



deeplearning.ai

Sentiment classification is the task of looking at a piece of text and telling if someone likes or dislikes the thing they're talking about. It is one of the most important building blocks in NLP and is used in many applications.

NLP and Word Embeddings

Sentiment classification

One of the challenges of sentiment classification is that you might not have a huge label training set for it. But with word embeddings, you're able to build good sentiment classifiers even with only modest-size label training sets.

Sentiment classification problem

x  y

The dessert is excellent.



Service was quite slow.



Good for a quick meal, but nothing special.



Completely lacking in good taste, good service, and good ambience.



10,000  100,000 words

Simple sentiment classification model 网易云课堂

The dessert is excellent



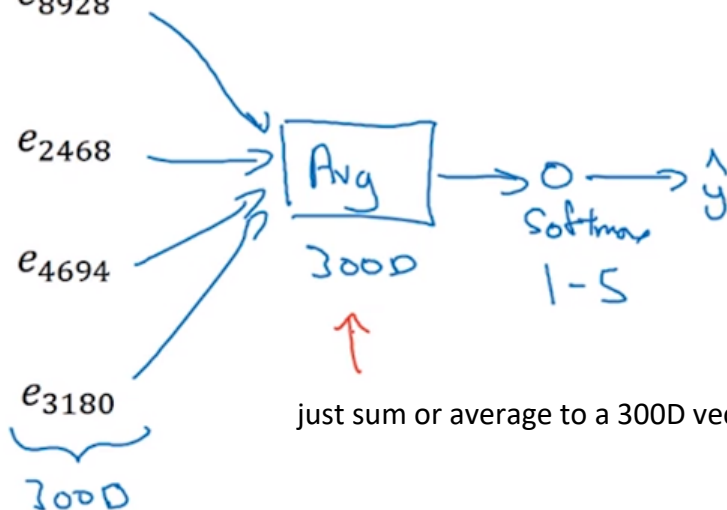
8928 2468 4694 3180

The o_{8928} \longrightarrow E \longrightarrow e_{8928}

dessert o_{2468} \longrightarrow E \longrightarrow e_{2468}

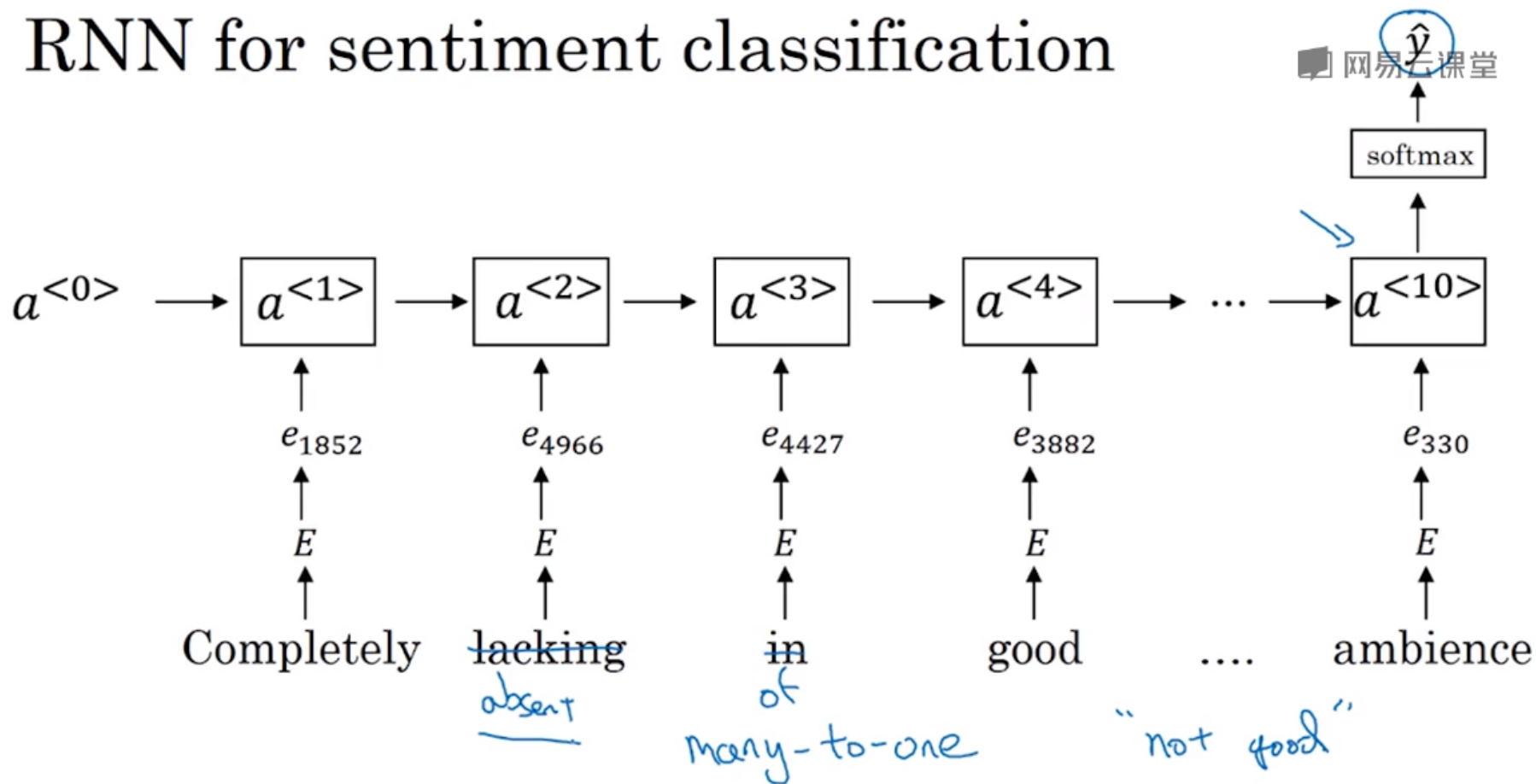
is o_{4694} \longrightarrow E \longrightarrow e_{4694}

excellent o_{3180} \longrightarrow E \longrightarrow e_{3180}



“Completely lacking in good taste, good service, and good ambience.”
↑
100 B words

RNN for sentiment classification



Andrew Ng

And because your word embeddings can be trained from a large dataset, this will do a better job generalizing to maybe even new words not in your training set. If it was in your 1 billion or 100 billion word corpus used to train the word embeddings, it might still get this right and generalize much better even to words that in the training set used to train word embeddings necessarily in the label training set that you hand for specifically the sentiment classification problem.