Contents

[1. Model Structure 2](#_Toc485371923)

[1.1 Input & Output 2](#_Toc485371924)

[1.2 Meaning of symbols 2](#_Toc485371925)

[2. Forward Propagation 4](#_Toc485371926)

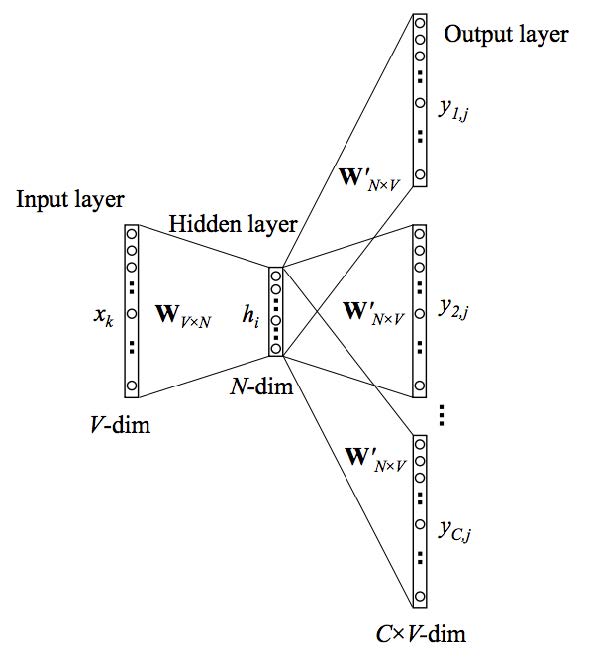
# Model Structure

## Input & Output

Input A single word .

OutputThe words **in ’s context** defined by a word window of size .

Format All of these words are one-hot encoded.



## Meaning of symbols

* represents the **one-hot encoded vector** corresponding to the input word in the training instance.
* are the one-hot encoded vectors corresponding to the output words in the training instance.
* is the weight matrix between the input layer and hidden layer, whose th row represents the weights corresponding to the th word in the vocabulary. **This weight matrix is what we are interested in learning because it contains the vector encodings of all the words in our vocabulary**.
* is the hidden layer.
  + Note that the exact size of is also a training parameter.
  + Since the input vector is one-hot encoded, we can observe that is simply the th row of . i.e.
* is the output matrix associated with each output vector.

# Forward Propagation