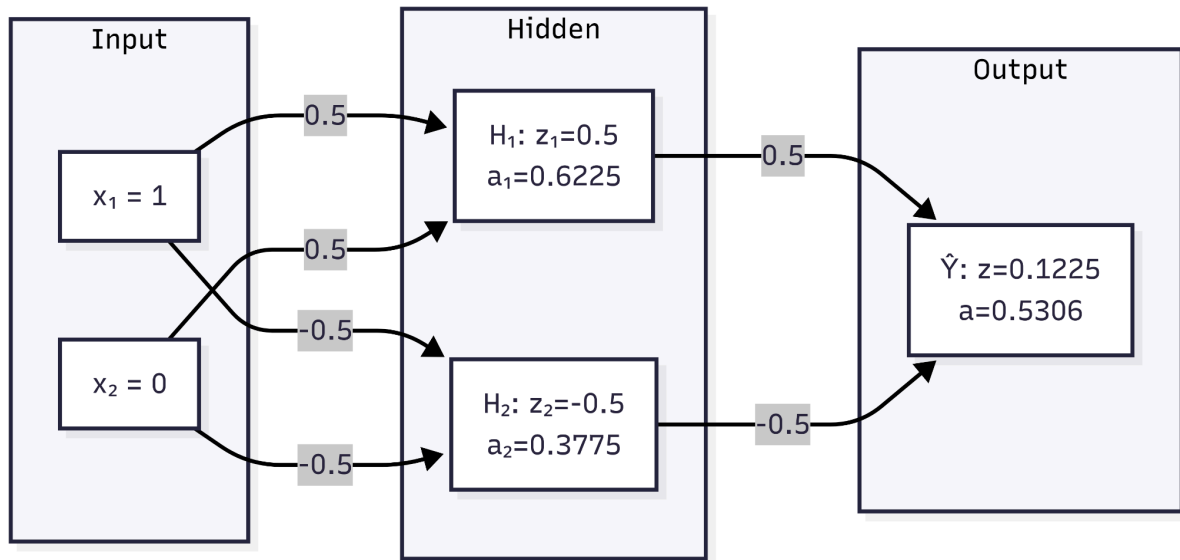


Diagram Arsitektur



1. Inisialisasi

Kita mulai dengan menetapkan parameter jaringan dan fungsi aktivasi. Ini adalah dasar sebelum training dimulai:

- **Bobot** input→hidden, dua neuron tersembunyi:

$$W^{(1)} = \begin{pmatrix} 0.5 & -0.5 \\ 0.5 & -0.5 \end{pmatrix}$$

- **Bias** hidden layer:
 $b^{(1)} = (0, 0)$
- **Bobot** hidden→output, satu neuron output:

$$W^{(2)} = (0.5, -0.5)$$

- **Bias** output layer:
 $b^{(2)} = 0$
- **Learning rate** untuk pembaruan bobot:
 $\eta = 0.1$

- **Fungsi aktivasi** sigmoid dan turunannya:

$$\sigma(z) = \frac{1}{1 + e^{-z}}, \sigma'(z) = \sigma(z)(1 - \sigma(z))$$

- **Data sampel** yang kita gunakan: input $(x_1, x_2) = (1, 0)$ dengan target $y = 1$.

2. Forward Pass

Di tahap ini kita hitung output prediksi dari jaringan:

1. Pre-aktivasi Hidden Unit 1

Masukkan (1,0)(1,0) ke bobot kolom pertama:

$$z_1^{(1)} = 1 \cdot 0.5 + 0 \cdot 0.5 + 0 = 0.5$$

2. Aktivasi Hidden Unit 1

Terapkan sigmoid untuk mendapatkan aktivasi:

$$a_1^{(1)} = \sigma(0.5) \approx 0.6225$$

3. Pre-aktivasi Hidden Unit 2

Masukkan (1,0)(1,0) ke bobot kolom kedua:

$$z_2^{(1)} = 1 \cdot (-0.5) + 0 \cdot (-0.5) + 0 = -0.5$$

4. Aktivasi Hidden Unit 2

$$a_2^{(1)} = \sigma(-0.5) \approx 0.3775$$

5. Pre-aktivasi Output

Gabungkan aktivasi hidden dan bobot output:

$$z^{(2)} = 0.6225 \cdot 0.5 + 0.3775 \cdot (-0.5) + 0 = 0.1225$$

6. Aktivasi Output (Prediksi)

$$a^{(2)} = \sigma(0.1225) \approx 0.5306$$

7. Hitung Loss (MSE)

Loss mengukur seberapa jauh prediksi kita dari target:

$$L = \frac{1}{2}, (y - a^{(2)})^2 = \frac{1}{2}, (1 - 0.5306)^2 \approx 0.1091$$

3. Backward Pass

3.1. Gradien di Output

Selisih prediksi dan target, dikali turunan sigmoid:

$$\delta^{(2)} = (a^{(2)} - y), \sigma'(z^{(2)}) = (0.5306 - 1) \times (0.5306 \cdot 0.4694) \approx -0.1170$$

3.2. Gradien Bobot Output

Gradien untuk masing-masing koneksi hidden→output:

$$\frac{\partial L}{\partial W^{(2)}_1} = a_1^{(1)}, \delta^{(2)} = 0.6225 \times (-0.1170) \approx -0.0728$$

$$\frac{\partial L}{\partial W^{(2)}_2} = a_2^{(1)}, \delta^{(2)} = 0.3775 \times (-0.1170) \approx -0.0441$$

3.3. Gradien Hidden Layer

Error dari output disebar ke masing-masing neuron hidden:

$$\delta_1^{(1)} = \delta^{(2)}, W^{(2)}_1, \sigma'(z_1^{(1)}) = -0.1170 \times 0.5 \times (0.6225 \cdot 0.3775) \approx -0.0138$$

$$\delta_2^{(1)} = \delta^{(2)}, W^{(2)}_2, \sigma'(z_2^{(1)}) = -0.1170 \times (-0.5) \times (0.3775 \cdot 0.6225) \approx +0.0138$$

3.4. Gradien Bobot Input→Hidden

Karena input kedua adalah nol, hanya bobot dari x1 yang berubah:

$$\frac{\partial L}{\partial W^{(1)}_{11}} = x_1, \delta_1^{(1)} = 1 \times (-0.0138) = -0.0138$$

$$\frac{\partial L}{\partial W^{(1)}_{12}} = x_1, \delta_2^{(1)} = 1 \times (+0.0138) = +0.0138$$

4. Update Bobot dan Bias (Learning Rate = 0.1)

4.1. Bobot output

$$W_1^{(2)} \leftarrow 0.5 - 0.1 \times (-0.0728) = 0.5073$$

$$W_2^{(2)} \leftarrow -0.5 - 0.1 \times (-0.0441) = -0.4956$$

4.2. Bias output

$$b^{(2)} \leftarrow 0 - 0.1 \times (-0.1170) = +0.0117$$

4.3. Bobot hidden

$$W_{11}^{(1)} \leftarrow 0.5 - 0.1 \times (-0.0138) = 0.5014$$

$$W_{12}^{(1)} \leftarrow -0.5 - 0.1 \times (+0.0138) = -0.5014$$

(bobot dari input 2 tetap)

4.4. Bias hidden

$$b_1^{(1)} \leftarrow 0 - 0.1 \times (-0.0138) = +0.0014$$

$$b_2^{(1)} \leftarrow 0 - 0.1 \times (+0.0138) = -0.0014$$