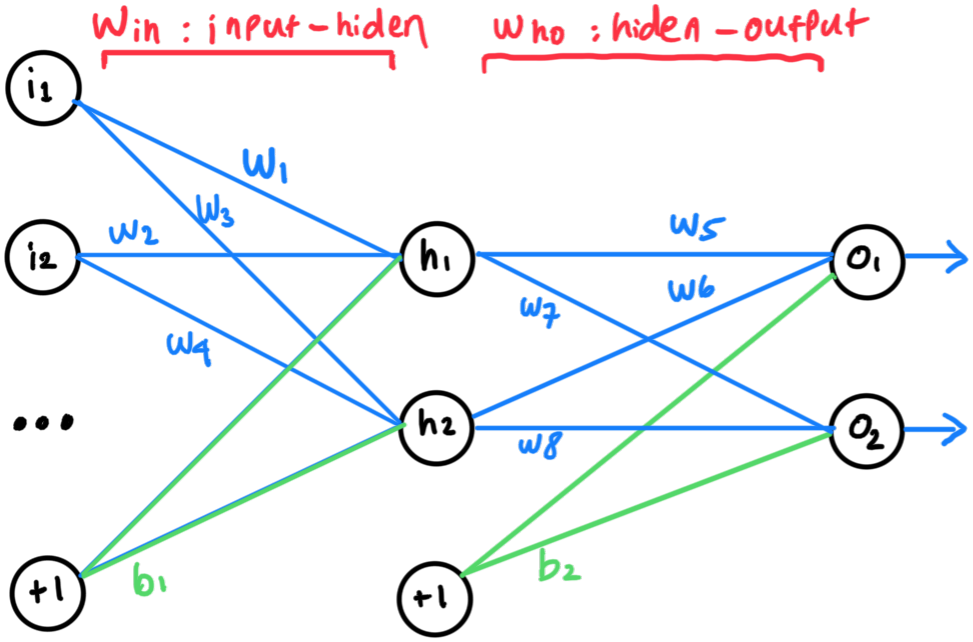
Neural Network From Scratch



Langkah dalam back propagation

- Forward Propagation
- Backward Propagation
- menggabung semua nilai kemudian memperbaharui weight

Langkah 1 - Forward Propagation

> menghitung net input h_1

$net\ h_1 = W_1 . i_1 + W_2 . i_2 + b_1 . 1$

> menghitung output h1 dengan mengenai fungsi aktivasi

$f(net\ h_1) = 1/1+e^{-net\ h1}$

> menghitung net input h_2

$net\ h_2 = W_3 . i_1 + W_4 . i_2 + b_1 . 1$

> menghitung output h_2 dengan mengenai fungsi aktivasi

$f(net\ h_2) = 1/1+e^{-net\ h_2}$

> menghitung net O_1

$net\ O_1 = W_5 . h_1 + W_6 . h_2 + b_2 . 1$

> menghitung output O_1 dengan mengenai fungsi aktivasi

$f(net\ O_1) = 1/1+e^{-net\ O_1}$

> menghitung net O_2

$net\ O_2 = w_7 . h_1 + w_8 . h_2 + b_2 . 1$

> menghitung output O_2 dengan mengenai fungsi aktivasi

$f(net\ O_2) = 1/1+e^{-net\ O_2}$

$$\begin{bmatrix} net\ h_1 \\ net\ h_2 \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} + \begin{bmatrix} W_1 & W_2 \\ W_3 & W_4 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} \quad \ldots (1)$$

net h b_1 W_{in} X

$$\begin{bmatrix} out\ h_1 \\ out\ h_2 \end{bmatrix} = \begin{bmatrix} f(net\ h_1) \\ f(net\ h_2) \end{bmatrix} \quad \ldots (2)$$

$$\begin{bmatrix} net\ O_1 \\ net\ O_2 \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} + \begin{bmatrix} W_5 & W_6 \\ W_7 & W_8 \end{bmatrix} \begin{bmatrix} out\ h_1 \\ out\ h_2 \end{bmatrix} \quad \ldots (3)$$

net O b_2 W_{HO} out h

$$\begin{bmatrix} out\ O_1 \\ out\ O_2 \end{bmatrix} = \begin{bmatrix} f(net\ O_1) \\ f(net\ O_2) \end{bmatrix} \quad \ldots (4)$$

Menghitung nilai error

Squared error func

$$\begin{bmatrix} E_{O1} \\ E_{O2} \end{bmatrix} = \frac{1}{2} \left(\begin{bmatrix} out\ O_1 \\ out\ O_2 \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right) \quad \ldots (5)$$

E_o out O target

$$E_{total} = sum(E_o)$$

Langkah 2 - Backward Propagation

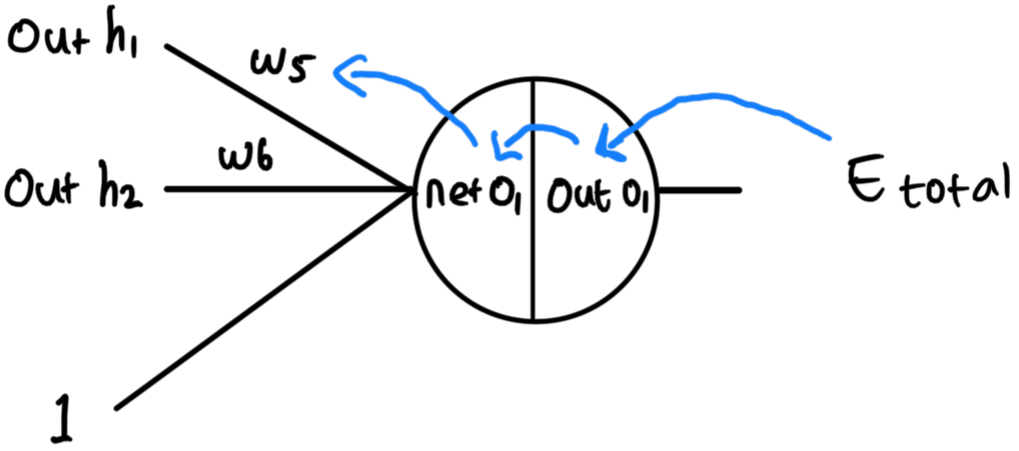
Goals : melakukan update parameter (weight dan bias) agar output yg dihasilkan NN mendekati target, dengan cara meminimalisir nilai error

Update Parameter diantara output - hidden (W_{ho} dan b_2)

karena backward maka update dimulai pada weight Who [[w_5, w_6], [w_7, w_8]]

Contoh update pada w_5. menghitung total error untuk setiap perubahan w_5, maka

$$\frac{\partial E_{total}}{\partial w_5} = \frac{\partial E_{total}}{\partial out_{o_1}} * \frac{\partial out_{o_1}}{\partial net_{o_1}} * \frac{\partial net_{o_1}}{\partial w_5} \quad \ldots (6)$$



Jabarkan suku pertama Pada Pers. (6)

$$E_{total} = \frac{1}{2}(y_1 - Out\ o_1)^2 + \frac{1}{2}(y_2 - Out\ o_2)^2$$

$$\frac{\partial E_{total}}{\partial Out\ o_1} = -(y_1 - Out\ o_1) \quad \ldots \quad (7)$$

Jabarkan suku kedua Pada Pers. (6)

$$Out\ o_1 = \frac{1}{1 + e^{-net\ o_1}}$$

$$\frac{\partial Out\ o_1}{\partial net\ o_1} = Out\ o_1\ (1 - Out\ o_1) \quad \ldots \quad (8)$$

Jabarkan suku ketiga pada pers. (6)

$$net\,o_1 = w_5 * Out\,h_1 + w_6 * Out\,h_2 + b_2 * 1$$

$$\frac{\partial net\,o_1}{\partial w_5} = Out\,h_1 \quad \ldots (9)$$

masukan pers (7), (8), dan (9) ke pers (6)

$$\frac{\partial E\,total}{\partial w_5} = -(y_1 - Out\,o_1) * Out\,o_1 (1 - Out\,o_1) * Out\,h_1 \quad \ldots (10)$$

atau dapat ditulis

$$\frac{\partial E\,total}{\partial w_5} = \delta o_1 * Out\,h_1 \quad \ldots (11)$$

Update weight dengan persamaan

$$w_5^{+} = w_5 - \alpha \frac{\partial E\,total}{\partial w_5}$$

$$w_5^{+} = w_5 - \alpha\, \delta o_1 * Out\,h_1$$

$$\begin{bmatrix} w_5^{+} & w_6^{+} \\ w_7^{+} & w_8^{+} \end{bmatrix} = \begin{bmatrix} w_5 & w_6 \\ w_7 & w_8 \end{bmatrix} - \alpha \begin{bmatrix} \delta o_1 \\ \delta o_2 \end{bmatrix} * \begin{bmatrix} Out\,h_1 & Out\,h_2 \end{bmatrix} \quad (12)$$

Weight baru — weight lama — learning Rate

Untuk Update Bias

$$\frac{\partial E_{total}}{\partial b_2} = \frac{\partial E_{total}}{\partial out_{O1}} * \frac{\partial out_{O1}}{\partial net_{O1}} * \frac{\partial net_{O1}}{\partial b2} \quad \ldots (13)$$

Pada suku Ketiga pers (13)

$$\frac{\partial net_{O1}}{\partial b_2} = \frac{\partial \left(w_5 * out_{h1} + w_6 * out_{h2} + b_2 * 1 \right)}{\partial b_2}$$

$$\frac{\partial net_{O1}}{\partial b_2} = 1$$

Sehingga, Perubahan error untuk tiap perubahan bias adalah,

$$\frac{\partial E_{total}}{\partial b_2} = \frac{\partial E_{total}}{\partial out_{O1}} * \frac{\partial out_{O1}}{\partial net_{O1}}$$

$$\frac{\partial E_{total}}{\partial b_2} = \delta O_1$$

Update bias dengan persamaan

$$b_2^{+} = b_2 - \alpha \frac{\partial E_{total}}{\partial b_2}$$

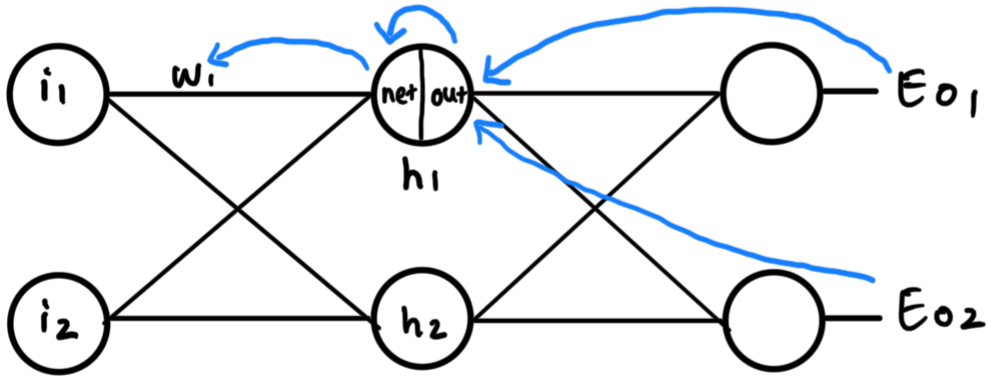
$$b_2^{+} = b_2 - \alpha * \delta O_1$$

$$\begin{bmatrix} b_{11}^{+} \\ b_{21}^{+} \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{12} \end{bmatrix} - \alpha \begin{bmatrix} \delta O_1 \\ \delta O_2 \end{bmatrix} \quad \ldots (14)$$

Update Parameter diantara hidden - input (W_{ih} dan b_r)

Selanjutnya akan dilakukan update untuk parameter
wih [[w1, w2], [w3, w4]]
Contoh update pada w1

$$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial out_{h_1}} * \frac{\partial out_{h_1}}{\partial net_{h_1}} * \frac{\partial net_{h_1}}{\partial w_1} \quad \ldots (15)$$



$$\frac{\partial E_{total}}{\partial out_{h_1}} = \frac{\partial E_{o1}}{\partial out_{h_1}} + \frac{\partial E_{o2}}{\partial out_{h_1}} \quad \ldots (16)$$

$$\frac{\partial E_{o1}}{\partial out_{h_1}} = \frac{\partial E_{o1}}{\partial net_{o1}} * \frac{\partial net_{o_1}}{\partial out_{h_1}}$$

$$\frac{\partial E_{o1}}{\partial net_{o1}} \rightarrow \delta o_1$$

$$\frac{\partial E_{o1}}{\partial out_{h_1}} = \delta o_1 * \frac{\partial net_{o_1}}{\partial out_{h_1}}$$

$$\frac{\partial E_{o1}}{\partial out_{h_1}} = \delta o_1 * \frac{\partial (w_5 * out_{h_1} + w_6 * out_{h_2} + b_2 * 1)}{\partial out_{h_1}}$$

$$\frac{\partial E_{o1}}{\partial out_{h_1}} = \delta o_1 * w_5$$

$$\frac{\partial E_{total}}{\partial out_h} = \delta o_1 * W_{ho} \quad \ldots (17)$$

Suku ke dua persamaan (15)

$$Out\,h_1 = \frac{1}{1 + e^{-net\,h_1}}$$

$$\frac{\partial out\,h_1}{\partial net\,h_1} = out\,h_1\,(1 - out\,h_1) \quad \ldots (18)$$

Suku ke tiga Persamaan (15)

$$net\,h_1 = w_1 * i_1 + w_3 * i_2 + b_1 * 1$$

$$\frac{\partial net\,h_1}{\partial w_1} = i_1 \quad \ldots (19)$$

masukan Pers. (17), (18), (19) ke Pers. (15)

$$\frac{\partial E\,total}{\partial w_1} = \underbrace{\delta o_1 * w_5 * out\,h_1\,(1 - out\,h_1)}_{\delta h_1} * i_1$$

$$\frac{\partial E\,total}{\partial w_1} = \delta h_1 * i_1 \quad \ldots (20)$$

Sekarang update w_1

$$w_1^{+} = w_1 - \alpha * \frac{\partial E\,total}{\partial w_1}$$

$$w_1^{+} = w_1 - \alpha * \delta h_1 * i_1$$

$$\begin{bmatrix} w_1^{+} & w_2^{+} \\ w_3^{+} & w_4^{+} \end{bmatrix} = \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} - \alpha \begin{bmatrix} \delta h_1 \\ \delta h_2 \end{bmatrix} * \begin{bmatrix} i_1 & i_2 \end{bmatrix} \ldots (21)$$

Weight baru — weight lama — learning Rate

untuk bias, serupa dengan penjabaran sebelumnya $\frac{\partial net\,h_1}{\partial b}$ akan

$= 1$ Sehingga $\frac{\partial E_{total}}{\partial b_2} = \delta h_1$

Update bias dengan persamaan

$$b_{12}^{+} = b_{12} - \alpha \frac{\partial E_{total}}{\partial b_{12}}$$

$$b_{12}^{+} = b_{12} - \alpha * \delta h_1$$

$$\begin{bmatrix} b_{11}^{+} \\ b_{21}^{+} \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{12} \end{bmatrix} - \alpha \begin{bmatrix} \delta h_1 \\ \delta h_2 \end{bmatrix} \quad \ldots (22)$$

b_2^{+} $\quad$ b_2

Semua parameter sudah terupdate, kemudian semua perhitungan di atas akan diiterasikan hingga didapatken error yang minimum

Kevin Boy Gunawan