Error Function

Dennis A. Christie

Machine Learning Course

22/02/2019

Mengenal Error Function

Apa dan Bagaimana?

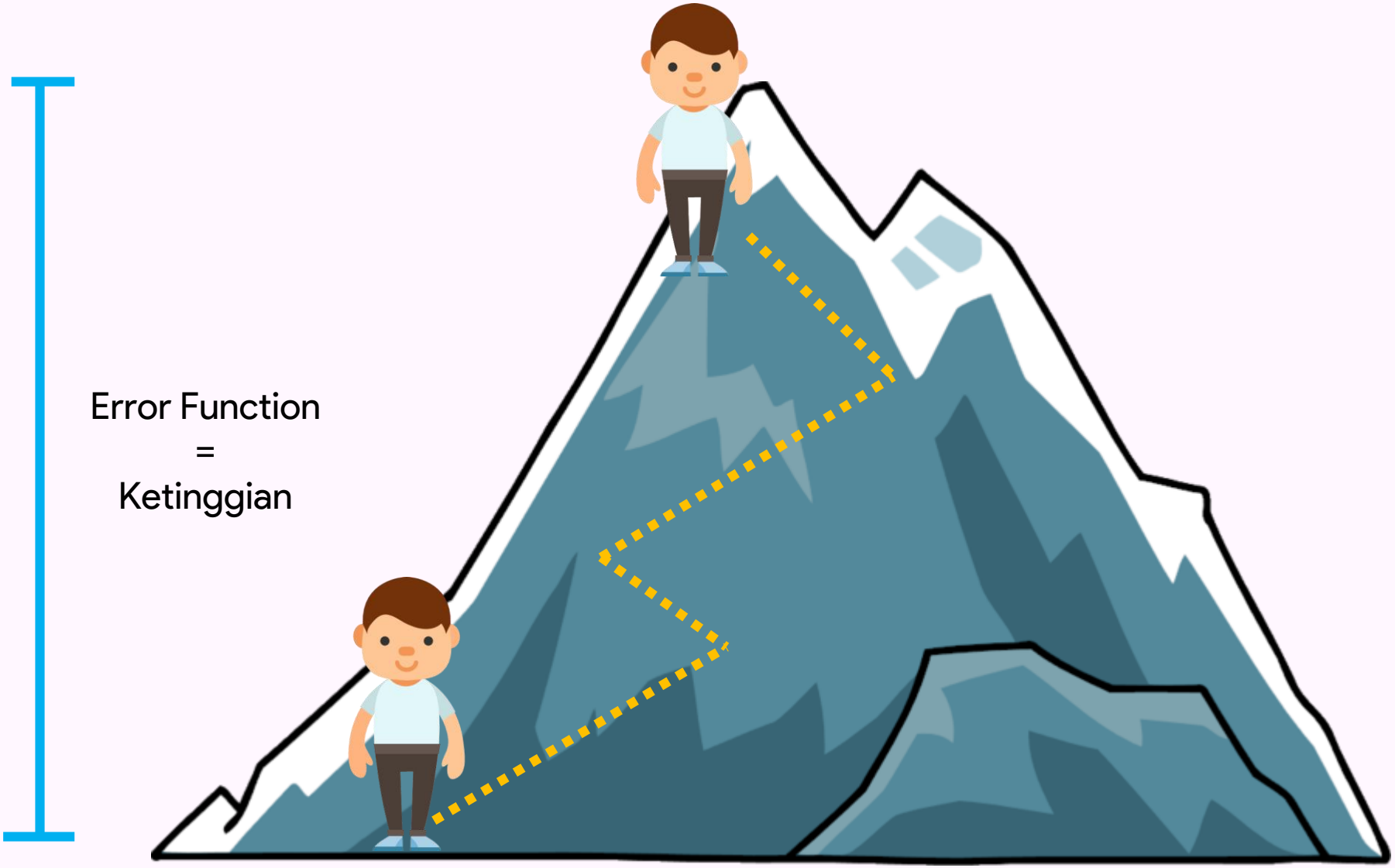
Apa itu Error Function

- Function yang menunjukkan seberapa buruk performa kita.

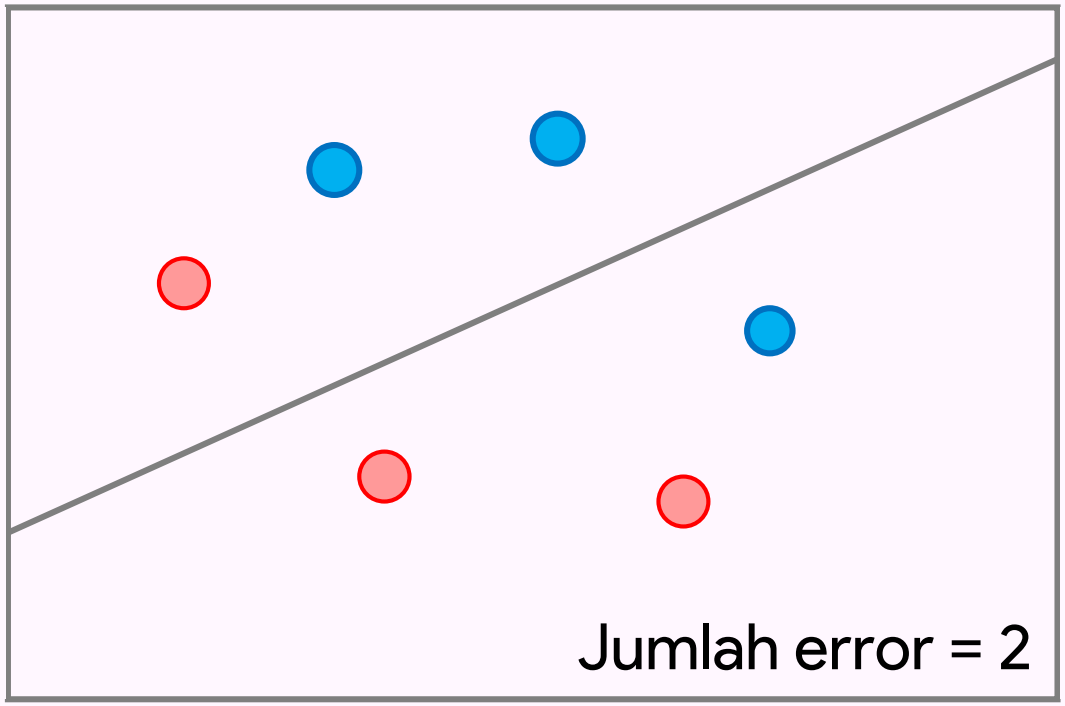
- Contoh:
  - Saya mau pergi ke meja terdekat
  - Error function saya = Jarak posisi saya ke posisi meja.

- Setelah mengetahui Error Function:
  - Setelah mengetahui bahwa posisi saya jauh, saya melihat sekitar untuk mengetahui arah mana yang terbaik menuju posisi meja.
  - Saya ambil langkah untuk bergerak mendekati meja.

Illustrasi Error Function

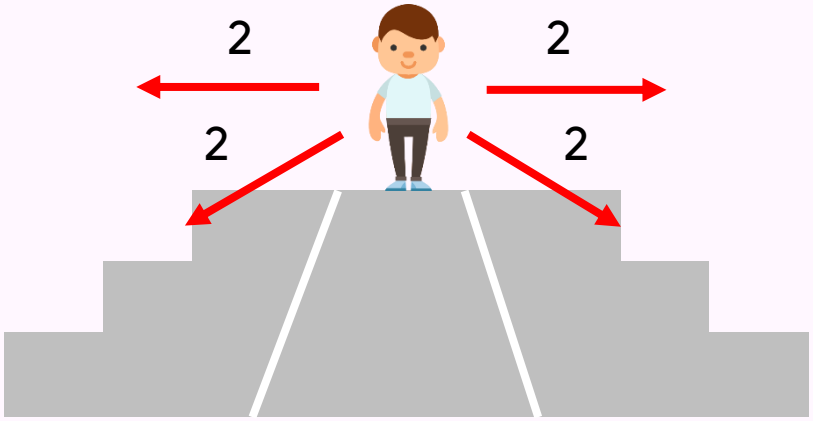


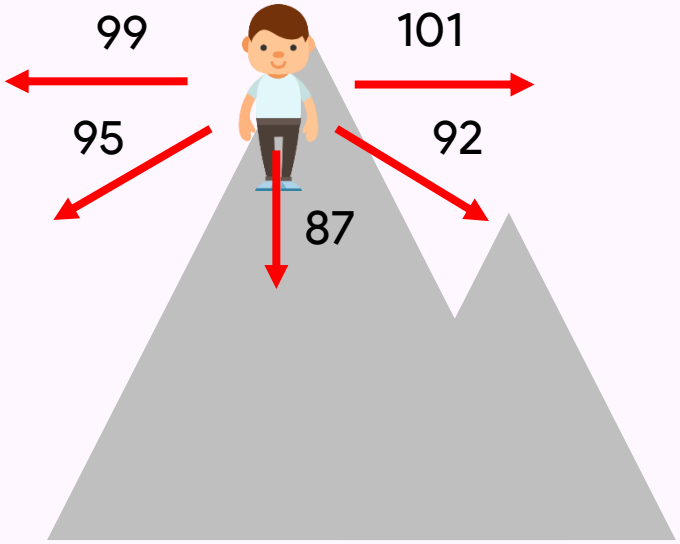
Permasalahan Error Function



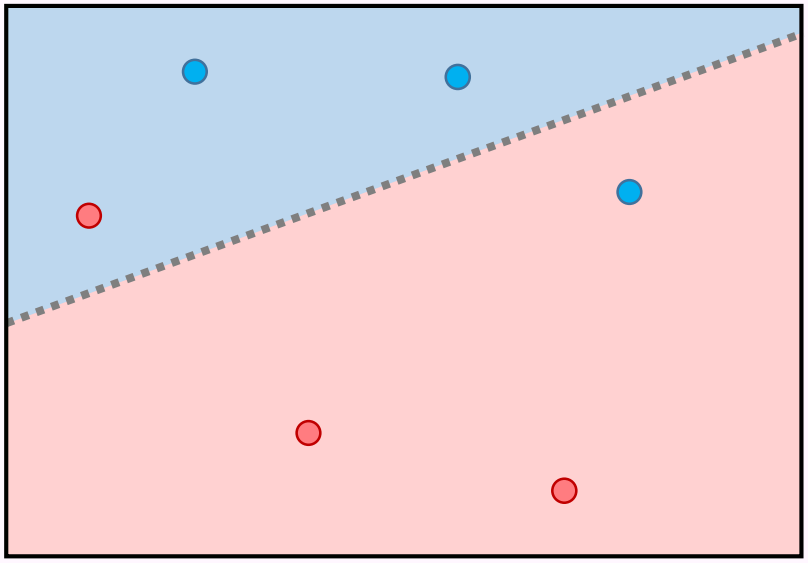
Kudu piye...

Diskrit vs Kontinyu



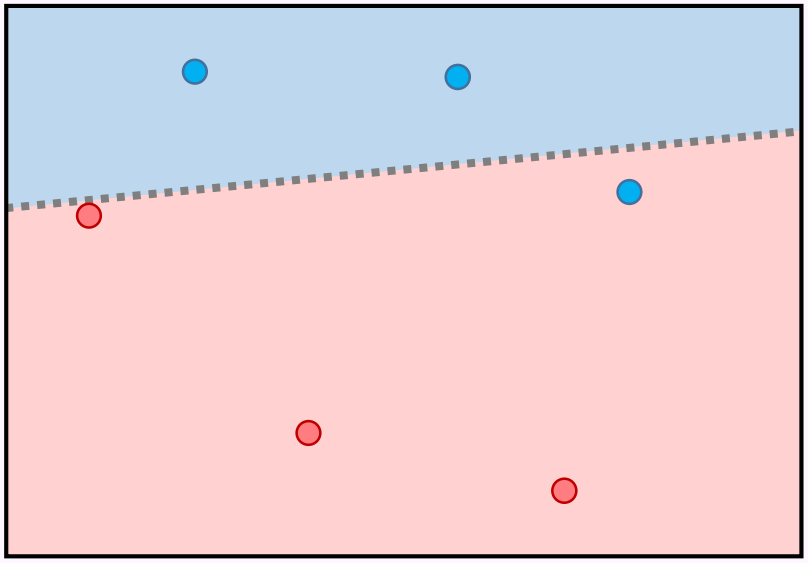


Error Function Kontinyu

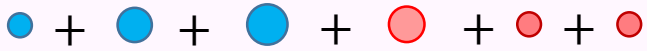


$E =$ 

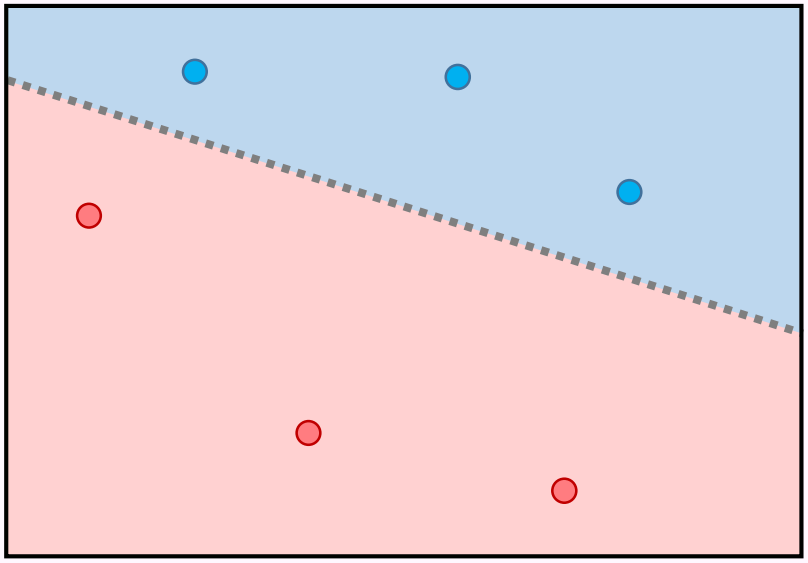
Error Function Kontinyu

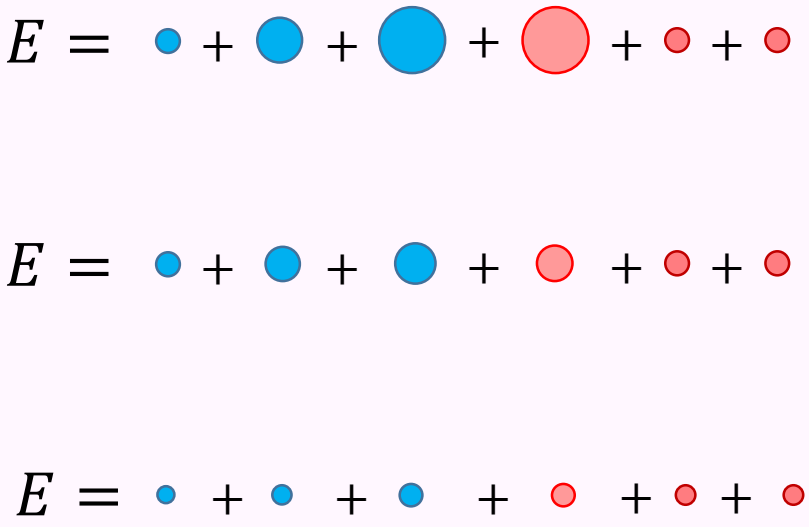


$E =$ 

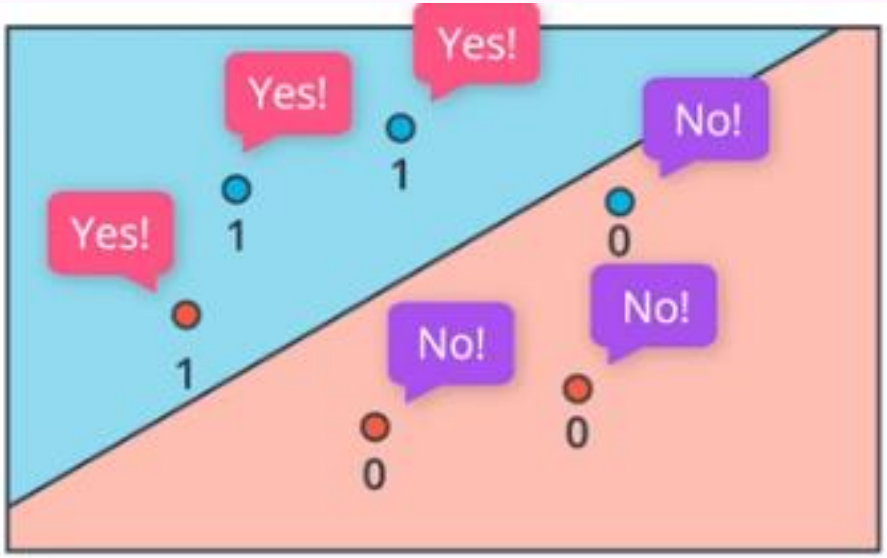
$E =$ 

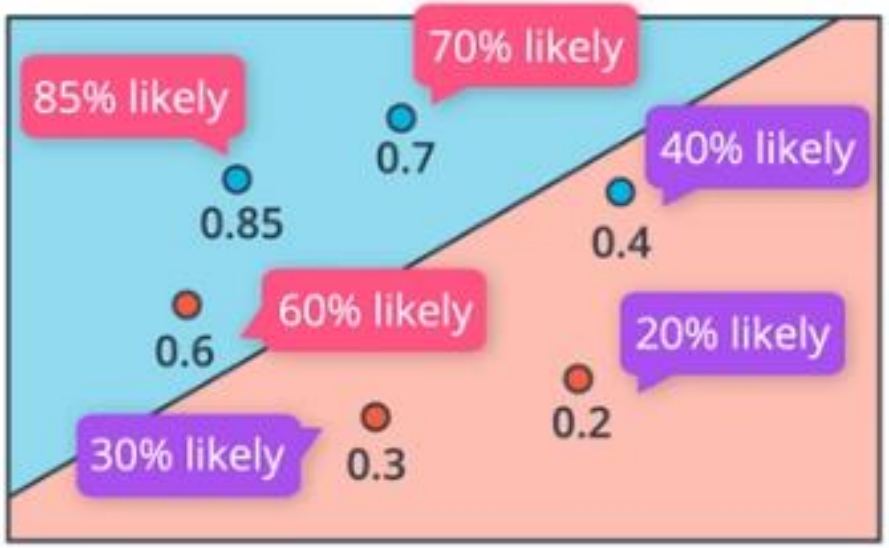
Error Function Kontinyu

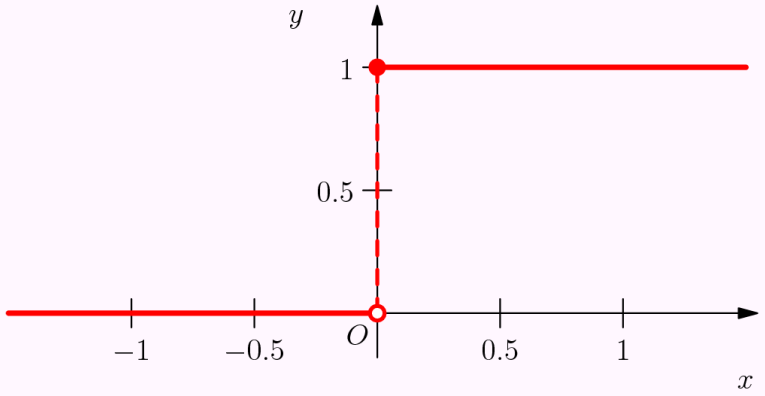




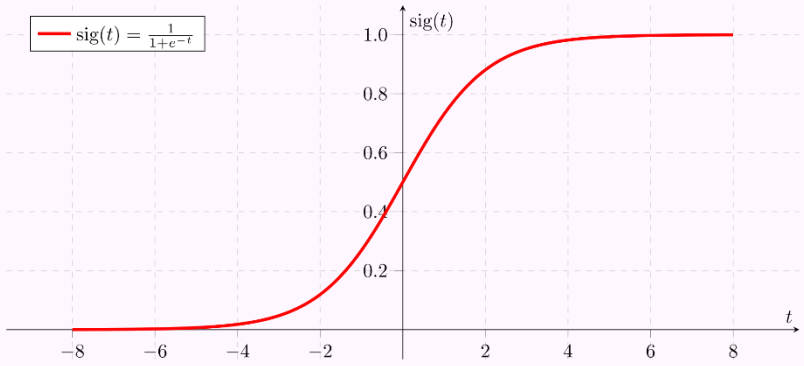
Error Function Kontinyu





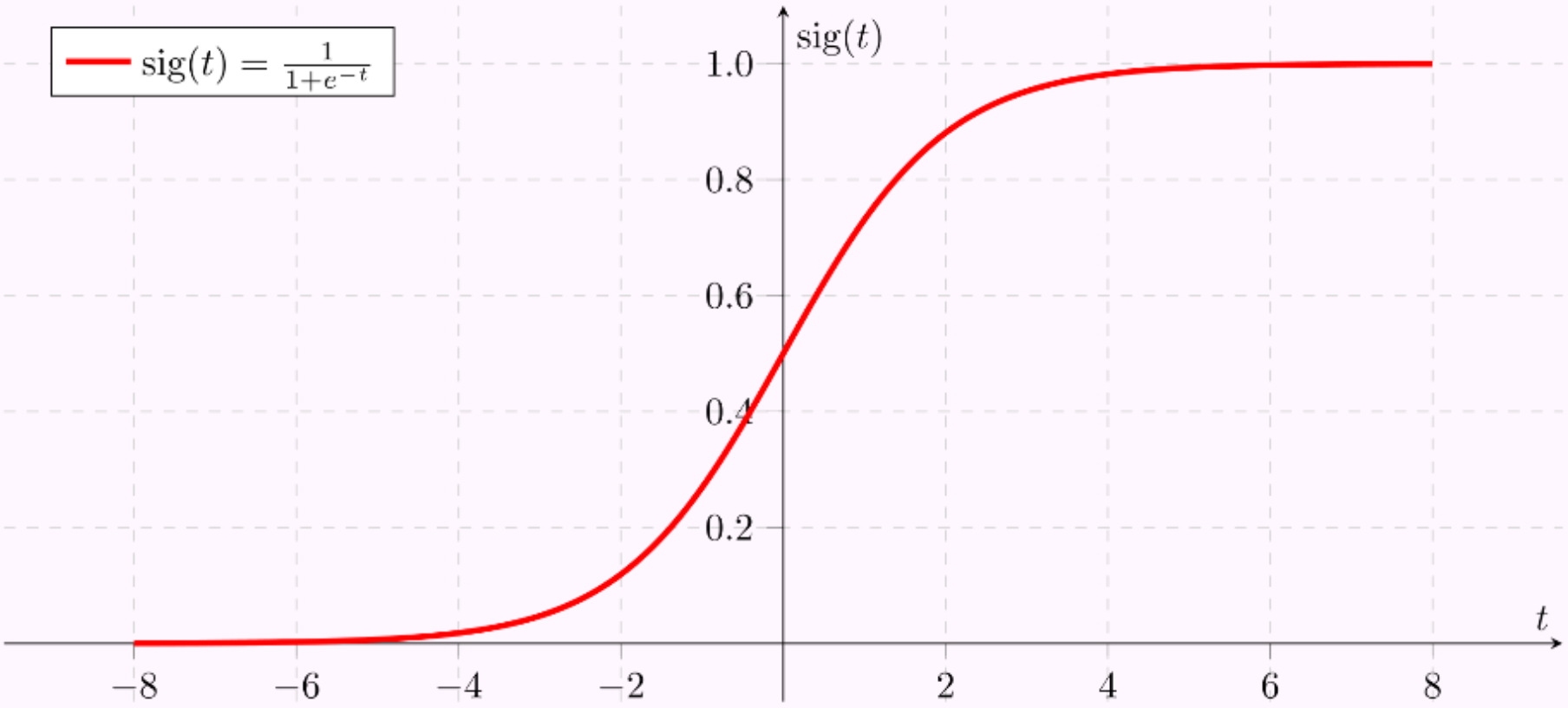


Step Function

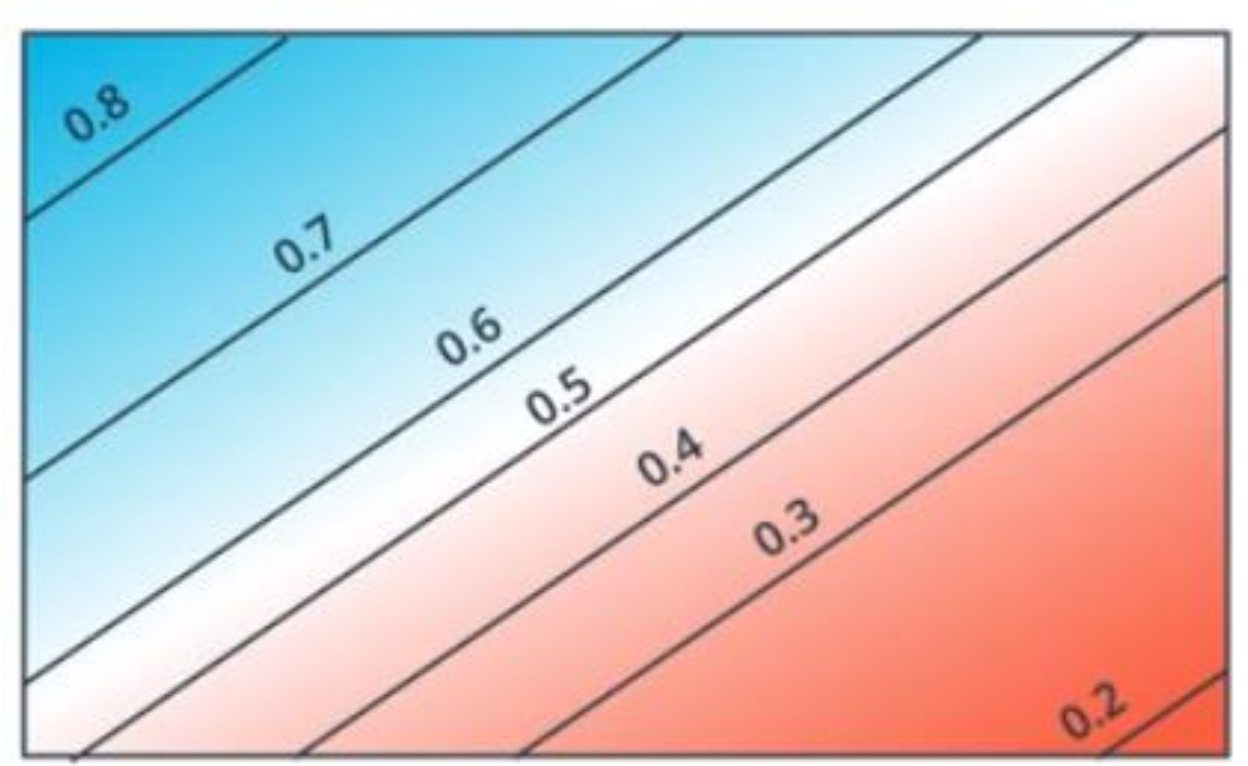


Sigmoid Function

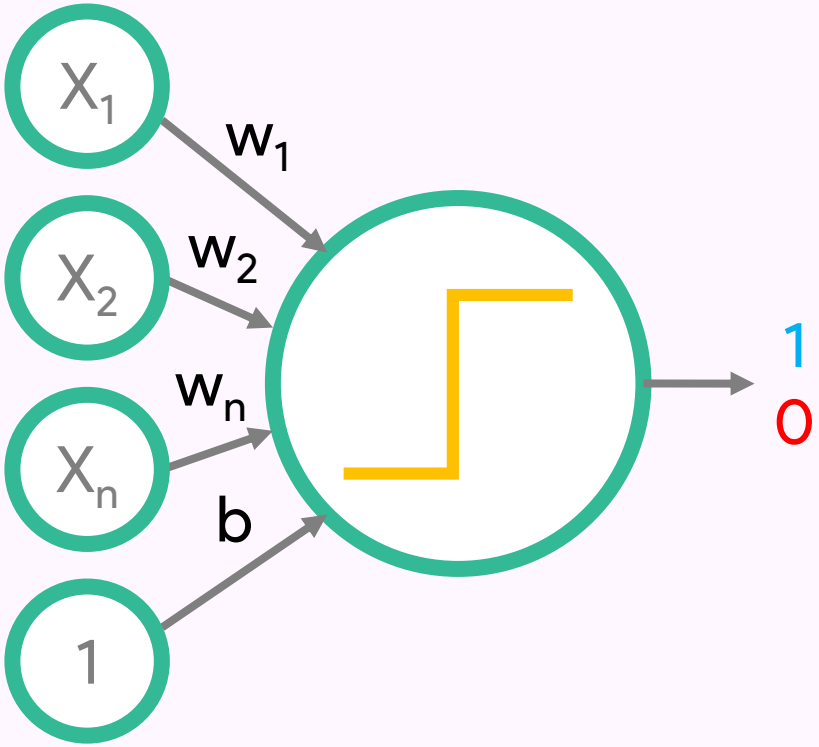
Error Function Kontinyu



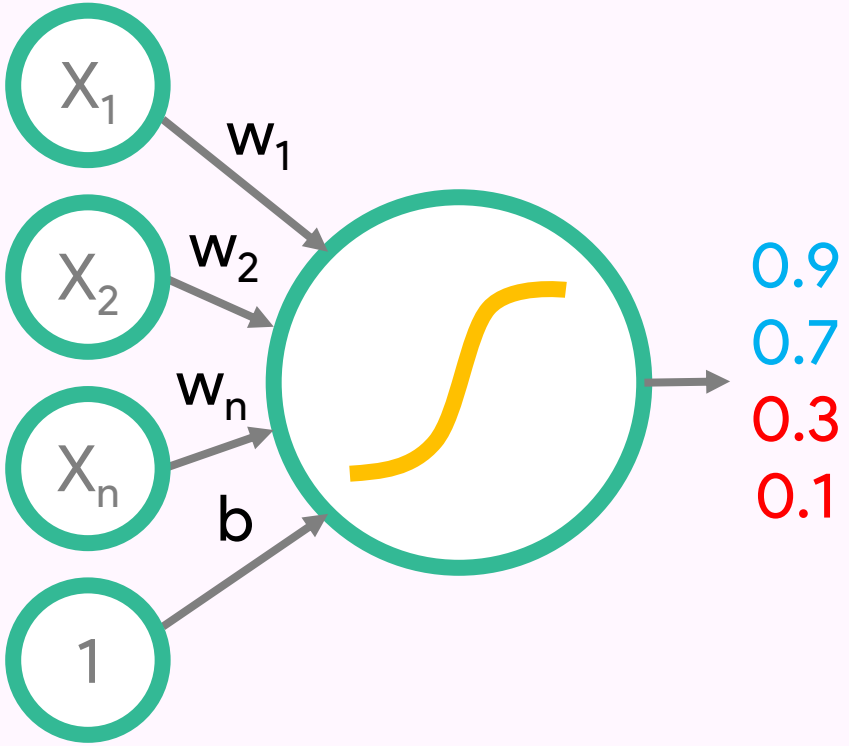
Pemetaan Nilai Error



Perubahan Perceptron



$$\text{step}(\text{W}x + b)$$



$$\sigma(\text{W}x + b)$$

Klasifikasi Multi-class

Dari Binary Classifier menjadi Multi-class Classifier

Ringkasan Sejauh Ini

Berat Badan x_1

Tinggi Badan x_2

Score

$w_1 x_1 + w_2 x_2 + b$

Probabilitas

$\sigma(\mathrm{W}\mathrm{x} + b)$

Total = 1

Date-able

0.1

Non-date-able

0.9

Klasifikasi Multiclass

Ikan Paus?

Ayam?

Rubah?

Total = 1

		Score	Probabilitas	
Ukuran Badan	x_1	$W_1x + b$	$?(W_1x + b)$	0.1
Lokasi tinggal	x_2	$W_2x + b$	$?(W_1x + b)$	0.7
...		$W_3x + b$	$?(W_1x + b)$	0.2

Klasifikasi Multiclass

$$0 \quad \frac{0}{0+2+1}$$

$$2 \quad \frac{2}{0+2+1}$$

$$1 \quad \frac{1}{0+2+1}$$

Bermasalah pada saat

$$\frac{1}{1+0+(-1)}$$

Fungsi yang selalu memberikan hasil positif?

a) Sin
b) Cos
c) Log
d) Exp

Klasifikasi Multiclass

0 $\frac{e^0}{e^0 + e^2 + e^1}$

2 $\frac{e^2}{e^0 + e^2 + e^1}$

1 $\frac{e^1}{e^0 + e^2 + e^1}$

Total = 1

0.09

0.24

0.67

Sigmoid dan Softmax

Sigmoid Function (Binary)

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

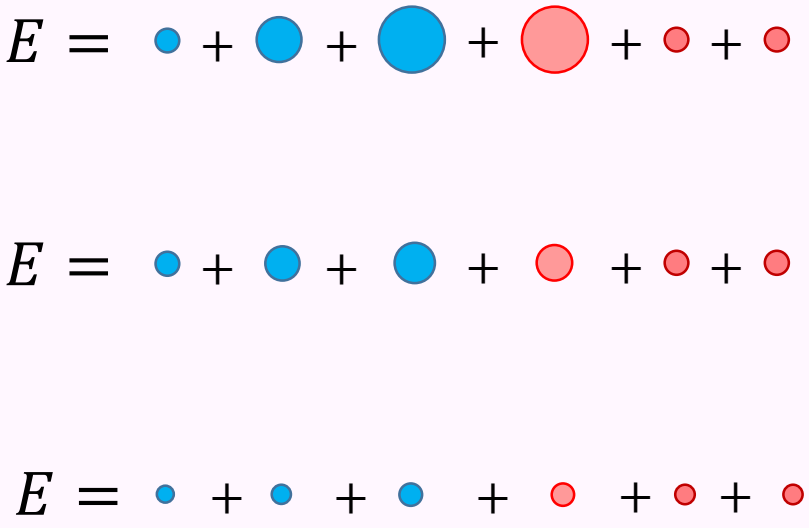
Softmax Function (Multiclass)

$$S(\hat{y}_i) = \frac{e^{\hat{y}_i}}{\sum_j e^{\hat{y}_j}}$$

Maximum Likelihood

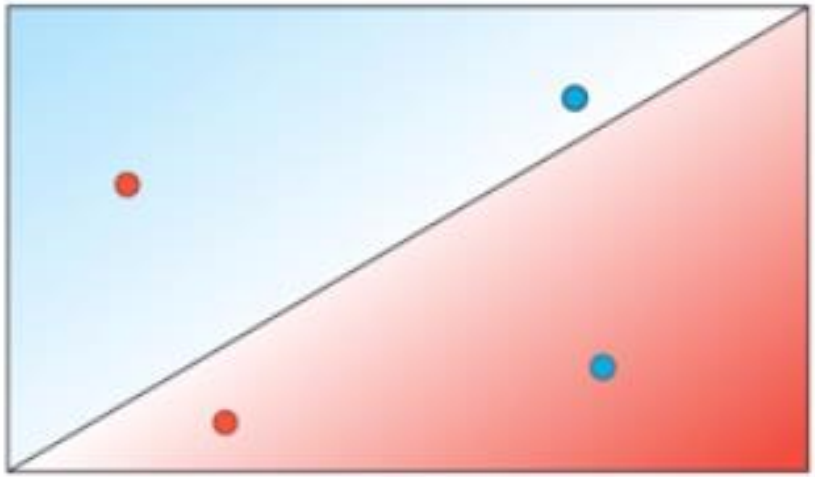
Memberi tahu komputer model yang baik dan buruk.

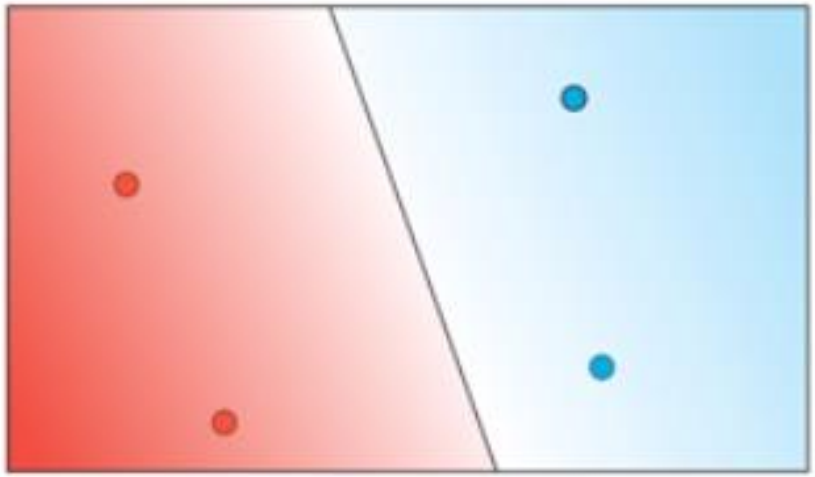
Error Function?



Maksudnya apa coba? Sangat tidak spesifik bagi komputer.

Menentukan Model yang Baik dan Buruk

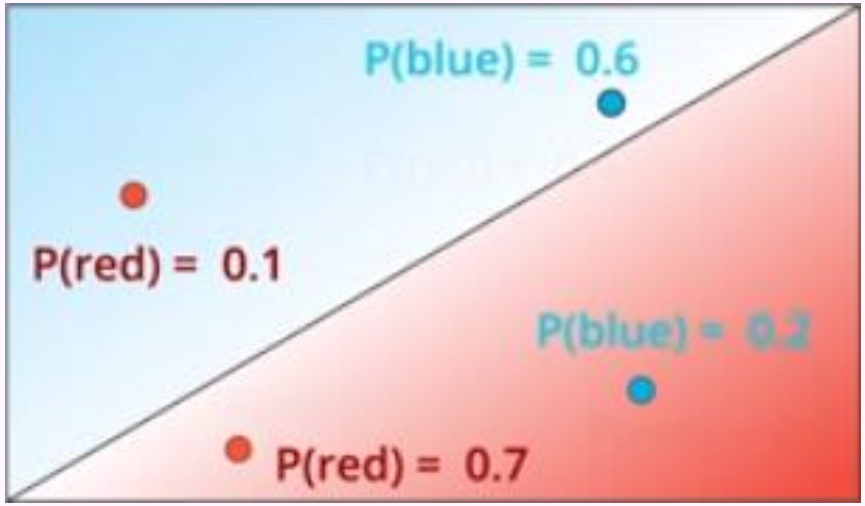




Yang ini buruk!

Tapi bagaimana cara kita memberitahu komputer bahwa ini model yang buruk?

Menentukan Model yang Baik dan Buruk



$$\hat{y} = \sigma(\mathrm{W}x + b)$$

$$\mathrm{P}(\mathrm{blue}) = \sigma(\mathrm{W}x + b)$$

Dengan asumsi warna-warna titik merupakan independent event

$$\mathrm{P}(\mathrm{red}) = 0.1$$

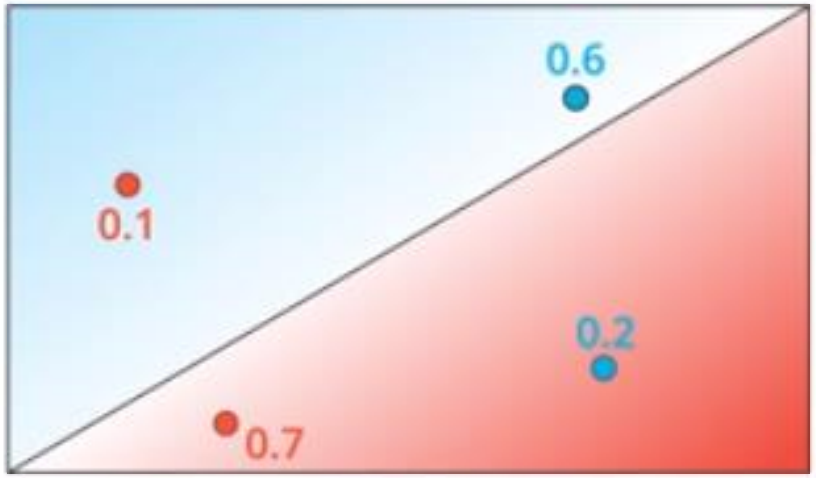
$$\mathrm{P}(\mathrm{blue}) = 0.6$$

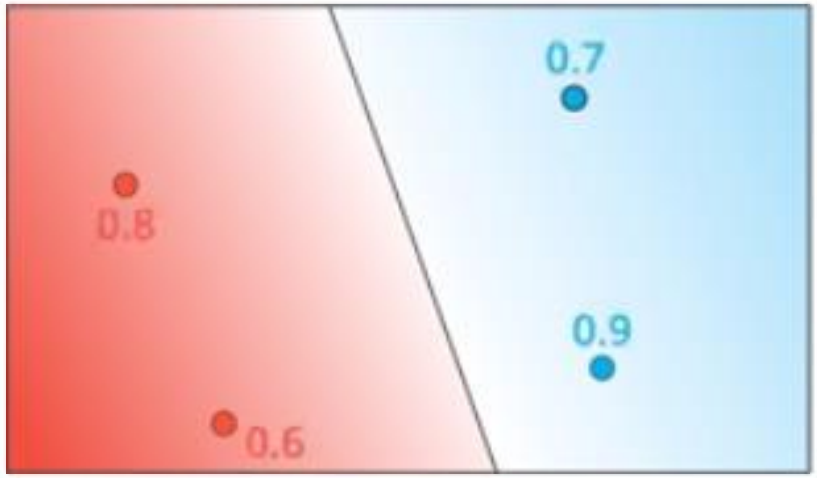
$$\mathrm{P}(\mathrm{red}) = 0.7$$

$$\mathrm{P}(\mathrm{blue}) = 0.2 \quad \times$$

$$\mathrm{P}(\mathbf{all}) = 0.0084$$

Menentukan Model yang Baik dan Buruk





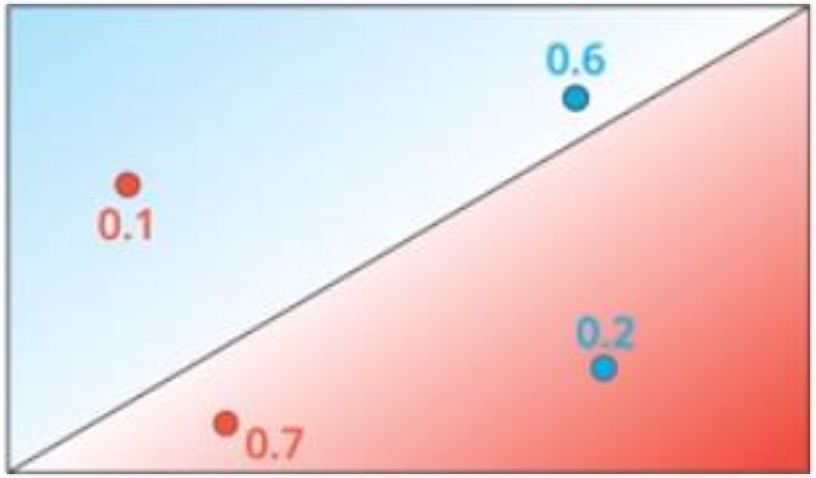
$0.1 \times 0.6 \times 0.7 \times 0.2 = 0.0084$

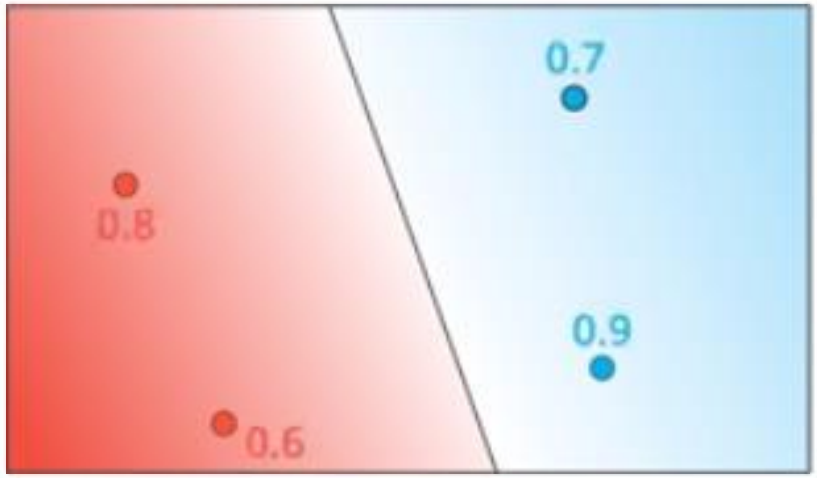
$0.8 \times 0.7 \times 0.9 \times 0.6 = 0.3024$

Spesifik dan mudah dimengerti komputer

Penentuan ini disebut dengan **Maximum Likelihood**

Menentukan Model yang Baik dan Buruk





$0.1 \times 0.6 \times 0.7 \times 0.2 = 0.0084$

$0.8 \times 0.7 \times 0.9 \times 0.6 = 0.3024$

- Sekarang kita tahu bahwa grafik kiri lebih buruk dari kanan berdarkan operasi product seluruh probabilitas point terklasifikasi sesuai warnanya.
- Semakin besar probabilitasnya semakin bagus model kita.
- Apakah kita berhasil menemukan Error Function yang kita cari?

Kelemahan Maximum Likelihood

$0.1 \times 0.6 \times 0.7 \times 0.2 = 0.0084$

$0.8 \times 0.7 \times 0.9 \times 0.6 = 0.3024$

- Sekarang bayangkan jika kita memiliki ribuan data poin.
- Masing-masing poin memiliki nilai probabilitas $0 < \mathrm{P} < 1$.
- $0.0 \ldots \times 0.0 \ldots \times 0.0 \ldots \times \cdots = 0.0000000 \ldots$
- Sangat tidak baik untuk perhitungan.

Cross Entropy

Error Function yang Sebenarnya

Formulasi Error Function

$0.1 \times 0.6 \times 0.7 \times 0.2 = 0.0084$

$\ln(0.1) + \ln(0.6) + \ln(0.7) + \ln(0.2)$

$-2.3 \quad -0.51 \quad -0.36 \quad -1.61$

$-\ln(0.1) - \ln(0.6) - \ln(0.7) - \ln(0.2)$

$2.3 \quad +0.51 \quad +0.36 \quad +1.61$

4.8

$0.8 \times 0.7 \times 0.9 \times 0.6 = 0.3024$

$\ln(0.8) + \ln(0.7) + \ln(0.9) + \ln(0.6)$

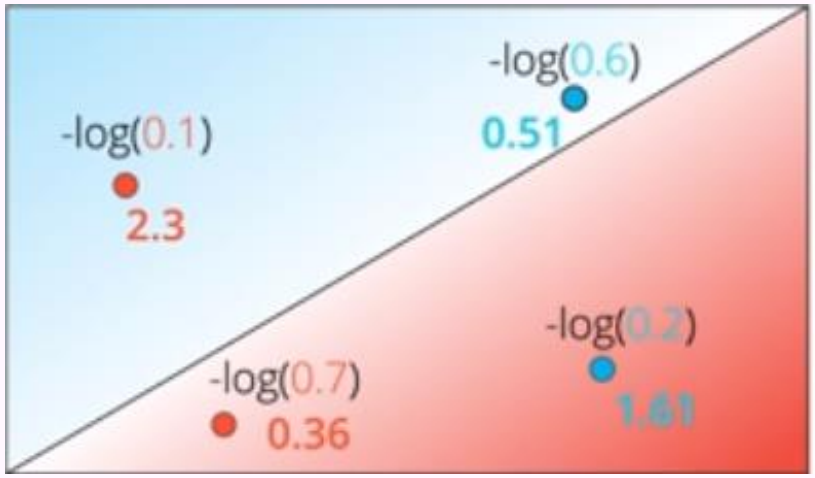
$-0.22 \quad -0.36 \quad -0.1 \quad -0.51$

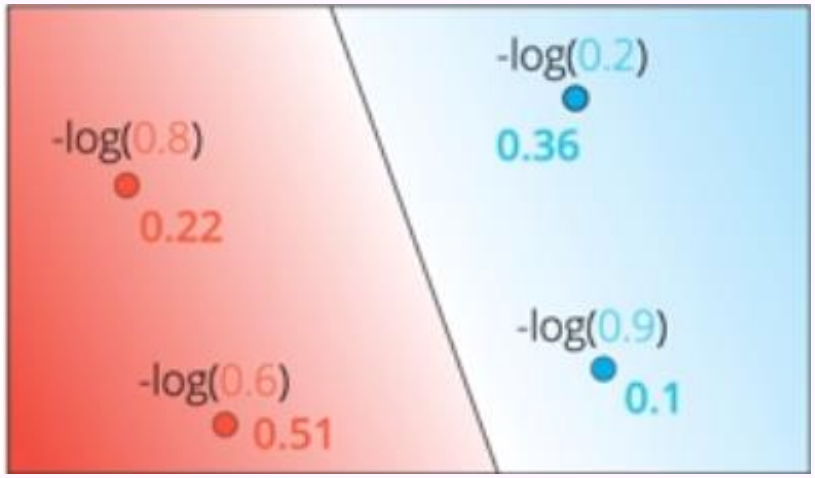
$-\ln(0.8) - \ln(0.7) - \ln(0.9) - \ln(0.6)$

$0.22 \quad +0.36 \quad +0.1 \quad +0.51$

1.2

Formulasi Error Function





$0.1 \times 0.6 \times 0.7 \times 0.2 = 0.0084$

$0.8 \times 0.7 \times 0.9 \times 0.6 = 0.3024$

$-\ln(0.1) - \ln(0.6) - \ln(0.7) - \ln(0.2)$

$-\ln(0.8) - \ln(0.7) - \ln(0.9) - \ln(0.6)$

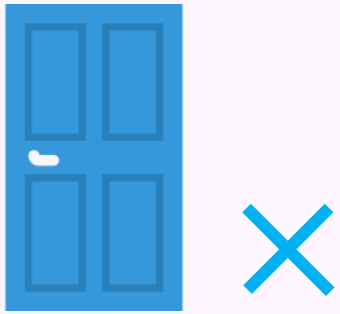
4.8

Cross Entropy

1.2

Formulasi Error Function

Formulasi Error Function

			
P(Gift)	0.8	0.7	0.1
P(No−gift)	0.2	0.3	0.9

Formulasi Error Function

			Probabilitas	-ln(Probabilitas)
0.8	0.7	0.1	0.0568	2.88
0.8	0.7	0.9	0.504	0.69
0.8	0.3	0.1	0.024	3.73
0.2	0.7	0.1	0.014	4.27
0.8	0.3	0.9	0.216	1.53
0.2	0.7	0.9	0.126	2.07
0.2	0.3	0.1	0.006	5.12
0.2	0.3	0.9	0.054	2.92

Formulasi Error Function

$p_1 = 0.8$

$p_2 = 0.7$

$p_3 = 0.1$



0.8 0.7 0.9

$\hat{y}_1$ $\hat{y}_2$ $1 - \hat{y}_3$

Label y_i $y_1 = 1$ $y_2 = 1$ $y_3 = 0$

Cross Entropy

$$-\ln(0.8) - \ln(0.7) - \ln(0.9)$$

```
sum = 0
for i in data:
    if (y==1):
        nilai = ln(ŷ_i)
    else:
        nilai = ln(1 − ŷ_i)
    sum = sum + nilai
CE = sum/length(sum)
```

Formulasi Error Function

$p_1 = 0.8$

$p_2 = 0.7$

$p_3 = 0.1$



0.8 0.7 0.9

$\hat{y}_1$ $\hat{y}_2$ $1 - \hat{y}_3$

Label y_i $y_1 = 1$ $y_2 = 1$ $y_3 = 0$

Cross Entropy

$$-\ln(0.8) - \ln(0.7) - \ln(0.9)$$

```
sum = 0
for i in data:
    if (y==1):
        nilai = ln(ŷ_i)
    else:
        nilai = ln(1 − ŷ_i)
    sum = sum + nilai
CE = sum/length(sum)
```

Formula Cross Entropy

$$-\frac{1}{m}\sum_{i=1}^{m}(y_i \ln(\hat{y}_i) + (1 - y_i)\ln(1 - y_i))$$

Formulasi Error Function

$p_1 = 0.8$

$p_2 = 0.7$

$p_3 = 0.1$

0.8 0.7 0.9

p_1 p_2 $1 - p_{13}$

Label y_i $y_i = 1$ $y_i = 1$ $y_i = 0$

Cross Entropy

$$-\ln(0.8) - \ln(0.7) - \ln(0.9)$$

```
sum = 0
for i in data:
    sum = sum + y_i * ln(ŷ_i) + (1 − y_i) * ln(1 − y_i)
CE = sum/length(sum)
```

Formula Cross Entropy

$$-\frac{1}{m}\sum_{i=1}^{m}(y_i \ln(\hat{y}_i) + (1 - y_i)\ln(1 - y_i))$$