Recurrent Neural Networks for Text Data

June 21, 2019

Block 3, Lecture 2
Applied Data Science
MMCi Term 4, 2019

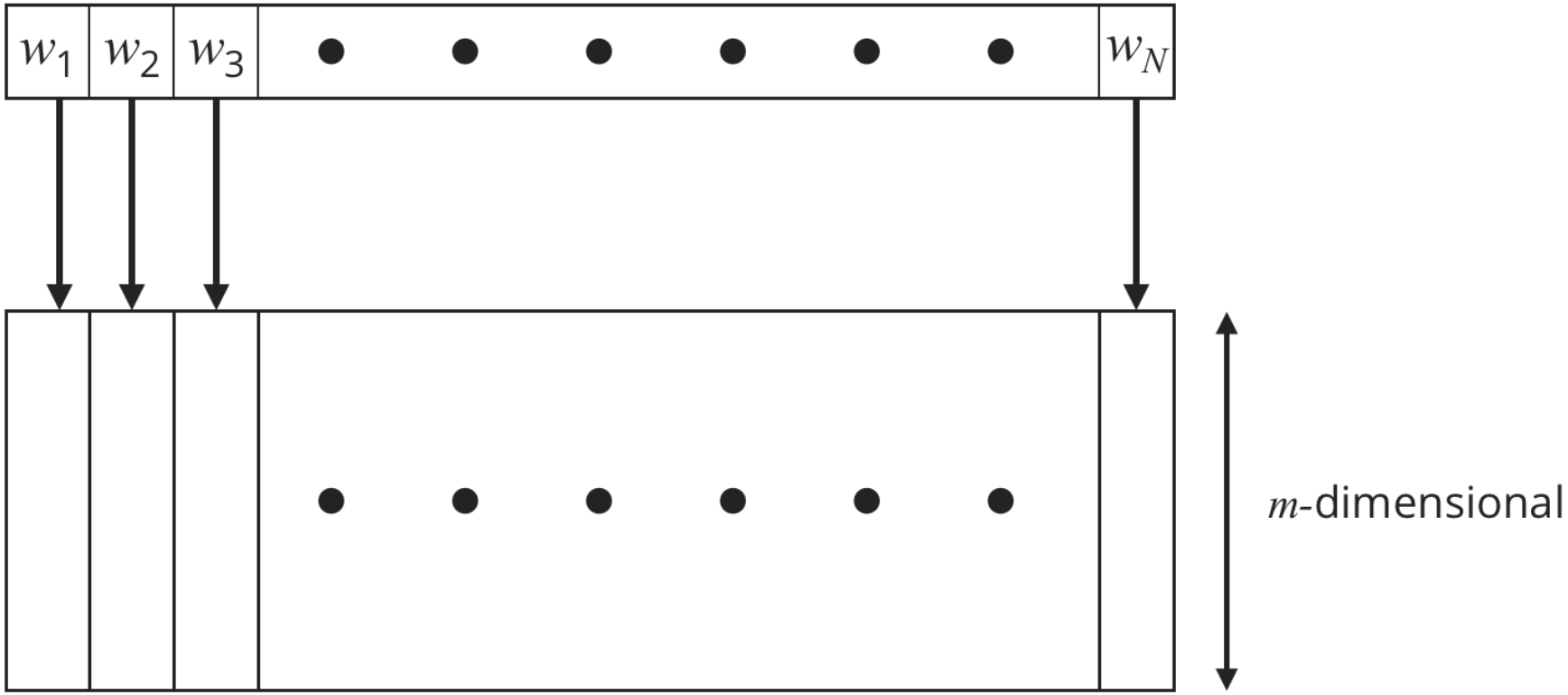
Matthew Engelhard

RECURRENT NEURAL NETWORKS

Generating Text

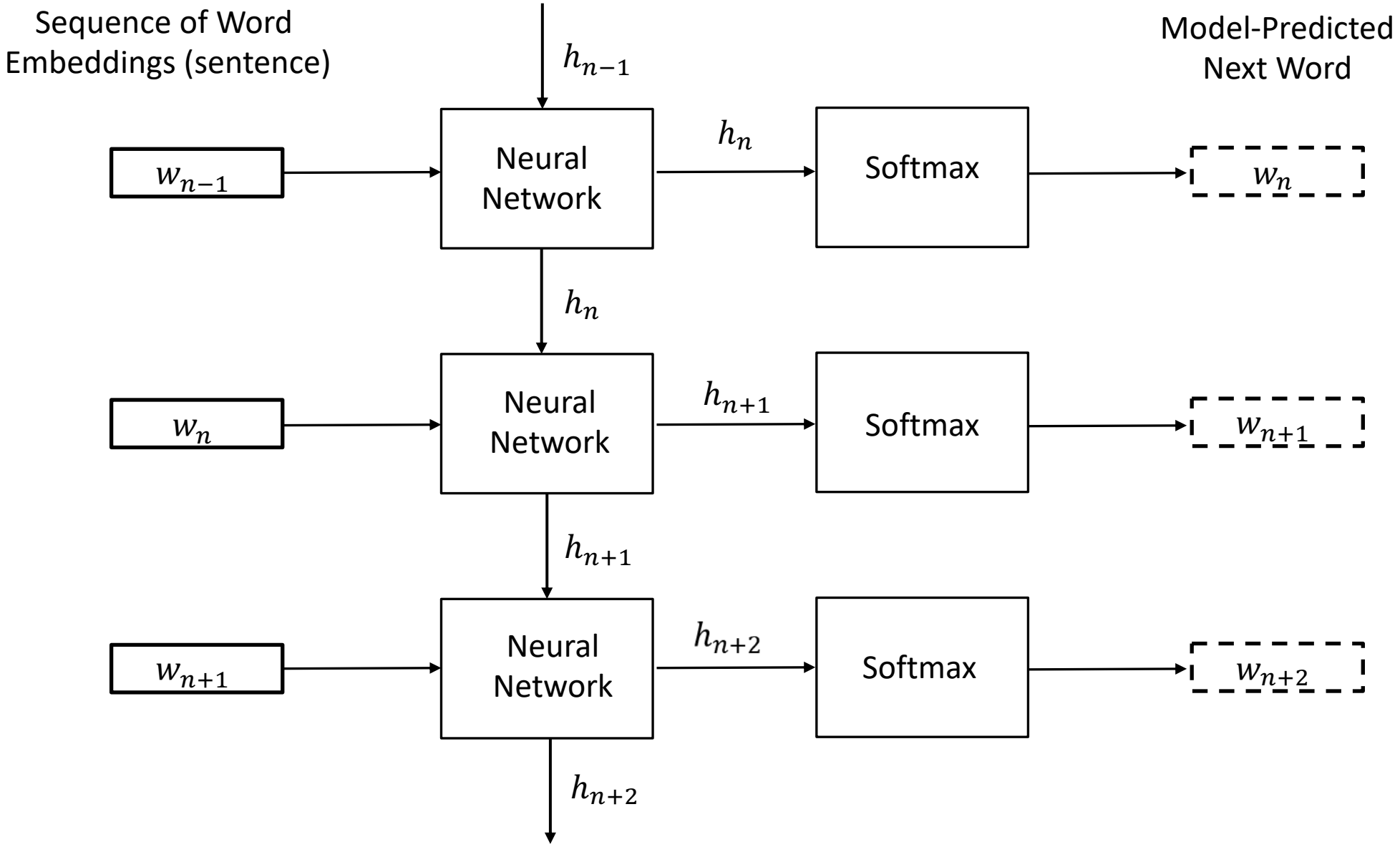
- Assume we have learned word embeddings (vectors) Want to use these embeddings in applications
- Text *synthesis* may be of interest for automatic captioning of images, and for translation from one language to another
- These tools are quite useful generally (will discuss in case studies)
- We require additional tools for text synthesis: The recurrent neural network (RNN)

Using Word Embeddings

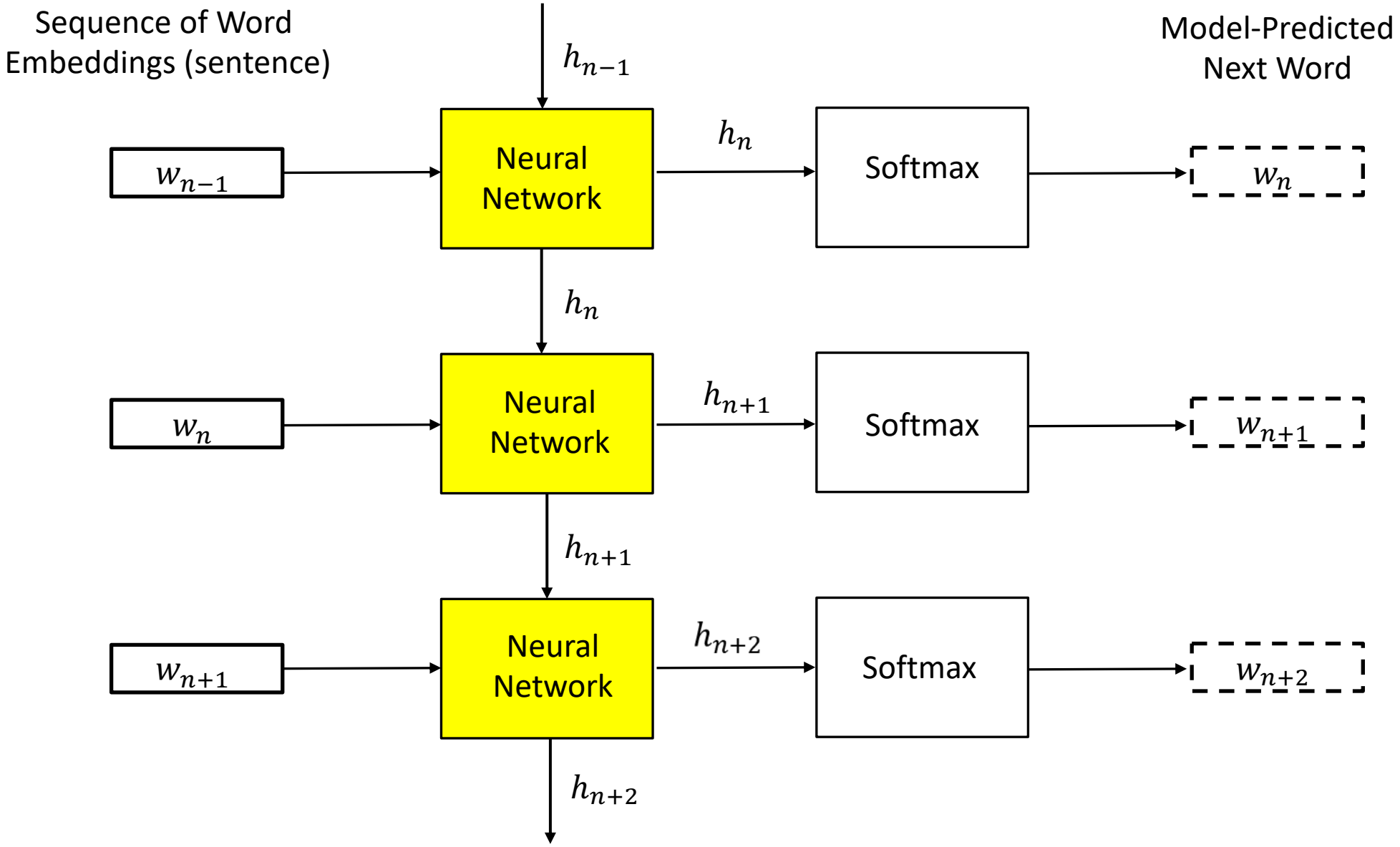


- Our representation depends on the number of words
  - Not a constant number of features!

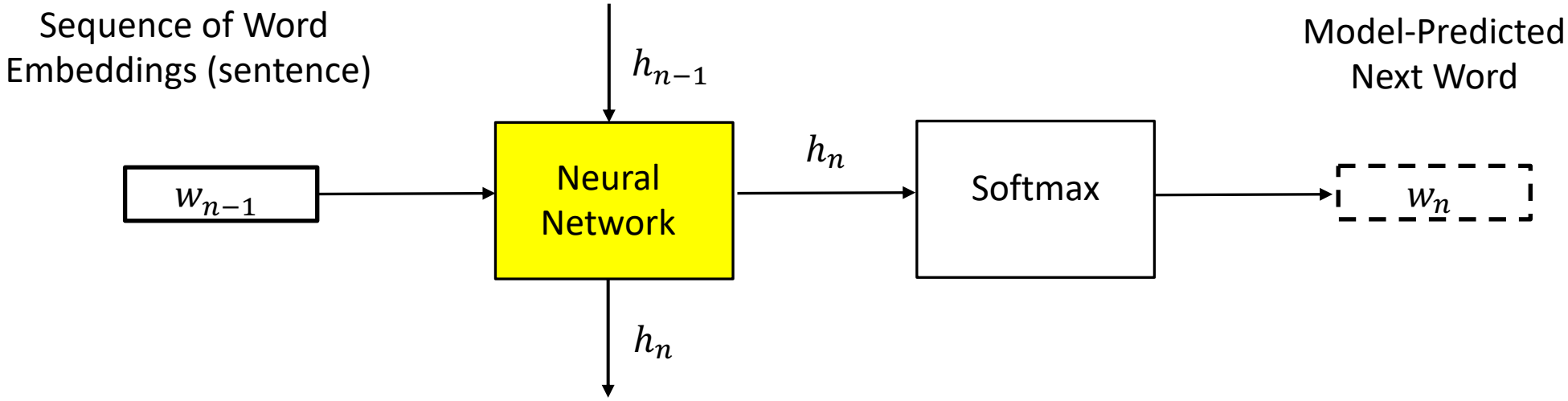
Recurrent Neural Network



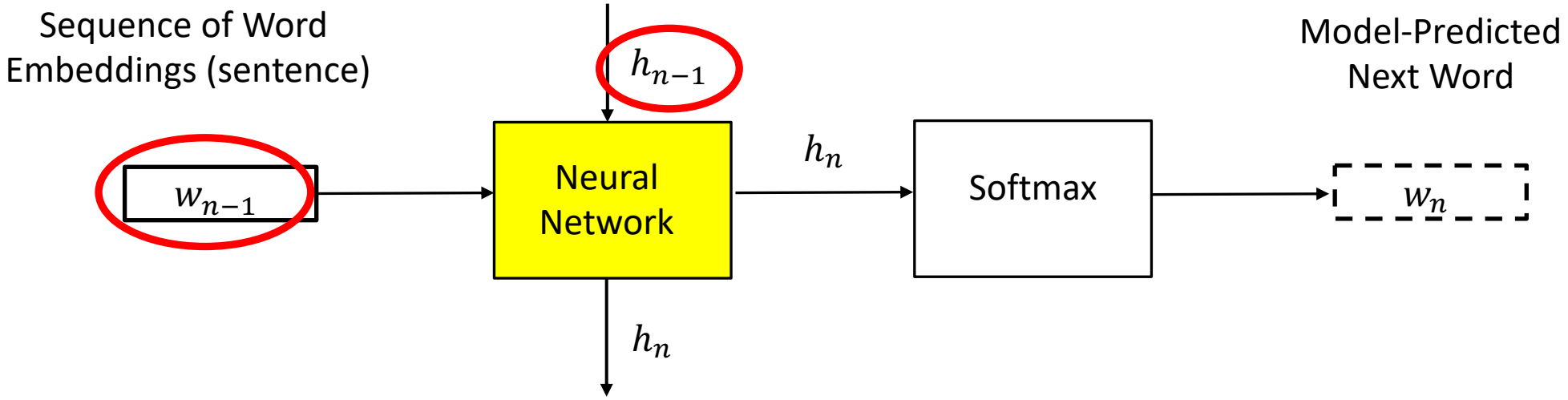
Recurrent Neural Network



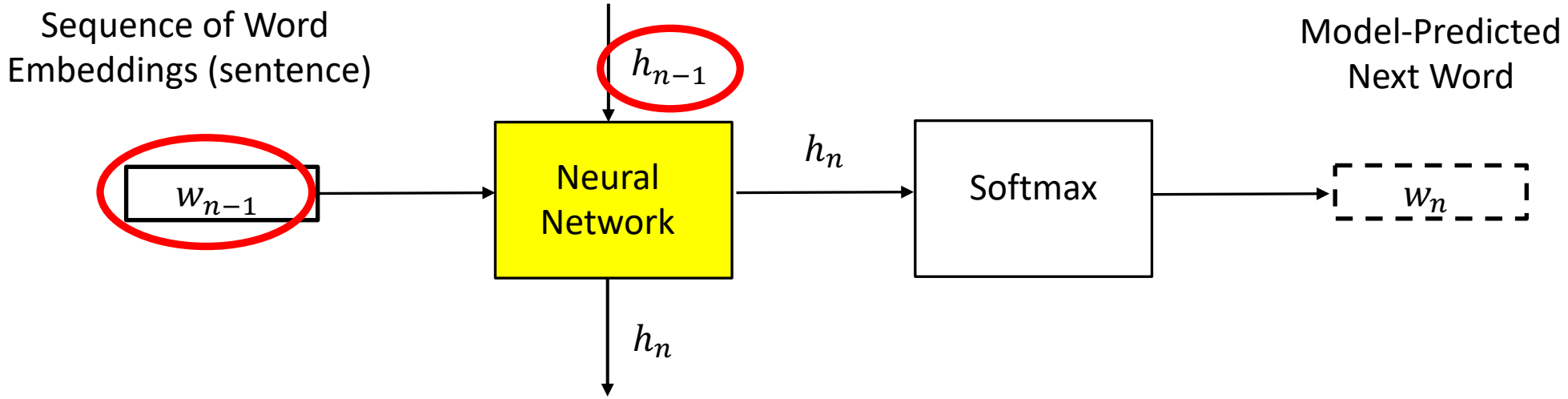
Recurrent Neural Network



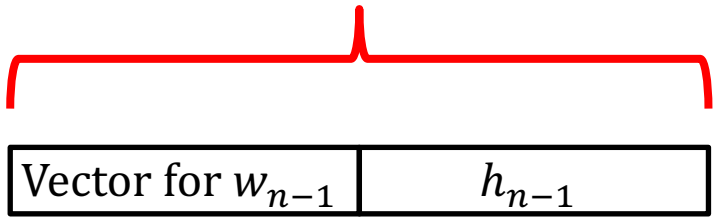
Recurrent Neural Network



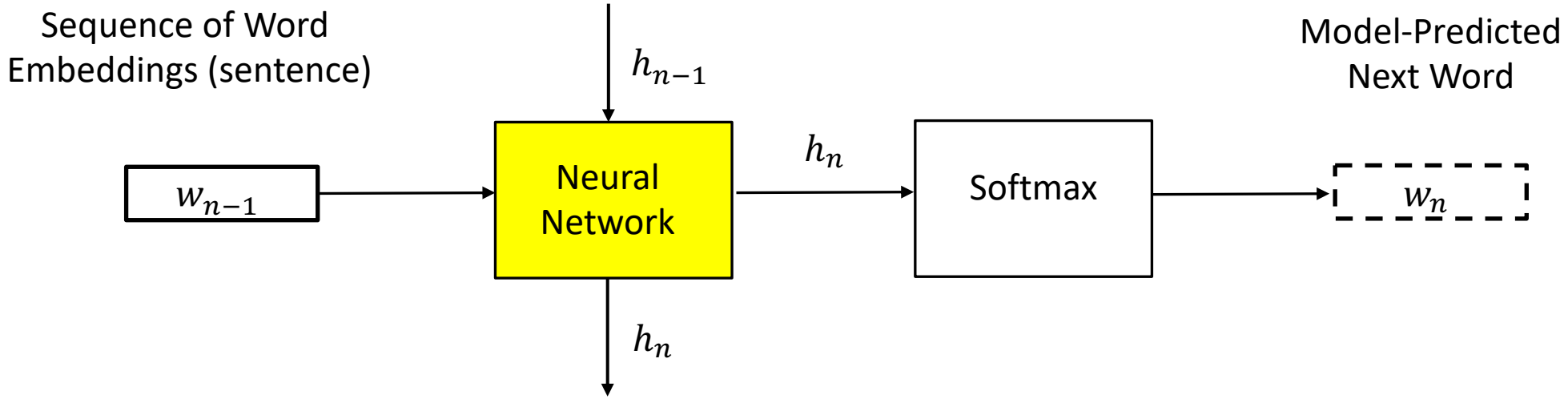
Recurrent Neural Network

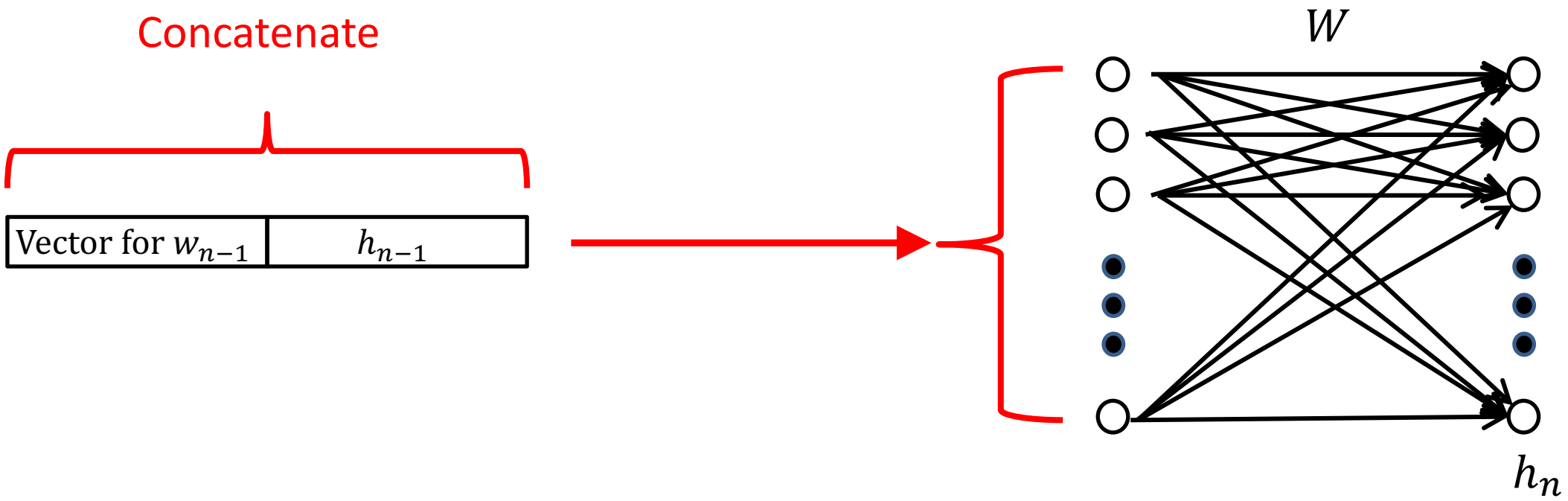


Concatenate

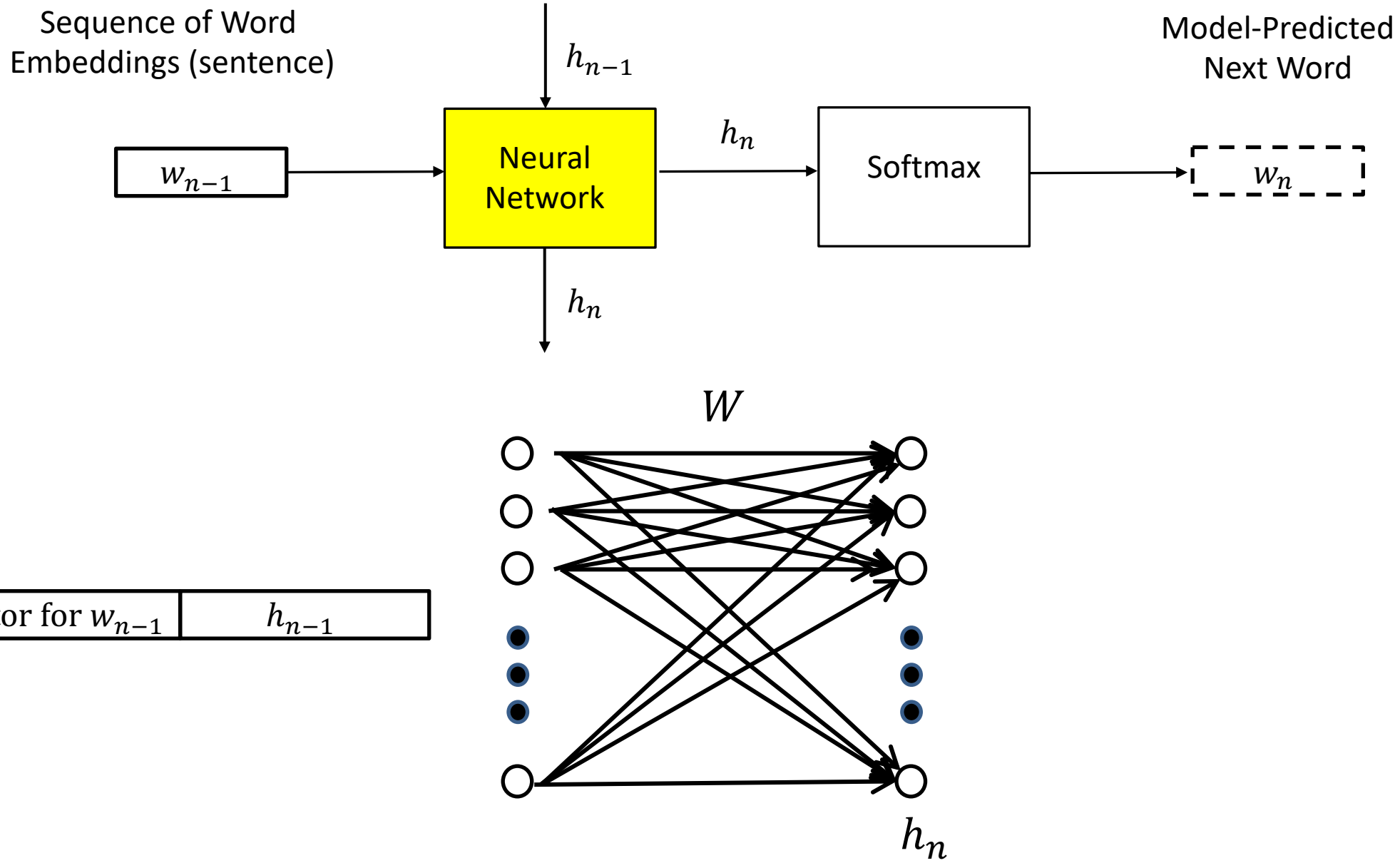


Recurrent Neural Network

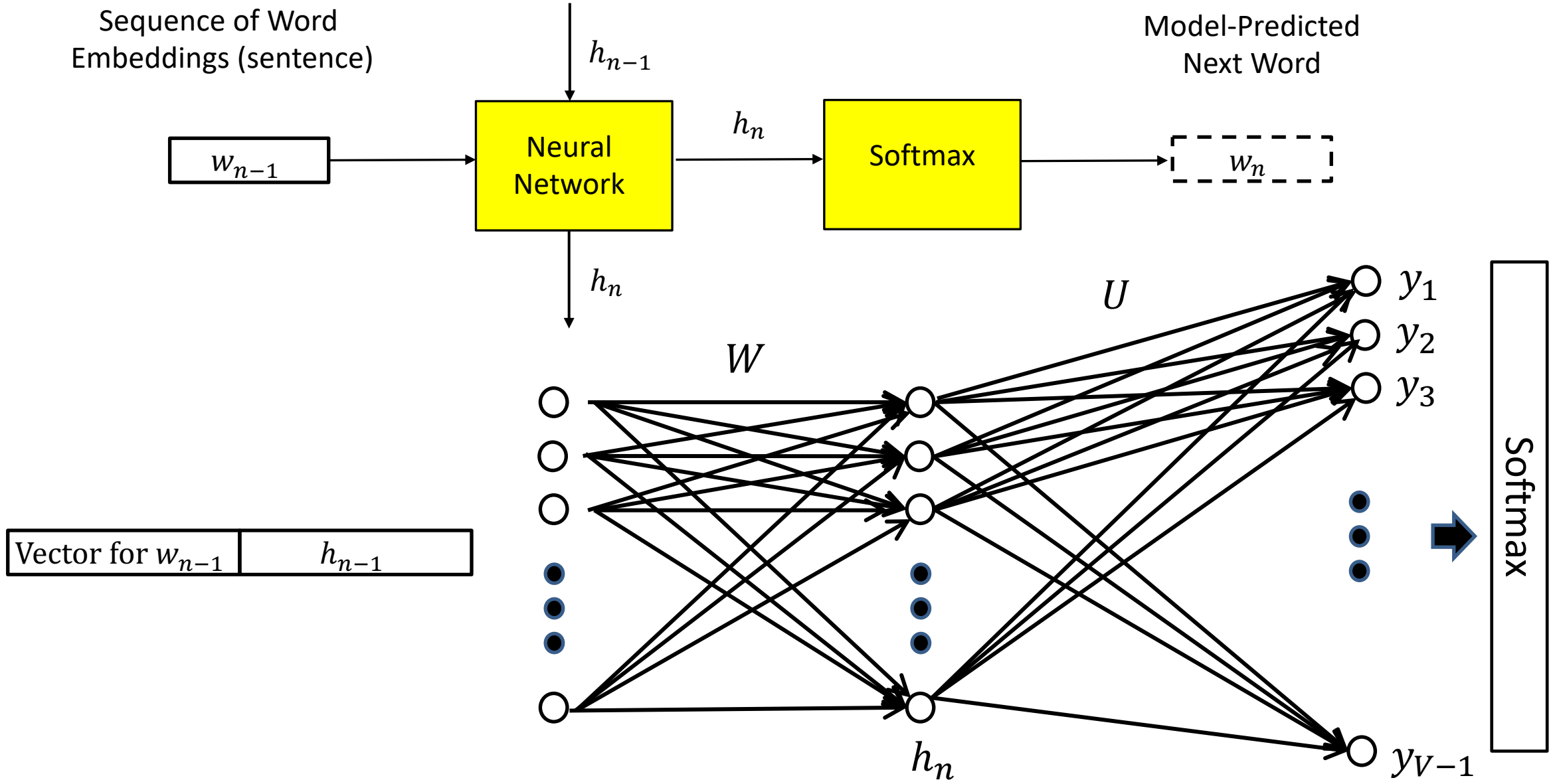




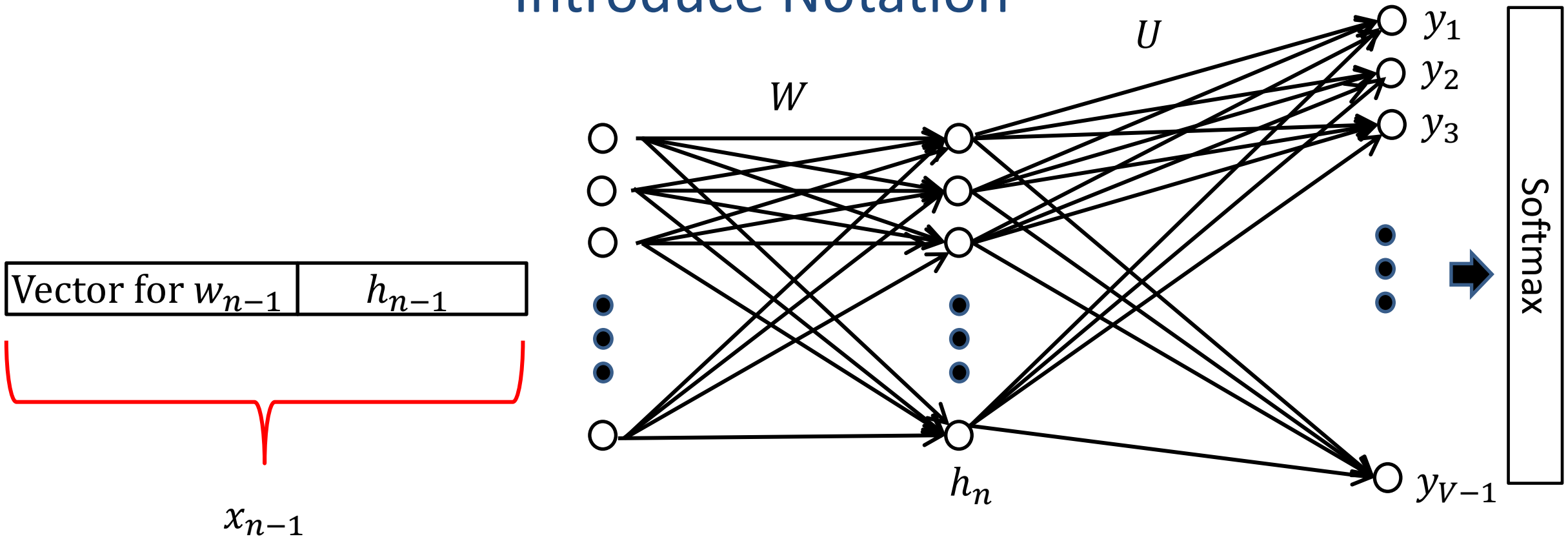
Recurrent Neural Network



Recurrent Neural Network



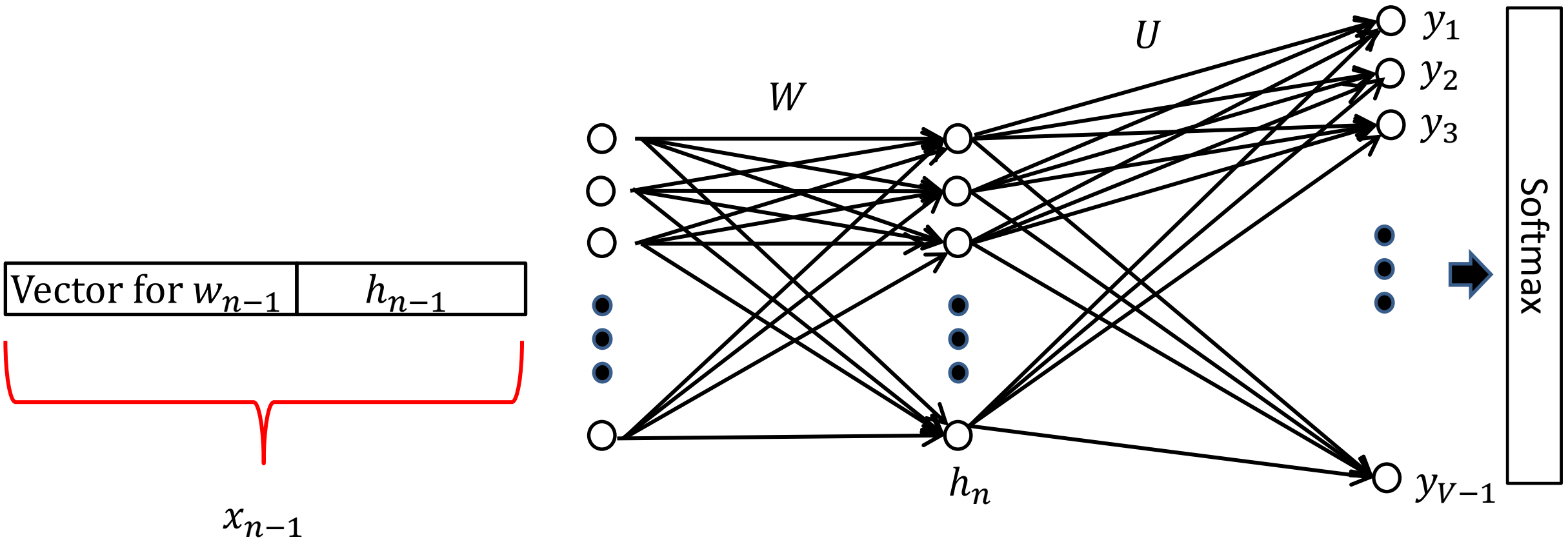
Introduce Notation



$$h_n = \tanh(W \cdot x_{n-1} + b)$$

$$p(w_n | w_{n-1}, h_{n-1}) = \text{softmax}(U \cdot h_n + \beta)$$

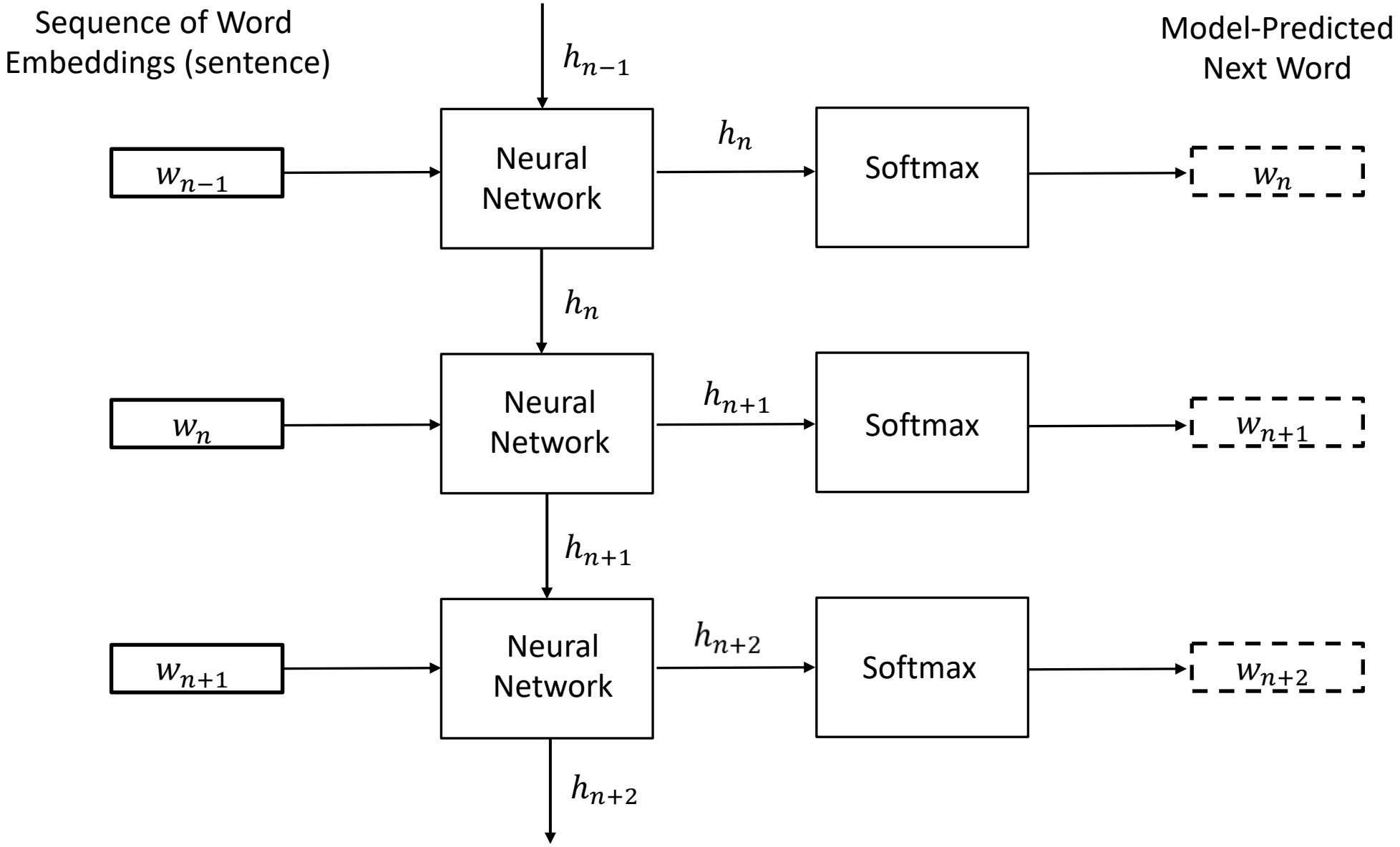
Intuition on Model for Predicting *n*th Word



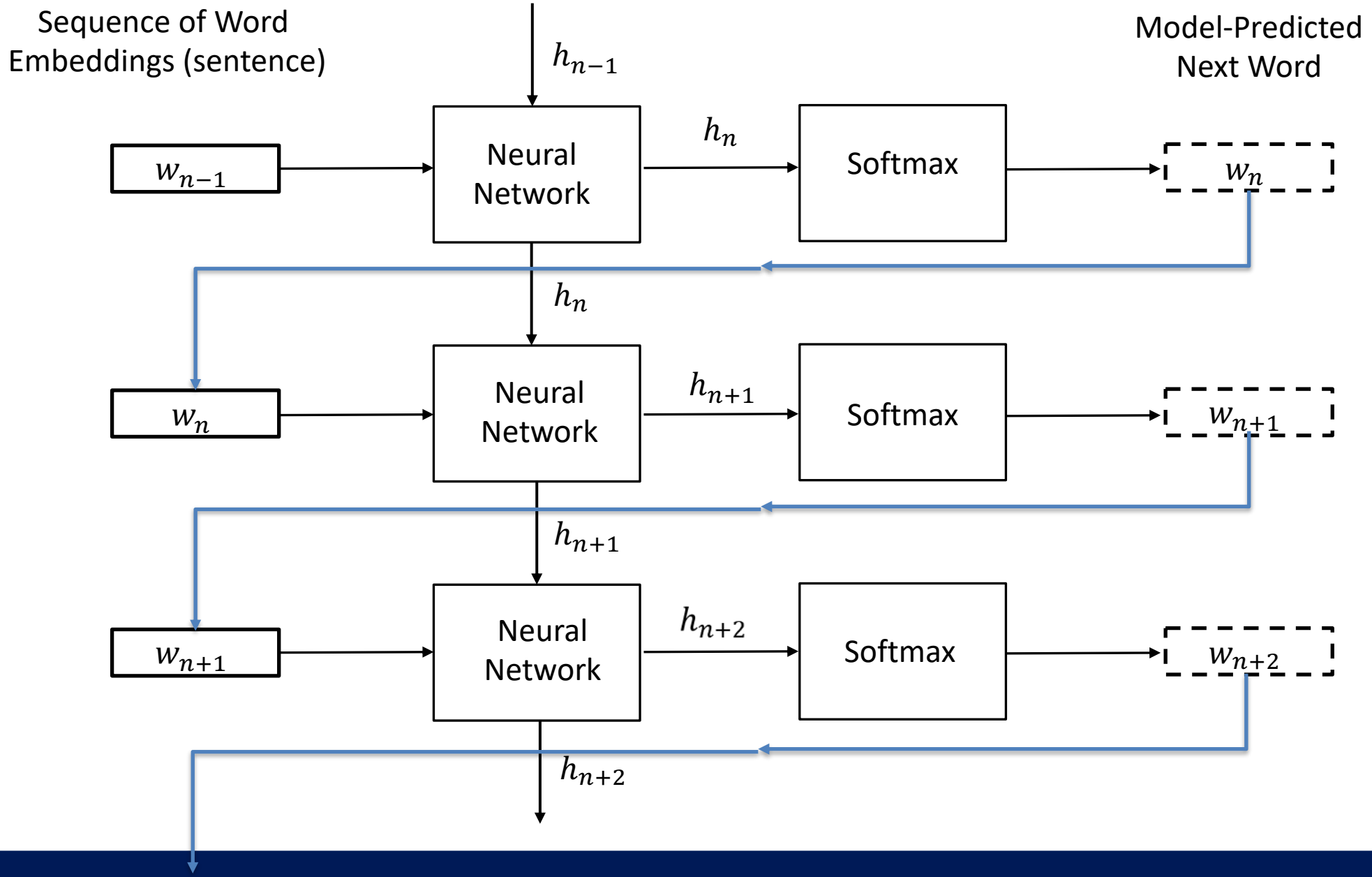
h_{n-1}: Tells us which words were likely prior to selection of previous word (context)

w_{n-1}: Tells us which word was used/selected at point $n-1$ in text, as we predict the nth word

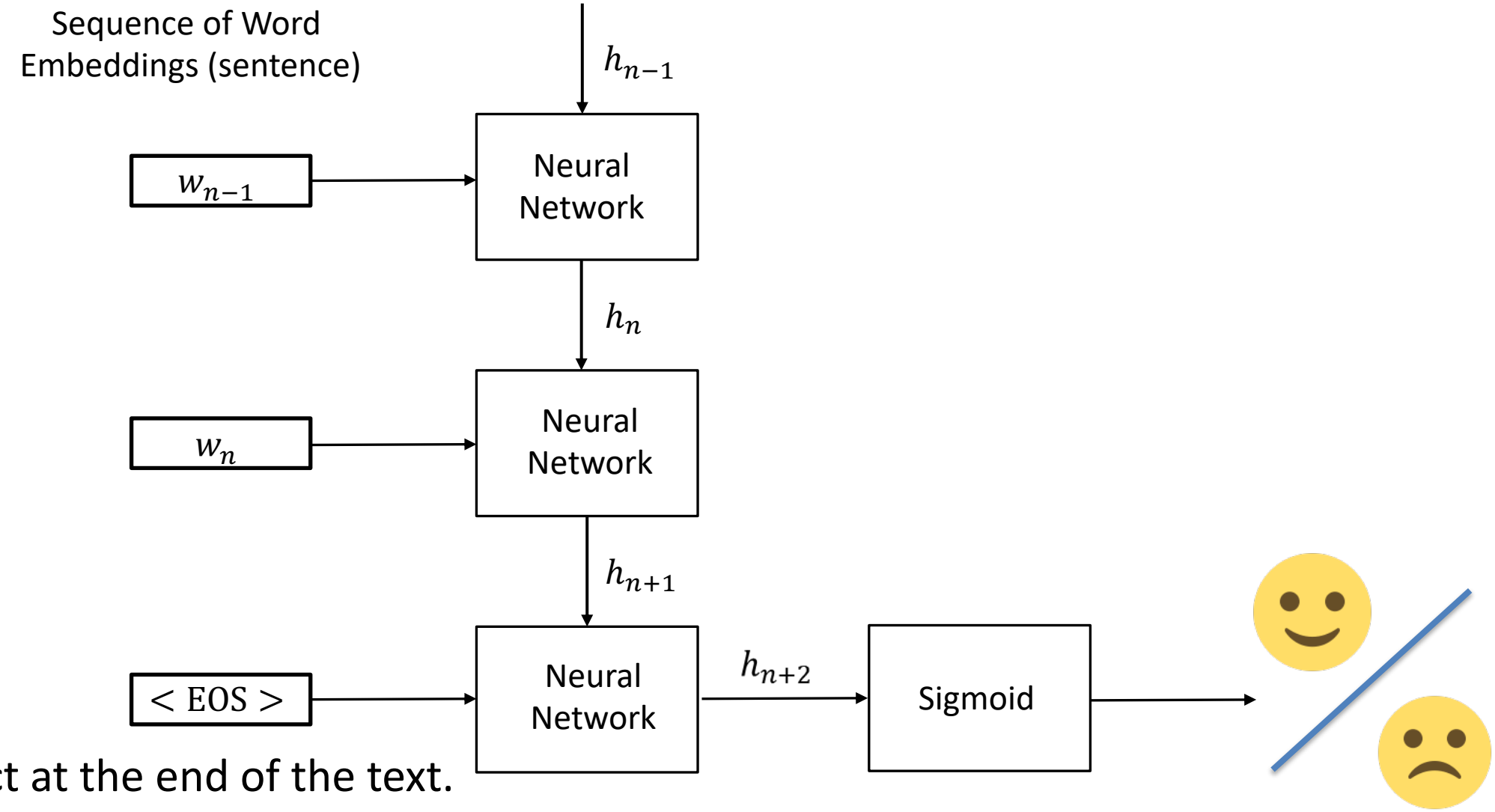
Recurrent Neural Network



Generating Text



Predicting a Single Output

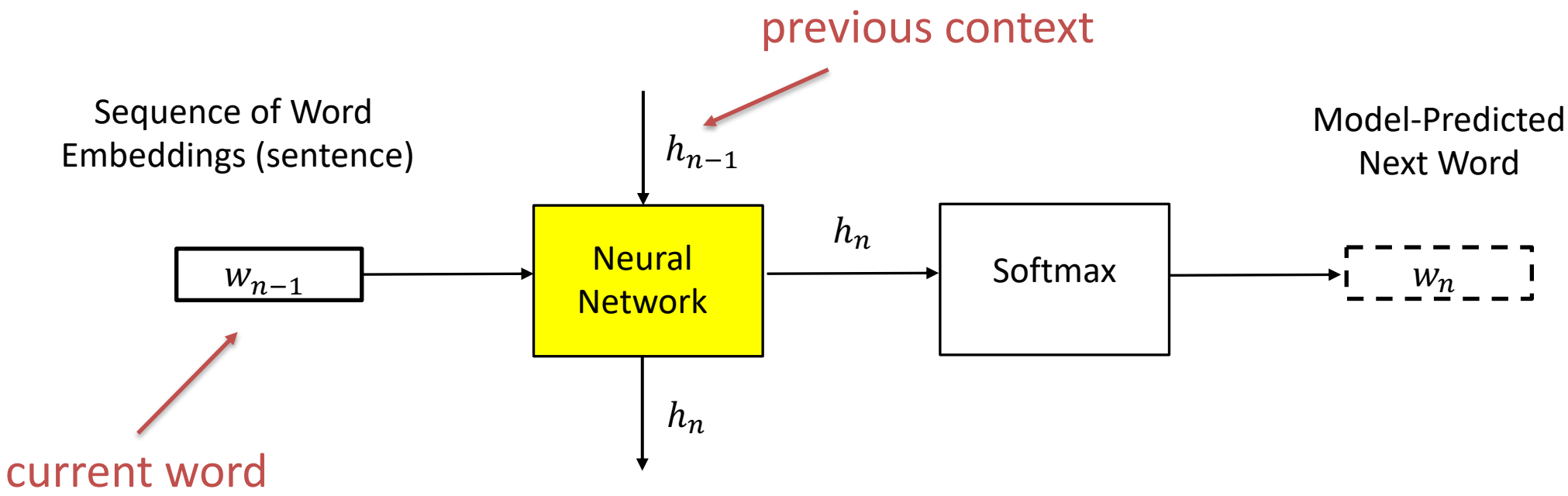


Predict at the end of the text.

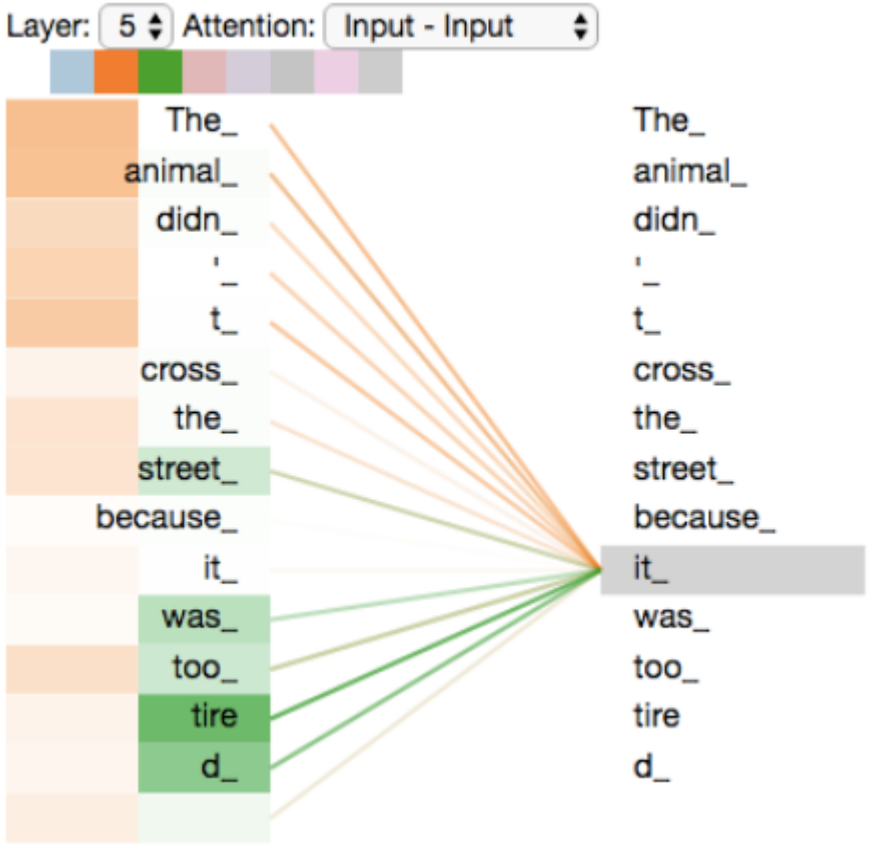
Conclusions

- Word embeddings and recurrent neural networks are currently the cornerstones of natural language processing
- Nearly all text systems are based on these techniques (e.g. google translate, chatbots, etc.)
- Many more versions of Recurrent Neural Networks
  - Long-Short Term Memory (LSTM) to build a “memory”
  - Stacking recurrent units to make deep recurrent networks
- Can combine with CNNs to analyze text and images

Semantic Context in RNNs



Self-attention: a more flexible mechanism for ‘putting words in context’



- Hierarchy of representations for each word/token
- At each level, representation is refined based on surrounding context

THANK YOU FOR YOUR ATTENTION!