Hide menu

Lecture: Word Embeddings

- **Video:** Week Introduction 1 min
- **Video:** Overview 2 min
- **Reading:** Overview
- **Video:** Basic Word Representations 3 min
- **Reading:** Basic Word Representations 5 min
- **Video:** Word Embeddings 3 min
- **Reading:** Word Embeddings 4 min
- **Video:** How to Create Word Embeddings 3 min
- **Reading:** How to Create Word Embeddings? 4 min
- **Video:** Word Embedding Methods 3 min
- **Reading:** Word Embedding Methods 4 min
- **Video:** Continuous Bag-of-Words Model 4 min
- **Reading:** Continuous Bag of Words Model 3 min
- **Video:** Cleaning and Tokenization 4 min
- **Reading:** Cleaning and Tokenization 5 min
- **Video:** Sliding Window of Words in Python 3 min
- **Reading:** Sliding Window of words in Python 10 min
- **Video:** Transforming Words into Vectors 3 min
- **Reading:** Transforming Words into Vectors 2 min
- **Lab:** Lecture Notebook - Data Preparation 30 min
- **Video:** Architecture of the CBOW Model 3 min
- **Reading:** Architecture for the CBOW Model 4 min

Week 4 > Training a CBOW Model: Forward Propagation

Previous Next

Training a CBOW Model: Forward Propagation

Forward propagation is defined as:

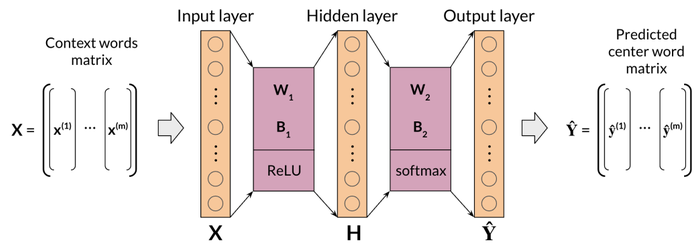
$$Z1 = W1X + B1$$

$$H = ReLU(Z1)$$

$$Z2 = W2H + B2$$

$$\hat{Y} = softmax(Z2)$$

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:

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

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

Predicted center word matrix

$$\hat{\mathbf{Y}} = \begin{bmatrix} \begin{bmatrix} \hat{y}^{(1)} \end{bmatrix} & \cdots & \begin{bmatrix} \hat{y}^{(m)} \end{bmatrix} \end{bmatrix}$$

Actual center word matrix

$$\mathbf{Y} = \begin{bmatrix} \begin{bmatrix} y^{(1)} \end{bmatrix} & \cdots & \begin{bmatrix} y^{(m)} \end{bmatrix} \end{bmatrix}$$

Mark as completed

Like Dislike Report an issue