



deeplearning.ai

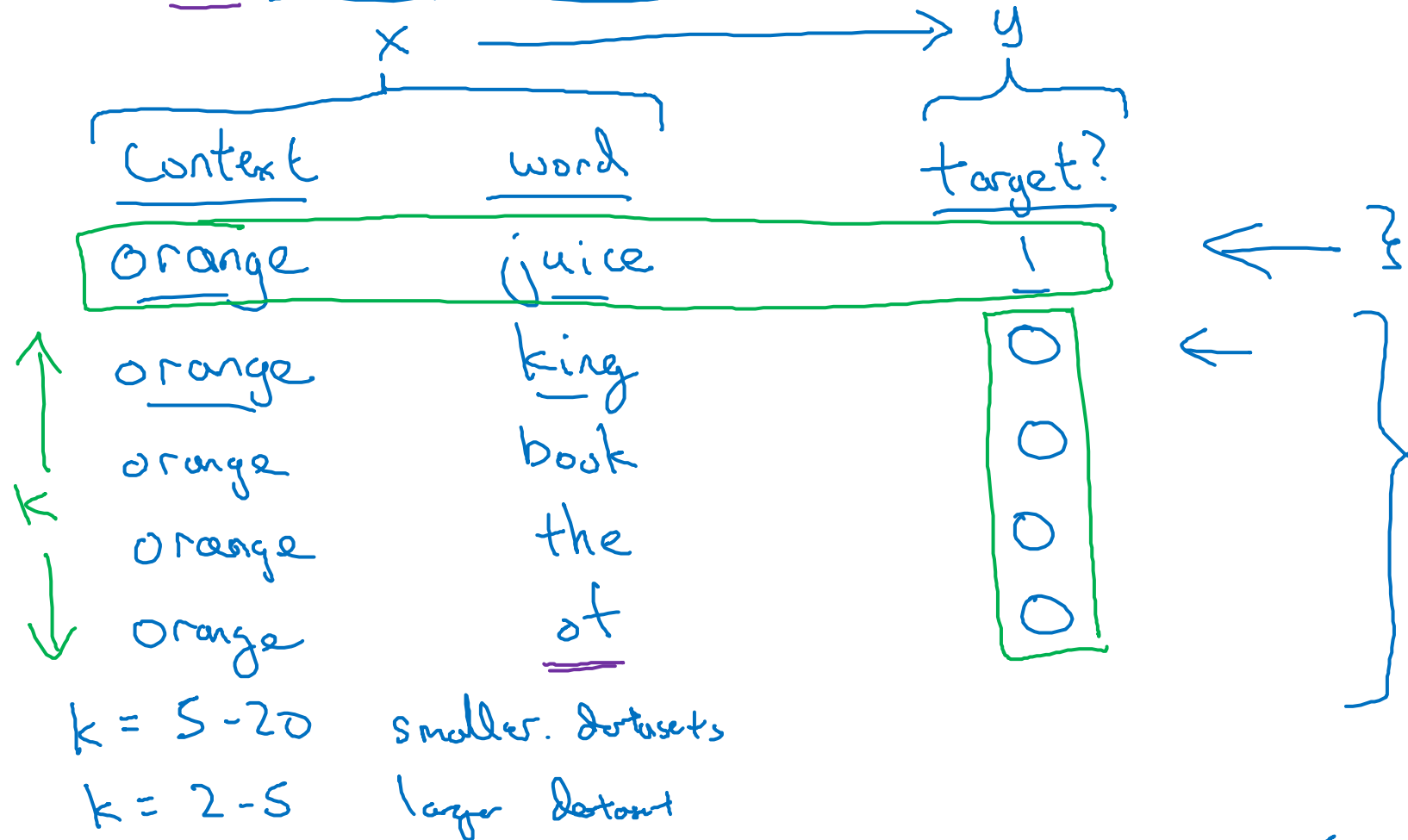
# NLP and Word Embeddings

---

## Negative sampling

# Defining a new learning problem

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



# Model

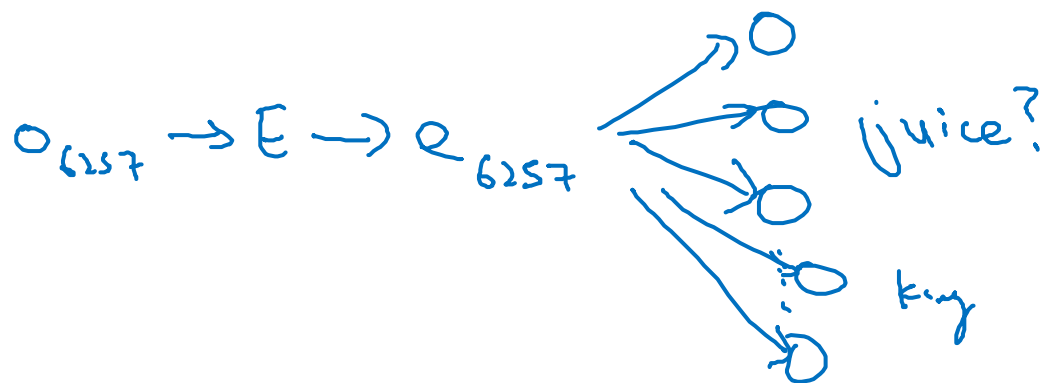
Softmax:

$$p(t|c) = \frac{e^{\theta_t^T e_c}}{\sum_{j=1}^{10,000} e^{\theta_j^T e_c}} \quad \left. \vphantom{\sum_{j=1}^{10,000}} \right\} \begin{array}{l} \downarrow \downarrow \\ 10,000\text{-way} \\ \text{softmax} \end{array}$$

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

$x$		$y$
<u>context</u>	<u>word</u>	<u>target?</u>
orange	juice	1
orange	king	0
orange	book	0
orange	the	0
orange	of	0
$\uparrow$ $c$	$\uparrow$ $t$	$\uparrow$ $y$

Orange  
6257



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

↑  
t

the, of, and, ...

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

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

↑