Hide menu

Introduction to Neural Networks and TensorFlow

Practice Assignment: Classification Using Deep Neural Networks

N-grams vs. Sequence Models

Video: Lesson Introduction
49 sec

Video: Traditional Language models
3 min

Reading: Traditional Language models
5 min

Video: Recurrent Neural Networks
4 min

Reading: Recurrent Neural Networks
4 min

Video: Applications of RNNs
3 min

Reading: Application of RNNs
3 min

Video: Math in Simple RNNs
3 min

Reading: Math in Simple RNNs
6 min

Lab: Hidden State Activation
20 min

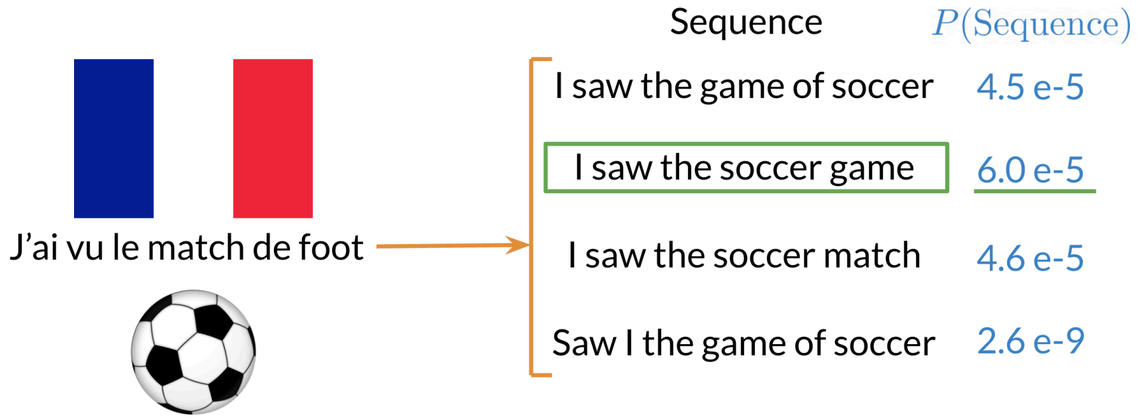
Video: Cost Function for RNNs
2 min

Reading: Cost Function for RNNs
5 min

Video: Implementation Note
1 min

Traditional Language models

Traditional language models make use of probabilities to help identify which sentence is most likely to take place.



In the example above, the second sentence is the one that is most likely to take place as it has the highest probability of happening. To compute the probabilities, you can do the following:

$$P(w_2|w_1) = \frac{\text{count}(w_1, w_2)}{\text{count}(w_1)} \longrightarrow \text{Bigrams}$$

$$P(w_3|w_1, w_2) = \frac{\text{count}(w_1, w_2, w_3)}{\text{count}(w_1, w_2)} \longrightarrow \text{Trigrams}$$

$$P(w_1, w_2, w_3) = P(w_1) \times P(w_2|w_1) \times P(w_3|w_2)$$

Large N-grams capture dependencies between distant words and need a lot of space and RAM. Hence, we resort to using different types of alternatives.