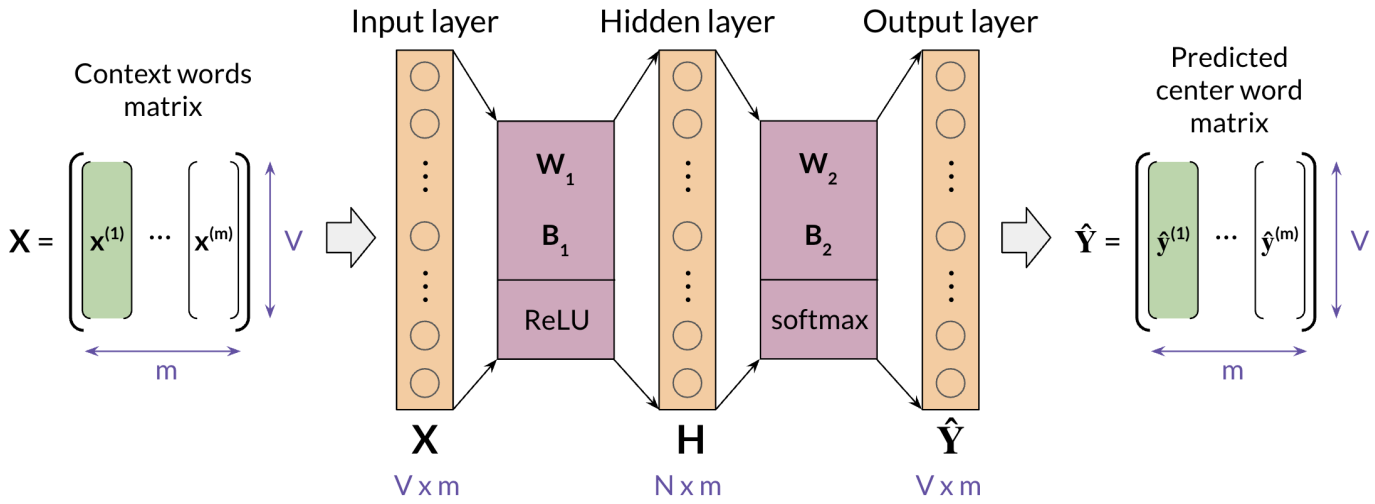
Architecture of the CBOW Model: Dimensions

When dealing with batch input, you can stack the examples as columns. You can then proceed to multiply the matrices as follows:



In the diagram above, you can see the dimensions of each matrix. Note that your $\hat{Y}$ is of dimension V by m. Each column is the prediction of the column corresponding to the context words. So the first column in $\hat{Y}$ is the prediction corresponding to the first column of *X*.