deeplearning.ai

NLP and Word Embeddings

Word representation

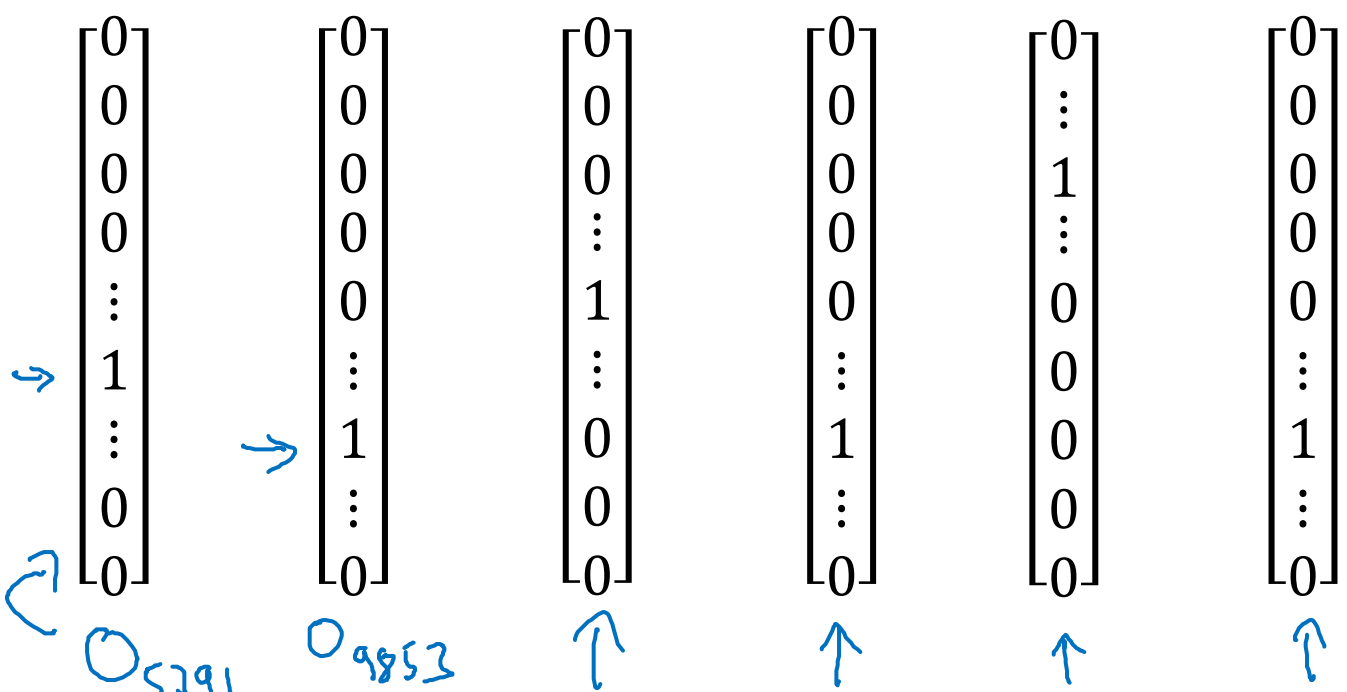
Word representation

V = [a, aaron, …, zulu, <UNK>]

$|V| = 10{,}000$

1-hot representation

Man (5391)	Woman (9853)	King (4914)	Queen (7157)	Apple (456)	Orange (6257)



I want a glass of orange juice.

I want a glass of apple ?.

Featurized representation: word embedding

	Man (5391)	Woman (9853)	King (4914)	Queen (7157)	Apple (456)	Orange (6257)
Gender	-1	1	-0.95	0.97	0.00	0.01
Royal	0.01	0.02	0.93	0.95	-0.01	0.00
Age	0.03	0.02	0.7	0.69	0.03	-0.02
Food	0.04	0.01	0.02	0.01	0.95	0.97
⋮	⋮	⋮				
Size						
Cost						
Alive						
Verb						

300

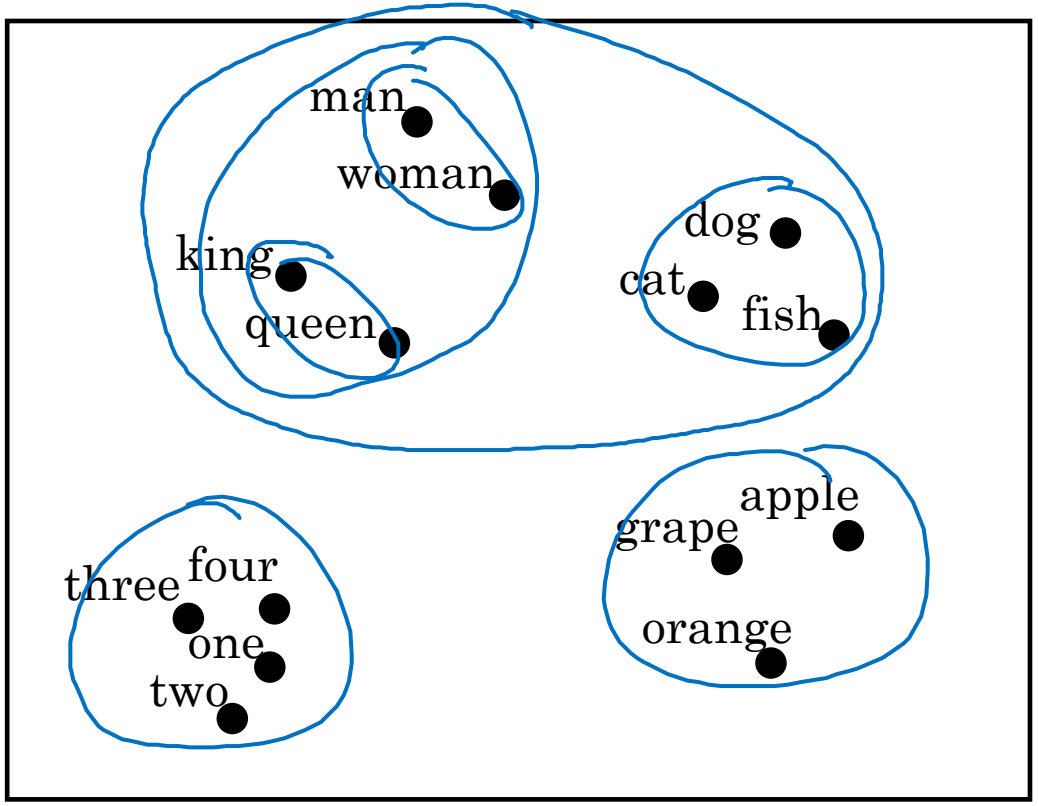
e_{5391} e_{9853}

I want a glass of orange juice.

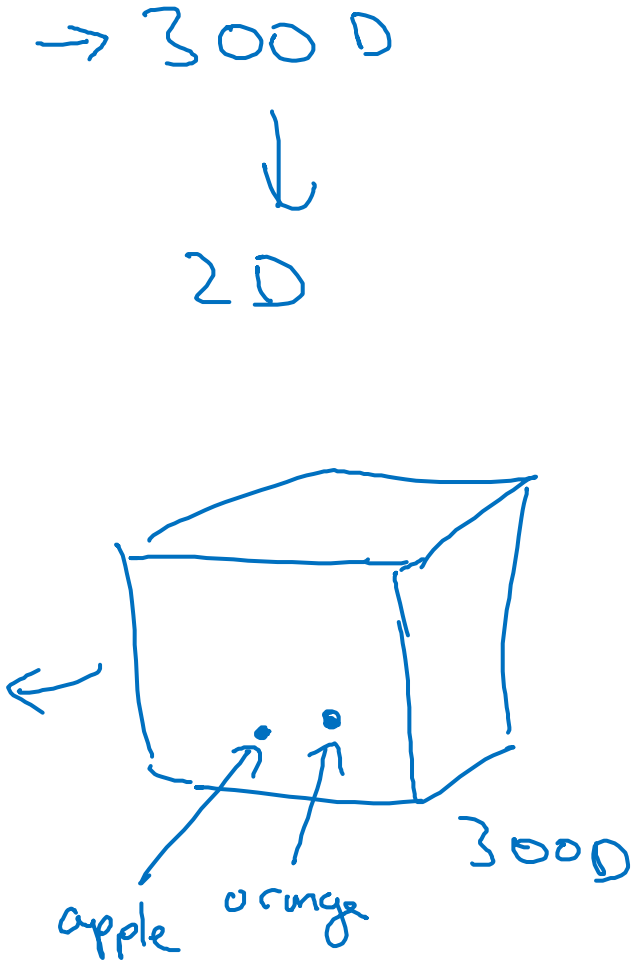
I want a glass of apple juice.

Andrew Ng

Visualizing word embeddings



t-SNE

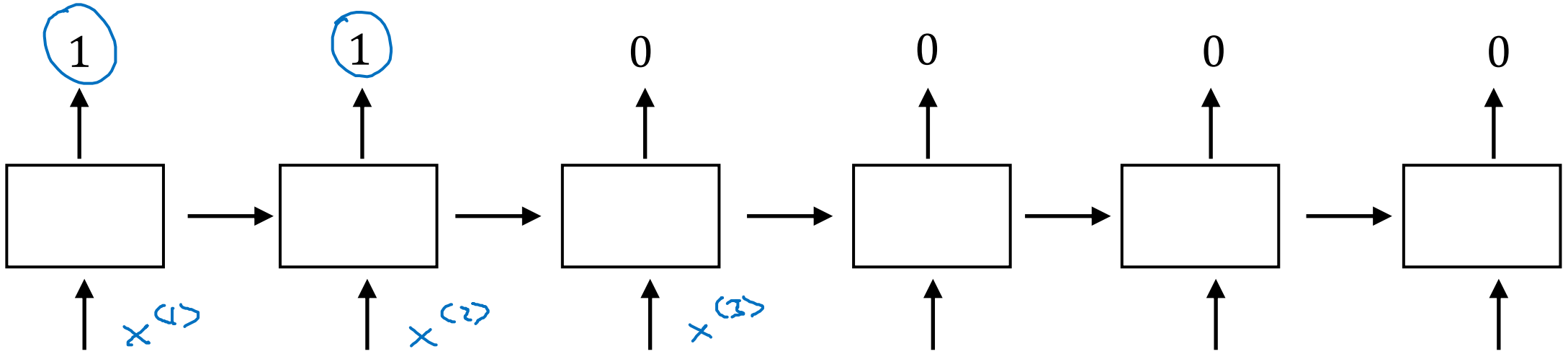


[van der Maaten and Hinton., 2008. Visualizing data using t-SNE]

NLP and Word Embeddings

Using word embeddings

Named entity recognition example



Robert Lin is ~~an~~ ~~apple~~ ~~farmer~~

durian cultivator.

→ 1B words – 100B words

→ 100k words

a BRNN

Transfer learning and word embeddings

A

1. Learn word embeddings from large text corpus. (1-100B words)

   (Or download pre-trained embedding online.)

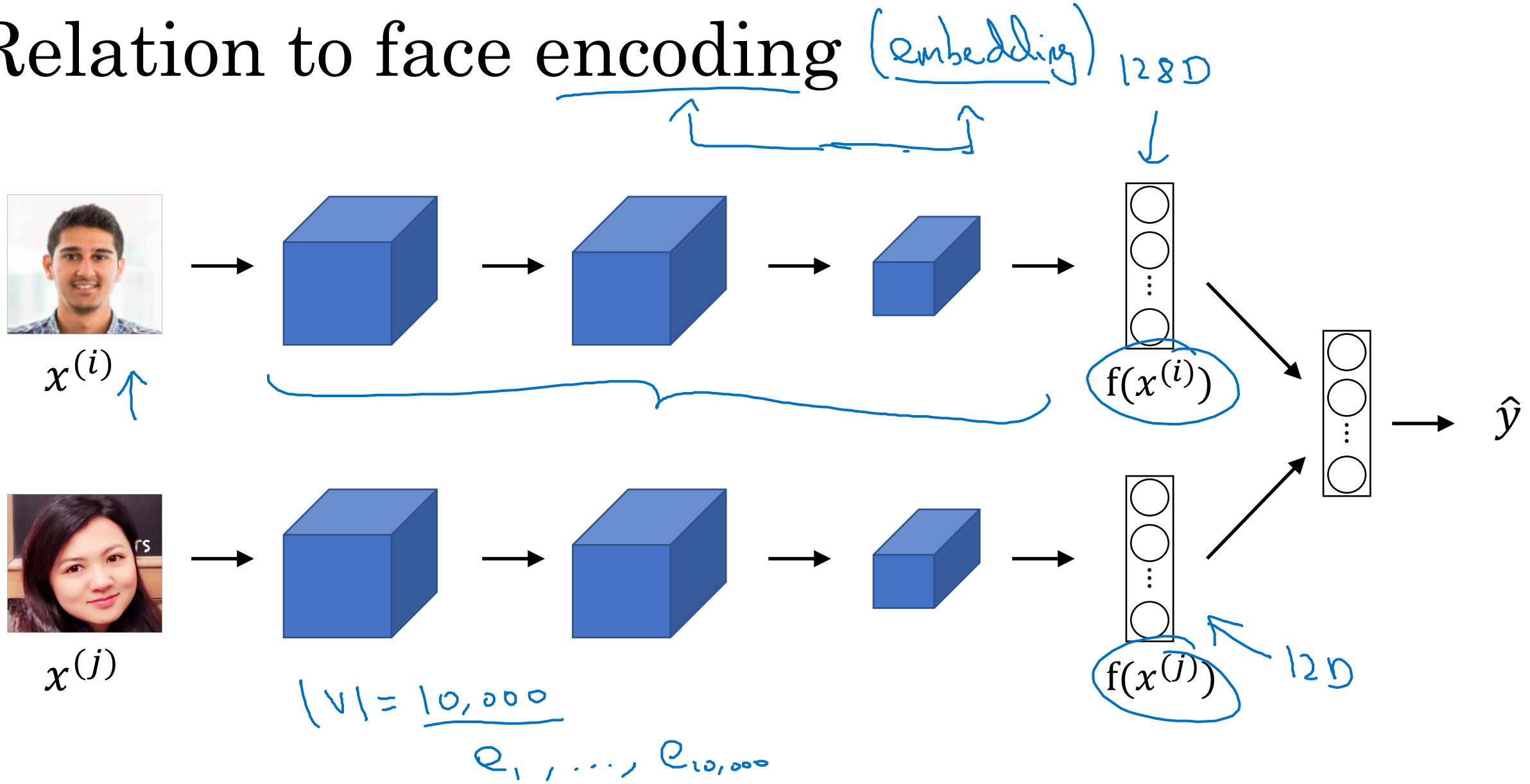
B

2. Transfer embedding to new task with smaller training set. (say, 100k words)

   → 10,000 → 300

3. Optional: Continue to finetune the word embeddings with new data.

Relation to face encoding



[Taigman et. al., 2014. DeepFace: Closing the gap to human level performance]

deeplearning.ai

NLP and Word Embeddings

Properties of word embeddings

Analogies

	Man (5391)	Woman (9853)	King (4914)	Queen (7157)	Apple (456)	Orange (6257)
Gender	−1	1	-0.95	0.97	0.00	0.01
Royal	0.01	0.02	0.93	0.95	-0.01	0.00
Age	0.03	0.02	0.70	0.69	0.03	-0.02
Food	0.09	0.01	0.02	0.01	0.95	0.97

e_{5391} e_{man} e_{woman}

Man → Woman as King → ? Queen

$e_{man} - e_{woman} \approx e_{king} - e_{?}$

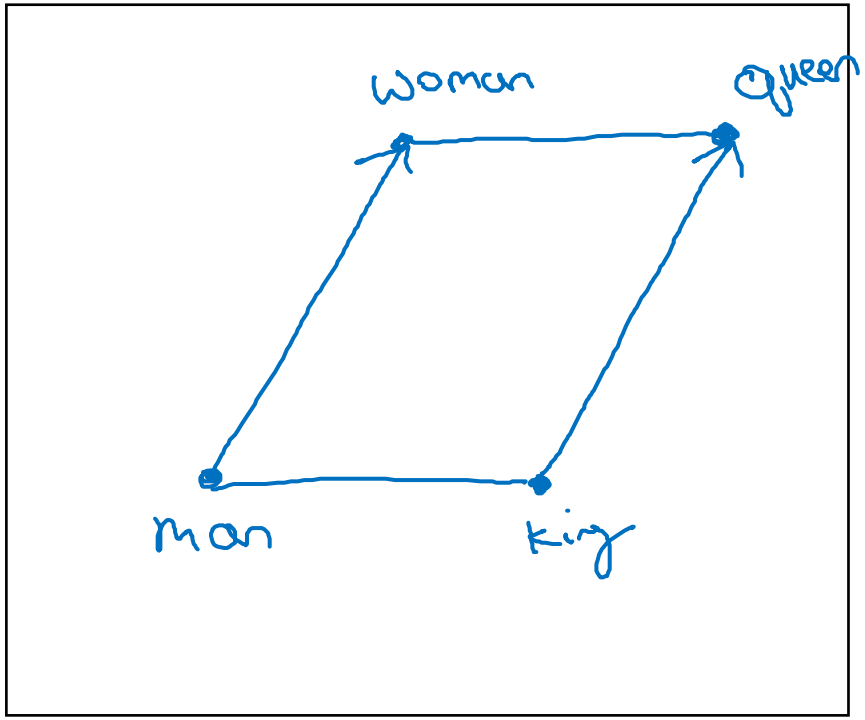
$e_{?} \approx e_{queen}$

$e_{man} - e_{woman} \approx \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

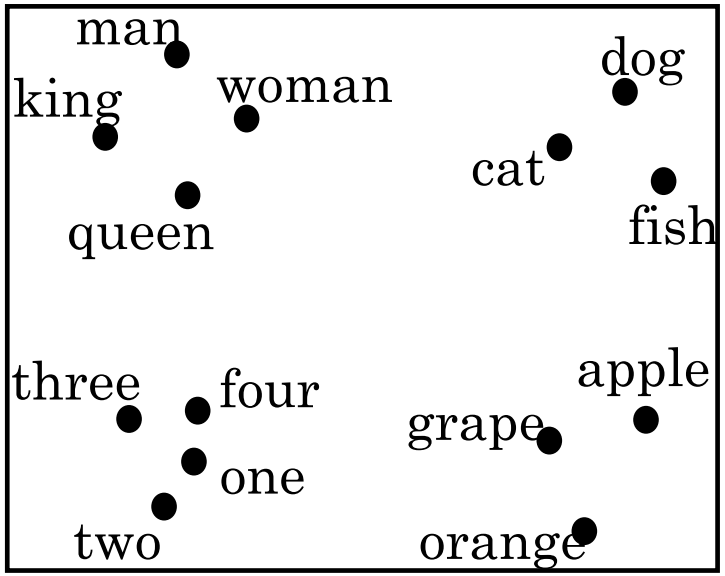
$e_{king} - e_{queen} \approx \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

[Mikolov et. al., 2013, Linguistic regularities in continuous space word representations]

Analogies using word vectors



300D



$300 \rightarrow 2D$

t-SNE

$$e_{man} - e_{woman} \approx e_{king} - e_?$$

e_w

Find word w: $\arg\max_w$ $sim(e_w, e_{king} - e_{man} + e_{woman})$

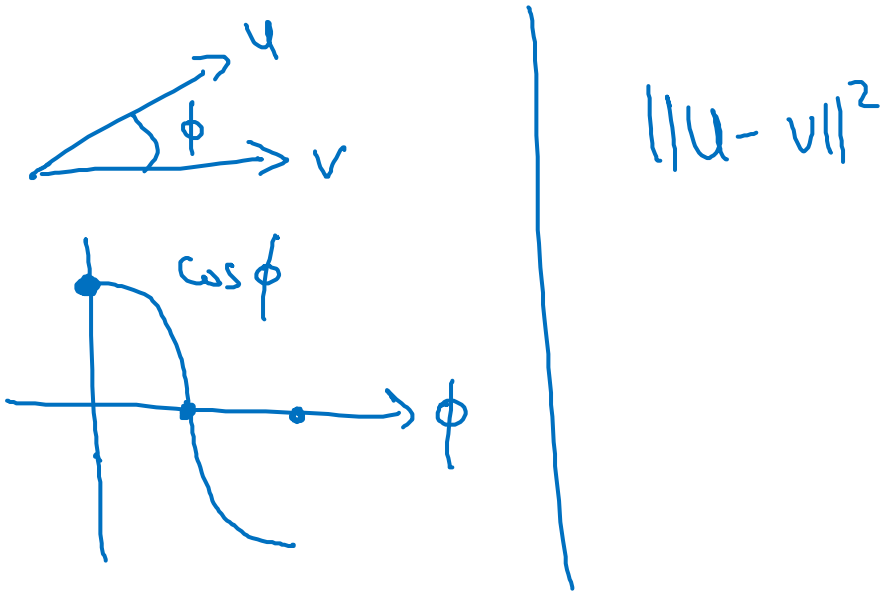
30 - 75%

Andrew Ng

Cosine similarity

$$sim(e_w, e_{king} - e_{man} + e_{woman})$$

$$sim(u, v) = \frac{u^T v}{\|u\|_2 \|v\|_2}$$



$$\|u - v\|^2$$

Man:Woman as Boy:Girl

Ottawa:Canada as Nairobi:Kenya

Big:Bigger as Tall:Taller

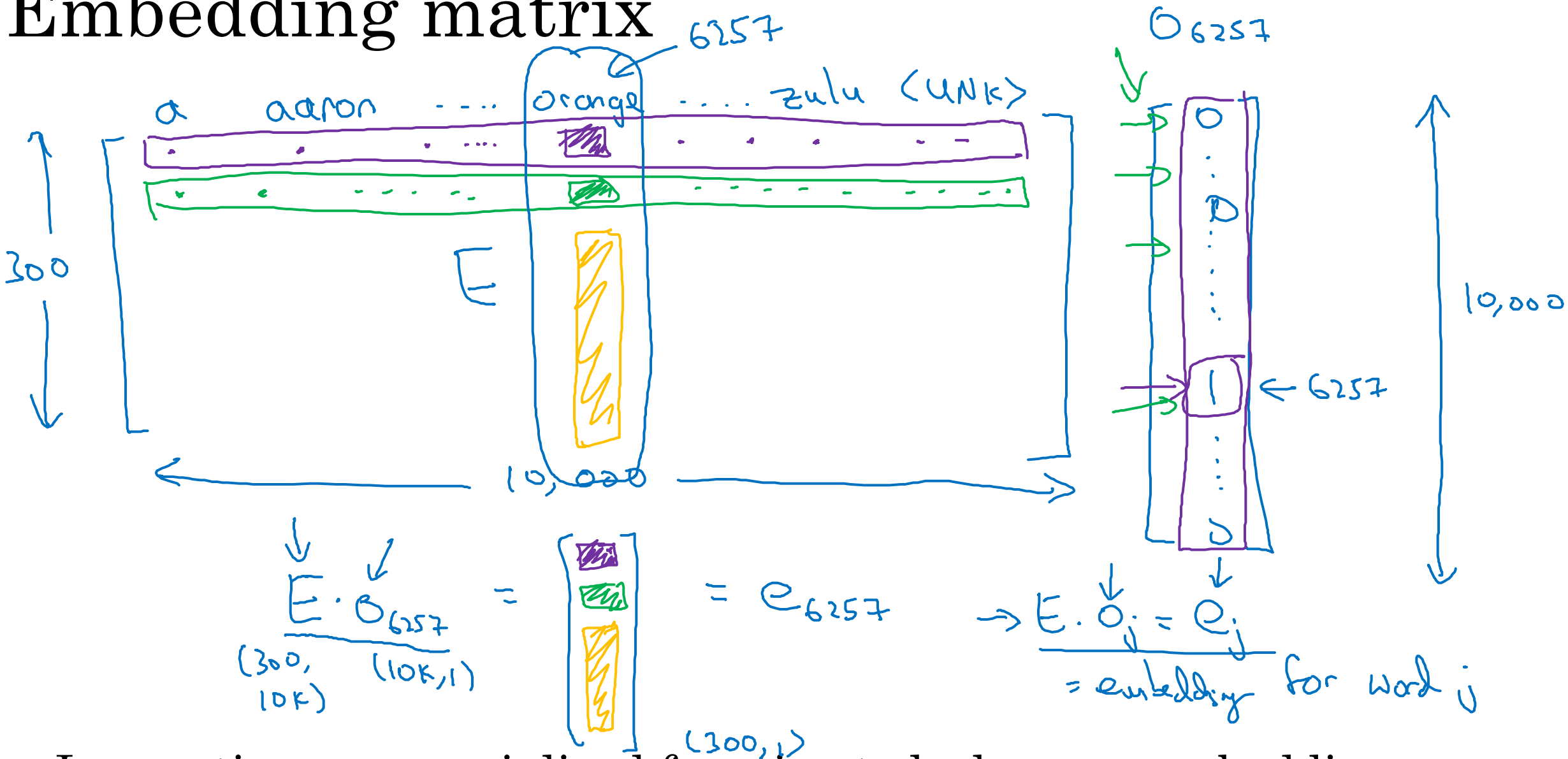
Yen:Japan as Ruble:Russia

deeplearning.ai

NLP and Word Embeddings

Embedding matrix

Embedding matrix



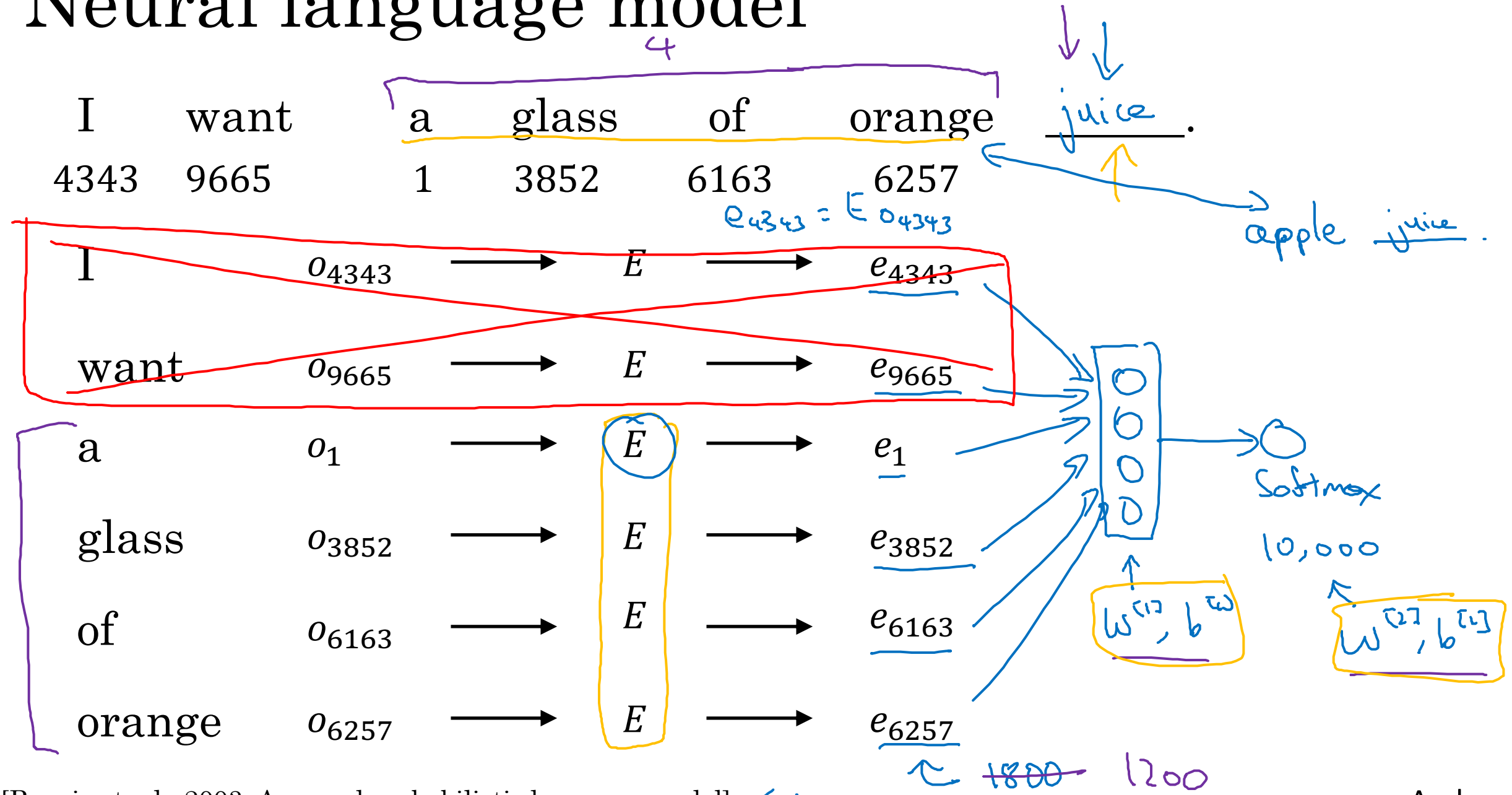
In practice, use specialized function to look up an embedding.

Andrew Ng

NLP and Word Embeddings

Learning word embeddings

Neural language model



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

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 alg with

orange ?

glass ?

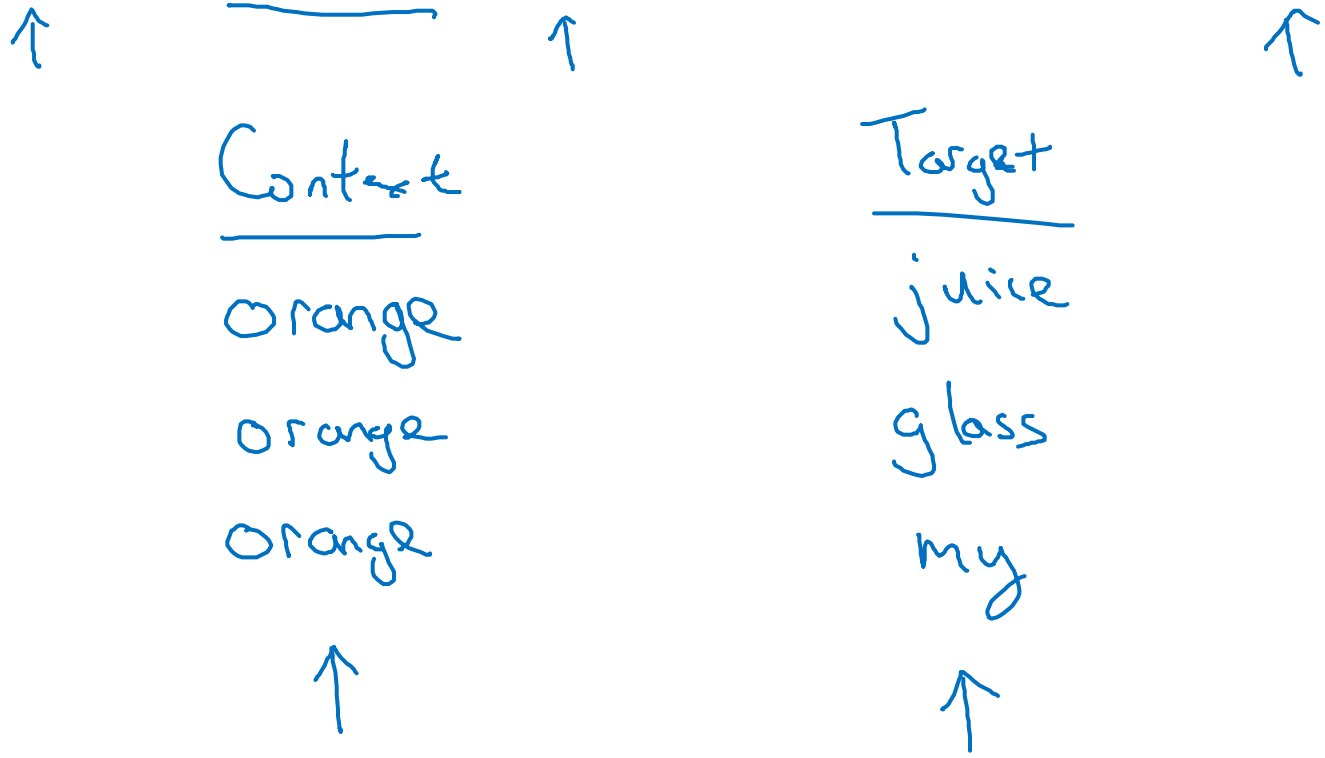
skip gram

NLP and Word Embeddings

Word2Vec

Skip-grams

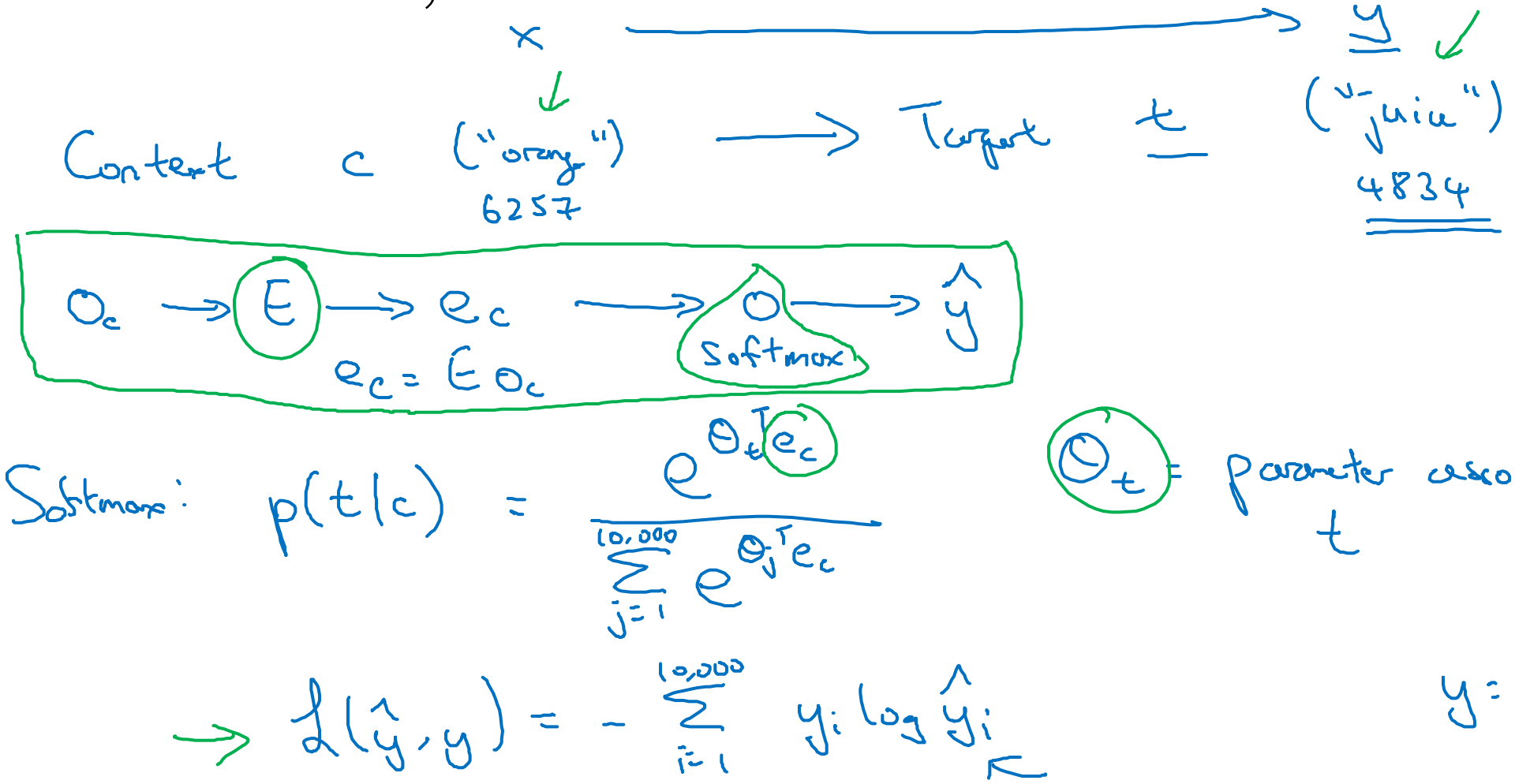
I want a glass of orange juice to go along with my cereal.



[Mikolov et. al., 2013. Efficient estimation of word representations in vector space.]

Model

Vocab size = 10,000k



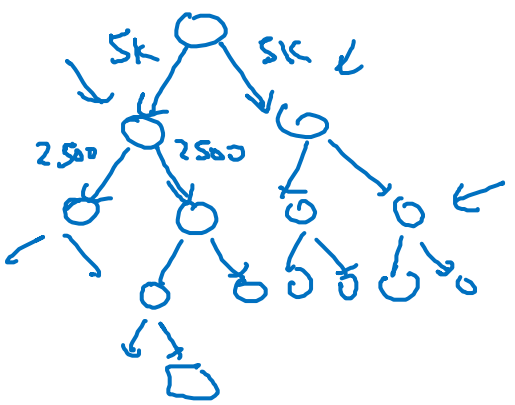
Andrew Ng

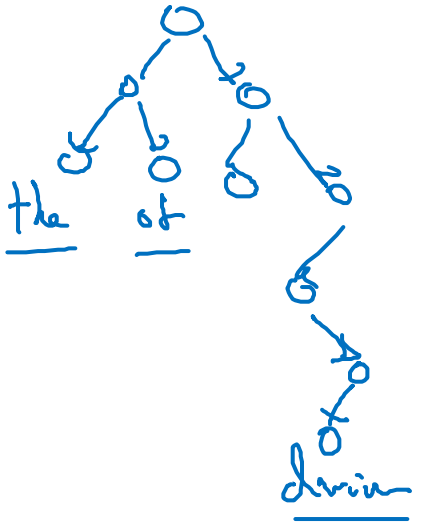
Problems with softmax classification

$$p(t|c) = \frac{e^{\theta_t^T e_c}}{\sum_{j=1}^{10,000} e^{\theta_j^T e_c}}$$

Hierarchical softmax.

$\log |V|$





How to sample the context c?

t

$c \to t$

→ the, of, a, and, to, ...

→ orange, apple, durian

$P(c)$

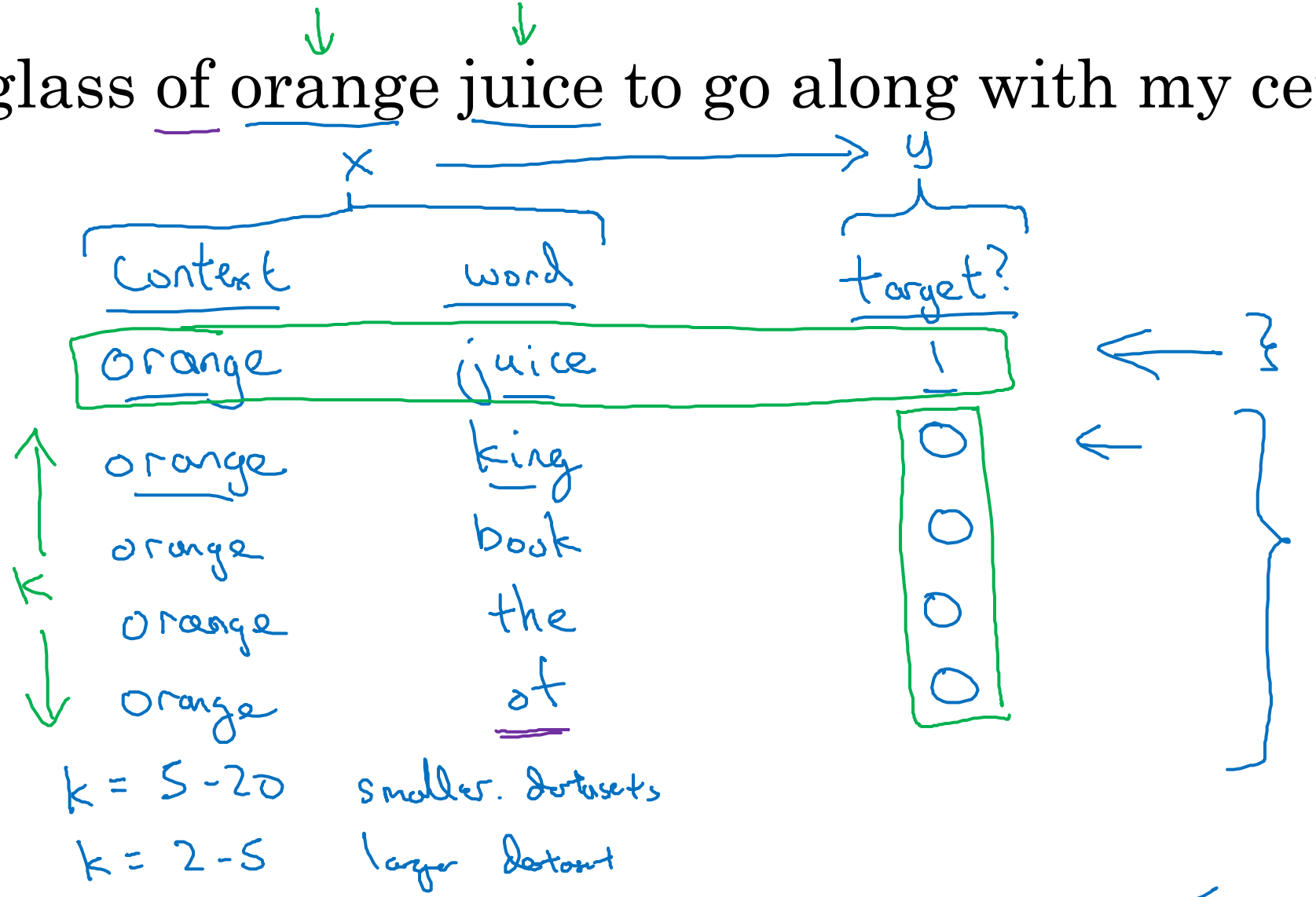
$\frac{1}{2}$ durian

NLP and Word Embeddings

Negative sampling

Defining a new learning problem

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



[Mikolov et. al., 2013. Distributed representation of words and phrases and their compositionality]

Model

Softmax: $p(t|c) = \frac{e^{\theta_t^T e_c}}{\sum_{j=1}^{10,000} e^{\theta_j^T e_c}}$ — 10,000-way softmax

$P(y=1 \mid c, t) = \sigma(\theta_t^T e_c)$

context	word	target?
orange	juice	1
orange	king	0
orange	book	0
orange	the	0
orange	of	0

k

c t y

Orange 6257

$o_{6257} \rightarrow E \rightarrow e_{6257}$

juice? king

10,000

10,000 binary classification problem

k+1

Selecting negative examples

context	word	target?
orange	juice	1
orange	king	0
orange	book	0
orange	the	0
orange	of	0

x y t

the, of, and, ...

$$P(w_i) = \frac{f(w_i)^{3/4}}{\sum_{j=1}^{10,000} f(w_j)^{3/4}}$$

$\frac{1}{|V|}$

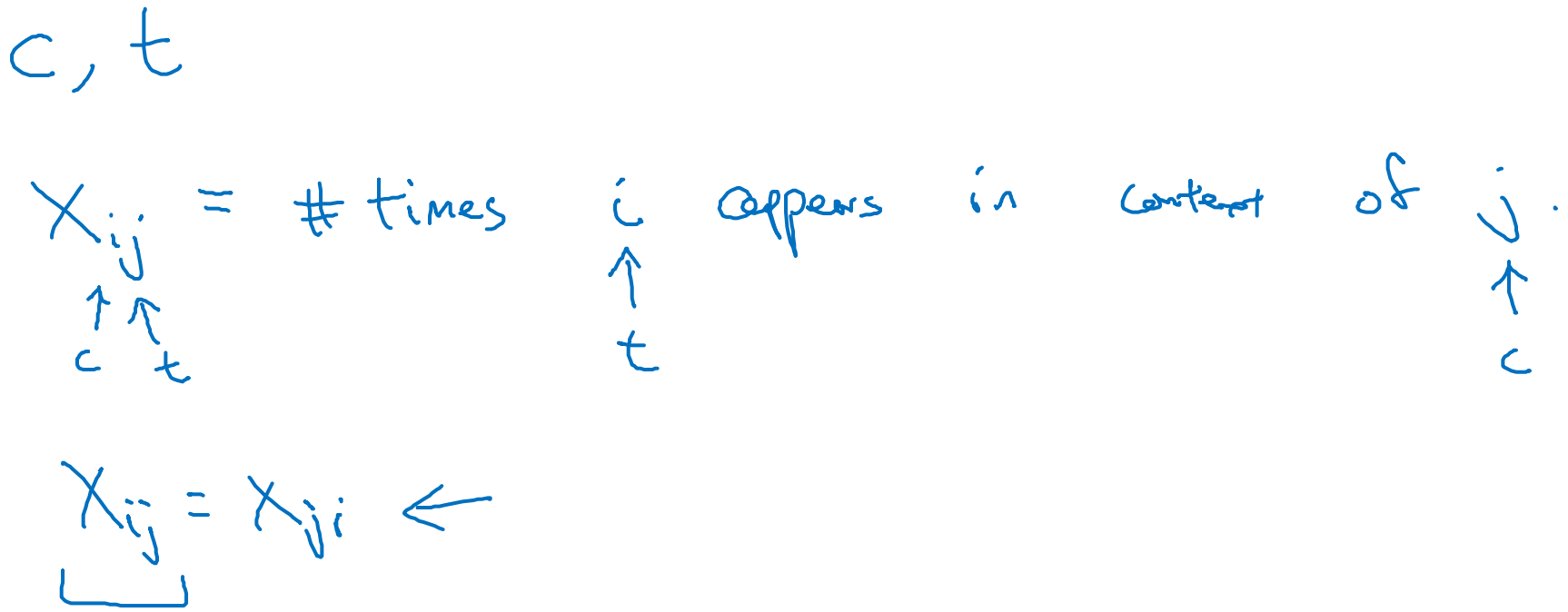
deeplearning.ai

NLP and Word Embeddings

GloVe word vectors

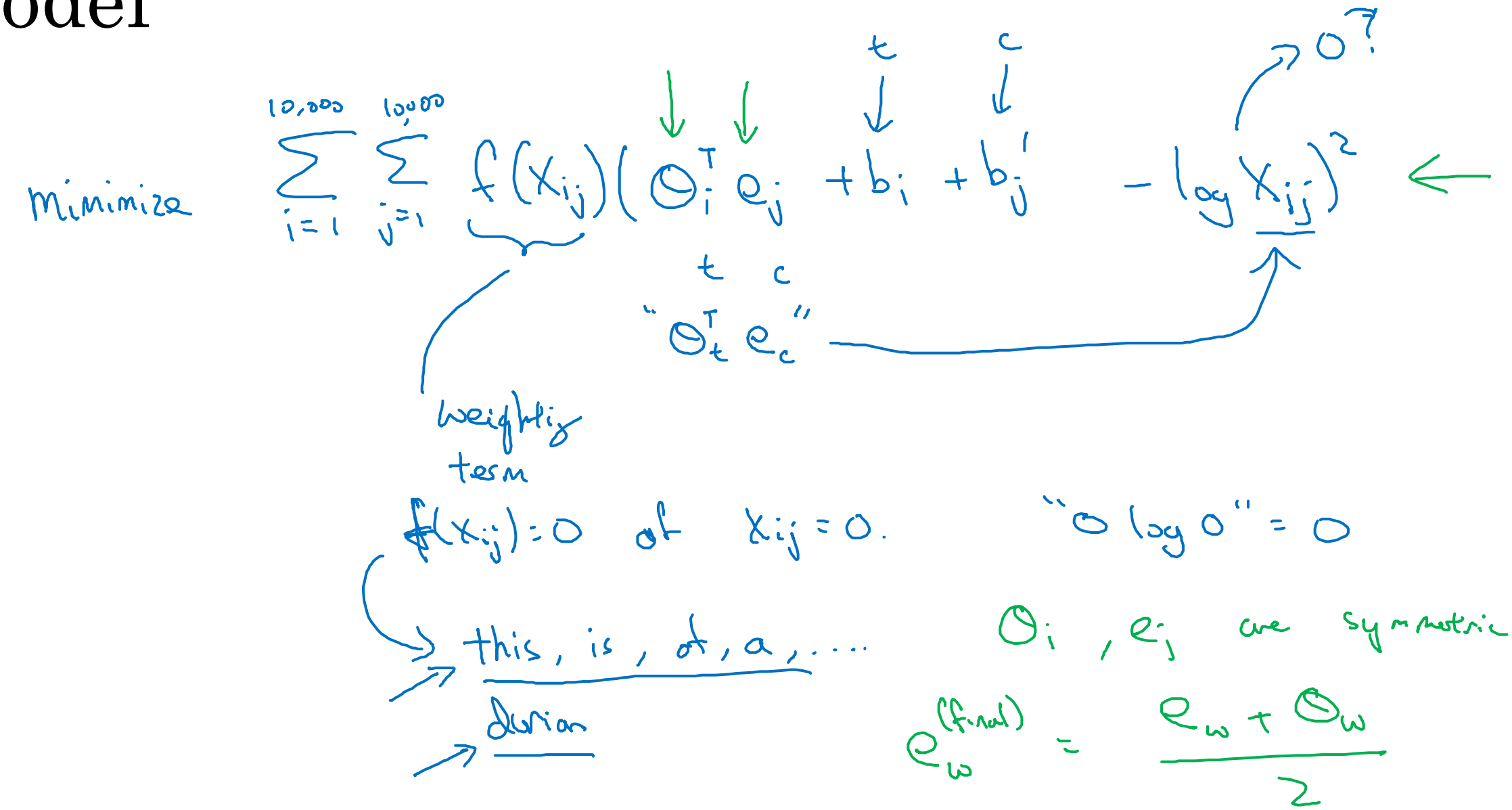
GloVe (global vectors for word representation)

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



[Pennington et. al., 2014. GloVe: Global vectors for word representation]

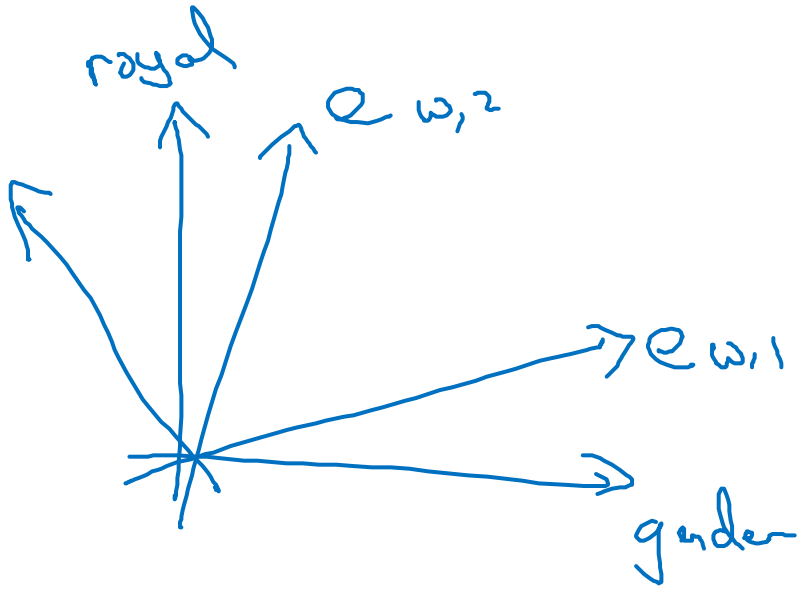
Model



Andrew Ng

A note on the featurization view of word embeddings

	Man (5391)	Woman (9853)	King (4914)	Queen (7157)	
Gender	−1	1	-0.95	0.97	←
Royal	0.01	0.02	0.93	0.95	←
Age	0.03	0.02	0.70	0.69	←
Food	0.09	0.01	0.02	0.01	←
...					



$$\text{minimize} \sum_{i=1}^{10,000} \sum_{j=1}^{10,000} f(X_{ij}) \left(\theta_i^T e_j + b_i - b_j' - \log X_{ij}\right)^2$$

$(A\theta_i)^T (A^{-T} e_j) = \theta_i^T A^T A^{-T} e_j$

deeplearning.ai

NLP and Word Embeddings

Sentiment classification

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. ★☆☆☆☆

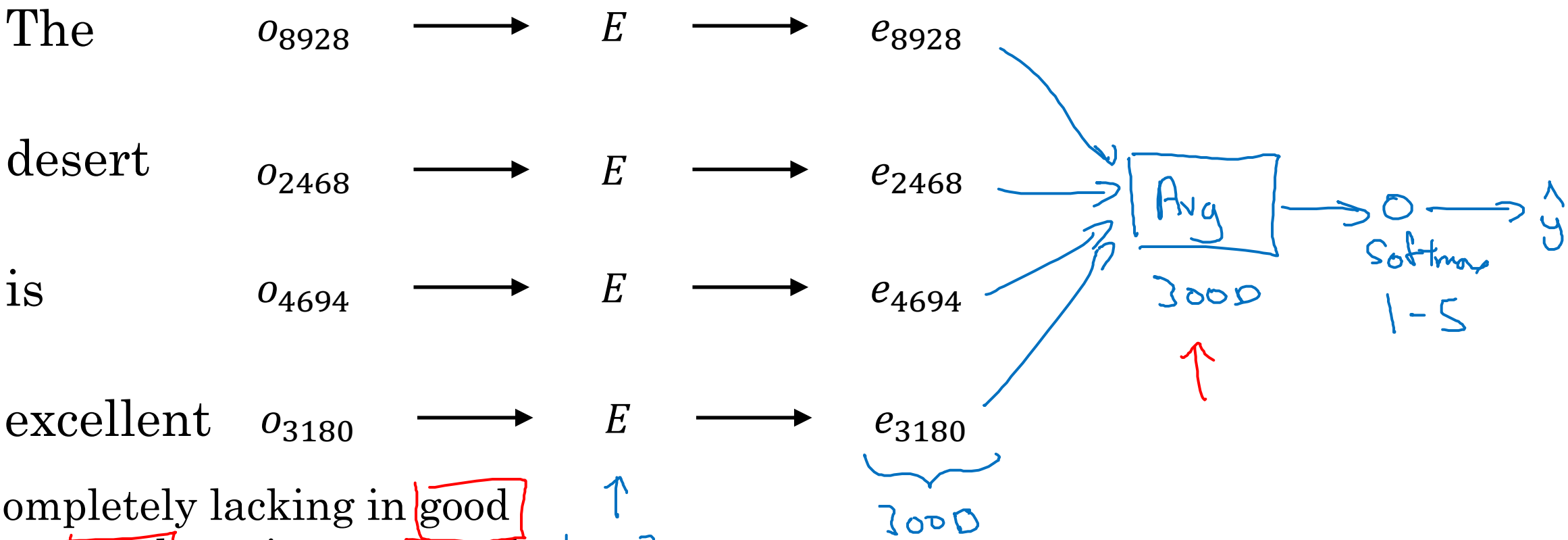
10,000 → 100,000 words

Simple sentiment classification model

The dessert is excellent

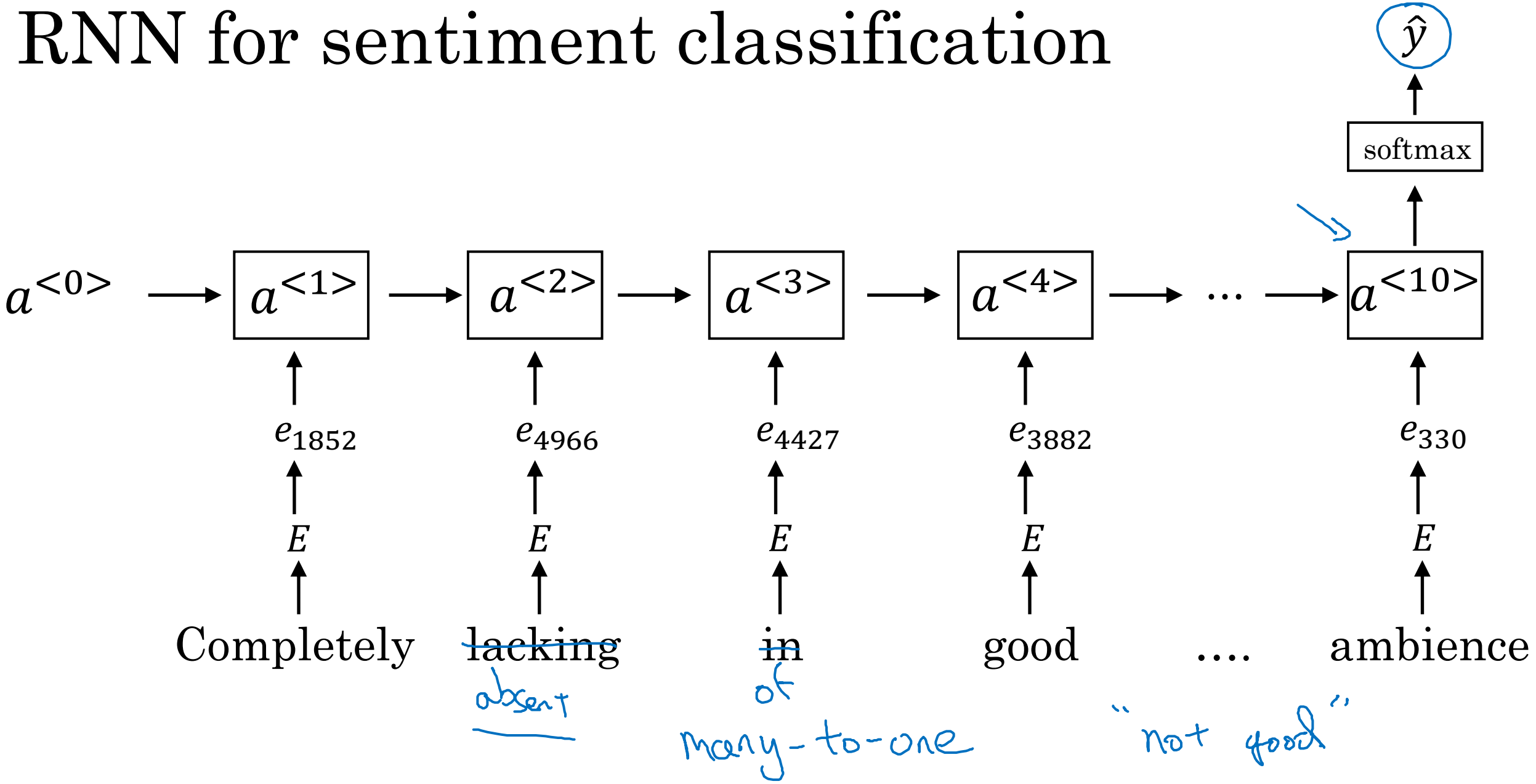
★★★★☆

8928 2468 4694 3180



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

RNN for sentiment classification



deeplearning.ai

NLP and Word Embeddings

Debiasing word embeddings

The problem of bias in word embeddings

Man:Woman as King:Queen

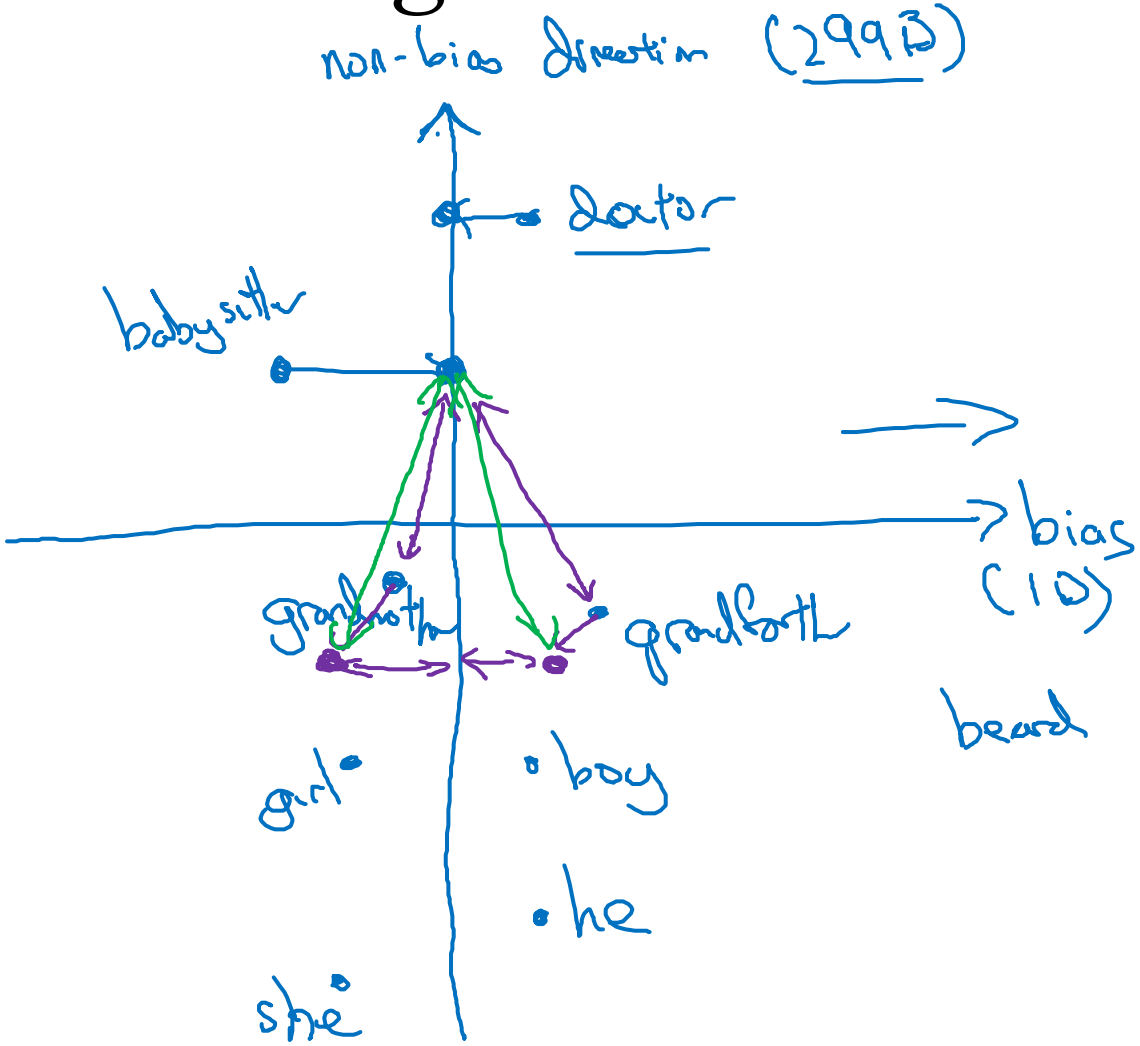
Man:Computer_Programmer as Woman:Homemaker

Father:Doctor as Mother:Nurse

Word embeddings can reflect gender, ethnicity, age, sexual orientation, and other biases of the text used to train the model.

[Bolukbasi et. al., 2016. Man is to computer programmer as woman is to homemaker? Debiasing word embeddings]

Addressing bias in word embeddings



1. Identify bias direction.

$e_{he} - e_{she}$

$e_{male} - e_{female}$

⋮

average

2. Neutralize: For every word that is not definitional, project to get rid of bias.

3. Equalize pairs.

→ grandmother — grandfather

girl — boy

[Bolukbasi et. al., 2016. Man is to computer programmer as woman is to homemaker? Debiasing word embeddings]