Logistic Regression

Linear Regression (Reg. Model)

$$\hat{y} = W^T X$$

↳ Cont. value $(-\infty, +\infty)$

Car price $\log(y) = W^T X$

Logistic Regression (clf Model)

$$\hat{y} = \phi(W^T X)$$

ϕ ↳ limits $(0,1)$

$\hat{y}$ ↳ class

Binary 0 1 → Hard clf

Multi 0, 1, 2, 3, . . .

Soft clf

$$\boxed{\hat{y} = P(\text{class } 1)}$$

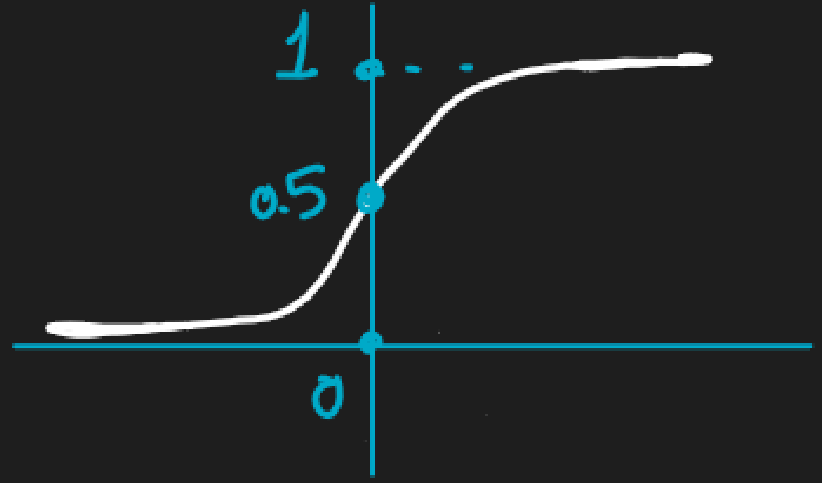
Cont. value $(0, 1)$

Examples

p(churn) = 0.8

p(spam) = 0.2

p(click) = 0.6

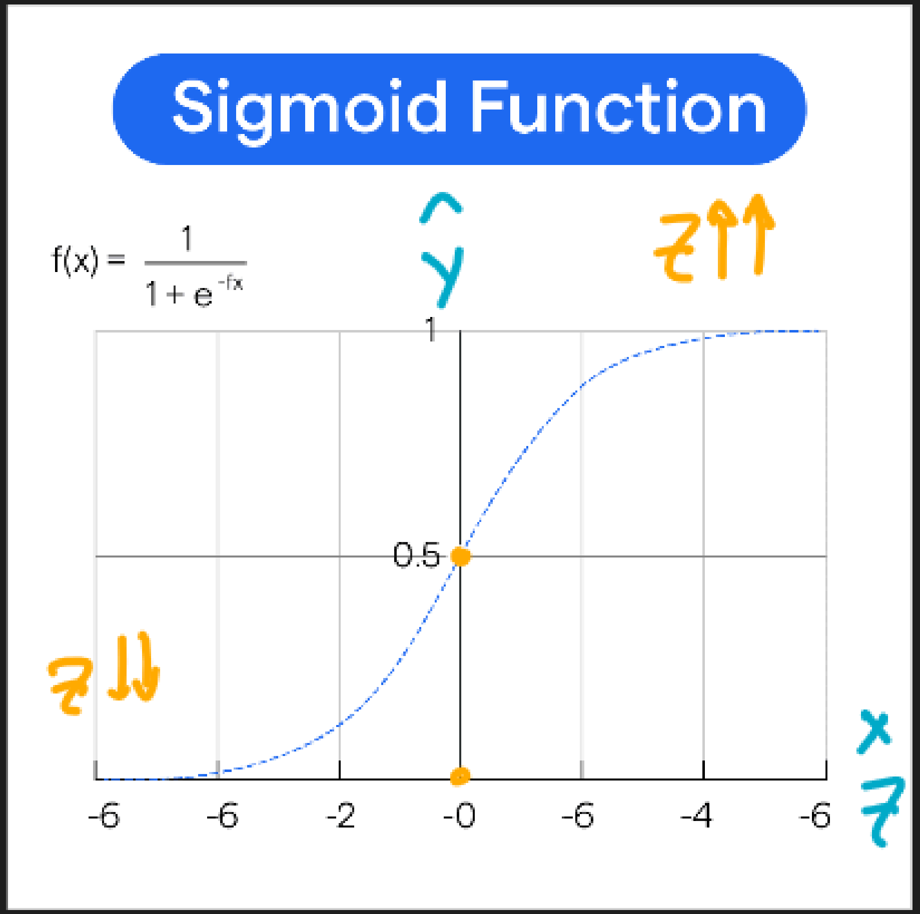


Soft clf → Hard clf

$$p(churn) = 0.8 \xrightarrow[0.5]{threshold} \text{churn (class 1)}$$

$$p(spam) = 0.2 \xrightarrow[0.4]{threshold} \text{Not spam (class 0)}$$

$$p(click) = 0.6 \xrightarrow[0.7]{threshold} \text{Not click (class 0)}$$



$$\underbrace{\hat{y}}_{p(cls1)} = \sigma(\underbrace{w^T x}_{z = w^T x})$$

Logistic Regression

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

Limits
$-\infty$
$\downarrow$
$+\infty$

$$z \underset{-\infty}{\downarrow\downarrow} \longrightarrow \sigma(z) \underset{0}{\downarrow\downarrow}$$

$$z \overset{+\infty}{\uparrow\uparrow} \longrightarrow \sigma(z) \overset{1}{\uparrow\uparrow}$$

$$z = 0 \longrightarrow \sigma(z) = 0.5$$

Limits
0
$\downarrow$
1

(prob)

$$\boxed{w_1 x_1 + w_2 x_2 = 0}$$

Decision Boundary

$$w^T x = 0 \longrightarrow \sigma(w^T x) = 0.5$$

$$w^T x > 0 \longrightarrow \sigma(w^T x) = 1$$

$$w^T x < 0 \longrightarrow \sigma(w^T x) = 0$$

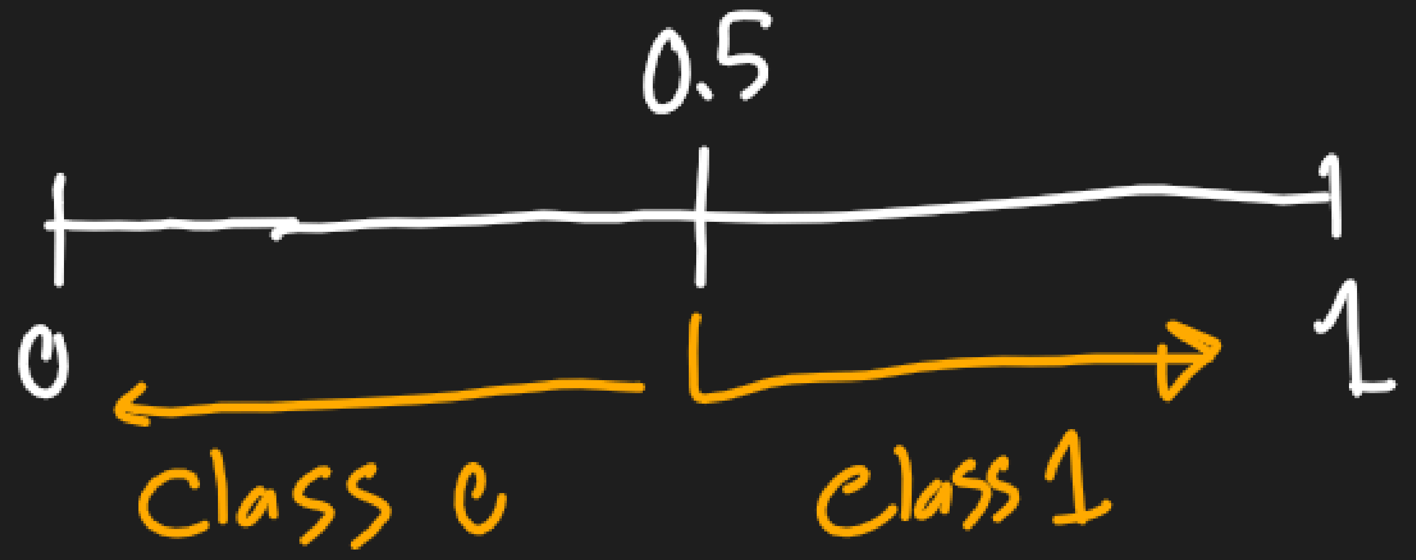
Logistic Regression

$$y = \sigma(w^T x)$$

$$y = \sigma(w_0 + w_1 x_1 + w_2 x_2 + \cdots + w_n x_n)$$

$$y = w^T x \geq 0$$

$$y = w_0 + w_1 x_1 + w_2 x_2 + \cdots + w_n x_n \geq 0$$



Logistic Regression Training → Get Weights

$$\underset{✓}{\underline{y}} = \sigma(\underset{\downarrow\ ??}{w}^T \underset{✓}{\underline{x}})$$

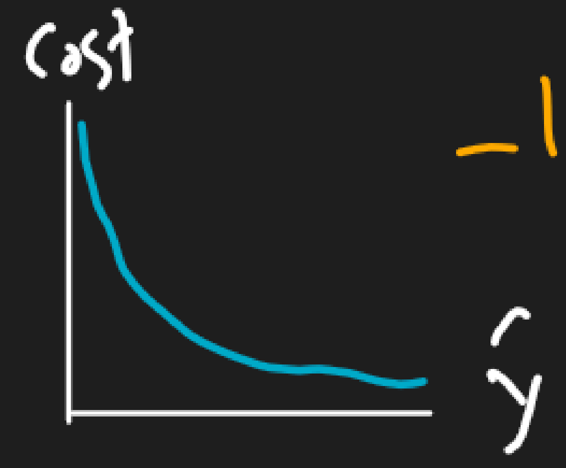
Binary clf

☑ Define Cost Function (Error)
↓
Misclassification

☑ optimization
↳ Get weights for Min Error

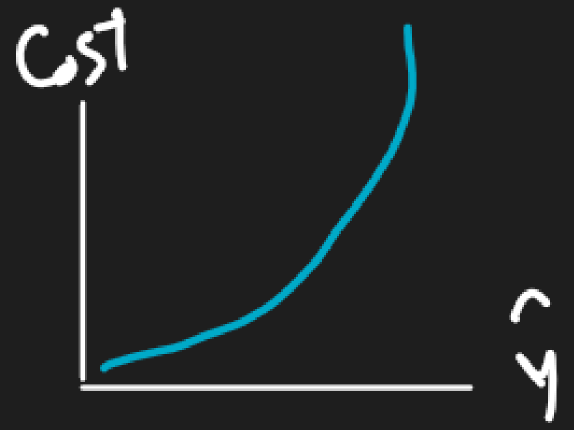
	(y) Actual	(ŷ) pred	Cost (Error)
y=0	0	0	↓
	0	1	↑
y=1	1	0	↑
	1	1	↓

$\boxed{y=1}$



$-\log(\hat{y})$

$\boxed{y=0}$



$-\log(1-\hat{y})$

$$\text{Cost } \underline{fn} = \begin{cases} -\log(\hat{y}) & y = 1 \\ -\log(1-\hat{y}) & y = 0 \end{cases}$$

J

Binary

Cross Entropy

log_loss function

$$\boxed{J = -y\log(\hat{y}) - (1-y)\log(1-\hat{y})}$$

$$J(w) = -\frac{1}{m}\sum_{i=1}^{m} y^{(i)}\log(\alpha(w^T x^{(i)})) + (1-y^{(i)})\log(1-\alpha(w^T x^{(i)}))$$

→ Get Weights for Min $\underline{\underline{J}}$

Linear Reg	Log. Reg.
$J = f_1(w)$	$J = f_2(w)$
MSE	Cross Entropy

→ **Gradient Descent** (optimization) → Get Best Weights

1- initialize Random weights

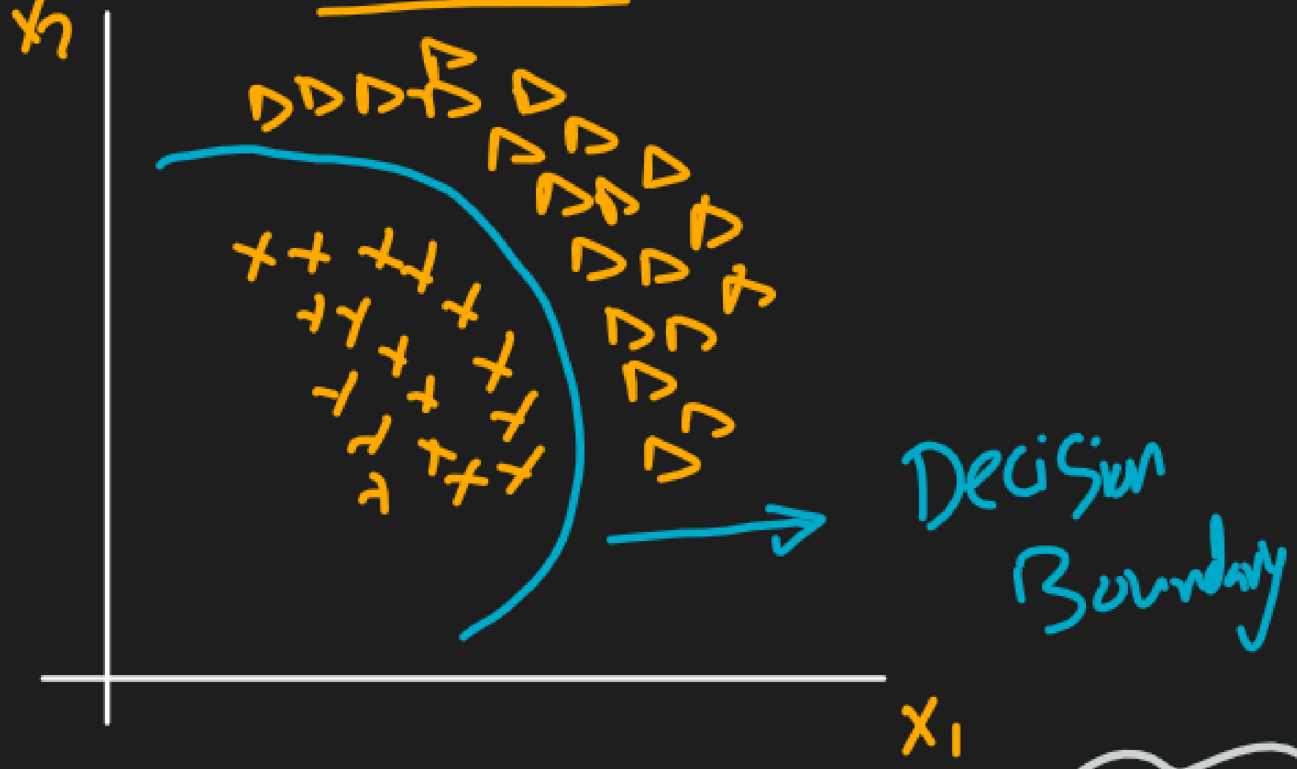
2- calculate gradients $\partial J / \partial w$

3- update weights $w^{new} = w^{old} - \eta \frac{\partial J}{\partial w}$

iter

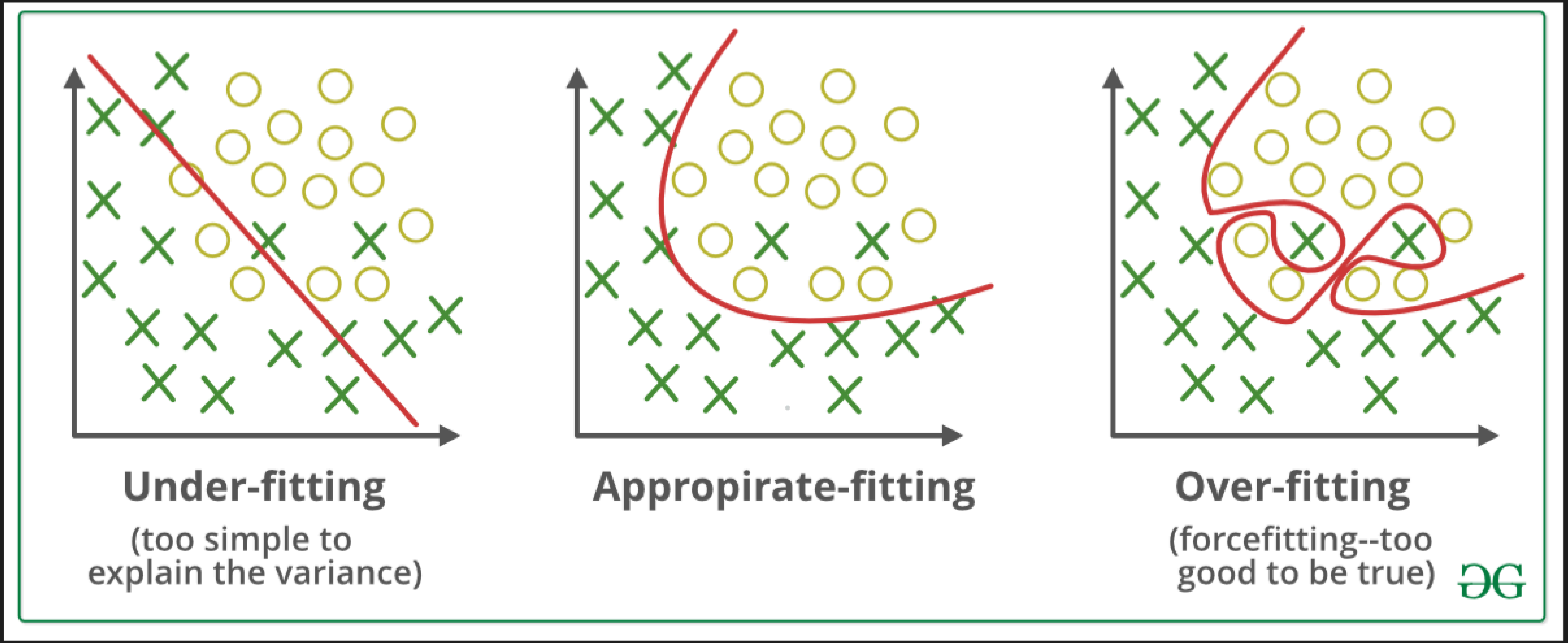
till Convergence $w^{new} \approx w^{old}$ or Max. iters

* Nonlinear



$$\hat{y} = \sigma\left(w_0 + w_1x_1 + w_2x_2 + \underbrace{w_3x_1^2}_{\text{polynomial}} + w_4x_2^2 + w_5x_1x_2\right)$$

Polynomial features + Logistic Regression



Train Error	↑	↗	↓
Test Error	↑	↗	↑

* To reduce overfitting: Apply penalty (L_1, l_2, Elastic)

→ Add Regularization term to Cost fn

→ Apply optimization

Control Regularization

Linear Regression

$$\text{alpha} \propto \text{penalty}$$

Logistic Reg

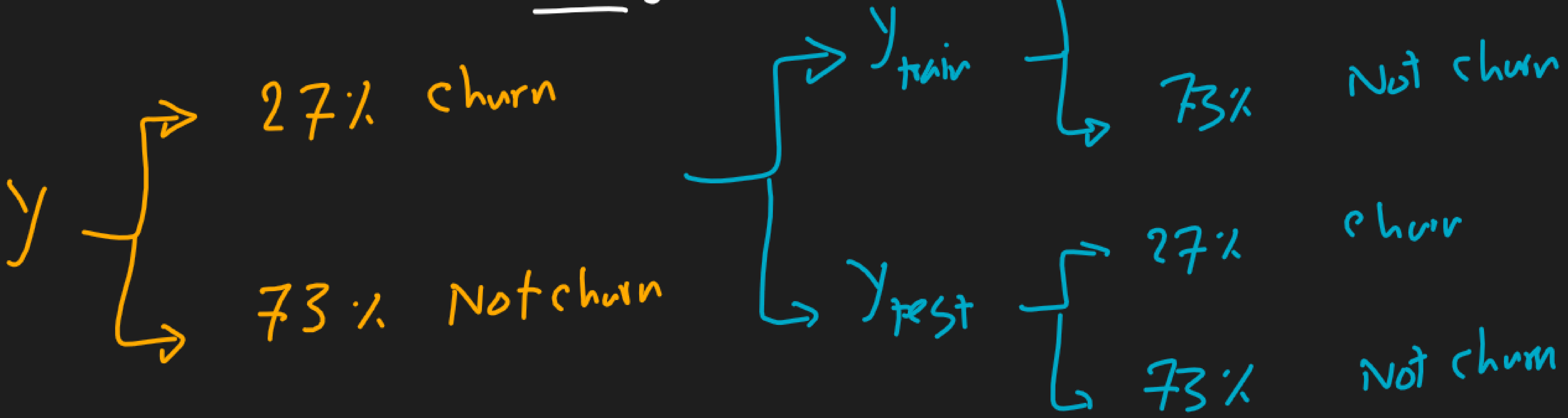
$$C \propto \frac{1}{\text{penalty}}$$

check

penalty compliance with Solver

Sklearn Docs

Stratify


y
27% churn
73% Notchurn
y_{train}
27% churn
73% Not churn
y_{test}
27% churn
73% Not churn