



deeplearning.ai

# NLP and Word Embeddings

---

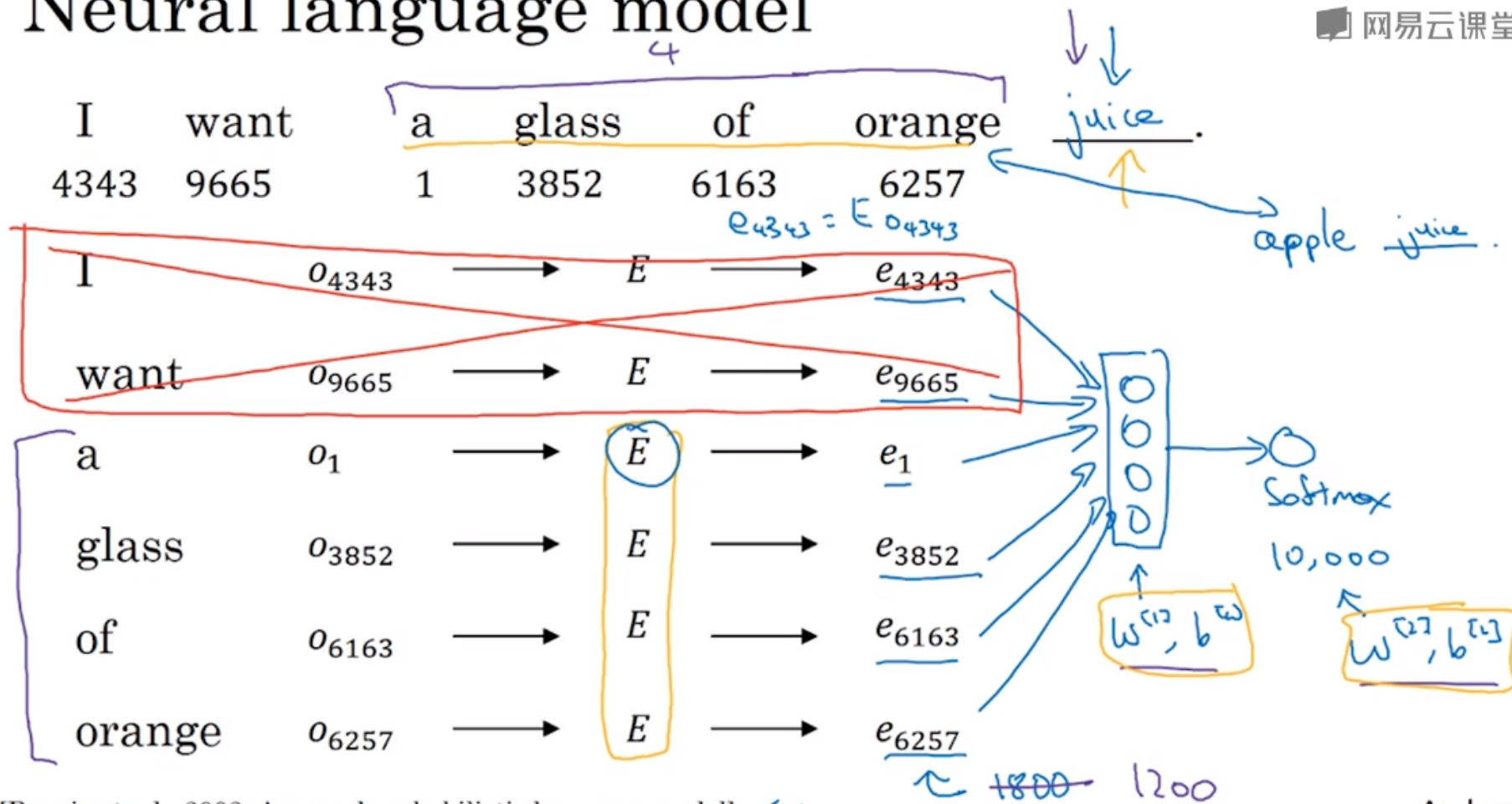
## Learning word embeddings

In the history of deep learning as applied to learning word embedding, people actually started off with relatively complex algorithms. And over time researchers discovered they can use simpler and simpler algorithms and still get very good results especially for a large dataset

# Neural language model

Let's start to formalize the problem of learning a good word embedding. When you implement an algorithm to learn a word embedding what you end up learning is an embedding matrix.

And again, variety of choices. And using a fixed history window, just means that you can deal with even arbitrarily long sentences because the input size is always fixed.



[Bengio et. al., 2003, A neural probabilistic language model]

Andrew Ng

And you can use backprop to perform gradient descent to maximize the likelihood of your training set, to just repeatedly predict given four words in a sequence, what is the next word in your text corpus.

# Other context/target pairs

I want a glass of orange juice to go along with my cereal.

Context

target

Context: Last 4 words.

4 words on left & right

Last 1 word

Nearby 1 word

a glass of orange ? to go along with

orange ?

glass ?

skip gram

Andrew Ng

So what the researches found is that if you really want to build a language model, it's natural to use the last few words as a context. But if your main goal is really to learn a word embedding, then you can use all of these other contexts and they will result in very meaningful word embeddings as well.

To summarize, in this video, you saw how the language modeling problem which causes the pose of a ML problem where you input the context like the last 4 words and predicts some target words, how posing that problem allows you to learn a good word embedding.