

Introduction to NLP (CS7.401)

Spring 2022, IIIT Hyderabad
22 Feb, Friday (Lecture 10)

Taught by Prof. Manish Shrivastava

Deep Learning and NLP

The outputs of vector representation learning models can be used as input to deep learning models for various tasks.

Neural Network Language Models

Neural networks language carry out exactly the same task as statistical models, except that they use weights and biases and operate on the *representations* of the previous $(n - 1)$ words. As output, they provide a probability distribution over the vocabulary, representing the likelihood of each word coming after the input $(n - 1)$ -gram.

The fact that the representations are dense makes it possible to carry out the computation on any input, seen or unseen.

A feed-forward NNLM simply uses the inputs of the $(n - 1)$ input words simultaneously to generate the distribution.

Recurrent NNLMs, on the other hand, use the output of each word to influence the output of the next. They aggregate the information from past decisions to inform future decisions.

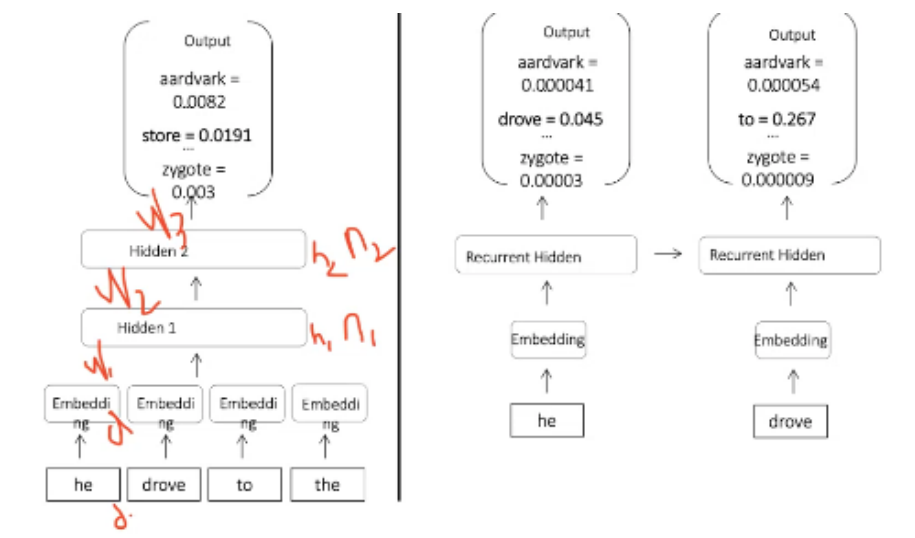


Figure 1: Feed-Forward and Recurrent NNLMs