

# Deep Dive into Deep Learning (Backpropagation)

Indonesia AI

Proprietary document of Indonesia AI 2023



# OBJECTIVE & OUTLINE

Proprietary document of Indonesia AI 2023



## Deep Dive into Deep Learning (Backpropagation)

Pahami konsep dasar melalui materi dan sesi langsung proses Backpropagation dalam algoritma Deep Learning.

### Outline:

1. Konsep Backpropagation
2. Loss Function
3. Learning Rate
4. Stochastic Gradient Descent (SGD)

# Konsep Backpropagation

# MULTI LAYER PERCEPTRON (REVIEW)

Proprietary document of Indonesia AI 2023



Indonesia AI

Ada 2 proses utama pada algoritma Deep Learning:

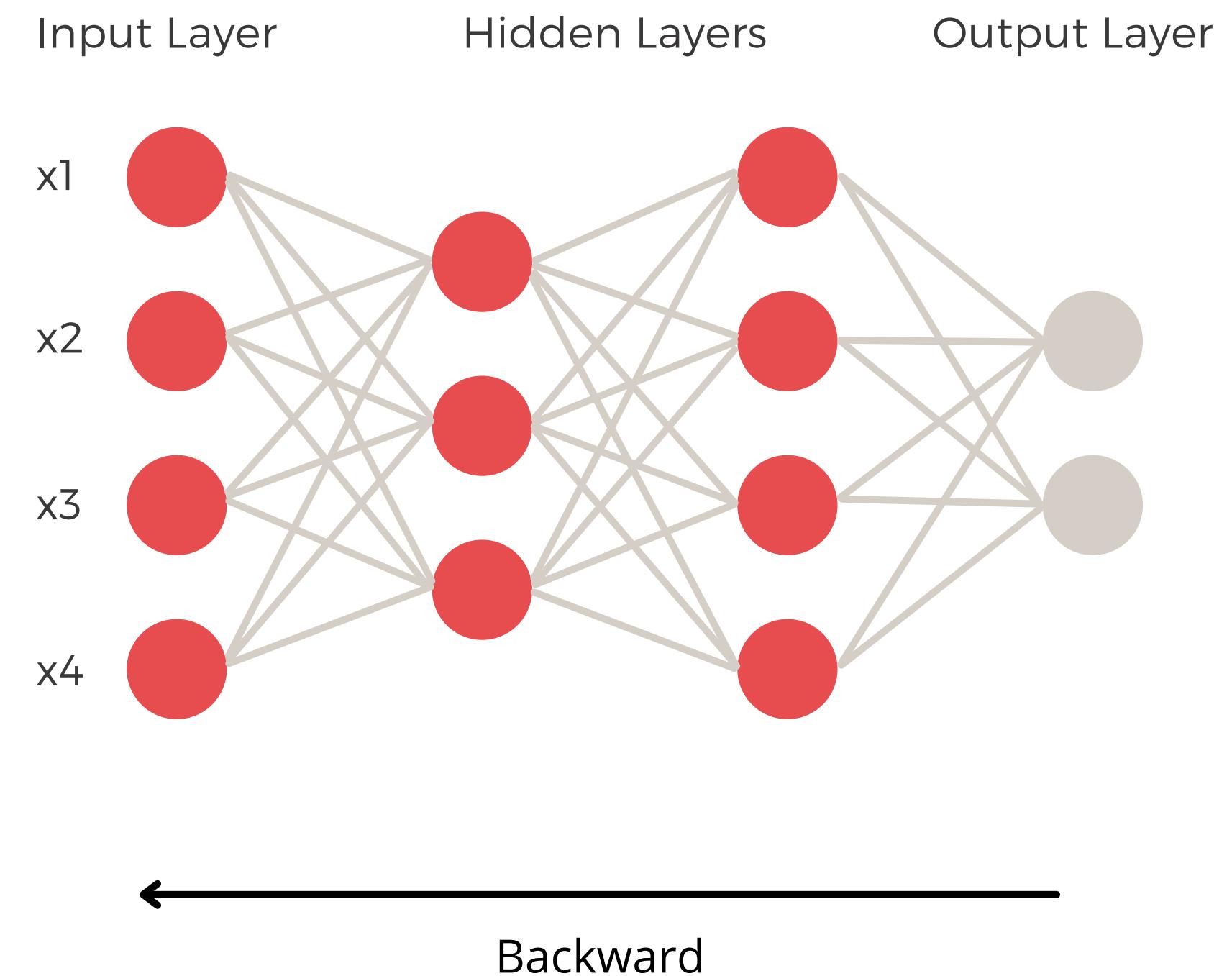
1. Proses Feedforward
2. Proses Backward Propagation

# PROSES BACKWARD

Proprietary document of Indonesia AI 2023

Backpropagation merupakan salah satu algoritma optimasi yang digunakan dalam Deep Learning untuk melakukan training pada model neural network.

Algoritma ini bekerja dengan menghitung gradien (gradient) dari error function (fungsi kesalahan) terhadap setiap parameter (bobot) yang ada pada jaringan saraf.



# PROSES BACKWARD

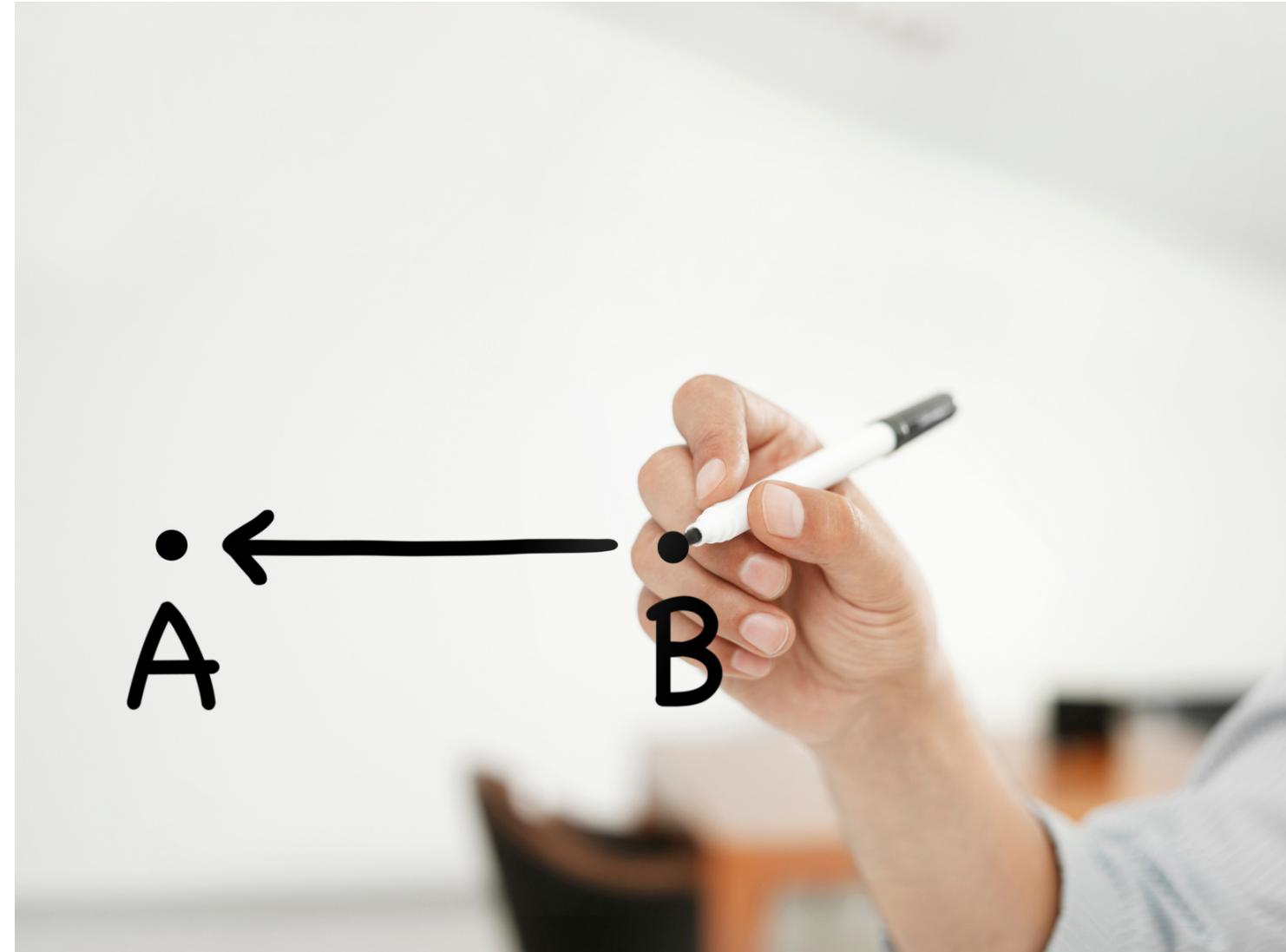
Proprietary document of Indonesia AI 2023

Secara umum, **backpropagation berarti berjalan mundur (atau "backwards")** melalui neural network, dari output layer menuju input layer.

Apa saja rangkaian proses yang terjadi disana?

# PROSES BACKWARD

Proprietary document of Indonesia AI 2023



Secara umum, proses yang terjadi adalah:

1. Perhitungan loss function
2. Perhitungan gradient
3. Update bobot/weight dan bias menggunakan algoritma optimizer (SGD, RMSProp, Adam etc.)

**Disclaimer:**  
**This is not math class so we will apply  
helicopter view technique**

# EXTRA CLASS (SOON!)

Proprietary document of Indonesia AI 2023



## Farah Qotrunnada

M.Sc Data Science, University of Sheffield UK

S.Si Matematika, Institut Teknologi Bandung (ITB)

Spesialisasi:

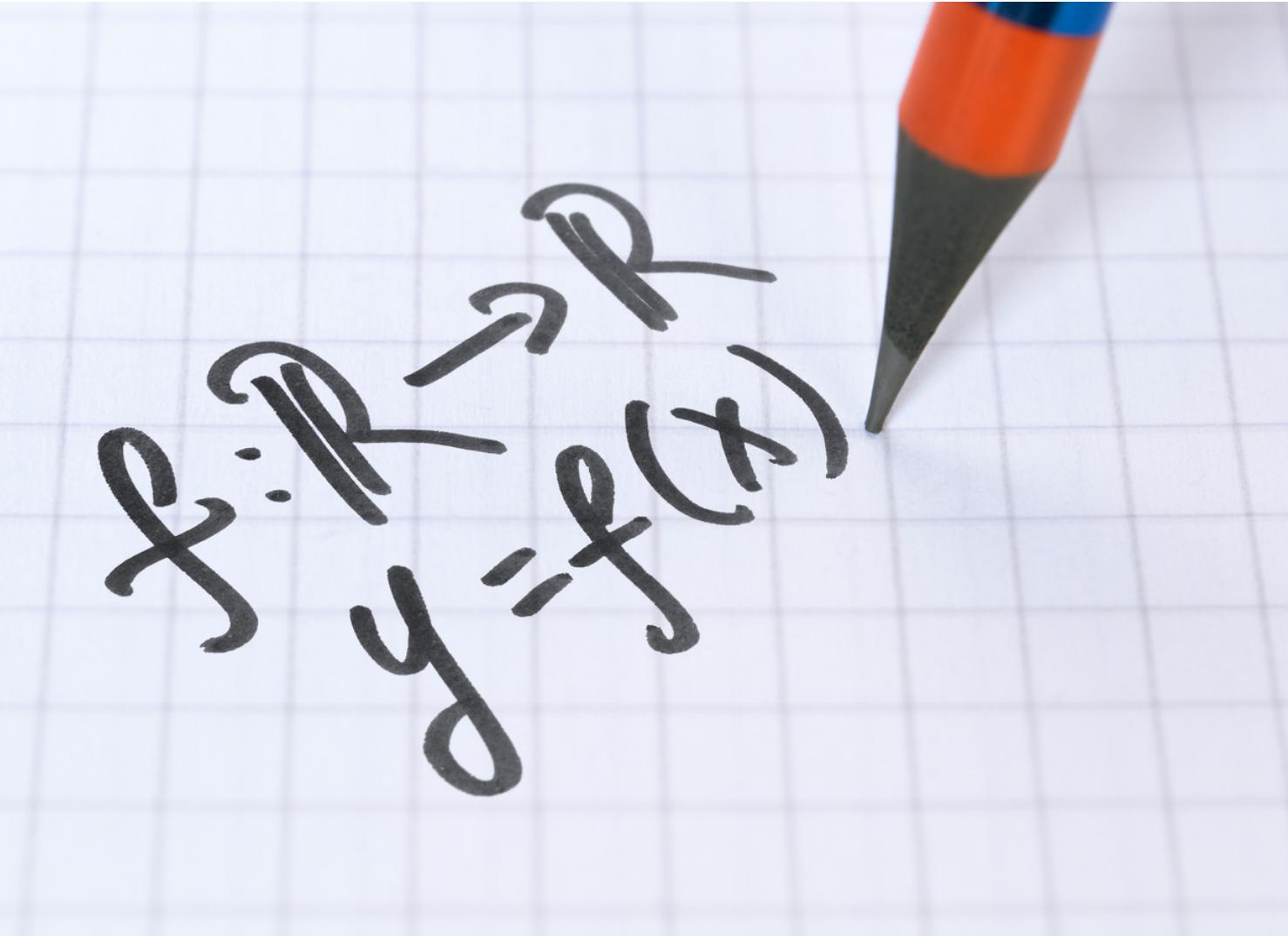
Math for Machine Learning & Deep Learning

**EXTRA CLASS:** 1 & 8 April 2023

# — Loss Function

# DEFINISI

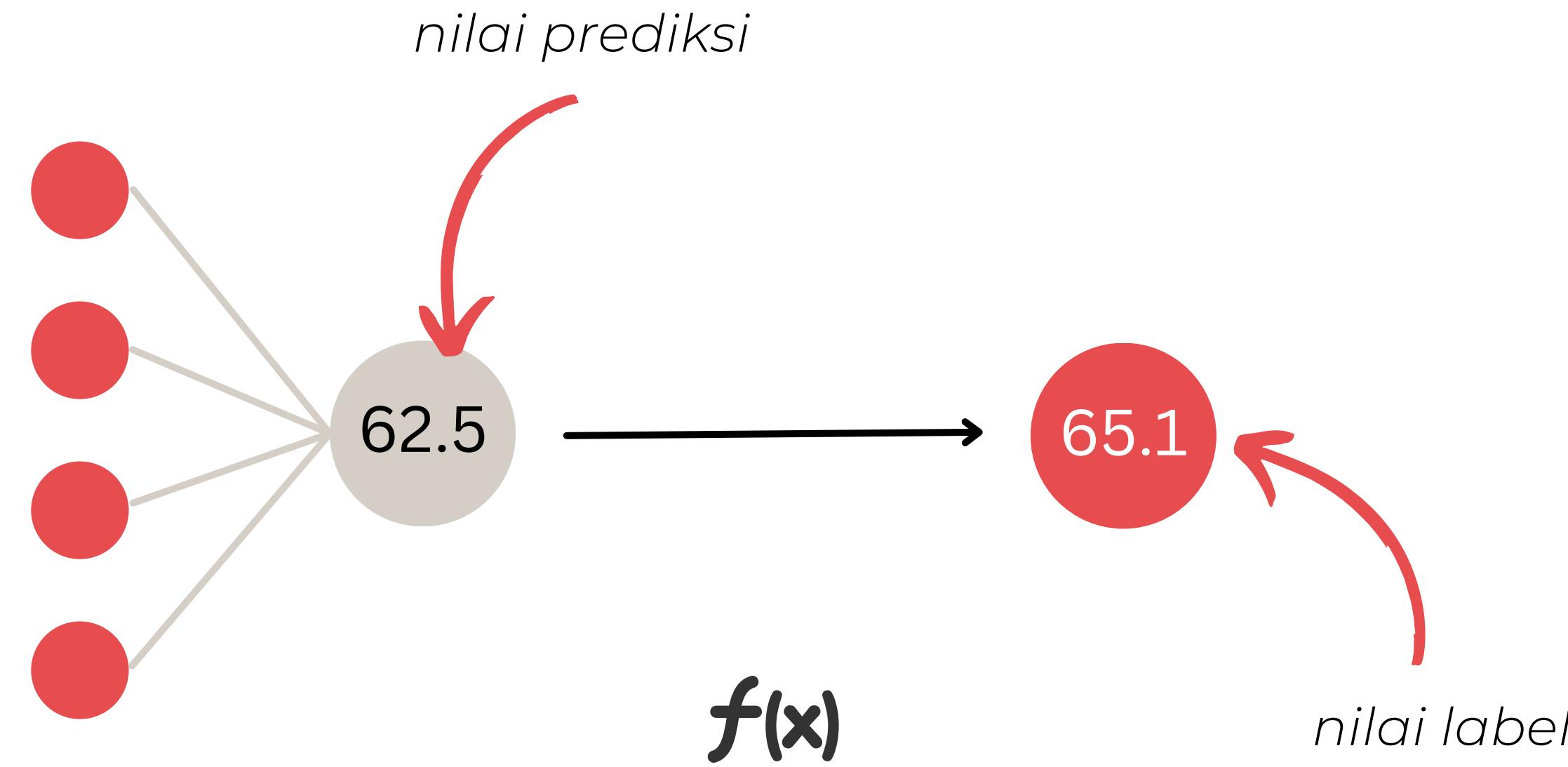
Proprietary document of Indonesia AI 2023



Loss function adalah suatu fungsi matematis yang digunakan sebagai metrik untuk mengukur seberapa baik performa model neural network dalam memprediksi. Pengukuran ini terjadi setiap iterasi saat proses training berlangsung.

# DEFINISI

Proprietary document of Indonesia AI 2023



Sederhananya sih, Loss function ini mesti bisa ngitung seberapa betul atau deket hasil nilai prediksi si neural network terhadap nilai yang ada pada label yang kita punya....

# DEFINISI

Proprietary document of Indonesia AI 2023

Nah, tujuan dari proses training model neural network adalah untuk **meminimalkan nilai loss** yang di dapat dari iterasi ke iterasi berikutnya agar model dapat memberikan prediksi yang pada akhirnya akurat.

*Nah ini namanya learning/belajar ~*

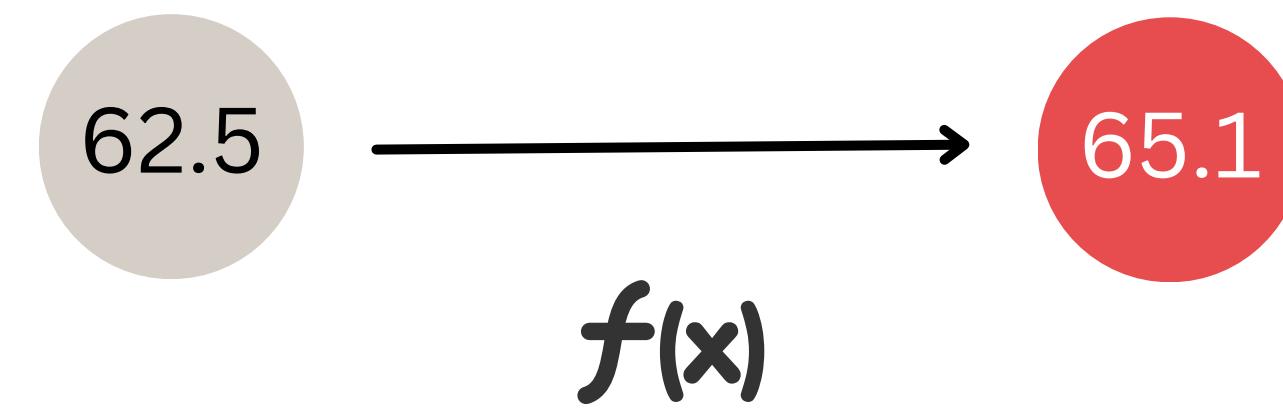
# LOSS FUNCTION

Proprietary document of Indonesia AI 2023

Simpelnya sih kalau ngitung seberapa betul atau deket hasil prediksi si neural network, tinggal dikurangin aja nilai keduanya yah.

Namun, para ilmuwan mengajukan beberapa loss function yang beragam dan cerdas:

1. RMSE
2. Binary Cross Entropy
3. Cross Entropy loss



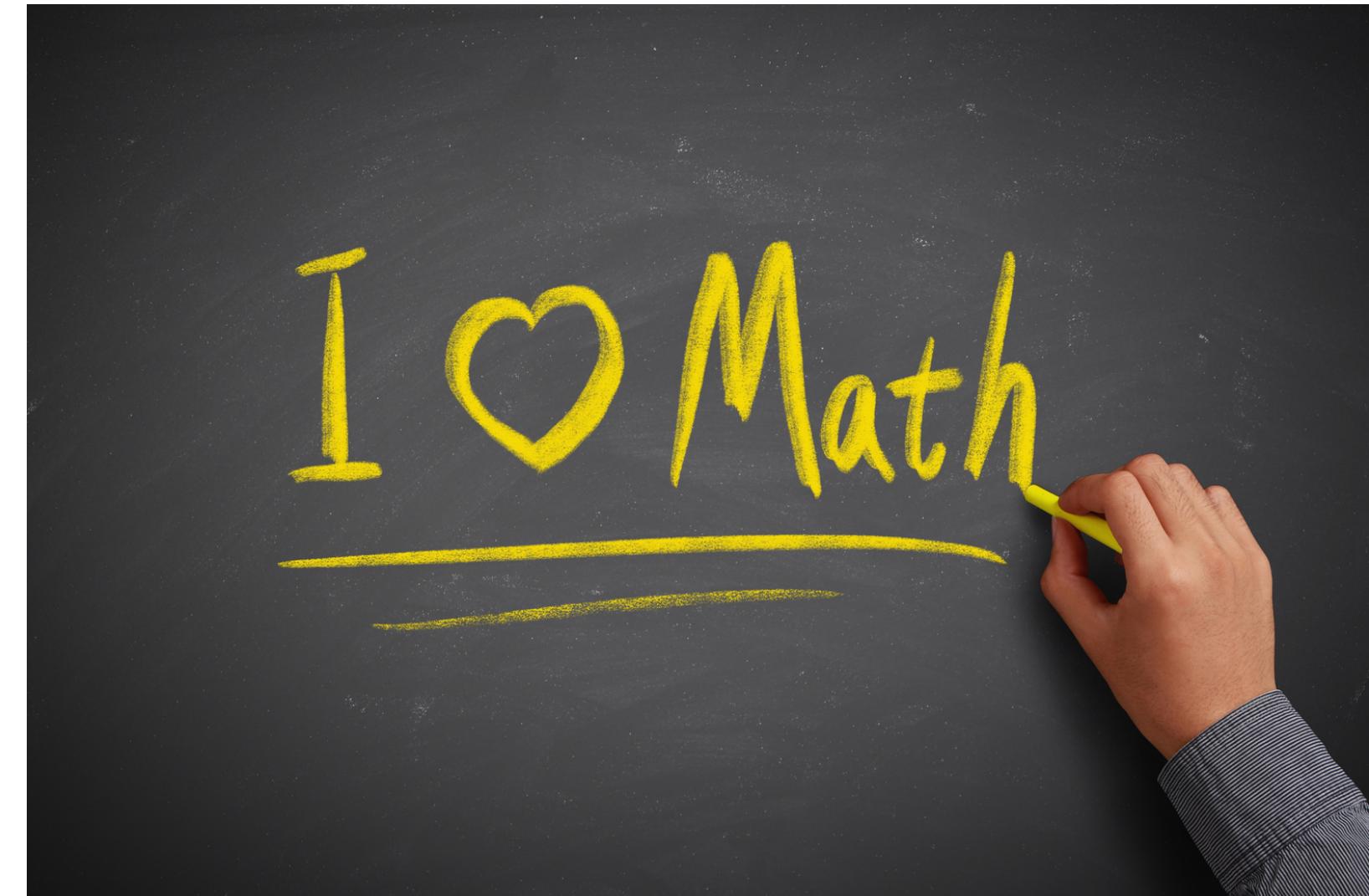
$$65.1 - 62.5 = 2.6?$$

# RMSE (ROOT MEAN SQUARRED ERROR)

Proprietary document of Indonesia AI 2023

RMSE (Root Mean Square Error) adalah salah satu metrik evaluasi yang digunakan untuk mengukur seberapa baik performa model regresi dalam memprediksi nilai target pada setiap iterasi pelatihan.

RMSE adalah akar dari nilai rata-rata dari kuadrat selisih antara nilai prediksi dengan nilai target yang sebenarnya.



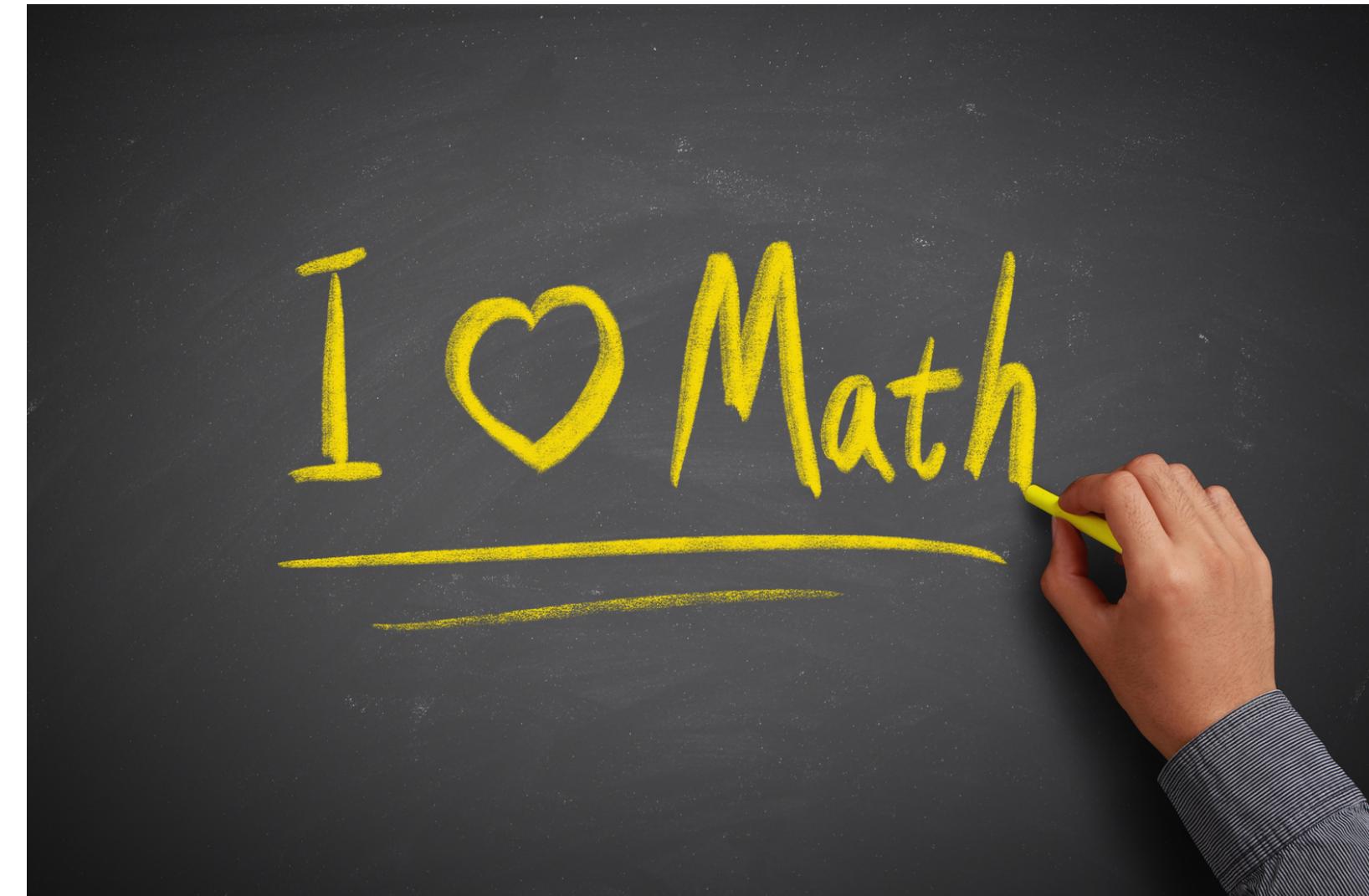
# RMSE (ROOT MEAN SQUARRED ERROR)

Proprietary document of Indonesia AI 2023

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum (y_{\text{pred}} - y_{\text{true}})^2}$$

Keterangan:

- N : jumlah sampel pada data
- y\_pred : nilai prediksi dari model
- y\_true : nilai target yang sebenarnya pada data

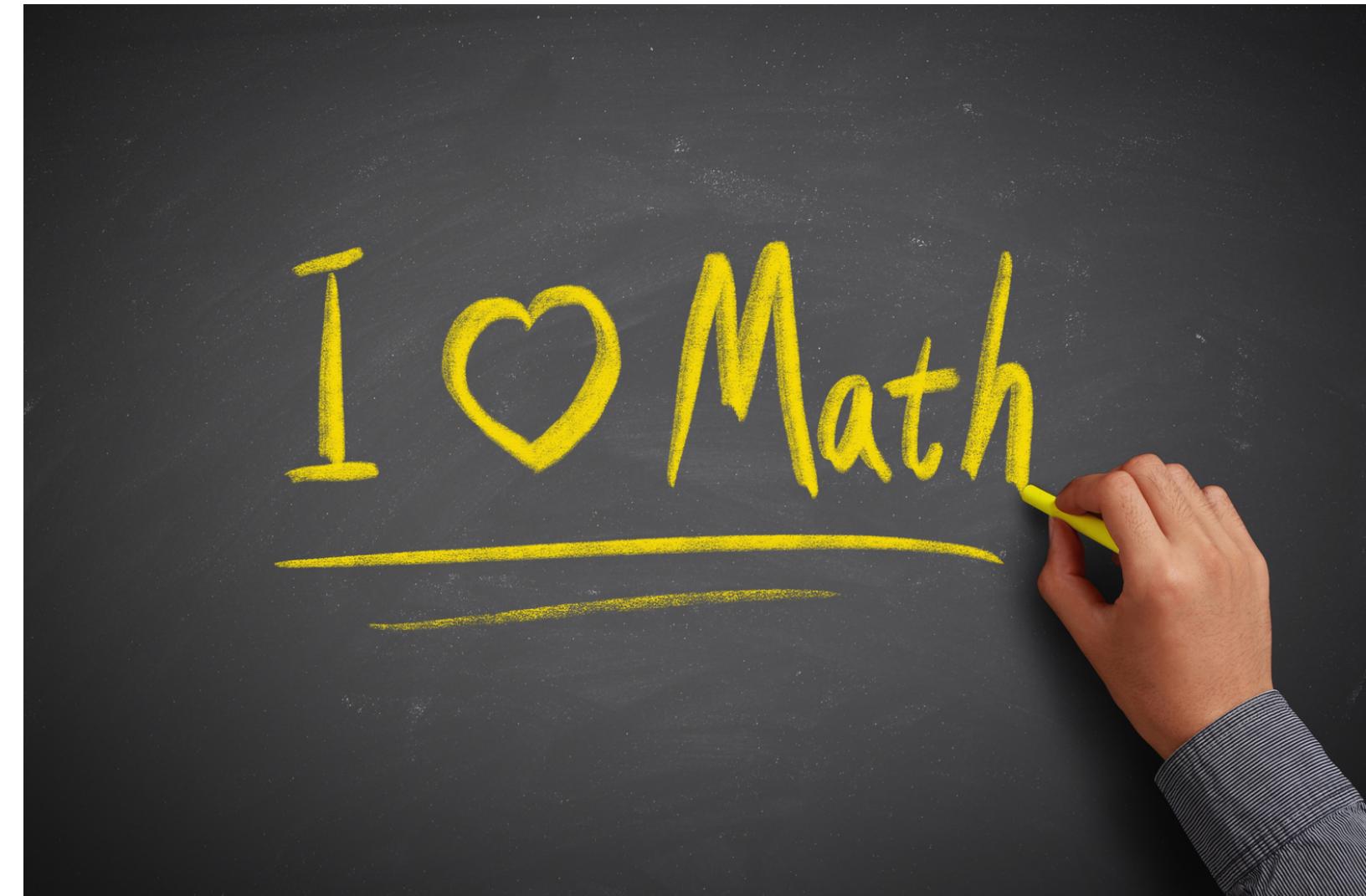


# BINARY CROSSENTROPY

Proprietary document of Indonesia AI 2023

Binary Cross-Entropy (BCE) adalah salah satu fungsi kerugian yang umum digunakan pada model neural network untuk masalah klasifikasi biner.

BCE digunakan untuk mengukur kesalahan atau perbedaan antara nilai prediksi dan nilai target pada setiap iterasi pelatihan.



# BINARY CROSSENTROPY

Proprietary document of Indonesia AI 2023

$$\text{BCE} = -(y_{\text{true}} * \log(y_{\text{pred}}) + (1 - y_{\text{true}}) * \log(1 - y_{\text{pred}}))$$

Keterangan:

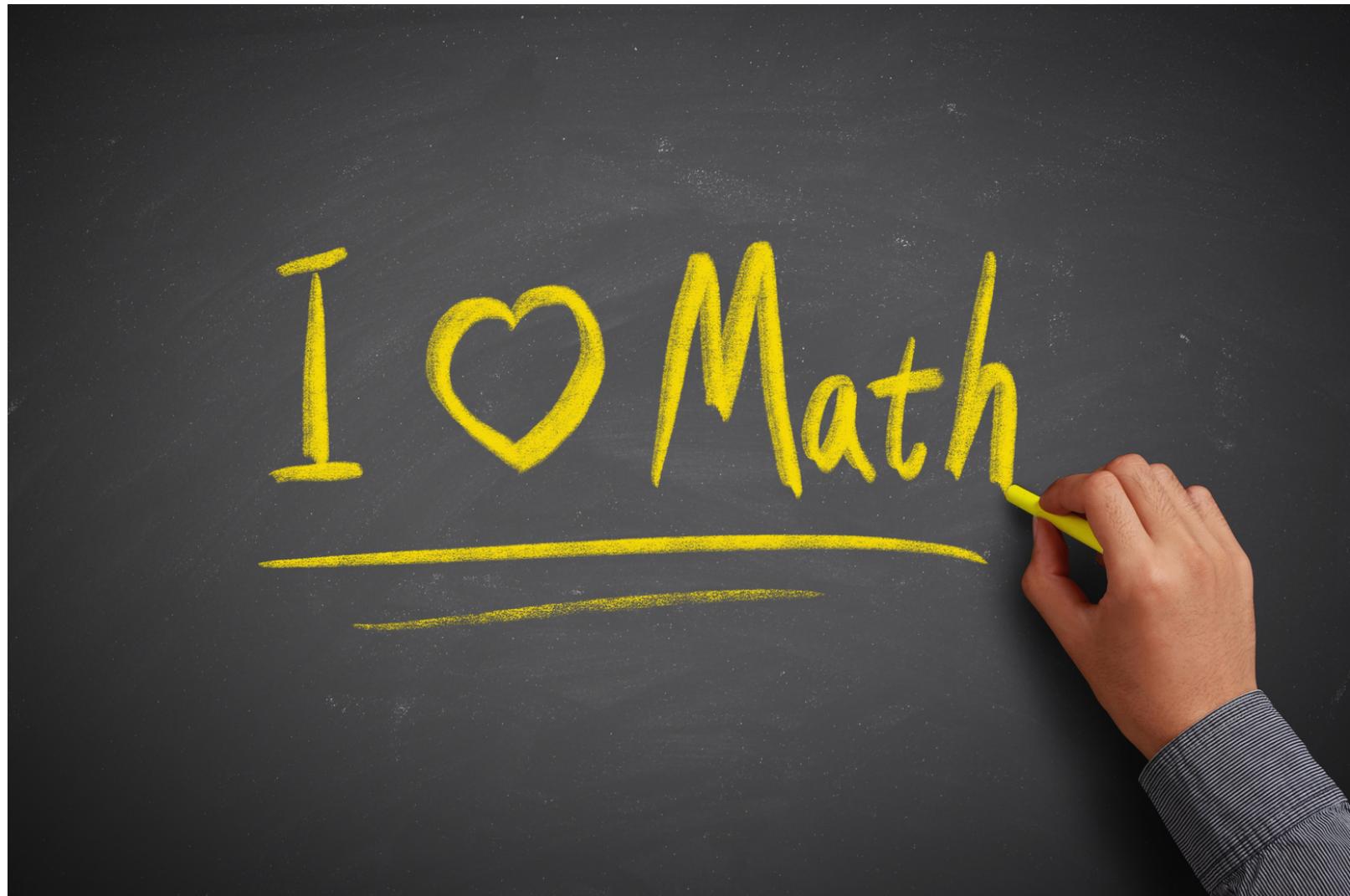
- $y_{\text{true}}$  : nilai target yang sebenarnya (0 atau 1)
- $y_{\text{pred}}$  : probabilitas prediksi dari model (dalam rentang 0 hingga 1)



# CROSSENTROPY LOSS

Proprietary document of Indonesia AI 2023

Cross-entropy loss adalah suatu metrik atau fungsi kerugian yang digunakan pada model neural network untuk mengukur kesalahan atau perbedaan antara distribusi probabilitas prediksi dan distribusi probabilitas nilai target pada setiap iterasi pelatihan.



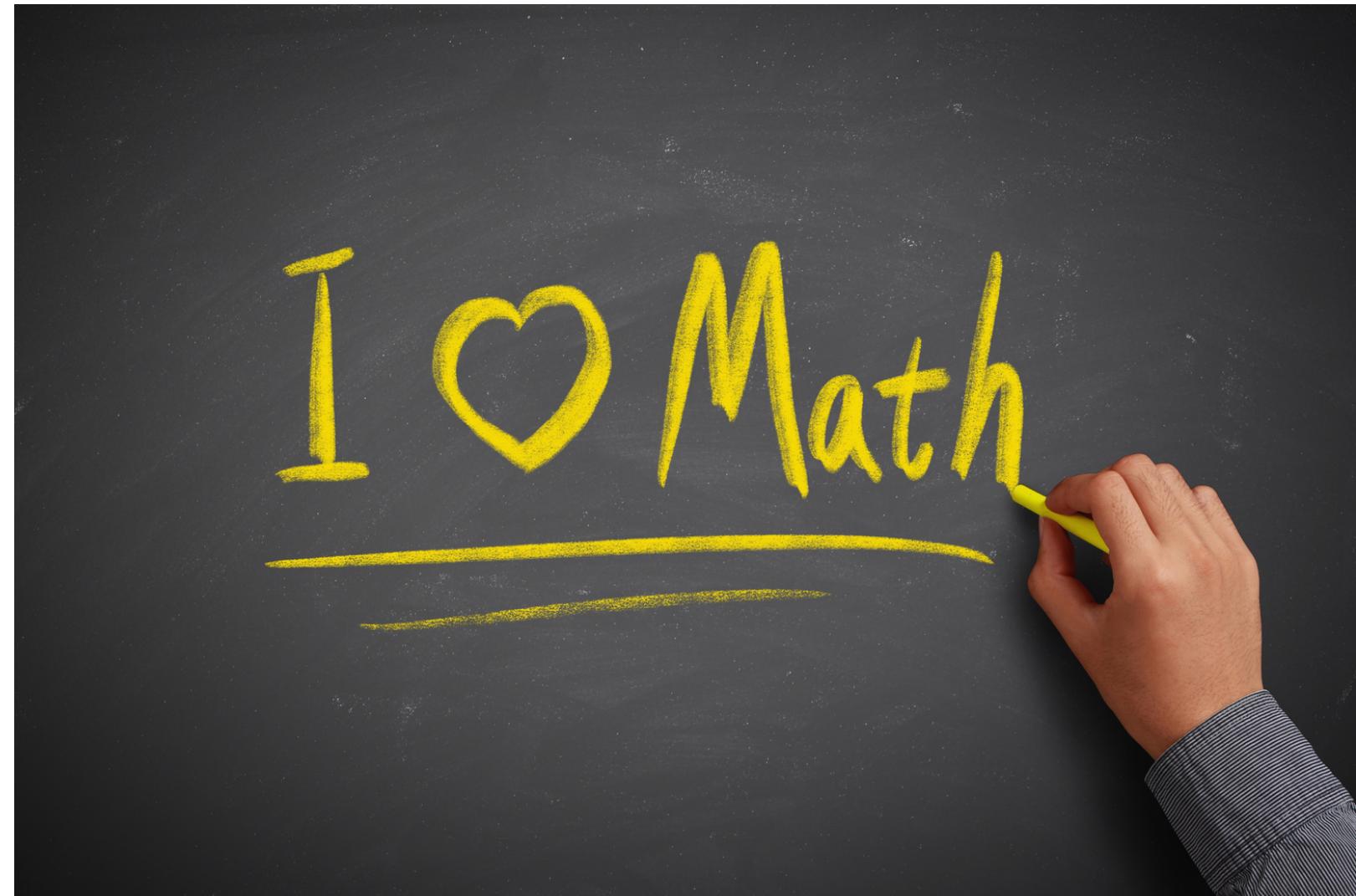
# CROSSENTROPY LOSS

Proprietary document of Indonesia AI 2023

$$\text{Cross-Entropy Loss} = - \sum (y_{\text{true}} * \log(y_{\text{pred}}))$$

Keterangan:

- $y_{\text{true}}$  : distribusi probabilitas nilai target atau label (dalam bentuk one-hot encoding atau integer)
- $y_{\text{pred}}$  : distribusi probabilitas prediksi dari model



# LOSS FUNCTION LAINNYA

Proprietary document of Indonesia AI 2023

## Regression Loss Functions

- Mean Squared Error Loss
- Mean Squared Logarithmic Error Loss
- Mean Absolute Error Loss



## Binary Classification Loss Functions

- Binary Cross-Entropy
- Hinge Loss
- Squared Hinge Loss

## Multi-Class Classification Loss Functions

- Categorical Cross-Entropy Loss
- Sparse Multiclass Cross-Entropy Loss
- KL (Kullback Leibler) Divergence Loss

# Any question guys ~

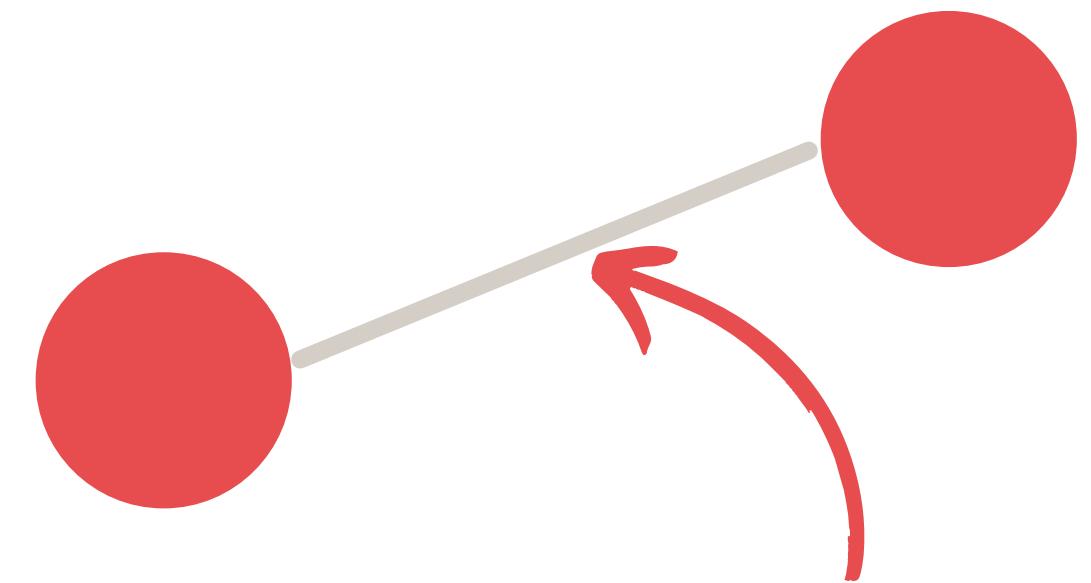
# — Learning Rate

# LEARNING RATE

Proprietary document of Indonesia AI 2023

Learning rate adalah salah satu parameter penting pada algoritma optimisasi yang digunakan pada pelatihan model neural network, termasuk pada algoritma Stochastic Gradient Descent (SGD).

Learning rate mengontrol seberapa besar perubahan bobot/weight pada setiap iterasi selama proses training.



bobot/weight

new weight = old weight - learning rate \* gradient of error

# LEARNING RATE

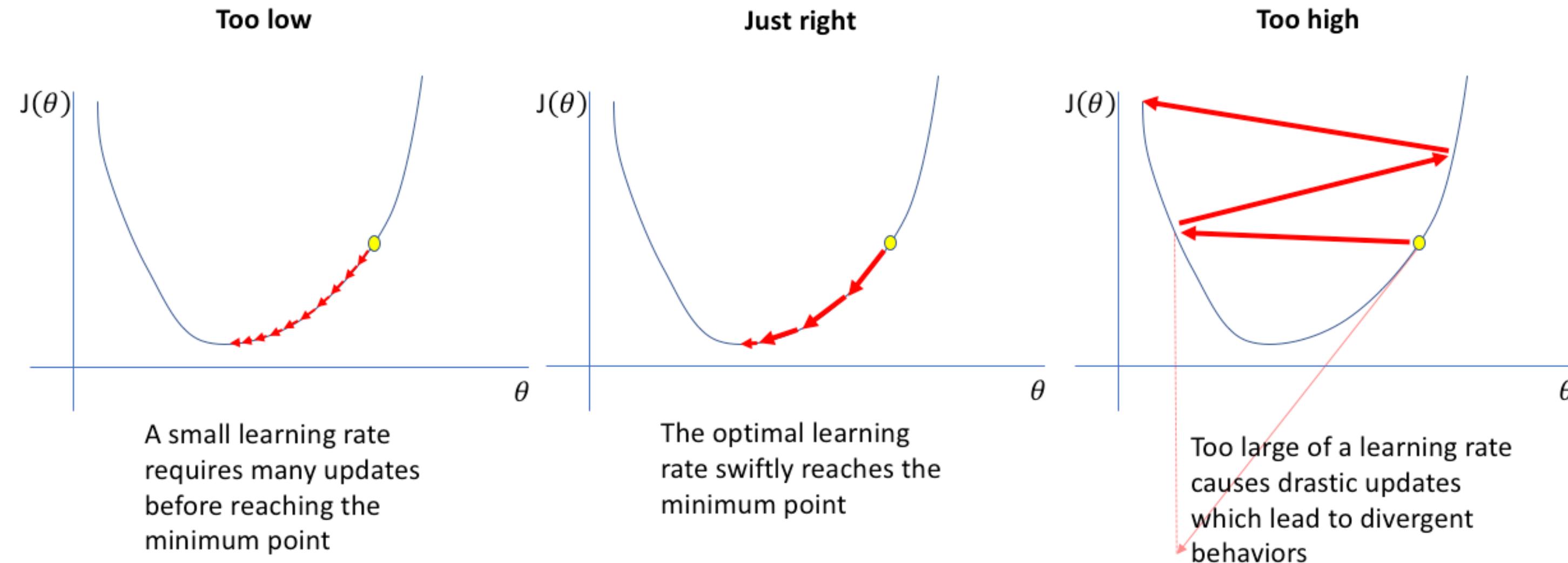
Proprietary document of Indonesia AI 2023

Sama seperti Hyperparameter lainnya seperti Epoch & Batch Size, nilai Learning rate juga perlu ditentukan dengan baik.

Learning rate yang terlalu besar dapat menyebabkan model menjadi tidak stabil atau bahkan tidak konvergen, sedangkan learning rate yang terlalu kecil dapat menyebabkan model memerlukan waktu yang lebih lama untuk mencapai hasil yang optimal.

# PEMILIHAN LEARNING RATE

Proprietary document of Indonesia AI 2023



Indonesia AI

Image source: [jeremyjordan.me](https://jeremyjordan.me)

# PEMILIHAN LEARNING RATE

Proprietary document of Indonesia AI 2023



*AI is empirical science ~*

Dalam prakteknya, nilai learning rate biasanya **ditentukan secara empiris** melalui eksperimen berulang pada model yang sama.

Beberapa teknik yang dapat diterapkan seperti learning rate scheduling, adaptive learning rate, dan cyclical learning rate.

# **Stochastic Gradient Descent (SGD)**

# DEFINISI

Proprietary document of Indonesia AI 2023

Stochastic Gradient Descent (SGD) adalah salah satu **algoritma optimisasi** yang digunakan untuk meminimalkan nilai dari loss function pada model neural network dan jenis algoritma ini sering digunakan dalam penerapan Deep Learning.

Karenanya SGD ini juga dikenal sebagai Optimizer!

$$\text{New weight} = \text{Old weight} - \text{Learning rate} \left( \frac{\partial \text{Error}}{\partial W_x} \right)$$

The diagram illustrates the SGD update rule with handwritten annotations:

- "Old weight" points to the term  $W_x$  in the equation.
- "Derivative of Error with respect to weight" points to the term  $\frac{\partial \text{Error}}{\partial W_x}$ .
- "Learning rate" points to the scalar value "a" in the equation.
- "New weight" points to the result of the subtraction in the equation.

# DEFINISI

Proprietary document of Indonesia AI 2023

Ada 2 jenis algoritma  
optimizer pada Deep  
Learning:

1. Gradient Descent Optimizer
2. Adaptive Optimizer

# DEFINISI

Proprietary document of Indonesia AI 2023

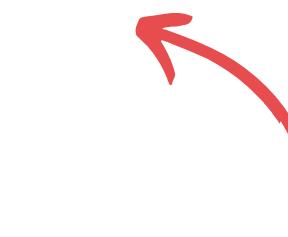


## Gradient Descent Optimizer

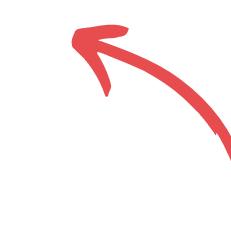
- Batch gradient descent
- Mini-batch gradient descent
- Stochastic gradient descent (SGD)

## Adaptive Optimizer

- Adagrad
- Adadelta
- RMSprop
- Adam



Tidak adaptif perlu LR



Adaptif gak perlu LR

# STOCHASTIC GRADIENT DESCENT (SGD)

Proprietary document of Indonesia AI 2023

---

Pada dasarnya, SGD bekerja dengan memperbarui bobot (weight) dalam neural network secara bertahap dengan menghitung gradien kesalahan terhadap bobot tersebut pada setiap iterasi.

Pada SGD, gradien kesalahan hanya dihitung pada **sebagian kecil data training secara acak** (random batch) yang disebut mini-batch, bukan pada seluruh data training.

# STOCHASTIC GRADIENT DESCENT (SGD)

Proprietary document of Indonesia AI 2023

---

Theoretically, sebetulnya seperti ini.

- Batch gradient descent = seluruh data training
- Mini-batch gradient descent = beberapa data training (random)
- Stochastic gradient descent = 1 data training

Namun pada praktiknya, penyebutan SGD bisa dipakai untuk ketiga teknik optimizer di atas.

# LANGKAH-LANGKAH PENGUNAAN

Proprietary document of Indonesia AI 2023

- Inisialisasi nilai awal untuk semua bobot pada neural network
- Atur hyperparameter SGD seperti learning rate, epoch, ukuran batch, dan momentum
- Acak data training dan pisahkan menjadi mini-batch
- Untuk setiap mini-batch, lakukan langkah-langkah berikut:
  - Hitung output neural network untuk setiap input data pada mini-batch
  - Hitung gradien of error terhadap bobot pada setiap layer (pakai backpropagation)
  - Update bobot pada setiap layer dengan menggunakan formula SGD



# LANGKAH-LANGKAH PENGUNAAN

Proprietary document of Indonesia AI 2023

- Hitung kesalahan (loss) pada model dengan menggunakan data validasi
- Cetak nilai kesalahan dan akurasi pada setiap epoch
- Ulangi langkah 3 sampai model mencapai konvergensi atau sudah mencapai jumlah epoch yang ditentukan.



# Any question guys ~

# Terimakasih!