

Penjelasan Matematika Model Deep Learning: RNN, LSTM, GRU

1. RNN (Recurrent Neural Network)

Persamaan:

$$h_t = \tanh(W_{hh} \cdot h_{t-1} + W_{xh} \cdot x_t + b_h)$$

$$\hat{y}_t = \sigma(W_{hy} \cdot h_t + b_y)$$

Penjelasan:

- h_t : Hidden state di waktu ke-t
- x_t : Input di waktu ke-t
- W_{xh} : Bobot input ke hidden
- W_{hh} : Bobot hidden ke hidden
- W_{hy} : Bobot hidden ke output
- σ : Fungsi aktivasi sigmoid
- \tanh : Fungsi aktivasi tanh

Masalah umum:

- Vanishing gradient
- Exploding gradient

2. LSTM (Long Short-Term Memory)

Persamaan:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\sim C_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \sim C_t$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t \odot \tanh(C_t)$$

Penjelasan:

- f_t : Forget gate
- i_t : Input gate
- $\sim C_t$: Candidate cell state
- C_t : Cell state
- o_t : Output gate
- h_t : Hidden state

Kelebihan:

- Menyelesaikan masalah vanishing gradient
- Cocok untuk data sequence panjang

3. GRU (Gated Recurrent Unit)

Persamaan:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r)$$

$$\sim h_t = \tanh(W_h \cdot [r_t \odot h_{t-1}, x_t] + b_h)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \sim h_t$$

Penjelasan:

- z_t : Update gate
- r_t : Reset gate
- $\sim h_t$: Candidate hidden state
- h_t : Hidden state

Kelebihan:

- Simpler dibanding LSTM
- Lebih cepat
- Performa hampir setara LSTM di banyak kasus

4. Loss Function

Binary Cross-Entropy Loss:

$$\text{Loss} = -(1/N) \sum [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

Penjelasan:

- y_i : Label asli
- \hat{y}_i : Prediksi model (probabilitas)

5. Tabel Ringkasan

Model	Gate	Memory Cell	Speed	Cocok untuk
RNN		Tidak ada	Cepat	Data sederhana
LSTM	Forget, Input, Output	Ada	Sedang	Data panjang, kompleks
GRU	Update, Reset	Tidak ada	Lebih cepat	Data panjang atau butuh efisiensi