CS 5525 - March 27, 2023

- given $X \xrightarrow{\text{predict}} Y$ output
  (input)

$\boxed{\text{generative model}}$

Ex: Naive bayes
GANs
Bayesian Network

$P(x, y)$ let y : red, blue

ball

$P(x, \text{red})$ → $P(x, \text{red}) > P(x, \text{blue})$ → red

$P(x, \text{blue})$ $P(x, \text{red}) < P(x, \text{blue})$ → blue

$$\underbrace{P(x,y)} = \underbrace{P(y)}_{\text{prior probability}} \cdot \underbrace{P(x \mid y)}_{\text{Likelihood}}$$

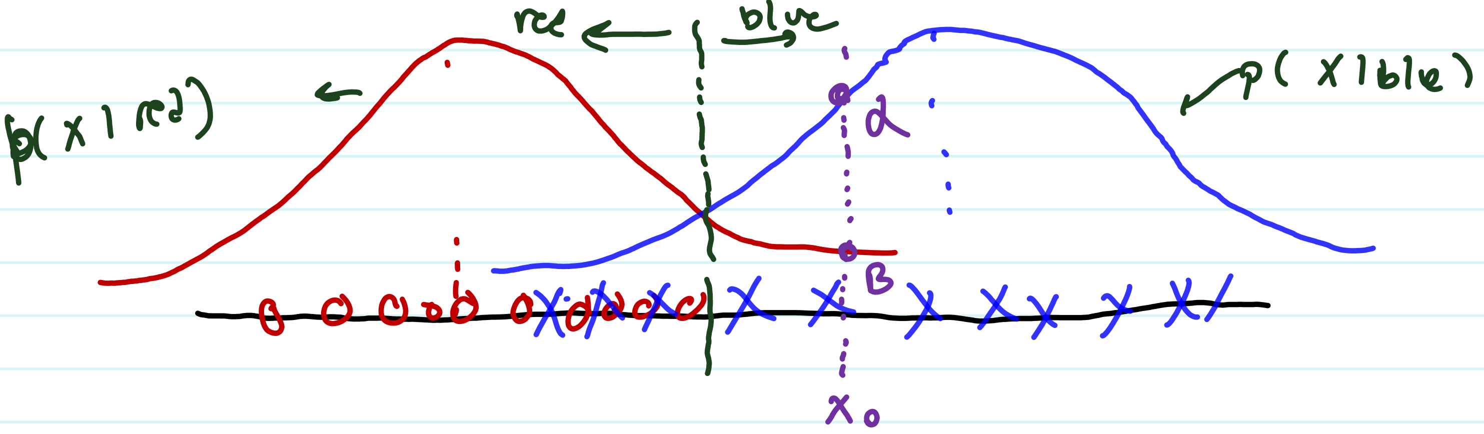
$$\underbrace{P(x \mid y)}_{\text{posterior}} = \frac{\overbrace{P(x, y)}}{P(y)}$$

$P(x, \text{red}) = P(\text{red}) \cdot P(x \mid \text{red})$

$\frac{1}{2} \cdot B$

threshold

$P(\text{red}) = P(\text{blue}) = 1/2$



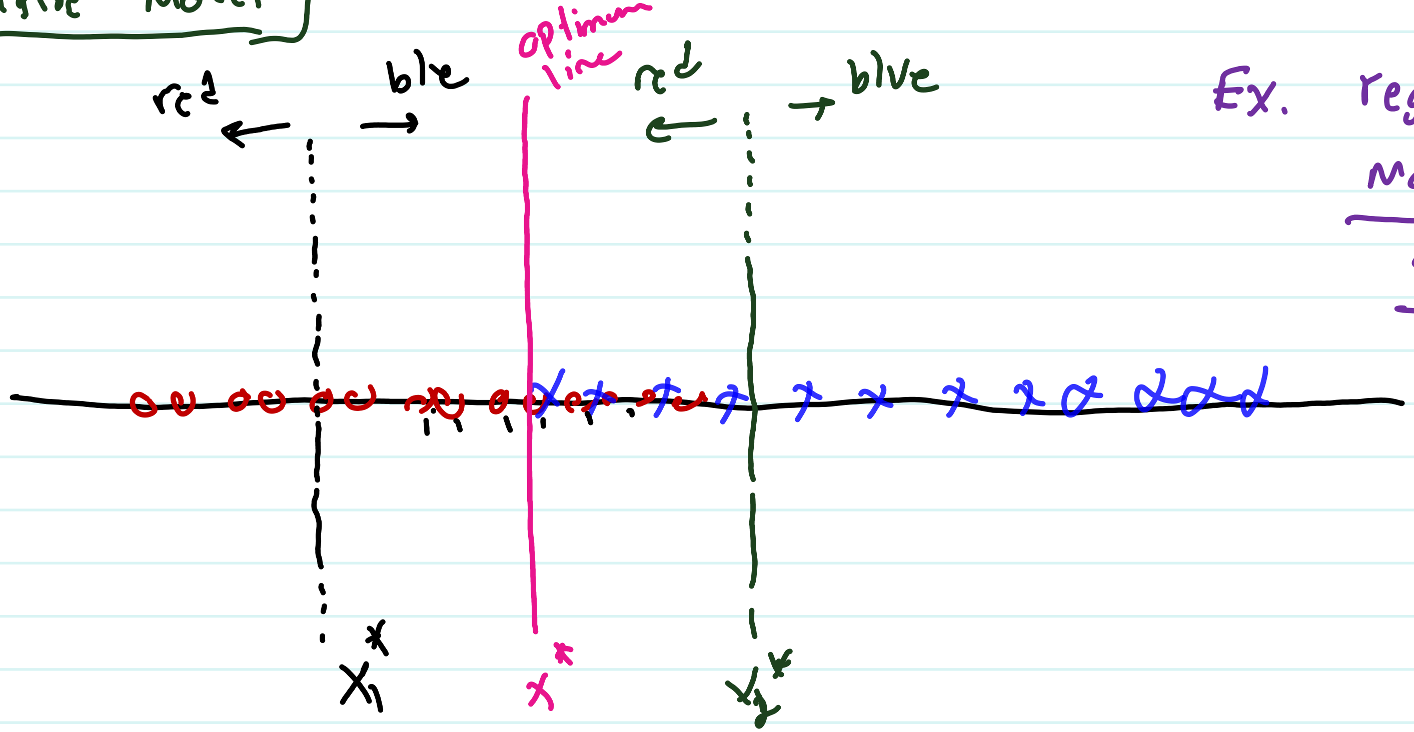
$p(x, \text{blue}) = \frac{1}{2} \cdot \alpha$

$\alpha > \beta \rightarrow x_0 \xrightarrow{\text{as}} \underbrace{p(x, \text{blue}) > p(x, \text{red})}$

$x_0 \rightarrow \text{blue}$

$\boxed{\text{discriminative Model}}$ → minimize the error within all possible thresholds

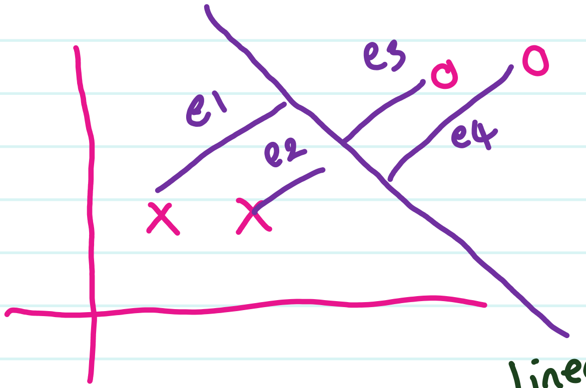
p(y | x)



Ex. regression model

SVM

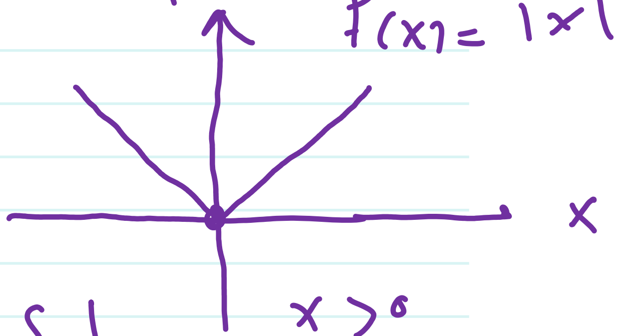
NN



$$SSE = e^T \cdot e$$

$$\sum |e_i|$$

$$f(x) = |x|$$



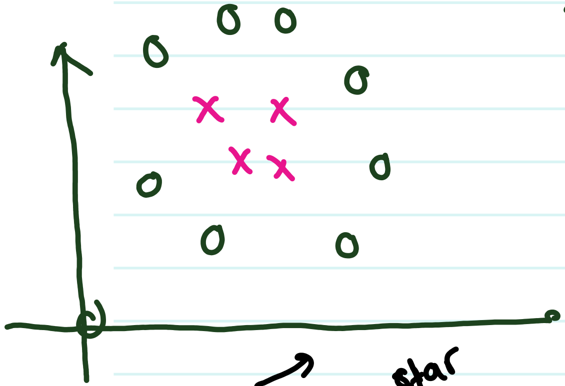
nonlinearly
seprable

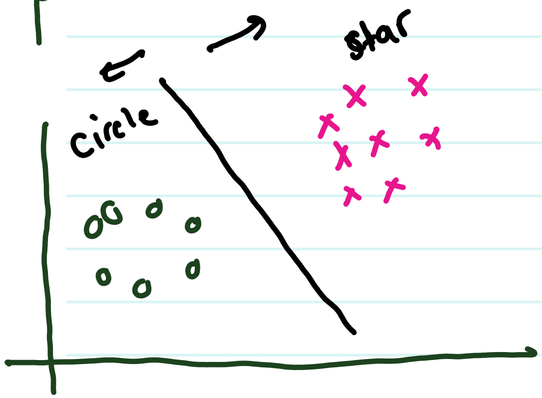
linear regression

$$z = w^T x + b$$

$$= w_1 x_1 + w_2 x_2 + \cdots + b$$

$$f'(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases}$$

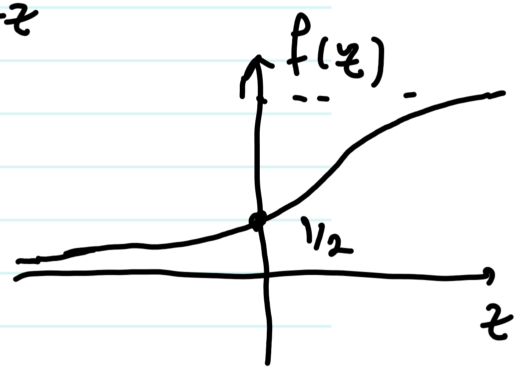




$$y = f(z) = \frac{1}{1+e^{-z}}$$

$$\lim_{z \to -\infty} f(z) = 0$$

$$\lim_{z \to \infty} f(z) = 1$$



$$f(0) = \frac{1}{1+e^0} = \frac{1}{1+1} = \frac{1}{2}$$

$$\ln\left(\frac{p}{q}\right) = z$$

$$\ln\left(\frac{p}{1-p}\right) = z \rightarrow \frac{p}{1-p} = e^{z} \rightarrow p = e^{z} - e^{z}p$$

$$p(1+e^{z}) = e^{z} \rightarrow p = \frac{e^{z}}{1+e^{z}} = \boxed{\frac{1}{1+e^{-z}}}$$

$$\boxed{p(z) = \frac{1}{1+e^{-z}}} = \frac{1}{1+e^{-(w_1x_1+w_2x_2+\cdots+b)}} \left\{ \begin{array}{l} 0 \\ 1 \end{array} \right.$$

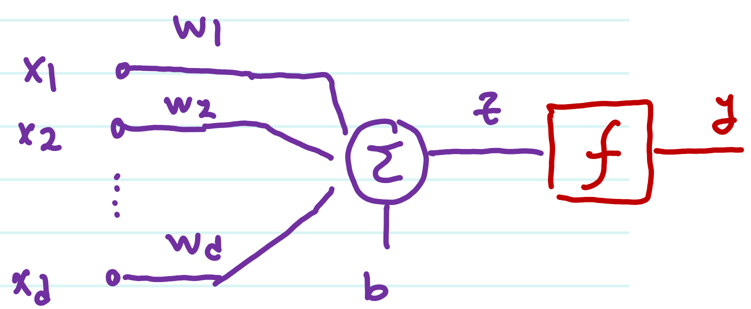
$$z = \beta_0 + \beta_1 x_1 + \cdots + \beta_d x_d$$

$$= \underbrace{\beta_0}_{b} + \underbrace{[\beta_1 \quad \beta_2 \quad \cdots \quad \beta_d]} \overbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}}^{X}$$

$$= b + w^{T} \cdot x$$

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

regression



Example 8

	Variable	Coefficient
c	constant	$C = -26.52$
x	Age	$\beta_1 = 0.78$

Model to subscribing magazine

$$z = c + \beta_1 x$$

so for 35 years old $z = -26.52 + 0.78 \times 35$

$$= 0.78$$

$$P = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-.78}} = \boxed{68\%}$$ this is probability

What about for 36 years old?

$$z = -26.52 + 0.78 \times 36 \rightarrow$$

$$= 1.56$$

$$P(\text{subs} \mid 36) = \frac{1}{1+e^{-1.56}} = \boxed{82\%}$$

So 35 68% → one year increment → 14% +

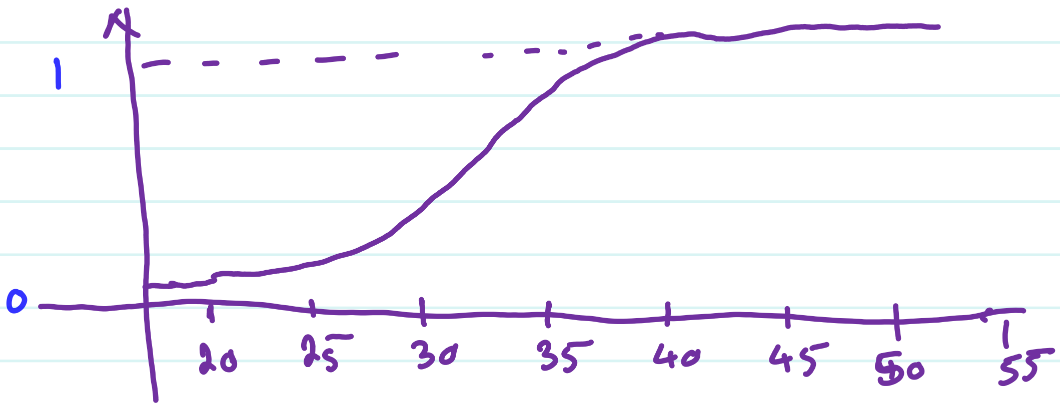
36 82%

For 25 years old $z = -7$

$p = 0.009$ or $\sim 1\%$

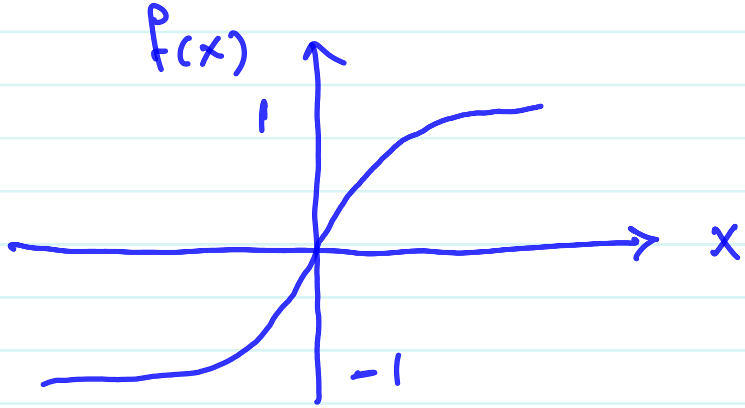
26 years old $z = -6.22$

$p = .002$ or $\sim 1\%$



$f(x) = \tanh(x)$

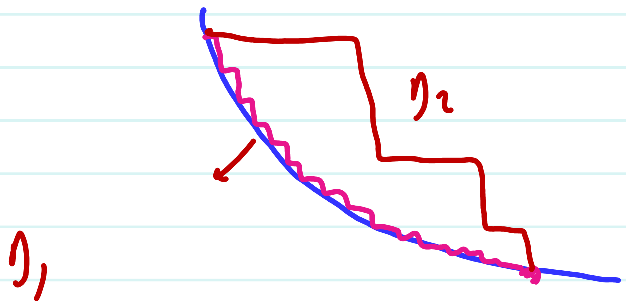
$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$



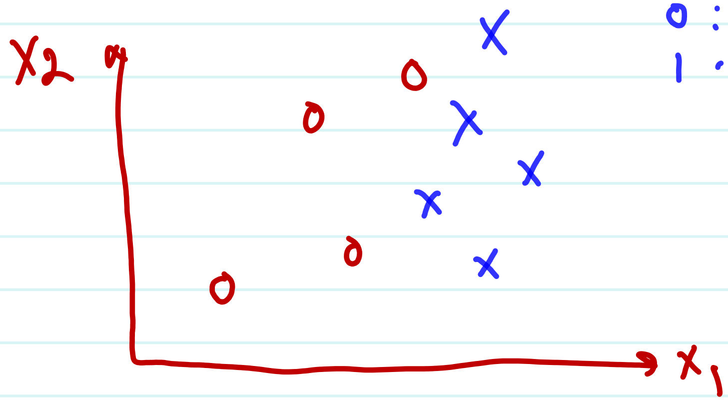
$\prod$

$\log(a \times b \times \dots) = \log(a) + \log(b) + \dots$

$\uparrow$

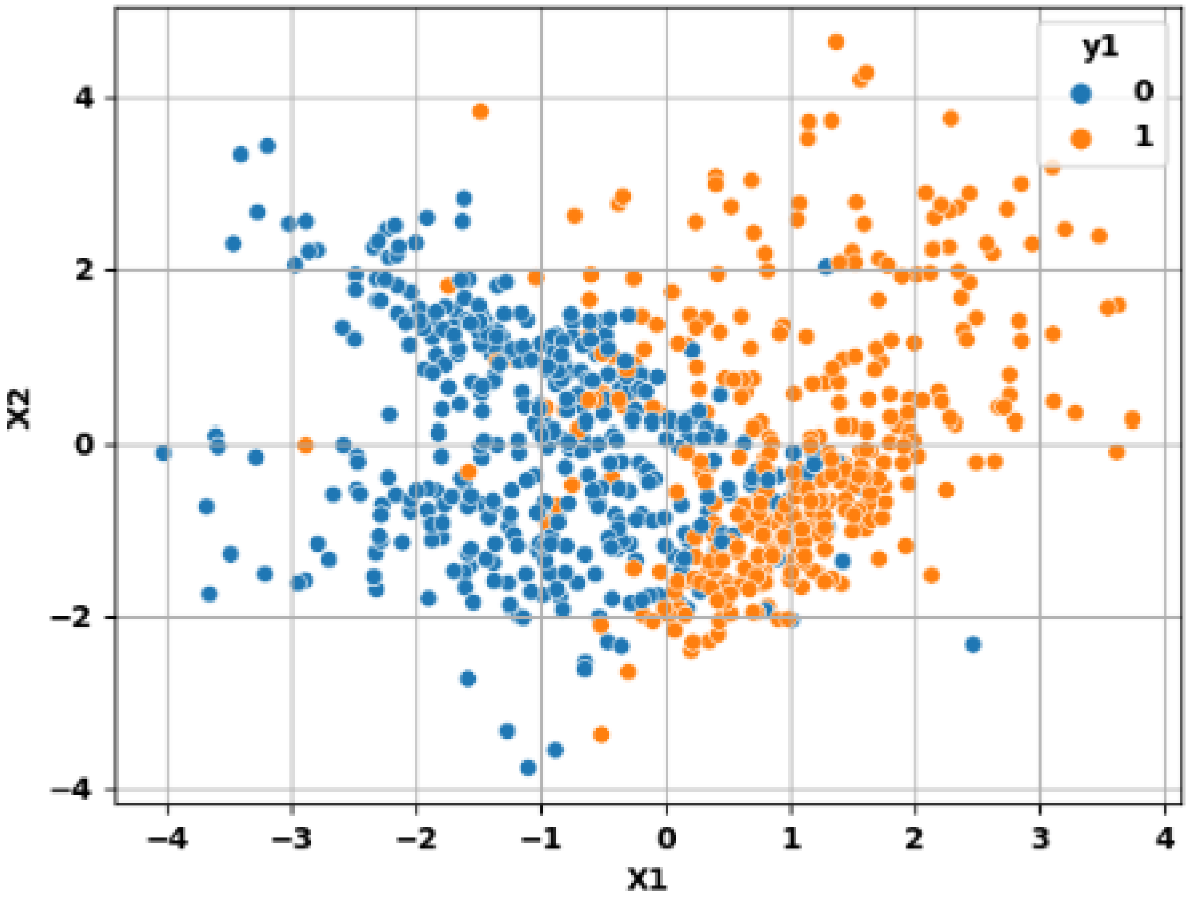


$\eta_1 < \eta_2$



0 : o

1 : x


y1
0
1
X2
X1
4
2
0
−2
−4
−4
−3
−2
−1
0
1
2
3
4

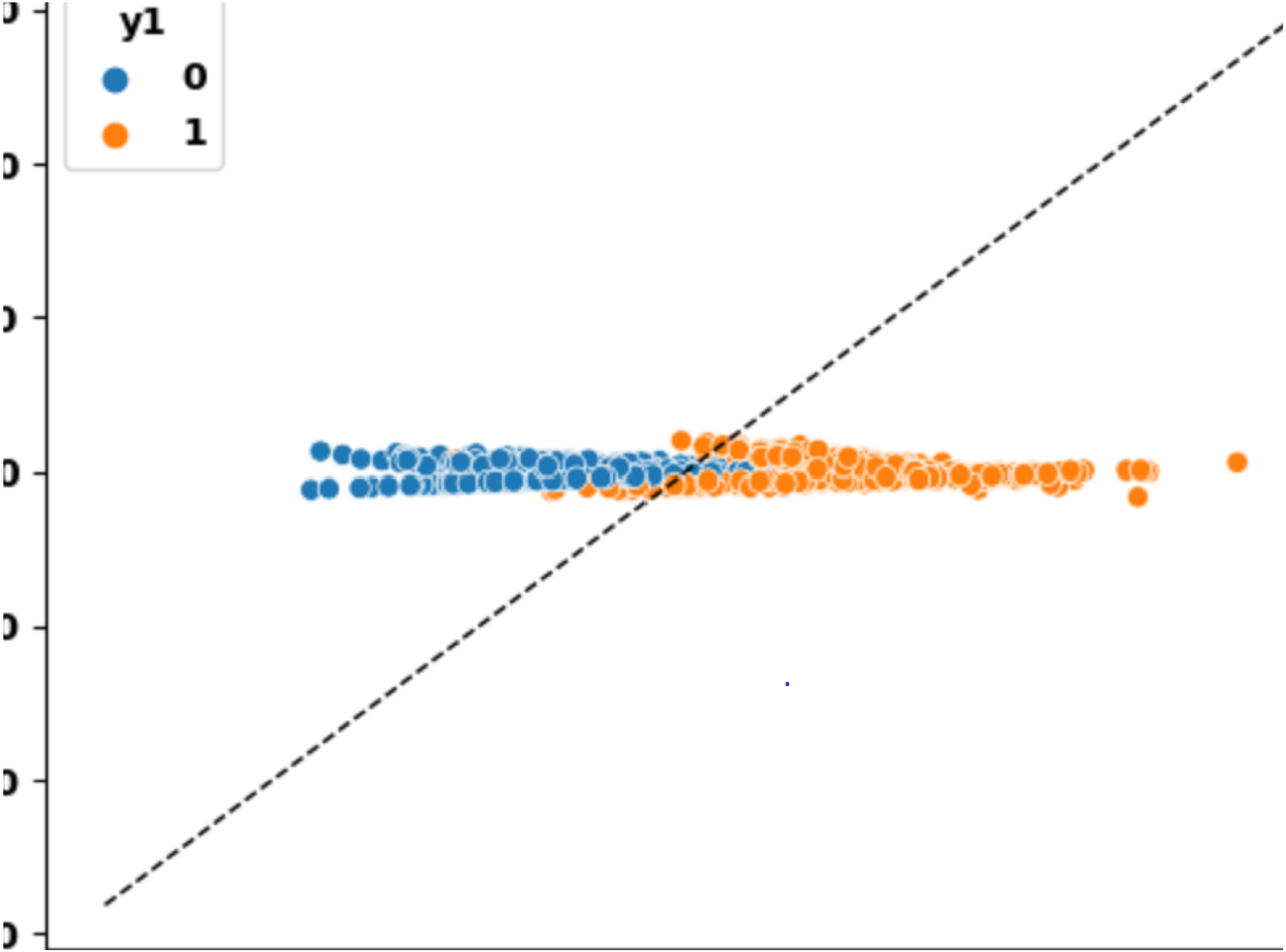
$$z = b + w_1 x_1 + w_2 x_2$$

$$w_1 x_1 + w_2 x_2 + b = 0 \rightarrow w_2 x_2 = -b - w_1 x_1$$

$$x_2 = \frac{-b}{w_2} - \frac{w_1}{w_2} x_1$$

slope $= \frac{-w_1}{w_2}$

bias $= \frac{-b}{w_2}$


y1
0
1