Training a CBOW Model: Forward Propagation

Forward propagation is defined as:

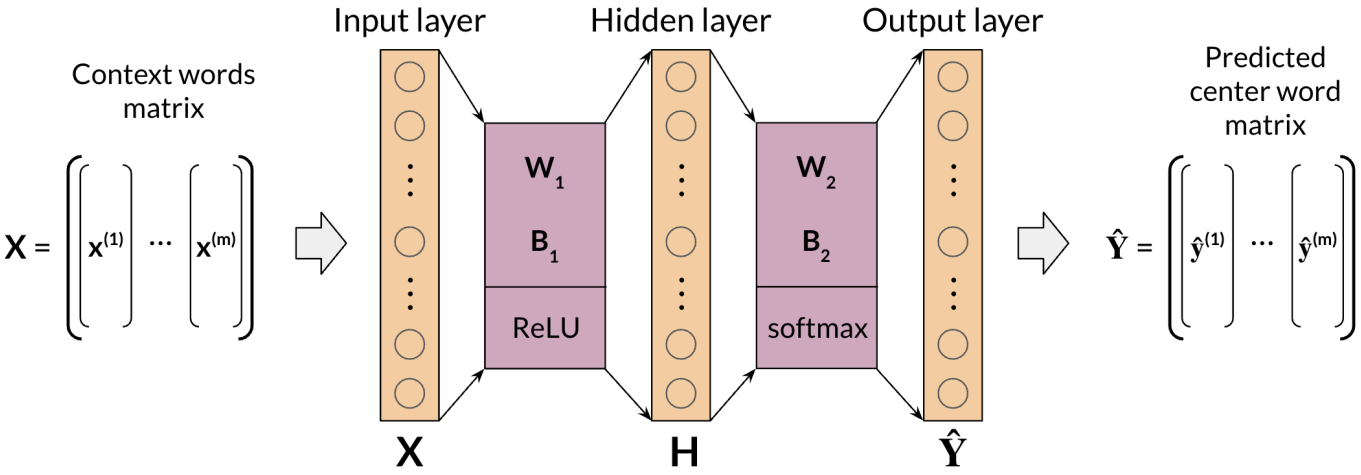
$$Z_1 = W_1 X + B_1$$

$$H = ReLU(Z_1)$$

$$Z_2 = W_2 H + B_2$$

$$\hat{Y} = softmax(Z_2)$$

In the image below you start from the left and you forward propagate all the way to the right.



To calculate the loss of a batch, you have to compute the following:

$$J_{batch} = -\frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{V} y_j^{(i)} \log \hat{y}_j^{(i)}$$

Given, your predicted center word matrix, and actual center word matrix, you can compute the loss.

✓ **Completed**