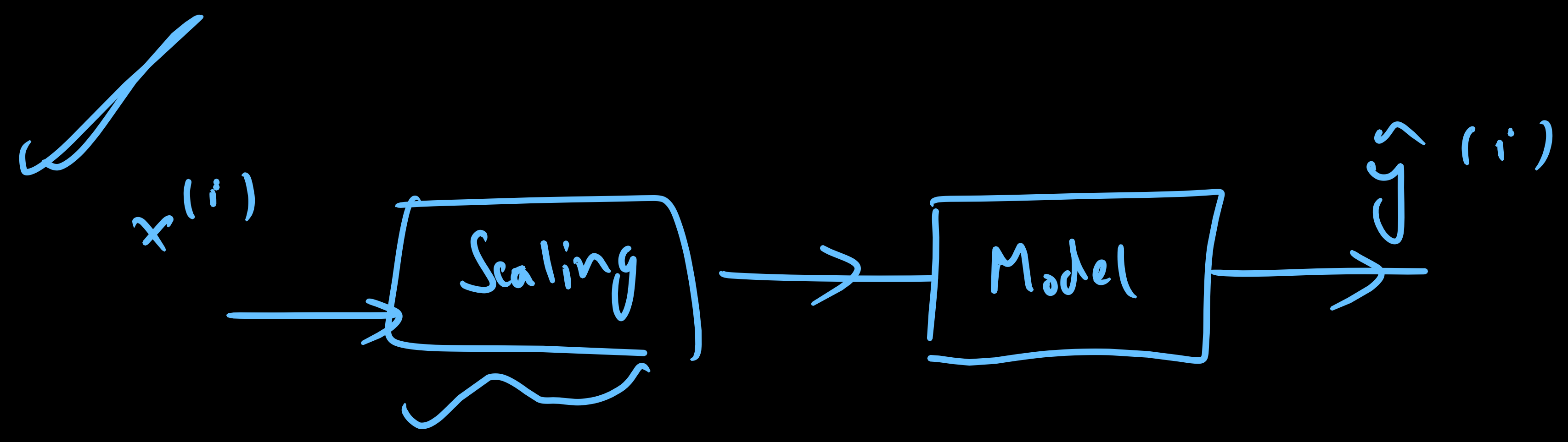
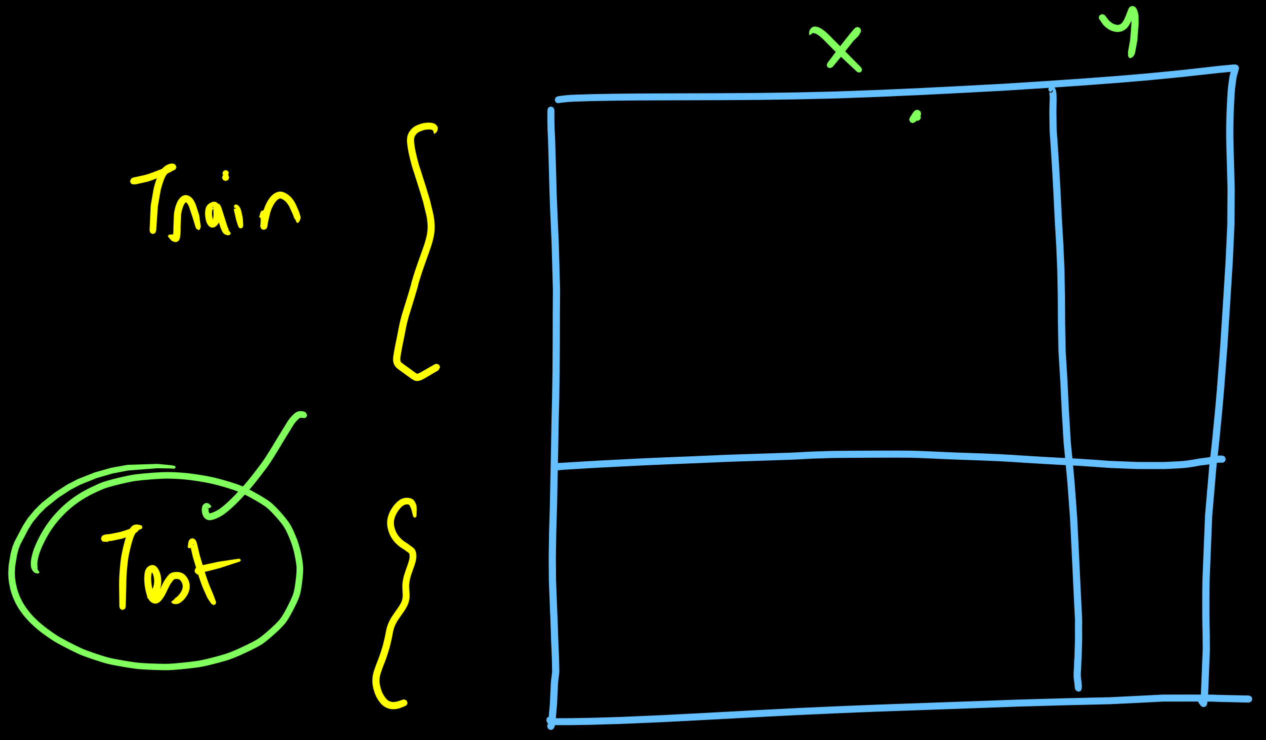
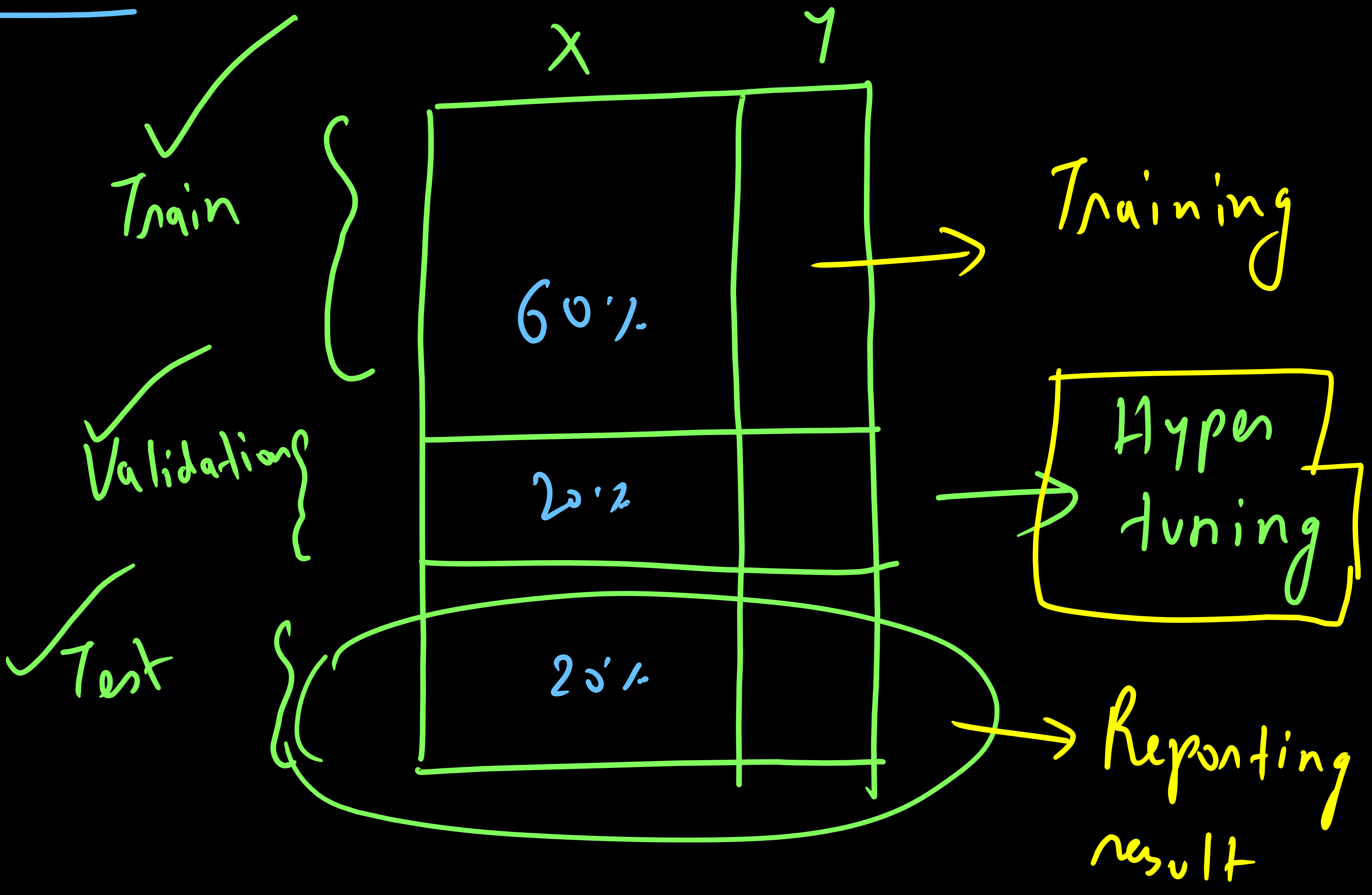

$x^{(i)}$
Scaling
Model
$\hat{y}^{(i)}$

Cross Validation





80%
10%

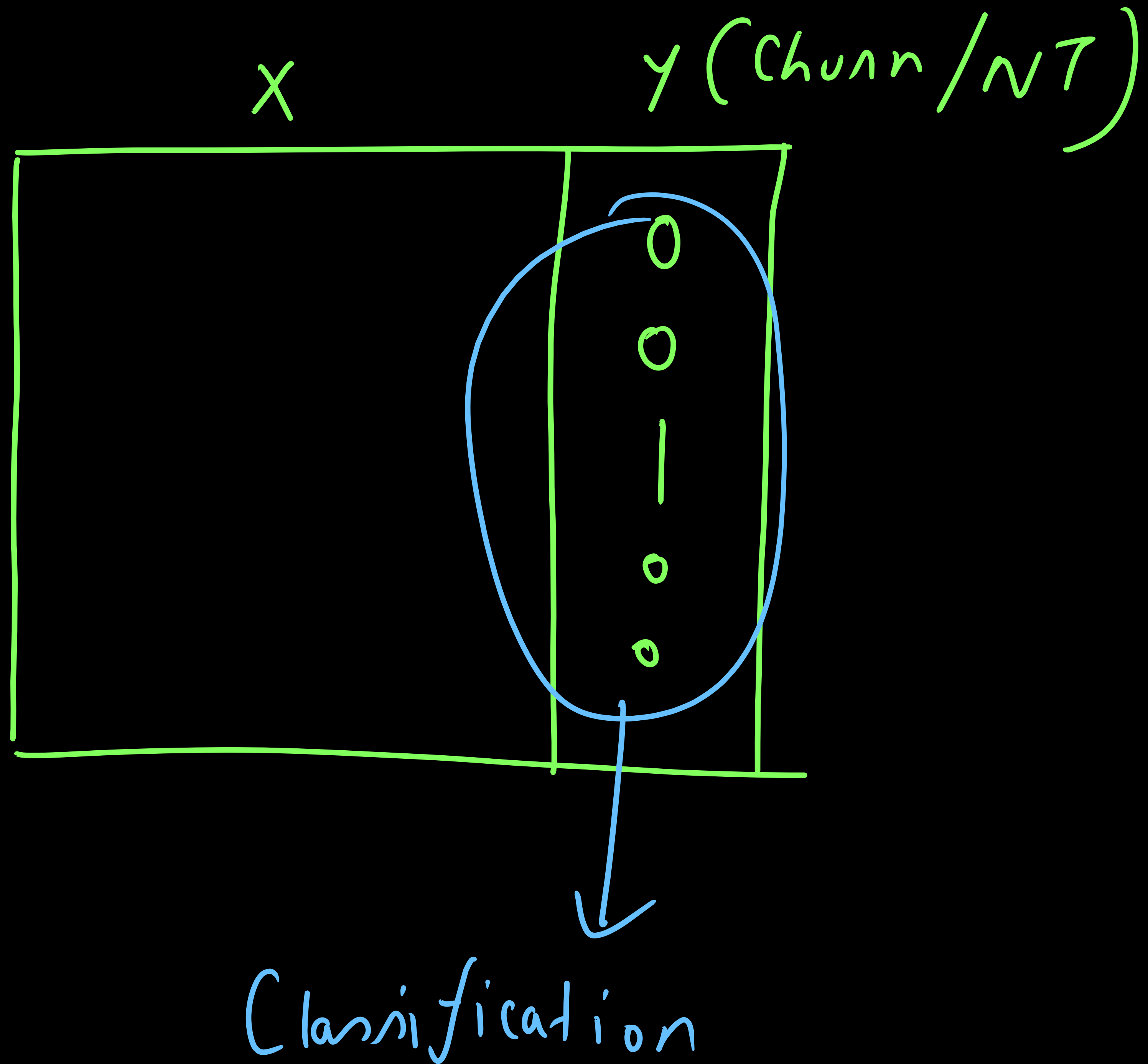
10%

→ Logistic Regression

A T & T

↳ Bad Net

↳ Poor Customer



```python
#0.6, 0.2, 0.2 split
from sklearn.model_selection import train_test_split
X_tr_cv, X_test, y_tr_cv, y_test = train_test_split(X, y, test_size=0.2, random_state=1)
X_train, X_val, y_train, y_val = train_test_split(X_tr_cv, y_tr_cv, test_size=0.25,random_state=1)
```

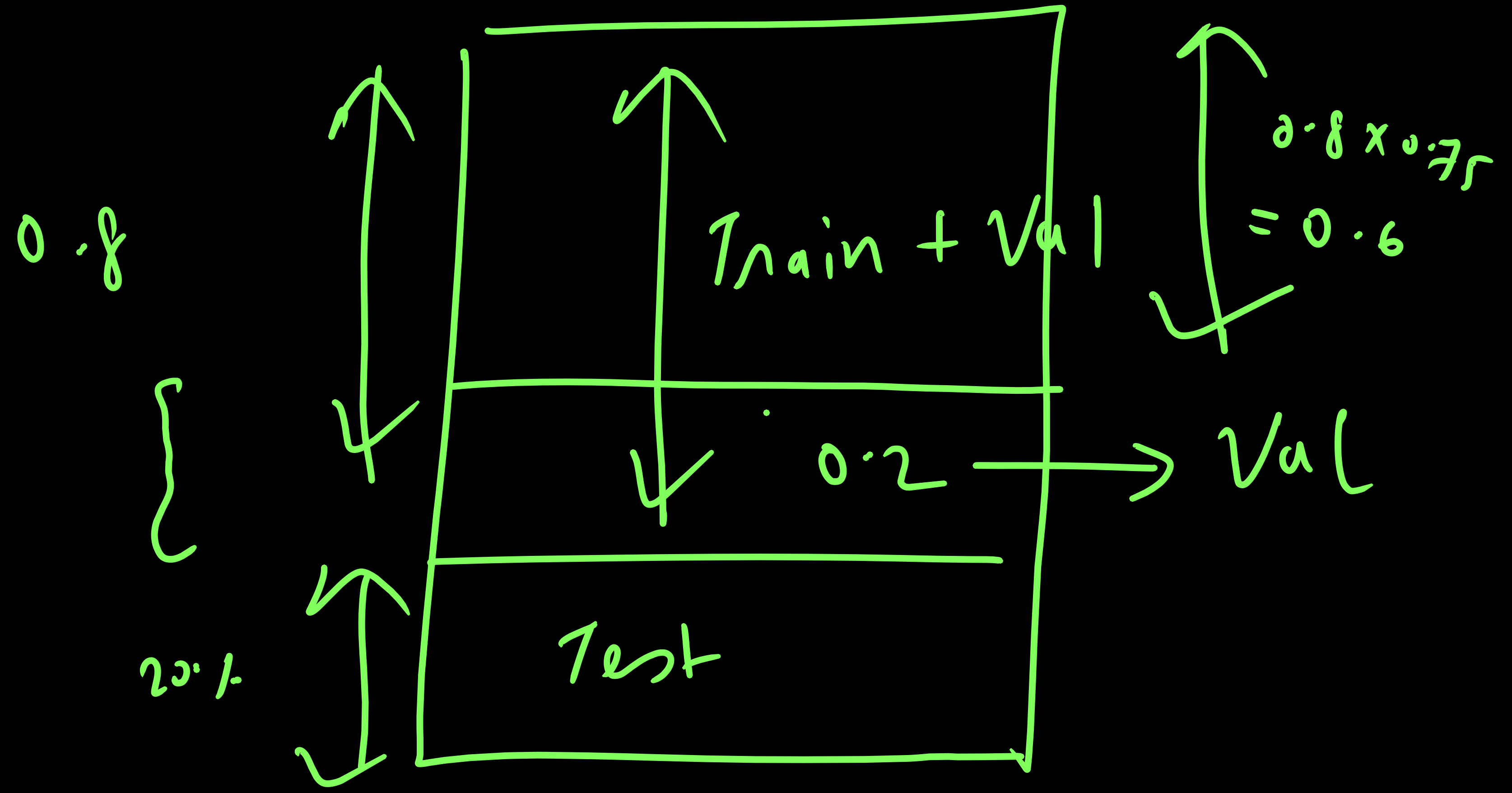
0.8

0.2

0.6

0.25×0.8

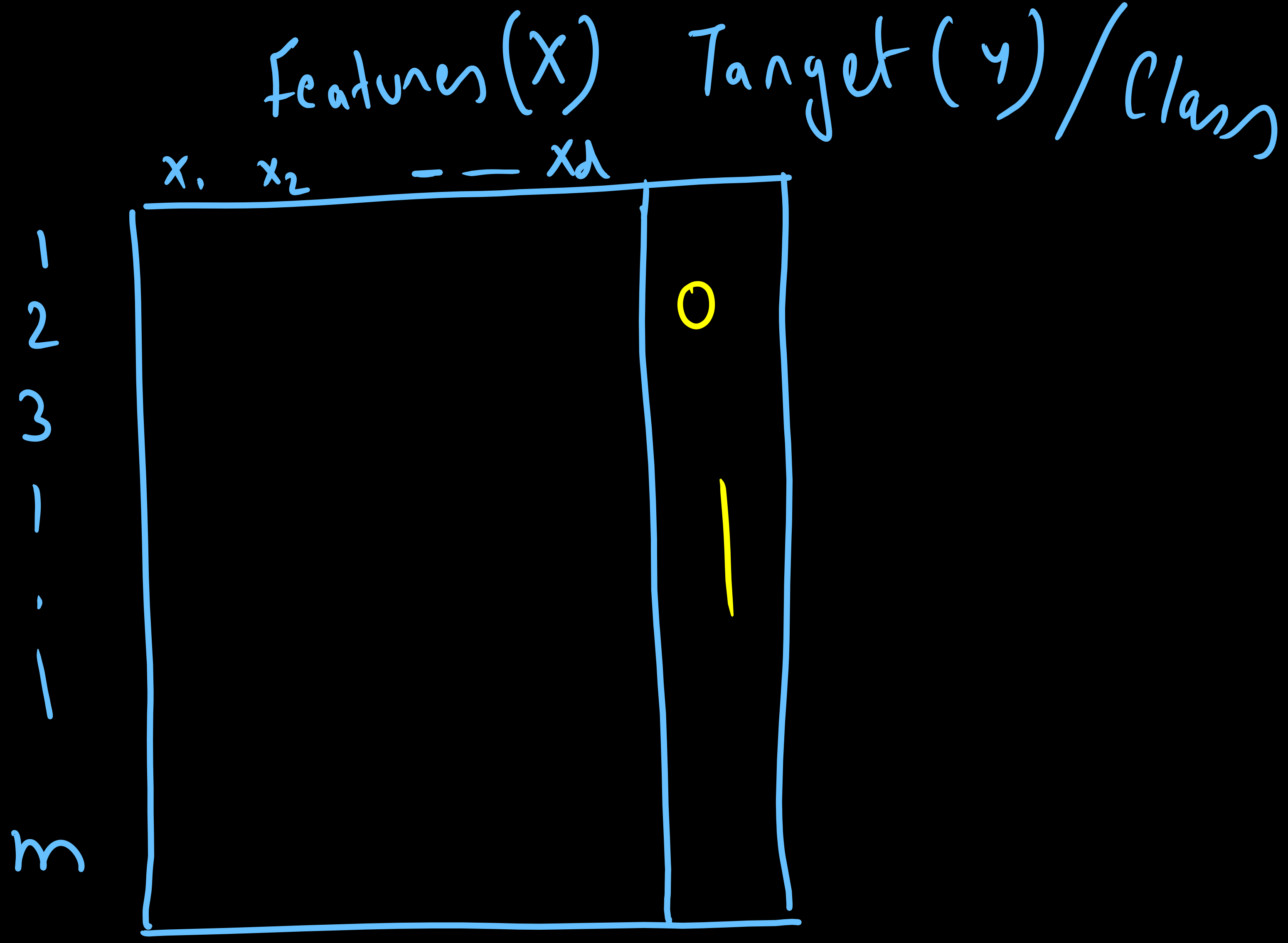
$= 0.2$



Classification

$\hookrightarrow$ Binary 0 or 1

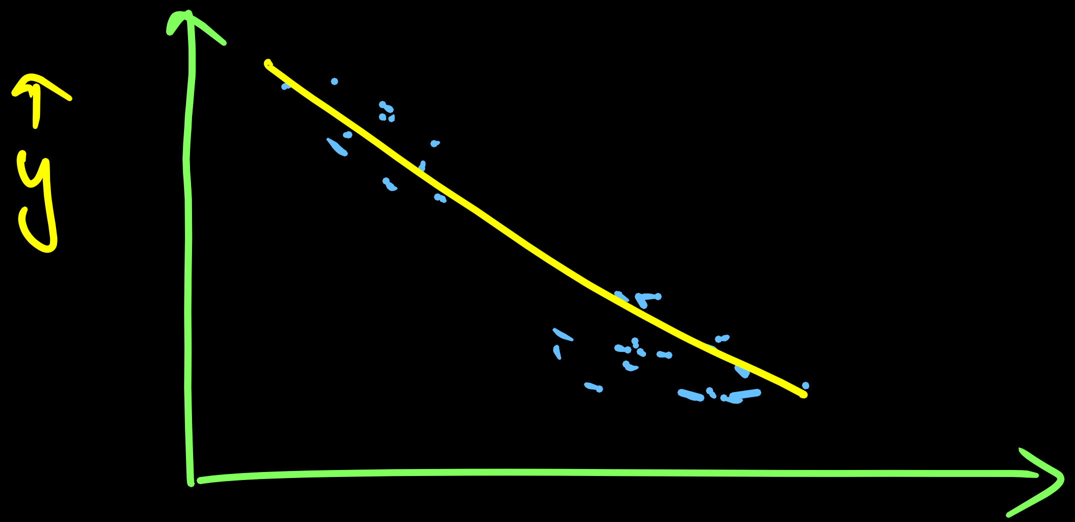
$\hookrightarrow$ Multi-class More than 2



$$X \rightarrow (m, d) \qquad y \rightarrow (m, 1)$$

Linear Regression

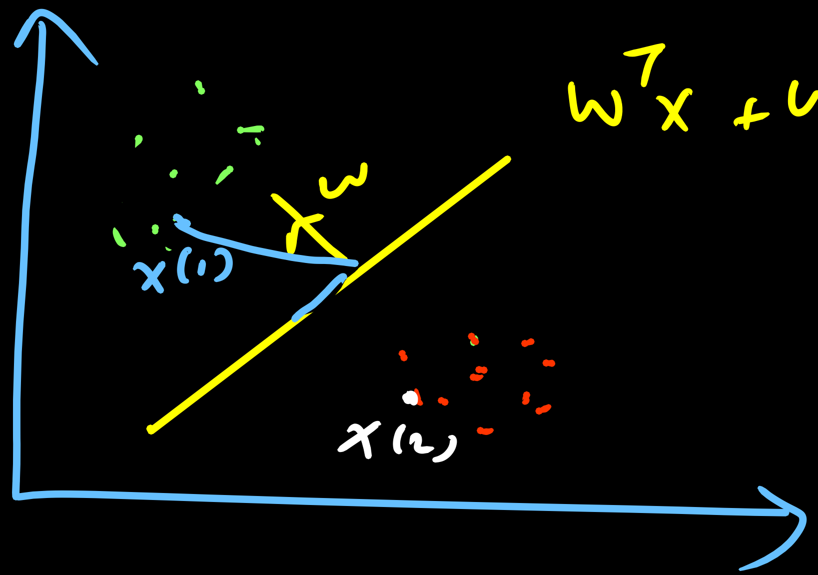
→ line of best fit



$\hat{y} \in R(-\alpha, \alpha)$

Logistic Regression

→ Best hyperplane to seperate the data



$w^T x + w_0 = 0$

$w^T x^{(1)} + w_0 > 0$

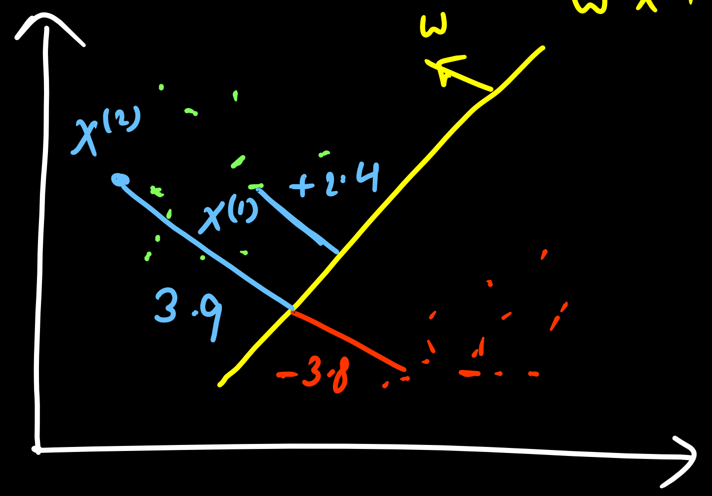
$w^T x^{(2)} + w_0 < 0$

$y \in \{0, 1\}$

$\perp$ dist $= \dfrac{w^T x + w_0}{||w||}$

$\underbrace{x^{(1)} \ \& \ x^{(2)}}$

+ve $=$ 1



$$\frac{(w^T x + w_0)}{\|w\|}$$

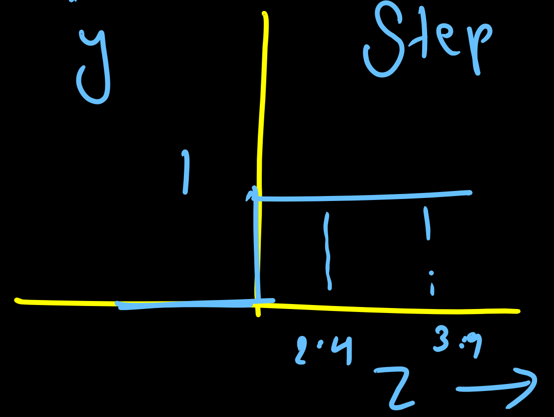
B, R

$$\boxed{w^T x + w_0}$$

$\rightarrow z = (w^T x + w_0) \rightarrow (-\infty, \infty)$

$\rightarrow$ $(1, 0)$

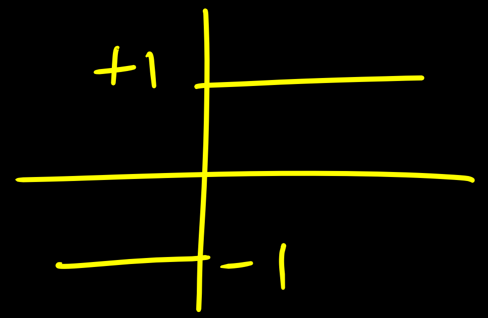
$\hat{y}$ Step



$\operatorname{sgn}((w^T x + w_0))$

$\hookrightarrow w^T x + w_0 > 0 \rightarrow 1$

$< 0 \rightarrow -1$

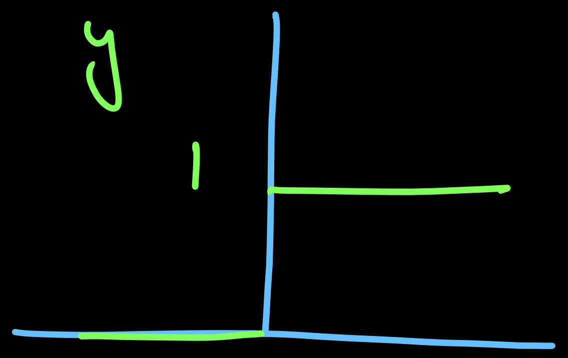


$\hat{y} = 1 \quad z > 0$

$\quad\ \ 0 \quad z < 0$

$x^{(2)} > x^{(1)}$

→ Step



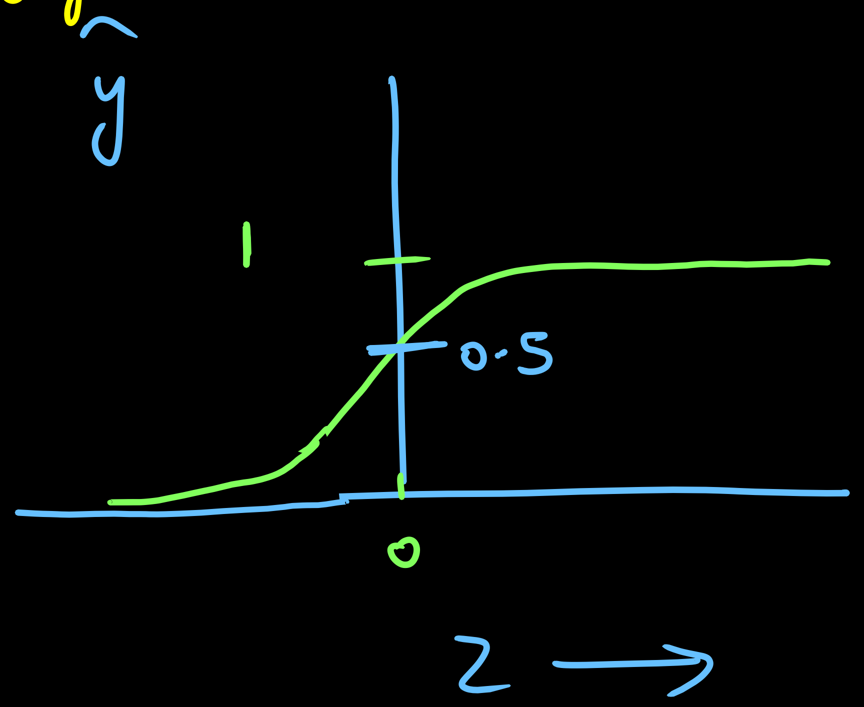
$z = w^T x + w_0$ $\quad z \rightarrow$

$$y = \begin{cases} 1 & z > 0 \\ 0 & z < 0 \end{cases}$$

1. Not diff $z = 0$
2. Even when $x^{(2)} > x^{(1)}$

$$\begin{array}{ccc} & \vert & \vert \\ \hat{y} & 1 & 1 \end{array}$$

→ Sigmoid



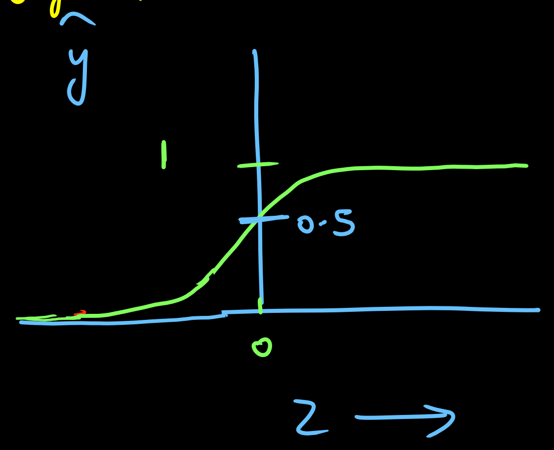
$z \rightarrow \infty \longrightarrow \hat{y} = 1$

$z \rightarrow -\infty \longrightarrow \hat{y} = 0$

$z = 0 \longrightarrow \hat{y} = 0.5$

$$\boxed{\hat{y} = \sigma(z) = \frac{1}{1 + e^{-z}}}$$

→ Sigmoid



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

$z = w^T x + b$

$z = 0 \quad \hat{y} = \frac{1}{1+e^{-0}} = \frac{1}{1+1} = 0.5$

$z = \infty \quad \hat{y} = \frac{1}{1+e^{-\infty}} = \frac{1}{1+0} = 1$

$z = -\infty \quad \hat{y} = \frac{1}{1+e^{\infty}} = \frac{1}{1+\infty} = \frac{1}{\infty} = 0$

$$\boxed{z = w^T x + w_0}$$

$z \to \infty \longrightarrow \hat{y} = 1$

$z \to -\infty \longrightarrow \hat{y} = 0$

$z = 0 \longrightarrow \hat{y} = 0.5$

$\hat{\ } \quad \sigma(\ ) \quad 1$

$\longrightarrow \quad \sigma(z) = 0 - - 1$

= Probabily of class 1

$= P(y=1/x)$

$$\boxed{P(y=1/x) = \sigma(z)}$$

$P(y=0/x) = 1 - P(y=1/x) = 1 - \sigma(z)$

$$x^{(i)} \longrightarrow z^{(i)} \xrightarrow{\text{Step}} \{0, 1\}$$

$$z^{(i)} = (w^T x^{(i)} + w_0)$$

$$[-\alpha, \alpha]$$

Binary Class.

$$x^{(i)} \longrightarrow z^{(i)} \xrightarrow{\sigma(z)} 0 \text{ - - - - } 1 \longrightarrow \{0, 1\}$$

$$z^{(i)} = (w^T x^{(i)} + w_0)$$

$$[-\alpha, \alpha]$$

0.1

0.2

0.99

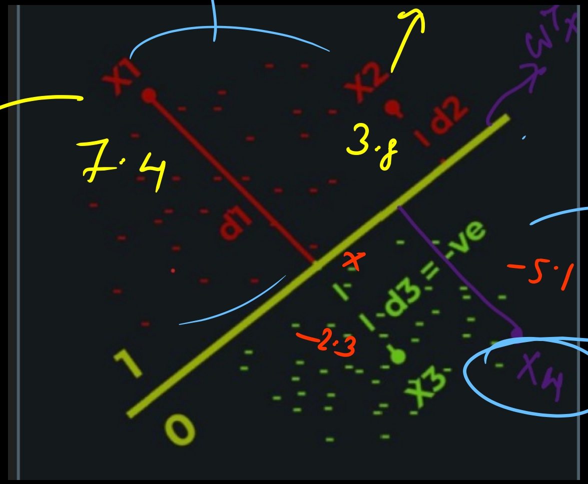
$\sigma(z) \geq 0.5 \longrightarrow 1$

$< 0.5 \longrightarrow 0$

$x^{(i)} \rightarrow z^{(i)} \longrightarrow \sigma(z^{(i)}) \longrightarrow$ Threshold $\rightarrow (1, 0)$

$\| $

$P(y=1/x)$



$\rightarrow 0$

$x^{(1)} \longrightarrow z^{(1)} = 7.4$

$\downarrow \sigma$

$\boxed{0.99}$

$x^{(2)} \longrightarrow z^{(2)} = 3.8$

$\hookrightarrow \boxed{0.76}$

$x^{(3)} \longrightarrow z^{(3)} \xrightarrow{\sigma(z^{(3)})} 0.39$

$z^{(3)} = -2.3$

$x^{(4)} \longrightarrow z^{(4)} \xrightarrow{\sigma(z^{(4)})} 0.08$

$z^{(4)} = -5.1$

$\boxed{P(y=1/x) = \sigma(z)}$

1

0

$\boxed{P(y=0/x) = 1 - P(y=1/x)}$

$P(y=0/x^{(3)}) = 1 - 0.39 = 0.61$

$P(y=0/x^{(4)}) = 1 - 0.08 = 0.92$

$z = \underline{w^T x + w_0}$

→ Break until 22:35

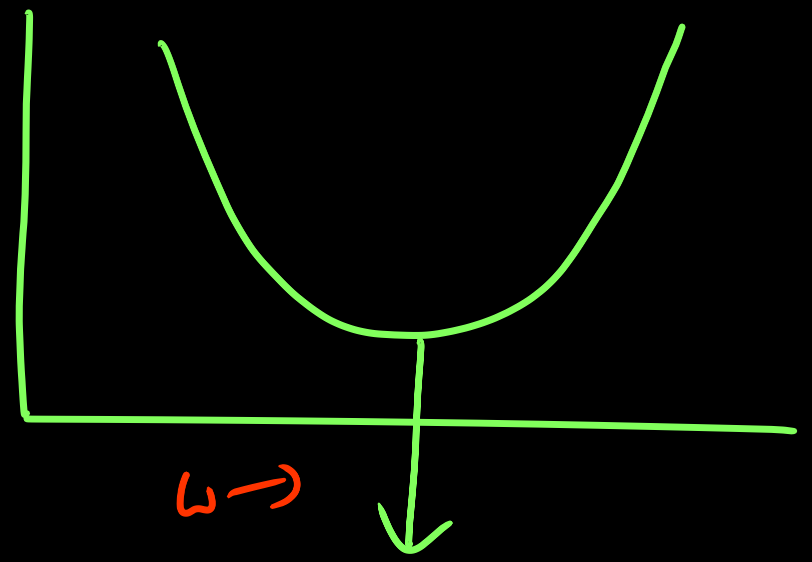
MSE

loss

$$MSE = \frac{1}{m}\sum\left(y^{(i)} - \hat{y}^{(i)}\right)^2$$

Line ← $\hat{y}^{(i)} = w^T x^{(i)} + w_0$

Convex



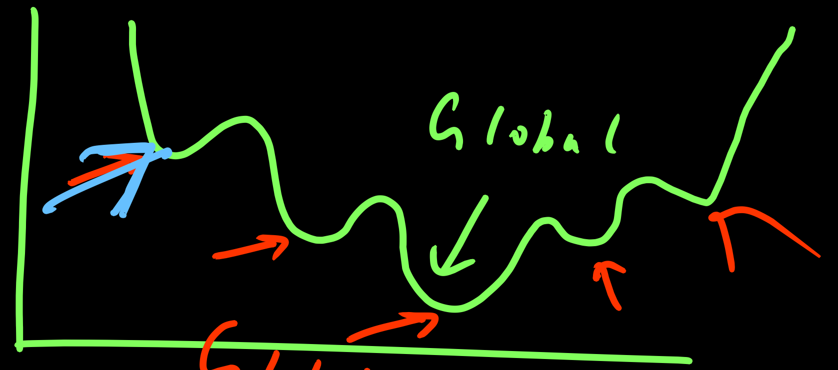
MSE

$$MSE = \frac{1}{m}\sum\left(y^{(i)} - \hat{y}^{(i)}\right)^2$$

$$\hat{y}^{(i)} = \sigma\left(z^{(i)}\right) \rightarrow \text{logistic}$$

$$= \sigma\left(w^T x^{(i)} + w_0\right)$$

Non-Convex



$$W = W - \eta \frac{\partial L}{\partial w}$$

$w^T x + w_0$

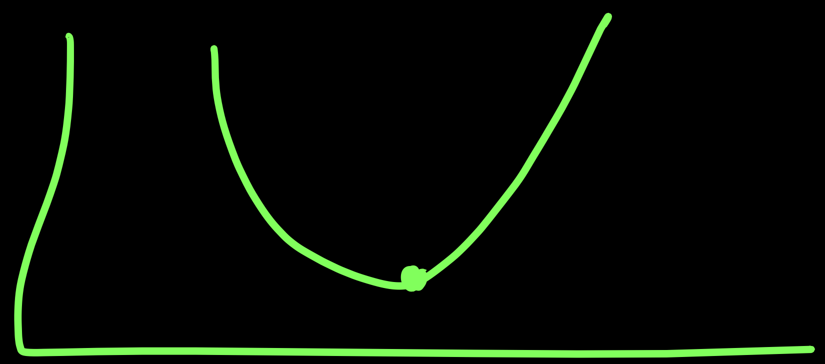
MSE + linear $\longrightarrow$ Convex

Sigmoid

MSE + logis $\longrightarrow$ Non-Convex

Log-loss + logistic $\longrightarrow$ Convex

$\sigma(w^T x + w_0)$



→ log_loss (loss for Logistic Reg.)

$$\text{log\_loss} = \begin{cases} -\log \hat{y}^{(i)} & y^{(i)} = 1 \\ -\log(1 - \hat{y}^{(i)}) & y^{(i)} = 0 \end{cases}$$

Probab $= \sigma(z^{(i)})$

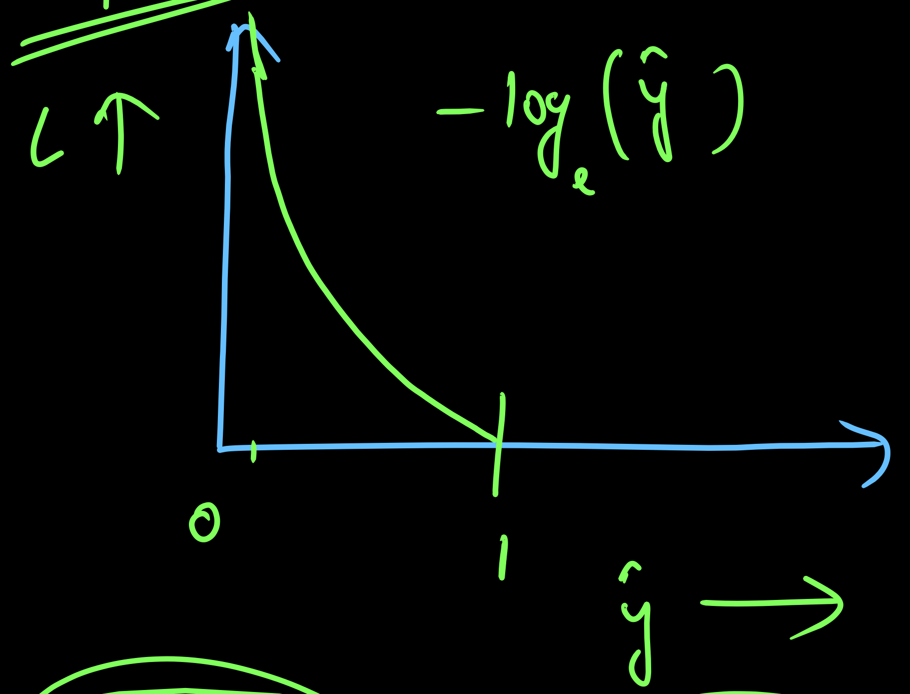
target

$$\hat{y}^{(i)} = \sigma(z^{(i)})$$

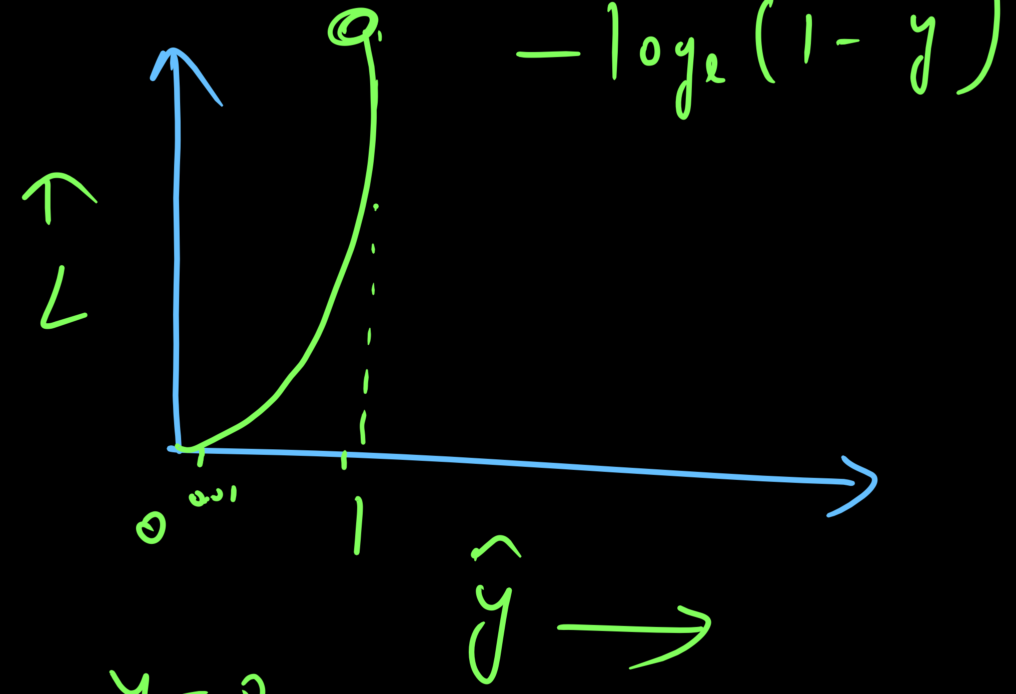
$(0 - 1)$

$$z^{(i)} = w^T x^{(i)} + w_0$$

$y = 1$



$y = 0$



$y = 1$
$\hat{y} = 0.999$

$y = 1$
$\hat{y} = 0.001$

$y = 0$
$\hat{y} = 0.01$

$-\log(1 - 0.01)$

$-\log_e(0.99)$

≈ 0

$y = 0$
$\hat{y} = 0.99$

$$L = -\left[y^{(i)} \log(\hat{y}^{(i)}) + (1 - y^{(i)}) \log(1 - \hat{y}^{(i)})\right]$$

$$\text{log\_loss}(L) = \begin{cases} -\log \hat{y}^{(i)} \xrightarrow{} \text{Probability} & y^{(i)} \xrightarrow{} \text{target} = 1 \\ -\log(1-\hat{y}^{(i)}) & y^{(i)} = 0 \end{cases}$$

$$L = \frac{-1}{m} \sum_{i=1}^{m} \left[\underbrace{y^{(i)} \log(\hat{y}^{(i)})}_{A} + \underbrace{(1-y^{(i)}) \log(1-\hat{y}^{(i)})}_{B} \right]$$

$y^{(i)} = 0$
1

$y^{(i)} = 1$

$L = -\log \hat{y}^{(i)}$

$y^{(i)} = 0$

$L = -\log(1-\hat{y}^{(i)})$

→ Optimization

→ Randomly initialize W, w_0

→ Repeat until convergence {

MSE

$$\frac{1}{2m} \sum_{i=1}^{m} \left(y^{(i)} - \hat{y}^{(i)}\right)^2$$

$$W = W - \eta \frac{\partial L}{\partial W}$$

}

Line → MSE

$f(w^T x + w_0)$

Log → log_loss

f

$\sigma(w^T x + w_0)$

$$\frac{\partial L}{\partial w_d} = \frac{1}{m} \sum_{i=1}^{m} \left(y^{(i)} - \hat{y}^{(i)}\right) x_d^{(i)}$$

$$\frac{\partial L}{\partial w_0} = \frac{1}{m} \sum_{i=1}^{m} \left(y^{(i)} - \hat{y}^{(i)}\right)$$

$w^T x + w_0$

$w_1 x_1 + w_2 x_2 - - - w_d x_d + w_0$

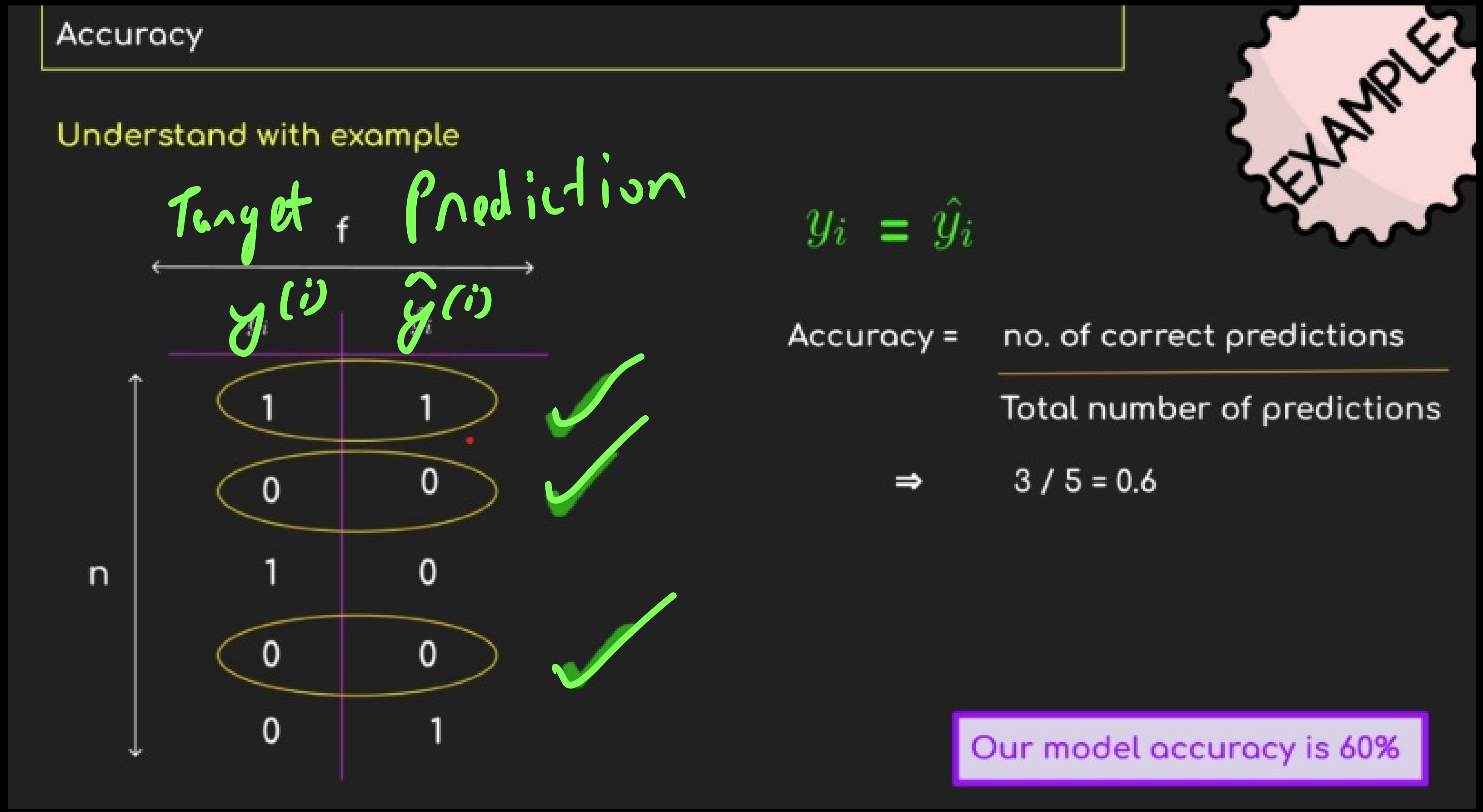
w_1 w_2

```
[23] model.coef_

array([[0.68445262, 0.29104301, 0.1363756 , 0.79630985, 0.06125924]])

model.intercept_

array([-0.01220319])
```

w_0

$R^2 \to \quad y^{(i)} \quad \hat{y}^{(i)} \quad (-\infty, \infty)$

Accuracy



$$\frac{3}{5} = 0.6$$

$$= 60\%$$

VIF_{thres} 5

$R_{thres} = 75.8$

> 5 < 10

85

> 10

MSE & logistic

1. Non convex

$y \rightarrow 1 \quad 1$

$\hat{y} \rightarrow 0.01 \quad 0.99 \quad 1$

2.

$y^{(i)} = 1$

$\hat{y}^{(i)} = 0.01$

$\left(y^{(i)} - \hat{y}^{(i)}\right)^2 = (1-0.01)^2$

$= 0.99^2 = \boxed{0.9801}$

$\text{log-loss} = -\left[y^{(i)} \log \hat{y}^{(i)} + (1-y^{(i)}) \log(1-\hat{y}^{(i)})\right]$

$= -1 \cdot \log(0.01) + 0$

$= 4.605$

For incorrect prediction, log-loss penalizes more than MSE