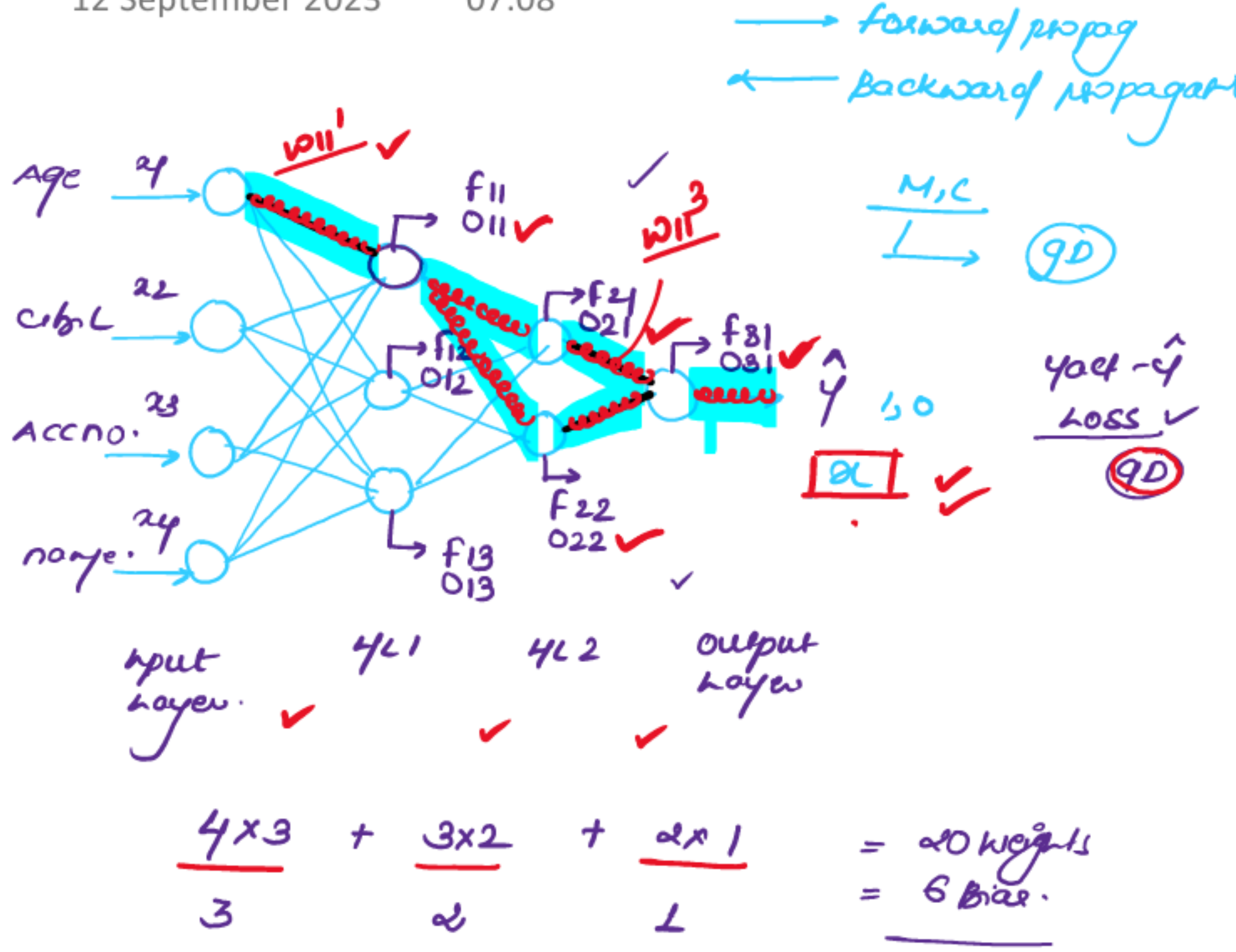


Chain Rule

12 September 2023

07:08



$$\frac{4 \times 3}{3} + \frac{3 \times 2}{2} + \frac{2 \times 1}{1} = 20 \text{ weights}$$

$$= 6 \text{ bias}$$

Forward propagation:
Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss

Backward propagation:
Loss \leftarrow Update of W, B

Matrix 1

$$\begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \\ w_{41} & w_{42} & w_{43} \end{bmatrix} \rightarrow w_{11}'$$

$$w_{11}' = w_{11} - \alpha \frac{\partial L}{\partial w_{11}}$$

Matrix 2

$$\begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \\ w_{31} & w_{32} \end{bmatrix} \rightarrow w_{11}''$$

$$w_{11}'' = w_{11}' - \alpha \frac{\partial L}{\partial w_{11}'}$$

Matrix 3

$$\begin{bmatrix} w_{11} \\ w_{21} \end{bmatrix} \rightarrow w_{11}'''$$

$$w_{11}''' = w_{11}'' - \alpha \frac{\partial L}{\partial w_{11}''}$$

$$\frac{\partial L}{\partial w_{11}'} = \left[\frac{\partial L}{\partial z_{11}} \wedge \frac{\partial z_{11}}{\partial w_{11}'} \right]$$

$$+ \left[\frac{\partial L}{\partial z_{21}} \wedge \frac{\partial z_{21}}{\partial w_{11}'} \right]$$

Randomly Forward propagation, Backward propagation

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B

Data \rightarrow NN \rightarrow \hat{y} \rightarrow Loss \leftarrow W, B