

James F R Tambunan
11323036

D3 Teknologi Informasi

~~100%~~

LSTM (Long Short Term Memory) adalah jenis jaringan RNN (Recurrent Neural Network) yang dirancang untuk mengatasi masalah Vanishing gradient dan mampu mengingat informasi penting dalam jangka panjang.

Rumus dan Implementasi Manual tiap Gate

Misalkan input pada waktu t adalah x_t dan hidden state sebelumnya adalah h_{t-1} .

1. Forget Gate

Menentukan bagian memori lama yang perlu dilupakan

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

- W_f = bobot forget gate

- b_f = bias

- σ = fungsi sigmoid (hasil antara 0-1)

Contoh Manual

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

$$f_t = \sigma (0,6 h_{t-1} + 0,3 x_t + 0,1)$$

Jika $h_{t-1} = 0,5$, $x_t = 0,8$

$$f_t = \sigma (0,6(0,5) + 0,3(0,8) + 0,1) = \sigma (0,77) \approx 0,68$$

Artinya 68% memori lama masih diingat.

2. Input Gate

Mengatur informasi baru yang akan ditambahkan ke cell state.

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

Dan Menghitung kandidat memori baru :

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

Contoh Manual :

$$i_t = \sigma(0,7h_{t-1} + 0,4x_t + 0,2)$$

$$\tilde{C}_t = \tanh(0,5h_{t-1} + 0,3x_t + 0,1)$$

Dengan $h_{t-1} = 0,5$, $x_t = 0,8$:

$$i_t = \sigma(0,7(0,5) + 0,4(0,8) + 0,2) = \sigma(0,9) \approx 0,71$$

$$\tilde{C}_t = \tanh(0,5(0,5) + 0,3(0,8) + 0,1) = \tanh(0,69) \approx 0,60$$

3. Cell State Update

Menggabungkan hasil forget gate dan input gate.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Jika $C_{t-1} = 0,4$:

$$C_t = (0,68)(0,4) + (0,71)(0,60) = 0,27 + 0,43 = 0,70$$

Jadi memori baru sebesar 0,70.

4. Output Gate

Menentukan informasi apa yang akan menjadi output berikutnya.

$$O_t = \sigma(W_o \cdot [h_t])$$

Dan hidden state baru dihitung sebagai:

$$h_t = O_t * \tanh(C_t)$$

Contoh manual :

$$O_t = \sigma(0,8h_{t-1} + 0,2x_t + 0,1) = \sigma(0,8(0,5) + 0,2(0,8) + 0,1) = \sigma(0,66) = 0,66$$

$$h_t = 0,66 * \tanh(0,70) = 0,66 * 0,604 = 0,40$$

Jadi, proses LSTM ini membuat jaringan mengingat sebagian informasi lama, menambah informasi baru, sehingga menghasilkan output yang lebih relevan berdasarkan konteks waktu tersebut.