



deeplearning.ai

# NLP and Word Embeddings

---

## Word representation

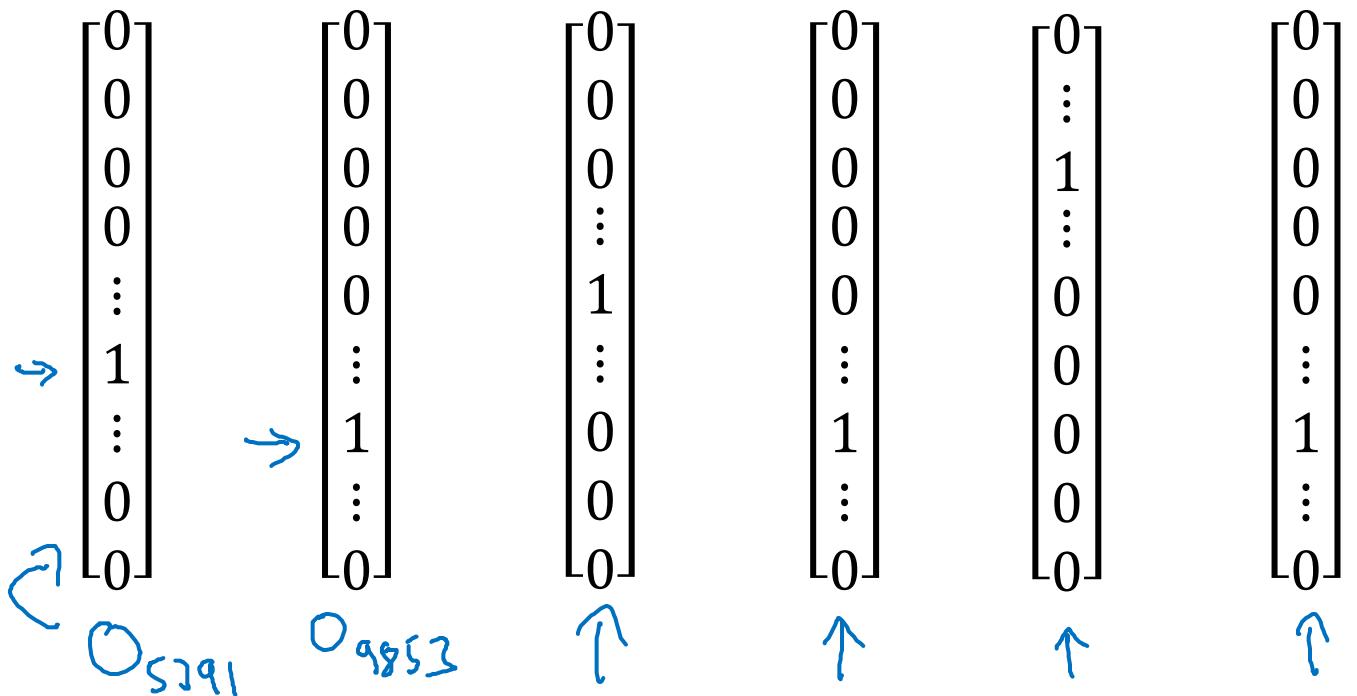
# Word representation

$$V = [a, \text{aaron}, \dots, \text{zulu}, \text{<UNK>}]$$

$$|V| = 10,000$$

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	0.62	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						
Color						
Verb						

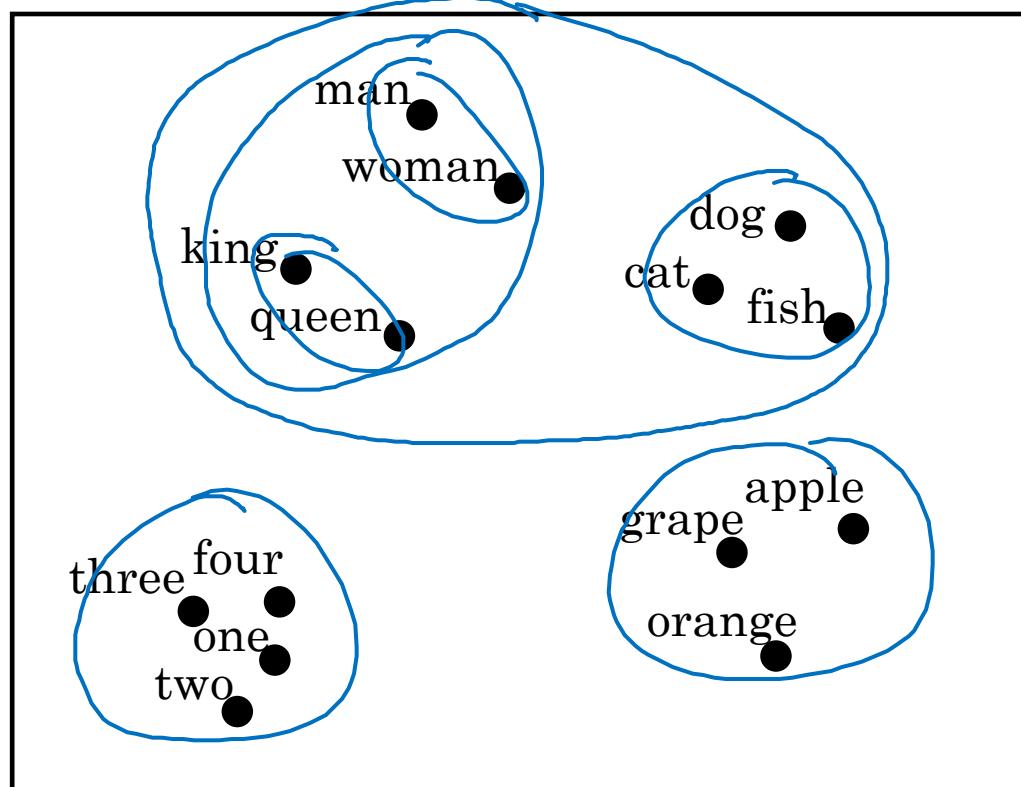
e<sub>5391</sub>      e<sub>9853</sub>

I want a glass of orange juice.

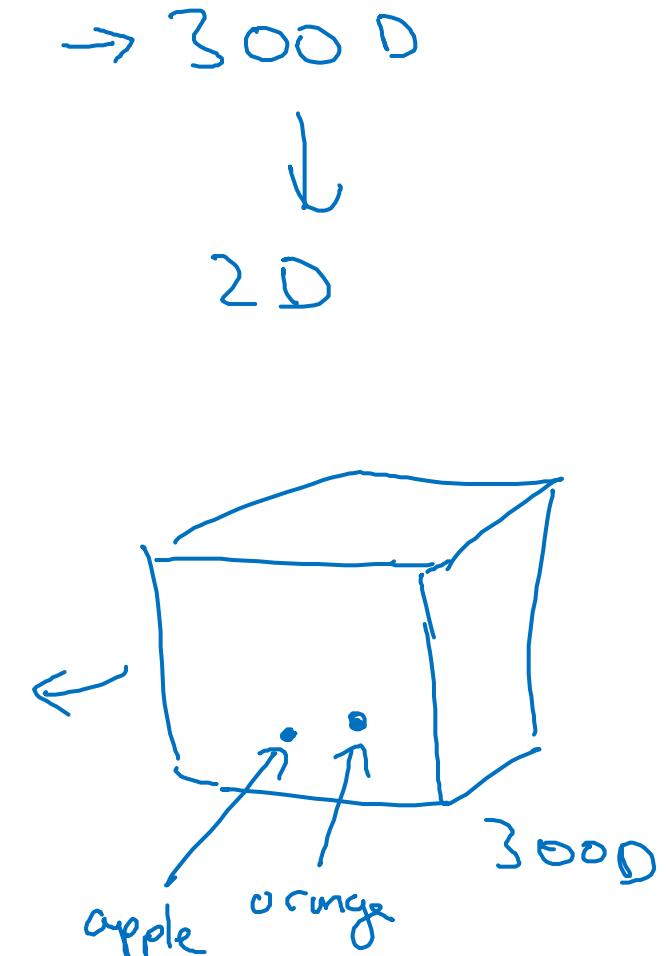
I want a glass of apple juice.

Andrew N

# Visualizing word embeddings



t-SNE





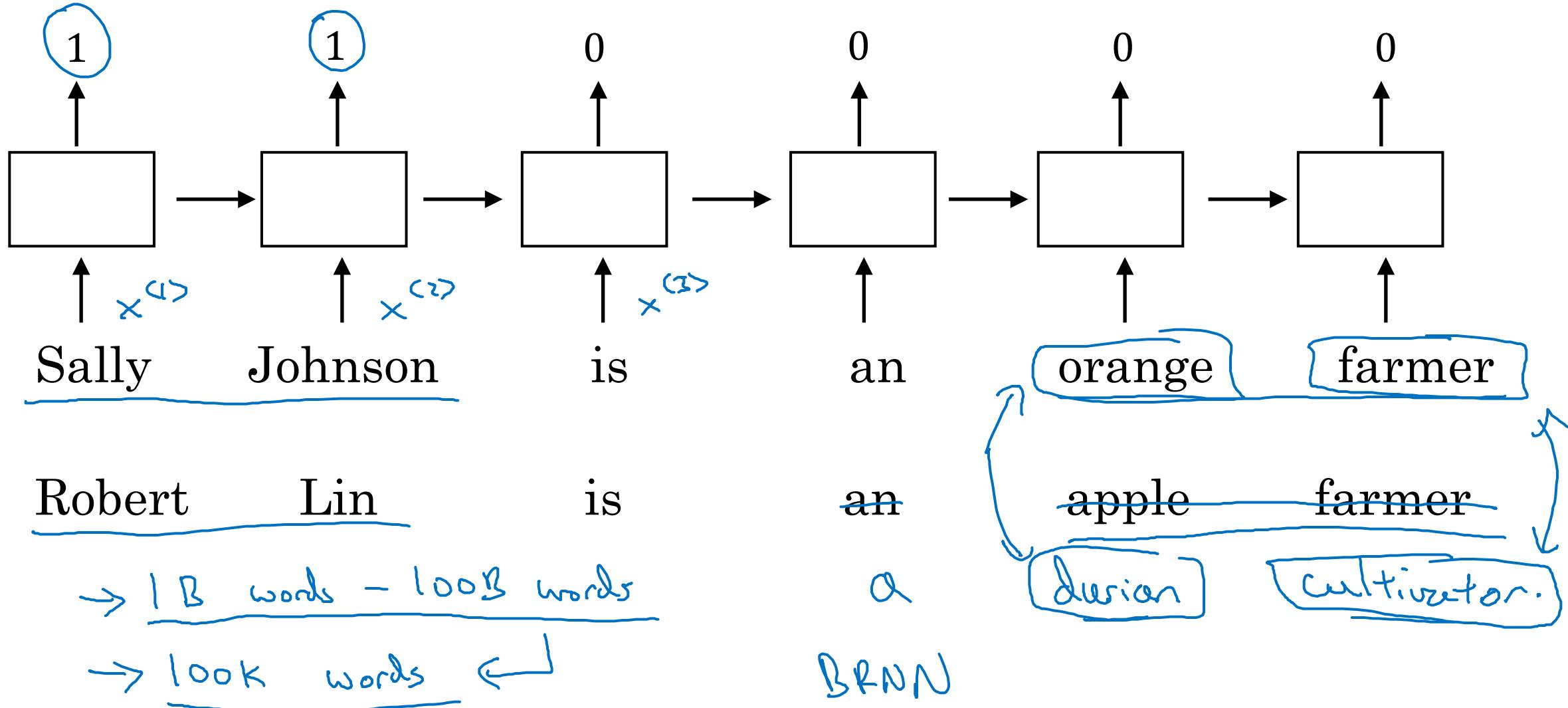
deeplearning.ai

# NLP and Word Embeddings

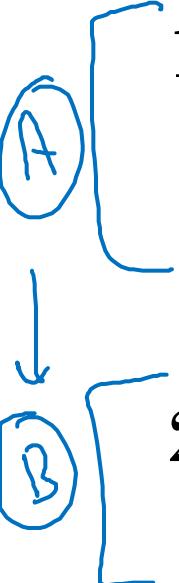
---

## Using word embeddings

# Named entity recognition example



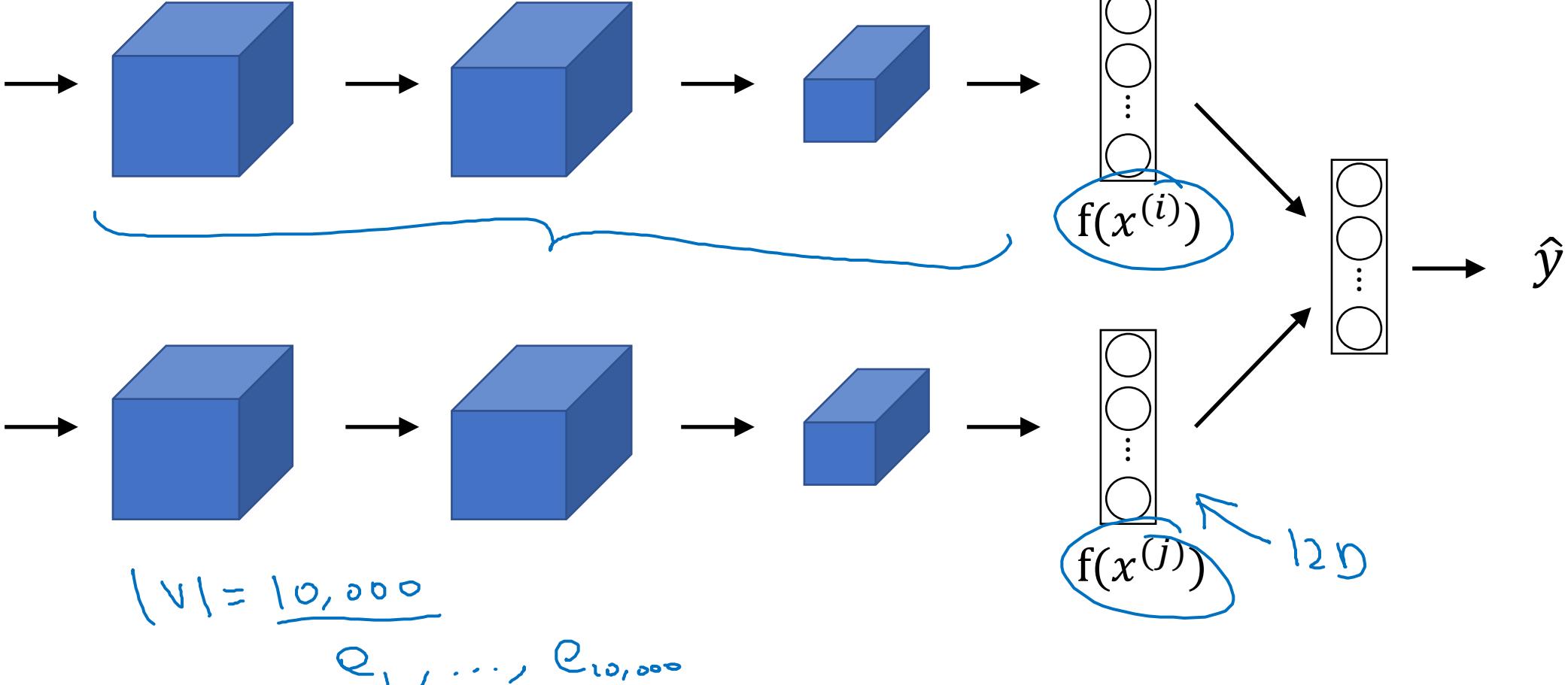
# Transfer learning and word embeddings

- 
1. Learn word embeddings from large text corpus. (1-100B words)  
(Or download pre-trained embedding online.)
  2. Transfer embedding to new task with smaller training set.  
(say, 100k words)  $\rightarrow 10,000 \quad \rightarrow 300$
  3. Optional: Continue to finetune the word embeddings with new data.

# Relation to face encoding (embedding) 128D



$x^{(i)}$





deeplearning.ai

# NLP and Word Embeddings

---

## Properties of word embeddings

# Analogy

	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

$$\begin{matrix} e_{5391} \\ e_{\text{man}} \end{matrix}$$

$$\underline{\text{Man} \rightarrow \text{Woman}}$$

$$e_{\text{man}} - e_{\text{woman}}$$

$$e_{\text{woman}}$$

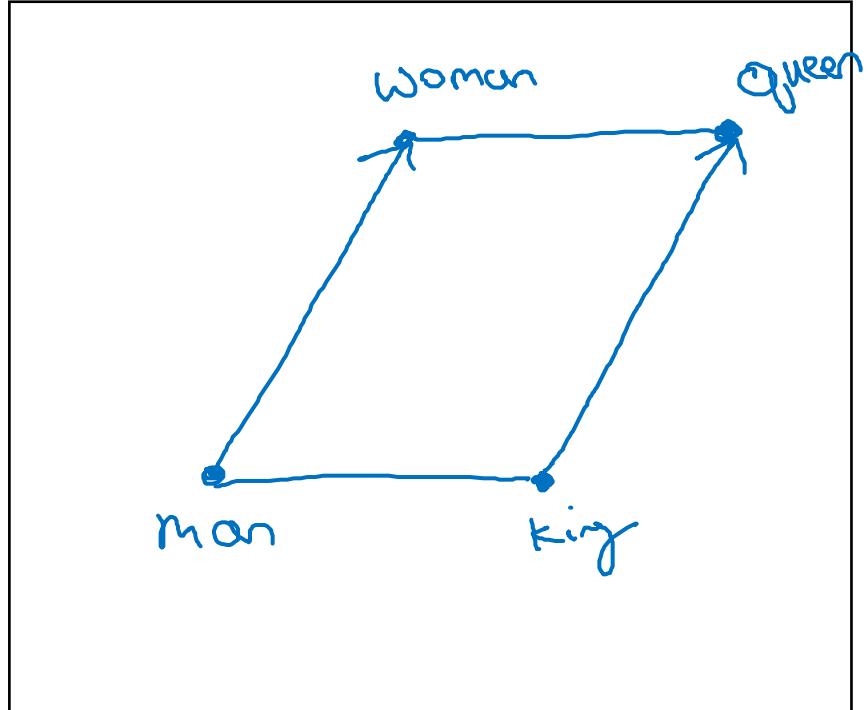
$$\underline{\text{King} \rightarrow ? \text{ Queen}}$$

$$e_{\text{king}} - e_{? \text{ Queen}}$$

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

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

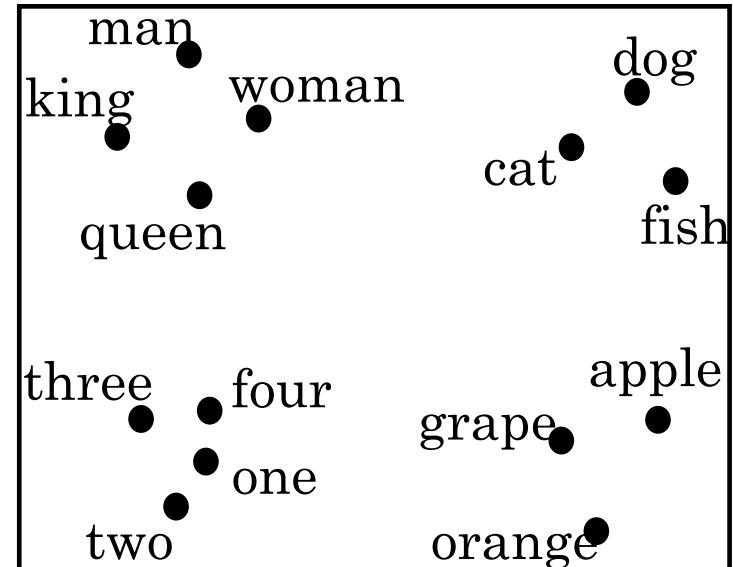
# Analogies using word vectors



300 D

Find word  $w_i : \arg \max_w$

300D  $\rightarrow$  2D  
↑



t-SNE

$$e_{\text{man}} - e_{\text{woman}} \approx e_{\text{king}} - e_{\underline{\text{?}}} e_w$$

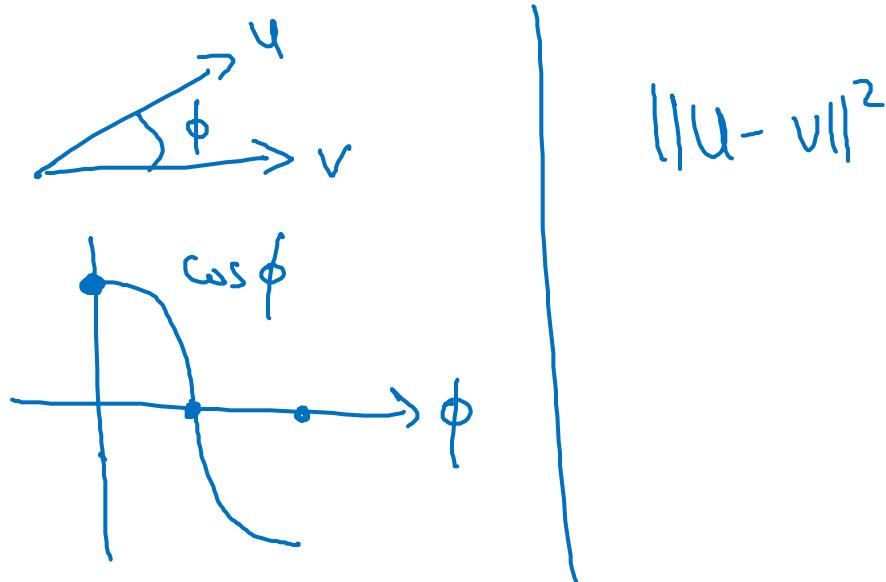
Sim( $e_w$ ,  $e_{\text{king}} - e_{\text{man}} + e_{\text{woman}}$ )

30 - 75%

# Cosine similarity

$$\rightarrow \boxed{\text{sim}(e_w, e_{king} - e_{man} + e_{woman})}$$

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



Man:Woman as Boy:Girl  
Ottawa:Canada as Nairobi:Kenya  
Big:Bigger as Tall:Taller  
Yen:Japan as Ruble:Russia



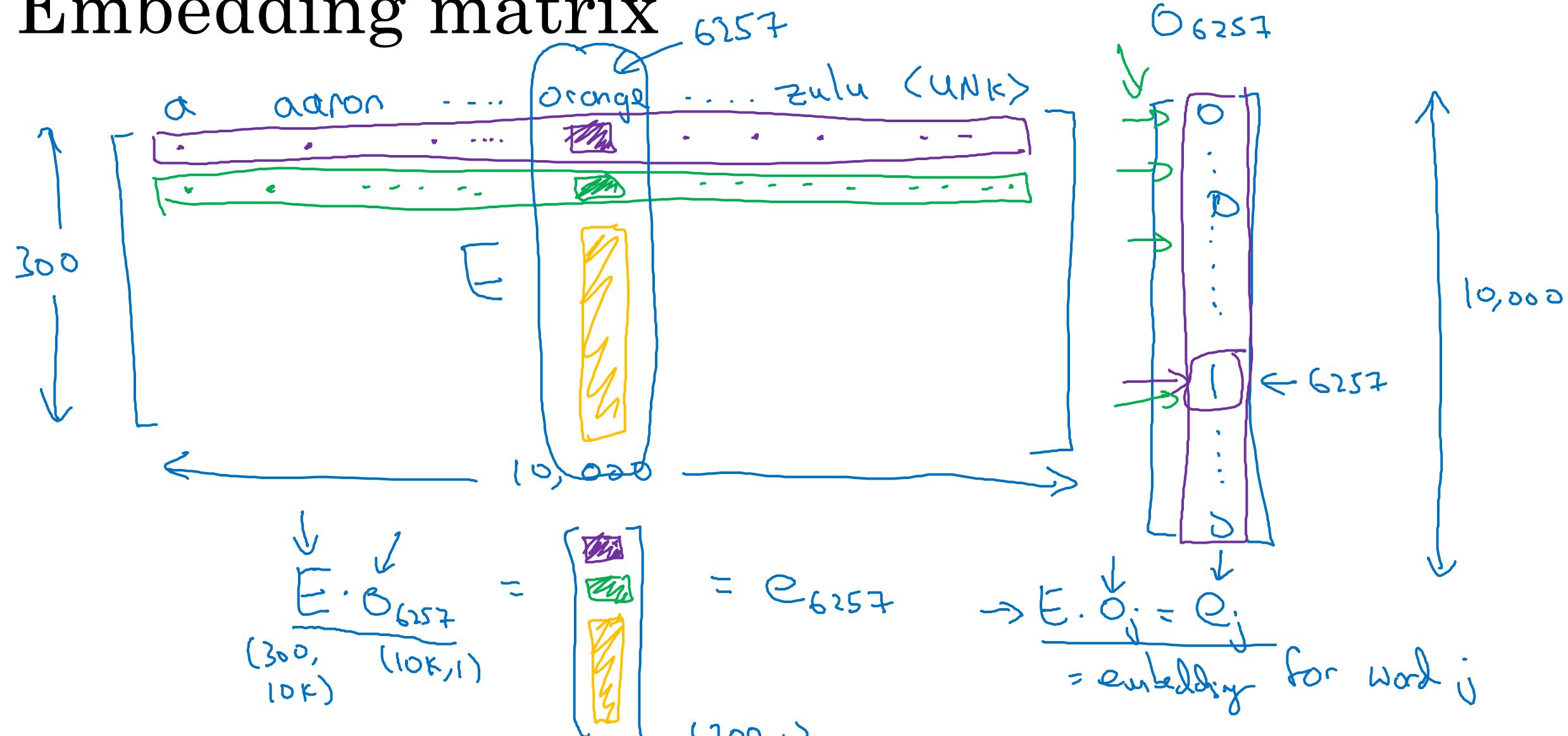
deeplearning.ai

# NLP and Word Embeddings

---

## Embedding matrix

# Embedding matrix



In practice, use specialized function to look up an embedding.  
 $\rightarrow$  Embedding



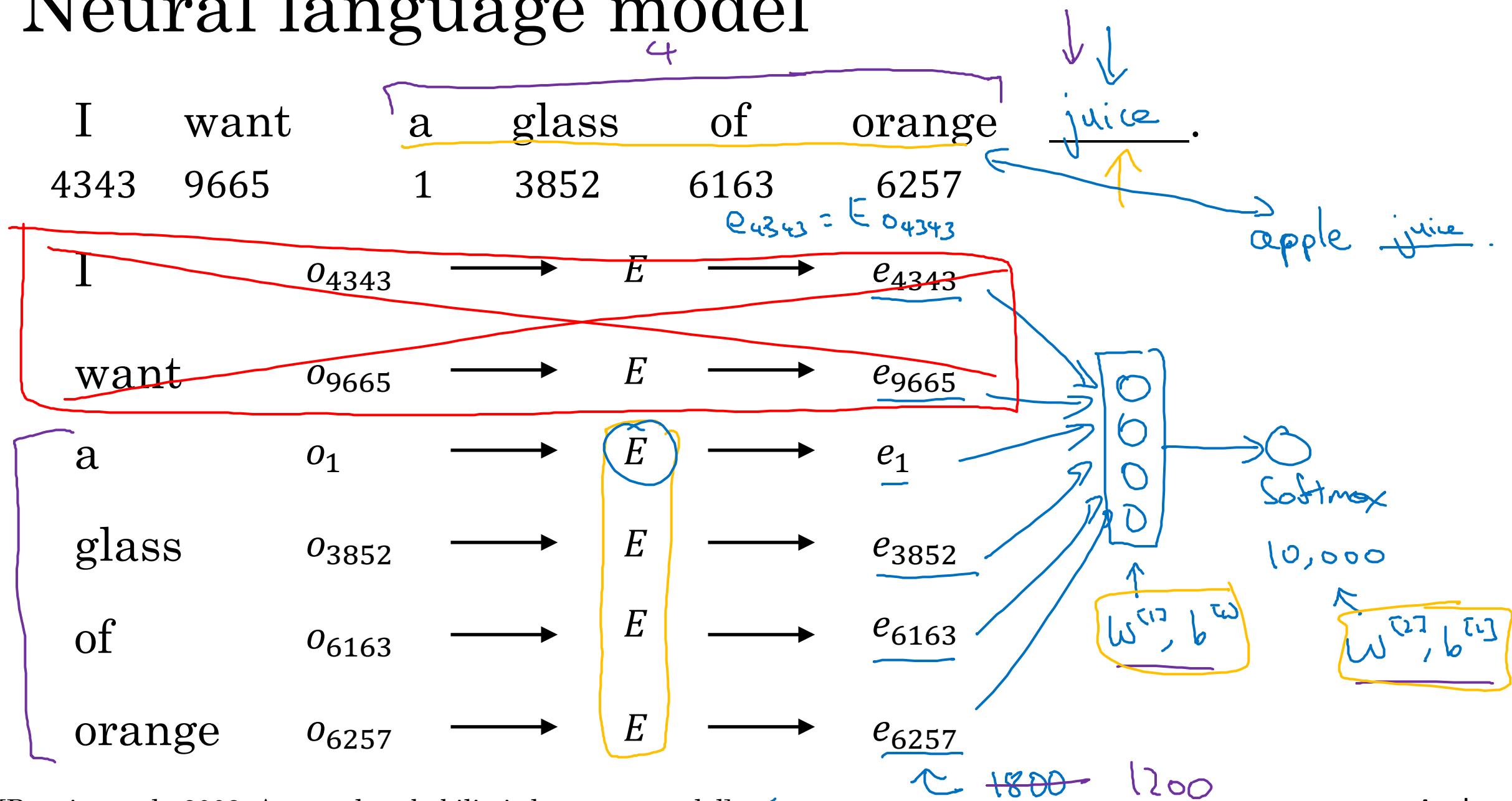
deeplearning.ai

# NLP and Word Embeddings

---

## Learning word embeddings

# Neural language model



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

4

# Andrew Ng

# Other context/target pairs

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

Context: Last 4 words.

4 words on left & right

Last 1 word

Nearby 1 word

skip gram

a glass of orange ? to go along with

orange ?

glass . ?



deeplearning.ai

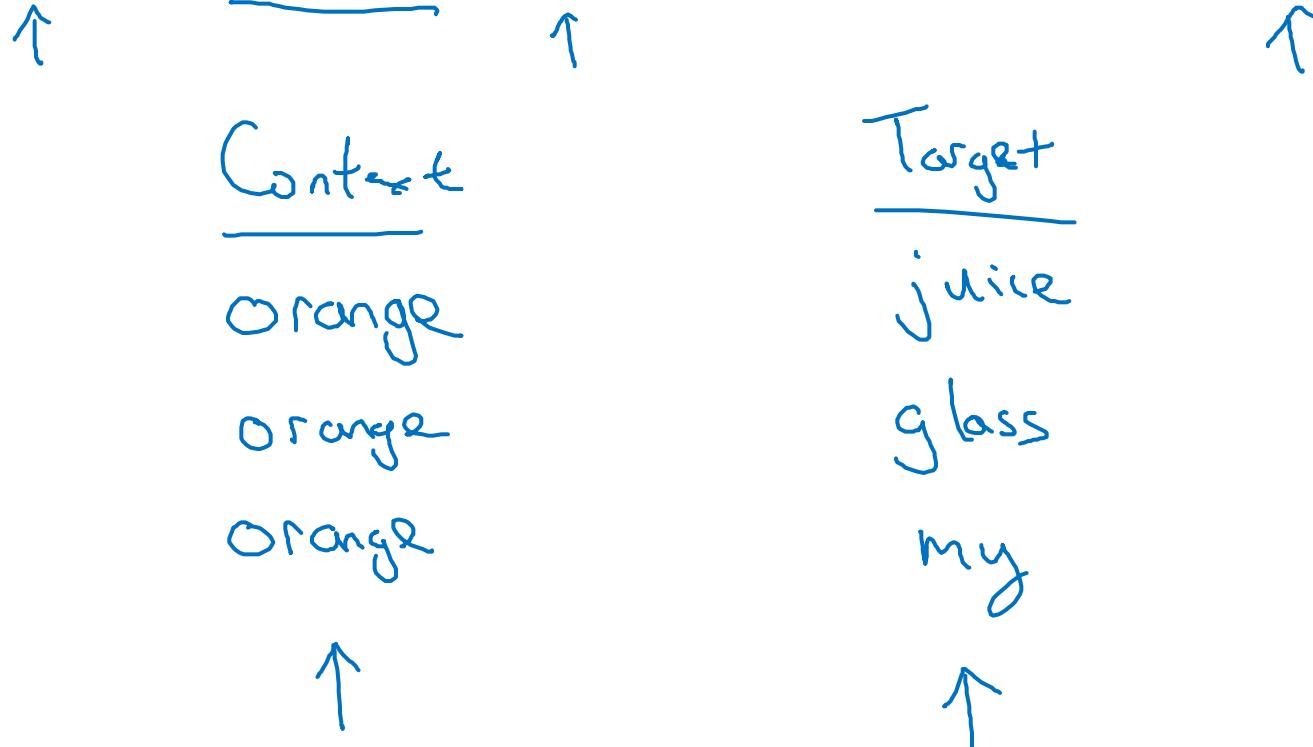
# NLP and Word Embeddings

---

## Word2Vec

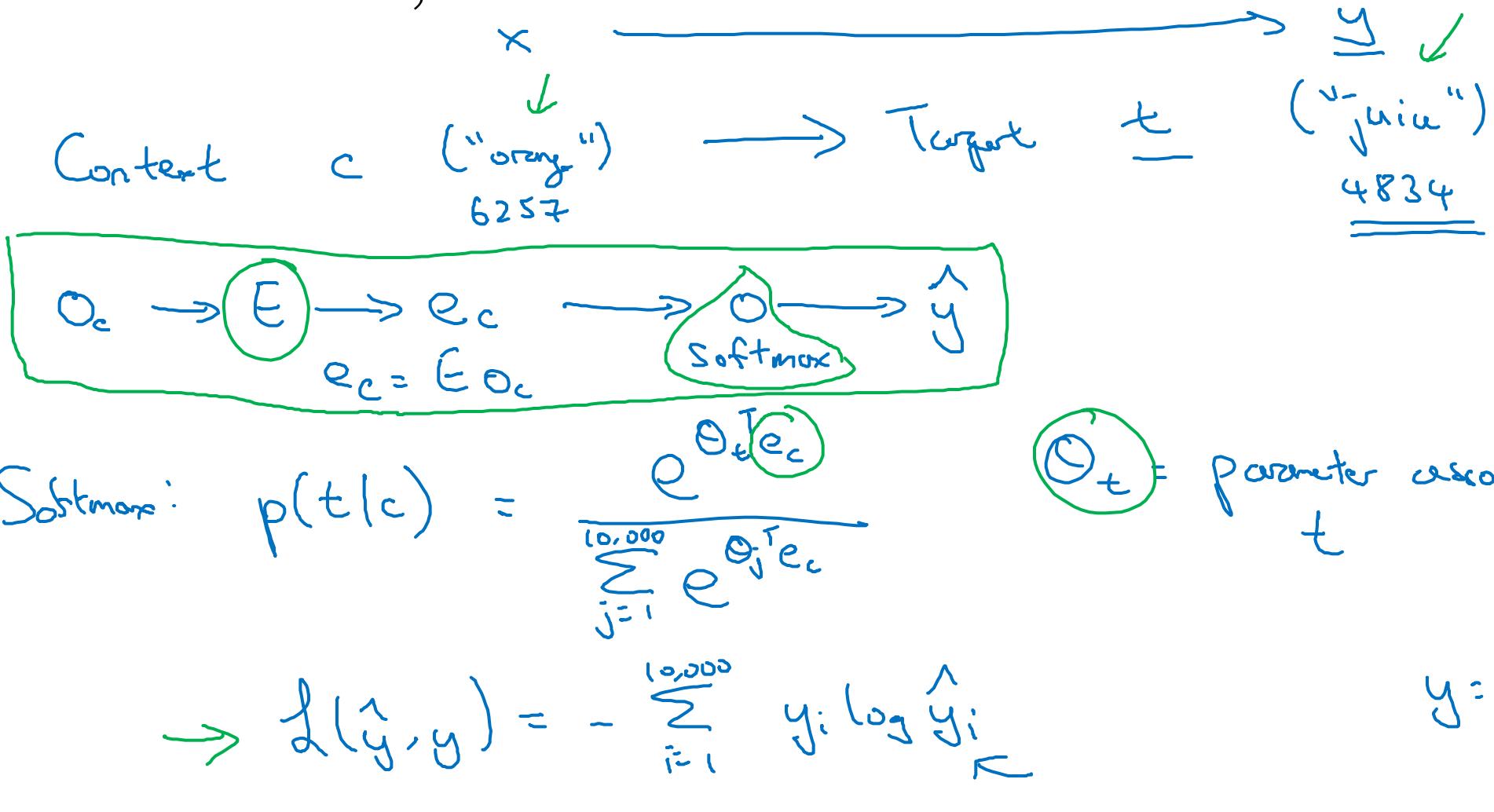
# Skip-grams

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



# Model

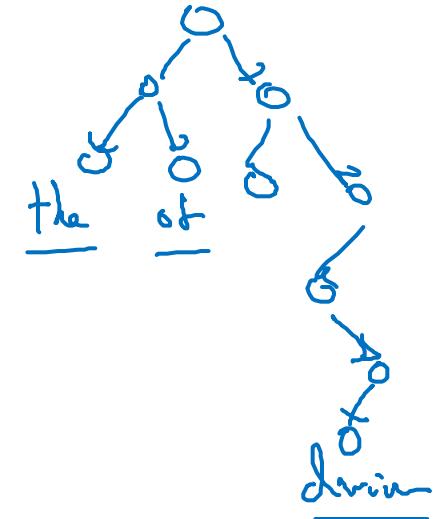
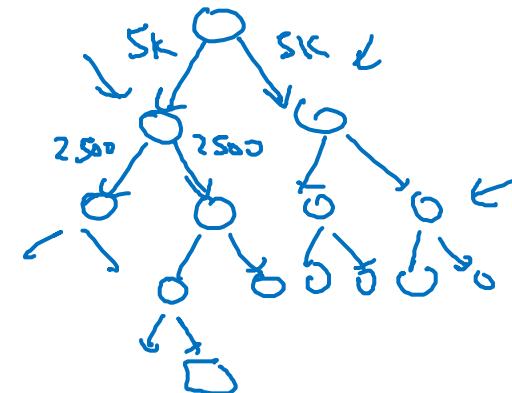
Vocab size = 10,000k



# 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.



How to sample the context  $c$ ?

→ the, of, a, and, to, ...

→ orange, apple, durian

$P_{\text{durian}}$

$t$   
 $c \rightarrow t$

$P(c)$



deeplearning.ai

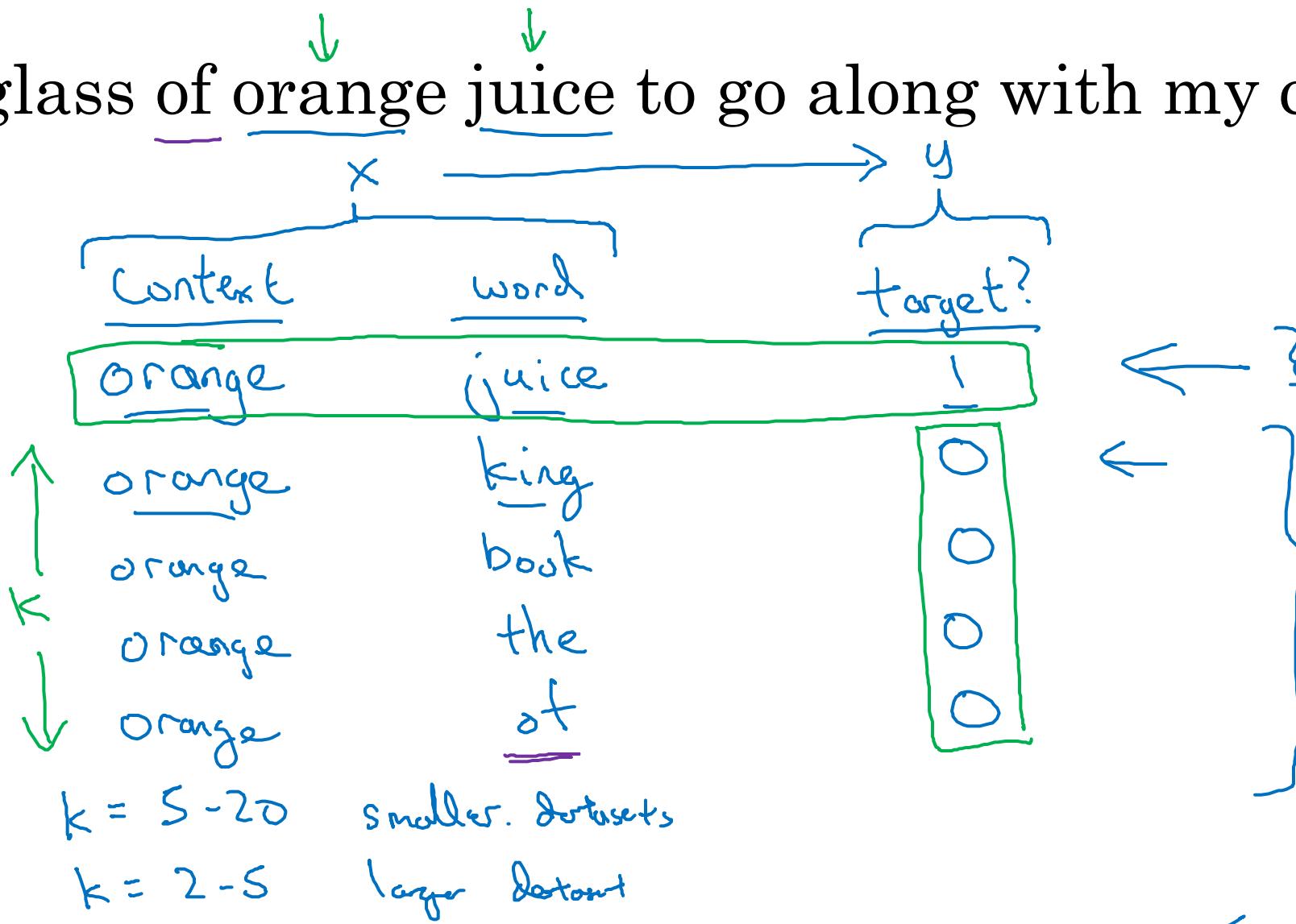
# NLP and Word Embeddings

---

## Negative sampling

# Defining a new learning problem

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



# Model

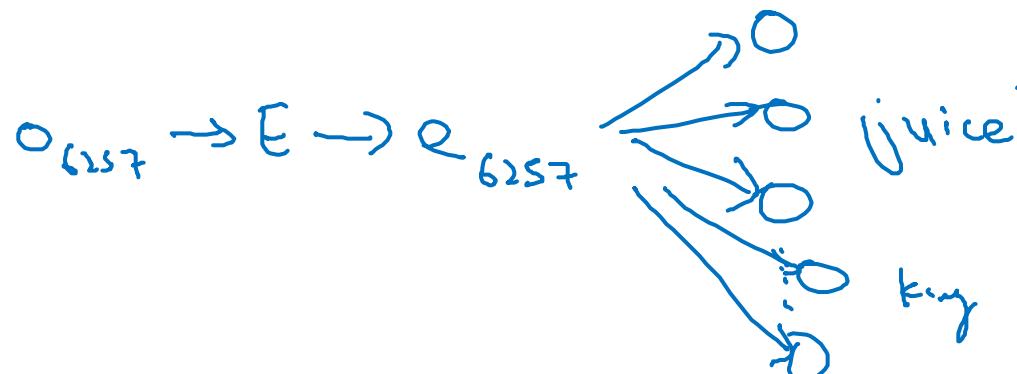
Softmax:

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

10,000-way softmax

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

Orange  
6257



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

10,000  
↓

10,000 binary  
classification  
problem

k+1

# Selecting negative examples

<u>context</u>	<u>word</u>	<u>target?</u>
orange	juice	1
orange	king	0
orange	book	0
orange	the	0
orange	of	0

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.

c, t

$x_{i,j} = \# \text{ times } i \text{ appears in context of } j.$

$x_{i,j}$        $i$        $j$   
↑      ↑      ↑  
c      t      c

$$x_{ij} = x_{ji} \leftarrow$$



## Model

Minimize

$$\sum_{i=1}^{10,000} \sum_{j=1}^{100,000} f(x_{ij}) (\Theta_i^T e_j + b_i + b_j - \log \frac{x_{ij}}{c})$$

$\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow$

$t \quad c \quad t \quad c \quad t \quad c$

$\Theta_t^T e_c$

*weightij term*

$f(x_{ij}) = 0 \text{ or } x_{ij} = 0. \quad "0 \log 0" = 0$

*this, is, of, a, ...*

*derian*

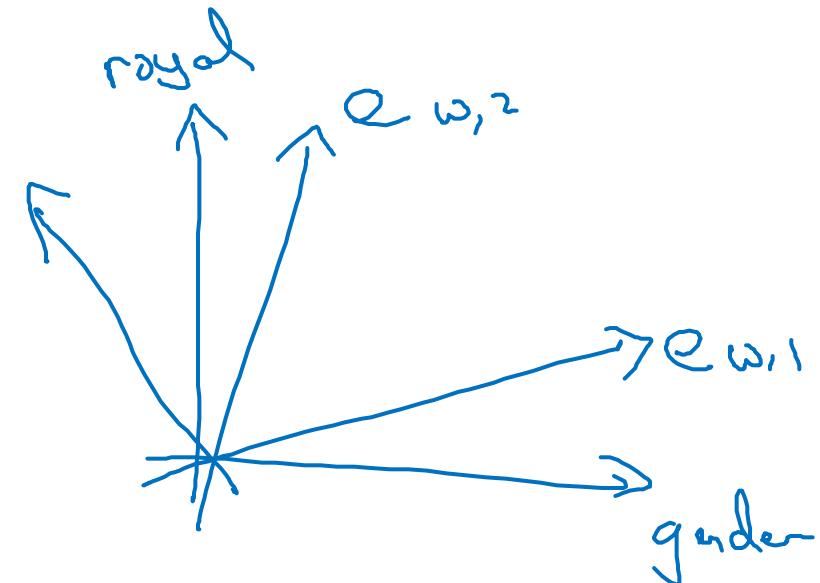
$\Theta_w^{(\text{final})} = \frac{\Theta_w + \Theta_w}{2}$

*$\Theta$ ?*

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



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

$$\left\langle (A\theta_i)^T (A^T e_j) \right\rangle = \cancel{\theta_i^T A^T A} \cancel{\theta_i^T} e_j$$



deeplearning.ai

# NLP and Word Embeddings

---

## Sentiment classification

# Sentiment classification problem



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



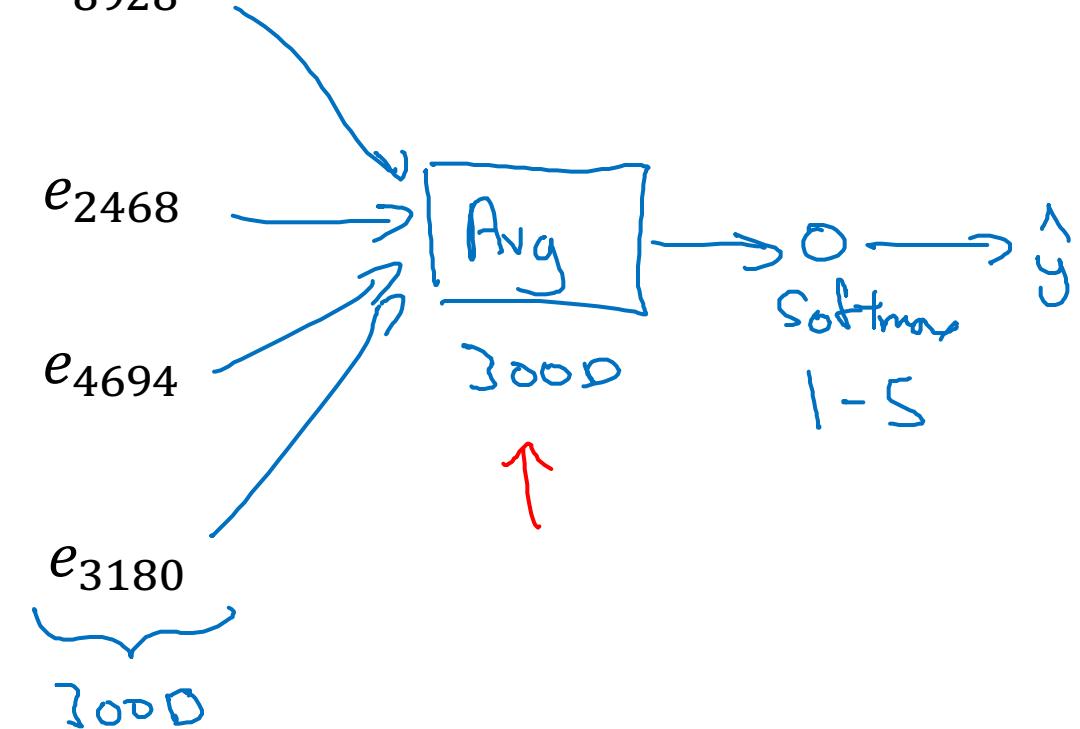
The  $o_{8928}$   $\rightarrow E \rightarrow e_{8928}$

desert  $o_{2468} \rightarrow E \rightarrow e_{2468}$

is  $o_{4694} \rightarrow E \rightarrow e_{4694}$

excellent  $o_{3180} \rightarrow E \rightarrow e_{3180}$

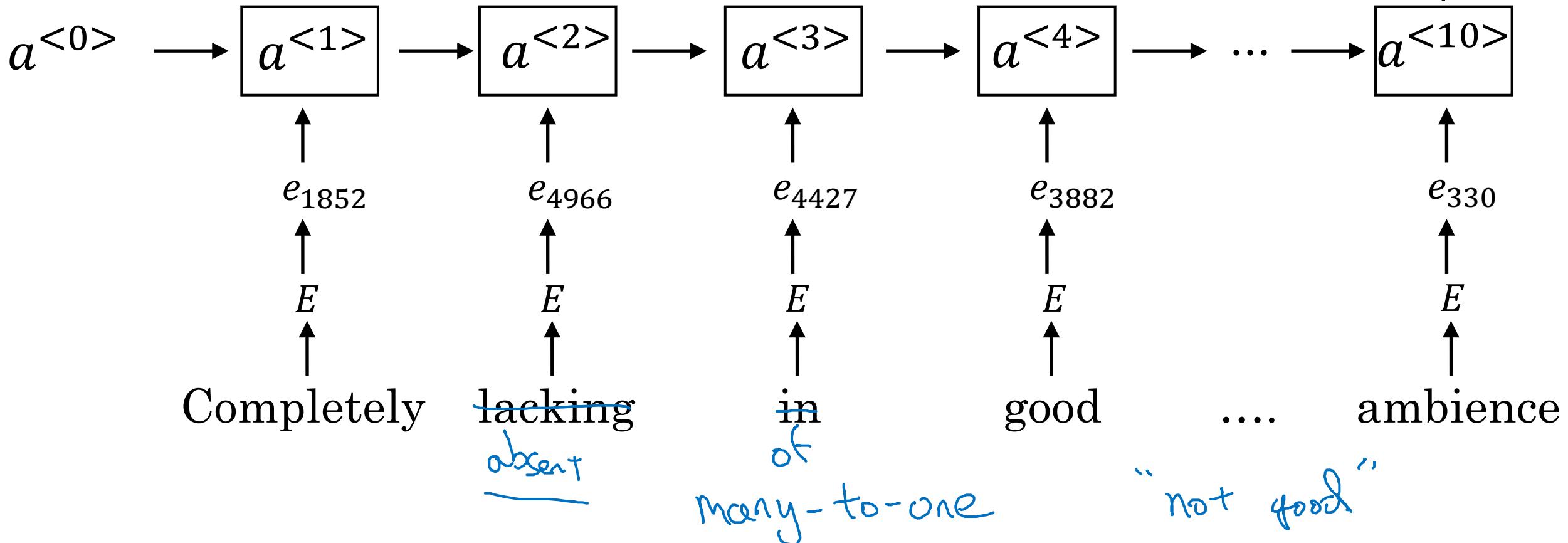
“Completely lacking in good taste, good service, and good ambience.”  
↑  
100 B words



# RNN for sentiment classification

$\hat{y}$

softmax





deeplearning.ai

# NLP and Word Embeddings

---

## Debiasing word embeddings

# The problem of bias in word embeddings

Man:Woman as King:Queen

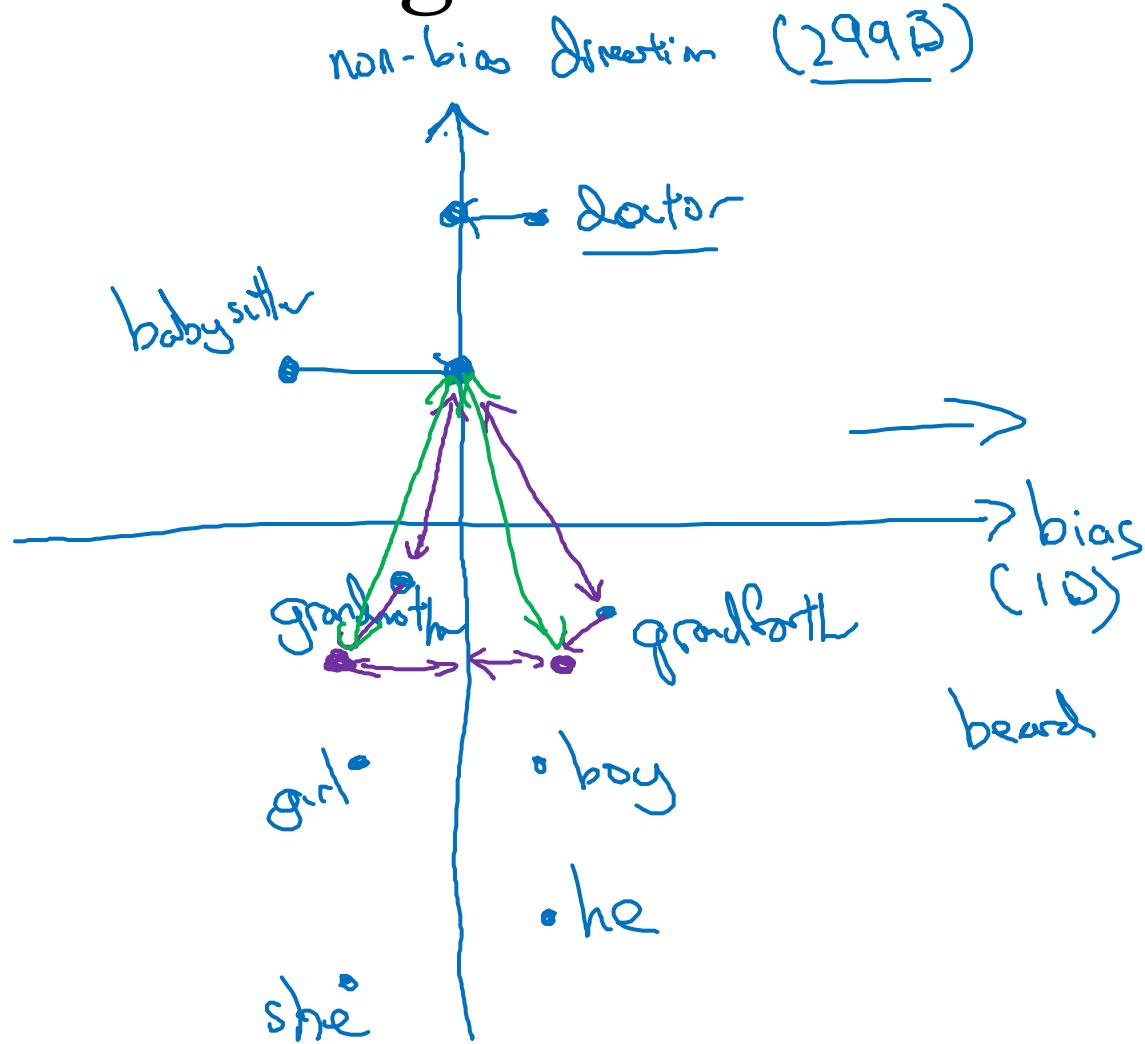
Man:Computer\_Programmer as Woman:Homemaker 

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.



# Addressing bias in word embeddings



1. Identify bias direction.

{ he - she  
male - female  
⋮  
average

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

3. Equalize pairs.

→ grandmother — grandfather  
girl boy