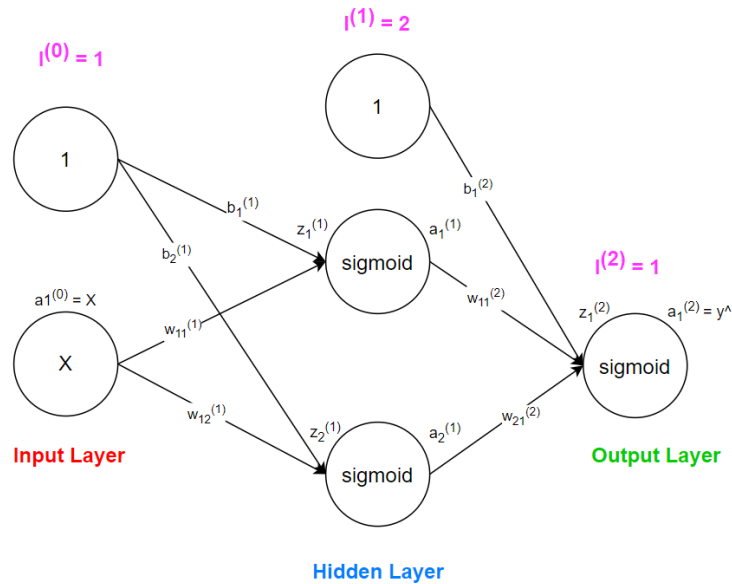


Neural network calculates by hand

Author : Nguyen Van Manh

Sơ đồ lớp



Tập dữ liệu đầu vào : Chuẩn hóa x và y bằng phương pháp Min-Max Scaling hoặc Standardization (z-score normalization)

nth	H(x)	W(y)	$x_{scaling}$	$y_{scaling}$
0	160	62	0	0
1	165	64	0.5	0.5
2	170	66	1	1

⇒ Ta thấy x và y quá lớn so với trọng số w chính vì thế khi tính ra z rồi cho qua sigmoid thì a gần như bằng 1, điều này dẫn đến kết quả sau đạo hàm sẽ có đoạn $a * (1 - a) = 0$

Dẫn đến việc không thể cập nhật trọng số. Nên chuẩn hóa cho x, y nhỏ lại.

$$x_{min} = 160; x_{max} = 170; y_{min} = 62; y_{max} = 66$$

$$x_{scaling} = \frac{x - x_{min}}{x_{max} - x_{min}}; y_{scaling} = \frac{y - y_{min}}{y_{max} - y_{min}}$$

Khởi tạo trọng số W, b

$w_{11}^{(1)}$	$w_{12}^{(1)}$	$b_1^{(1)}$	$b_{12}^{(1)}$	$w_{11}^{(2)}$	$w_{21}^{(2)}$	$b_1^{(2)}$
0.5	1	0	0	1	0.5	1

Learning rate : 0.1

Activation function : $\sigma(x) = \frac{1}{1 + e^{-x}}$ (Sigmoid)

Loss function : $L = -(y_i * \log_e(\hat{y}_i) + (1 - y_i) * \log_e(1 - \hat{y}_i))$ (Log loss – Cross Entropy)

$$J = \frac{1}{N} * \sum_{i=1}^N L = -\frac{1}{N} * \sum_{i=1}^N x_m^{(i)} * (y_i * \log_e(\hat{y}_i) + (1 - y_i) * \log_e(1 - \hat{y}_i))$$

Tóm tắt quá trình

- Lặp qua epochs lần để tối ưu W và b
- Lặp qua n bảng ghi có trong tập dữ liệu

- Qua mỗi bảng ghi
 - Khởi tạo W, b cho các lớp fully connected kết nối giữa các layer
 - Thực hiện quá trình Feedforward để tính z, a của mỗi layer và cuối cùng được y
 - Thực hiện quá trình Backpropagation : Mục đích là để cập nhật lại các W, b, hay nói cách khác là tối ưu W, b
 - Lưu ý : Ta đã biết ở các bài trước để tối ưu một giá trị bằng gradient sẽ thực hiện .
Ex : $w = w - \text{learning_rate} * \frac{dL}{dw}$ với L là một hàm mất mát
 - Vậy có thể hiểu đơn giản là tính $\frac{dL}{dw}$ và $\frac{dL}{db}$ sau đó cập nhật lại W và b .
- Sau n bảng ghi ta thu được ma trận W, b và các ma trận liên quan khác , từ đây ta sẽ tiến hành tính toán độ mất mát cho toàn tập dữ liệu để biết được có dừng lại hay không với mỗi ma trận W và b mới .

Thực hiện quá trình tính toán trên bảng ghi thứ hai của tập dữ liệu

- Tương tự với các bảng ghi khác
- Lặp qua số epochs lần cũng tương tự

Input : x = 0.5 , y = 0.5

Feedforward

$$z_1^{(1)} = x * w_{11}^{(1)} + b_1^{(1)} = 0.5 * 0.5 + 0 = 0.25$$

$$z_2^{(1)} = x * w_{12}^{(1)} + b_2^{(1)} = 0.5 * 1 + 0 = 0.5$$

$$a_1^{(1)} = \sigma(z_1^{(1)}) = \frac{1}{1 + e^{-z_1^{(1)}}} \approx 0.562$$

$$a_2^{(1)} = \sigma(z_2^{(1)}) = \frac{1}{1 + e^{-z_2^{(1)}}} \approx 0.622$$

$$z_1^{(2)} = a_1^{(1)} * w_{11}^{(2)} + a_2^{(1)} * w_{21}^{(2)} + b_1^{(2)} = 0.562 * 1 + 0.622 * 0.5 + 1 = 1.873$$

$$a_1^{(2)} = \hat{y}_1 = \sigma(z_1^{(2)}) = \frac{1}{1 + e^{-z_1^{(2)}}} \approx 0.867$$

Backpropagation (Xem đạo hàm chi tiết ở cuối file)

Ta có $\hat{y}_1 = \sigma(z_1^{(2)}) = \sigma(a_1^{(1)} * w_{11}^{(2)} + a_2^{(1)} * w_{21}^{(2)} + b_1^{(2)})$; $L = -(y_i * \log_e(\hat{y}_i) + (1 - y_i) * \log_e(1 - \hat{y}_i))$

$$\frac{\partial L}{\partial b_1^{(2)}} = \frac{\partial L}{\partial \hat{y}_1} * \frac{\partial \hat{y}_1}{\partial z_1^{(2)}} * \frac{\partial z_1^{(2)}}{\partial b_1^{(2)}} = \left(\frac{\hat{y}_1 - y_1}{\hat{y}_1 * (1 - \hat{y}_1)} \right) * (\hat{y}_1 * (1 - \hat{y}_1)) * 1 = (\hat{y}_1 - y_1) = 0.867 - 0.5 \approx \mathbf{0.367}$$

$$\frac{\partial L}{\partial w_{11}^{(2)}} = \frac{\partial L}{\partial \hat{y}_1} * \frac{\partial \hat{y}_1}{\partial z_1^{(2)}} * \frac{\partial z_1^{(2)}}{\partial w_{11}^{(2)}} = \left(\frac{\hat{y}_1 - y_1}{\hat{y}_1 * (1 - \hat{y}_1)} \right) * (\hat{y}_1 * (1 - \hat{y}_1)) * (a_1^{(1)}) = (\hat{y}_1 - y_1) * a_1^{(1)} = 0.367 * 0.562 \approx \mathbf{0.206}$$

$$\frac{\partial L}{\partial w_{21}^{(2)}} = \frac{\partial L}{\partial \hat{y}_1} * \frac{\partial \hat{y}_1}{\partial z_1^{(2)}} * \frac{\partial z_1^{(2)}}{\partial w_{21}^{(2)}} = \left(\frac{\hat{y}_1 - y_1}{\hat{y}_1 * (1 - \hat{y}_1)} \right) * (\hat{y}_1 * (1 - \hat{y}_1)) * (a_2^{(1)}) = (\hat{y}_1 - y_1) * a_2^{(1)} = 0.367 * 0.622 \approx \mathbf{0.228}$$

Ta có : Cứ tính lùi theo như các biểu thức trên Feedforward

$$\begin{aligned} \frac{\partial L}{\partial b_1^{(1)}} &= \frac{\partial L}{\partial \hat{y}_1} * \frac{\partial \hat{y}_1}{\partial z_1^{(2)}} * \frac{\partial z_1^{(2)}}{\partial a_1^{(1)}} * \frac{\partial a_1^{(1)}}{\partial z_1^{(1)}} * \frac{\partial z_1^{(1)}}{\partial b_1^{(1)}} = \left(\frac{\hat{y}_1 - y_1}{\hat{y}_1 * (1 - \hat{y}_1)} \right) * (\hat{y}_1 * (1 - \hat{y}_1)) * (w_{11}^{(2)}) * (a_1^{(1)} (1 - a_1^{(1)})) * 1 \\ &= (\hat{y}_1 - y_1) * w_{11}^{(2)} * (a_1^{(1)} (1 - a_1^{(1)})) = 0.367 * 1 * 0.562 * (1 - 0.562) = \mathbf{0.09} \end{aligned}$$

$$\frac{\partial L}{\partial b_2^{(1)}} = \frac{\partial L}{\partial \hat{y}_1} * \frac{\partial \hat{y}_1}{\partial z_1^{(2)}} * \frac{\partial z_1^{(2)}}{\partial a_2^{(1)}} * \frac{\partial a_2^{(1)}}{\partial z_2^{(1)}} * \frac{\partial z_2^{(1)}}{\partial b_2^{(1)}} = \left(\frac{\hat{y}_1 - y_1}{\hat{y}_1 * (1 - \hat{y}_1)} \right) * (\hat{y}_1 * (1 - \hat{y}_1)) * (w_{21}^{(2)}) * (a_2^{(1)} (1 - a_2^{(1)})) * 1$$

$$= (\hat{y}_1 - y_1) * w_{21}^{(2)} * (a_2^{(1)} (1 - a_2^{(1)})) = 0.367 * 0.5 * 0.622 * (1 - 0.622) = \mathbf{0.043}$$

$$\frac{\partial L}{\partial w_{11}^{(1)}} = \frac{\partial L}{\partial \hat{y}_1} * \frac{\partial \hat{y}_1}{\partial z_1^{(2)}} * \frac{\partial z_1^{(2)}}{\partial a_1^{(1)}} * \frac{\partial a_1^{(1)}}{\partial w_{11}^{(1)}} * \frac{\partial z_1^{(1)}}{\partial b_1^{(1)}} = \left(\frac{\hat{y}_1 - y_1}{\hat{y}_1 * (1 - \hat{y}_1)} \right) * (\hat{y}_1 * (1 - \hat{y}_1)) * (w_{11}^{(2)}) * (a_1^{(1)} (1 - a_1^{(1)})) * x$$

$$= (\hat{y}_1 - y_1) * w_{11}^{(2)} * (a_1^{(1)} (1 - a_1^{(1)})) * x = 0.367 * 1 * 0.562 * (1 - 0.562) * 0.5 = \mathbf{0.045}$$

$$\frac{\partial L}{\partial w_{12}^{(1)}} = \frac{\partial L}{\partial \hat{y}_1} * \frac{\partial \hat{y}_1}{\partial z_1^{(2)}} * \frac{\partial z_1^{(2)}}{\partial a_2^{(1)}} * \frac{\partial a_2^{(1)}}{\partial z_2^{(1)}} * \frac{\partial z_2^{(1)}}{\partial w_{12}^{(1)}} = \left(\frac{\hat{y}_1 - y_1}{\hat{y}_1 * (1 - \hat{y}_1)} \right) * (\hat{y}_1 * (1 - \hat{y}_1)) * (w_{21}^{(2)}) * (a_2^{(1)} (1 - a_2^{(1)})) * x$$

$$= (\hat{y}_1 - y_1) * w_{21}^{(2)} * (a_2^{(1)} (1 - a_2^{(1)})) * x = 0.367 * 0.5 * 0.622 * (1 - 0.622) * 0.5 = \mathbf{0.0215}$$

Áp dụng Gradient descent để cập nhật trọng số

$$\frac{\partial L}{\partial b_1^{(2)}} = 0.367 ; \frac{\partial L}{\partial w_{11}^{(2)}} = 0.206 ; \frac{\partial L}{\partial w_{21}^{(2)}} = 0.228 ; \frac{\partial L}{\partial b_1^{(1)}} = 0.09 ; \frac{\partial L}{\partial b_2^{(1)}} = 0.043 ; \frac{\partial L}{\partial w_{11}^{(1)}} = 0.045 ; \frac{\partial L}{\partial w_{12}^{(1)}} = 0.0215$$

$$w_{11}^{(1)} = w_{11}^{(1)} - learning_rate * \frac{\partial L}{\partial w_{11}^{(1)}} = 0.5 - 0.1 * 0.045 = 0.4955$$

$$w_{12}^{(1)} = w_{12}^{(1)} - learning_rate * \frac{\partial L}{\partial w_{12}^{(1)}} = 1 - 0.1 * 0.0215 = 0.99785$$

$$b_1^{(1)} = b_1^{(1)} - learning_rate * \frac{\partial L}{\partial b_1^{(1)}} = 0 - 0.1 * 0.09 = -0.009$$

$$b_2^{(1)} = b_2^{(1)} - learning_rate * \frac{\partial L}{\partial b_2^{(1)}} = 0 - 0.1 * 0.043 = -0.0043$$

$$w_{11}^{(2)} = w_{11}^{(2)} - learning_rate * \frac{\partial L}{\partial w_{11}^{(2)}} = 1 - 0.1 * 0.206 = 0.9794$$

$$w_{21}^{(2)} = w_{21}^{(2)} - learning_rate * \frac{\partial L}{\partial w_{21}^{(2)}} = 0.5 - 0.1 * 0.228 = 0.4772$$

$$b_1^{(2)} = b_1^{(2)} - learning_rate * \frac{\partial L}{\partial b_1^{(2)}} = 1 - 0.1 * 0.367 = 0.9633$$

Sau epochs thứ nhất

$w_{11}^{(1)}$	$w_{12}^{(1)}$	$b_1^{(1)}$	$b_2^{(1)}$	$w_{11}^{(2)}$	$w_{21}^{(2)}$	$b_1^{(2)}$
0.4955	0.99785	-0.009	-0.0043	0.9794	0.4772	0.9633

Áp dụng tương tự với các bảng ghi còn lại của tập dữ liệu .