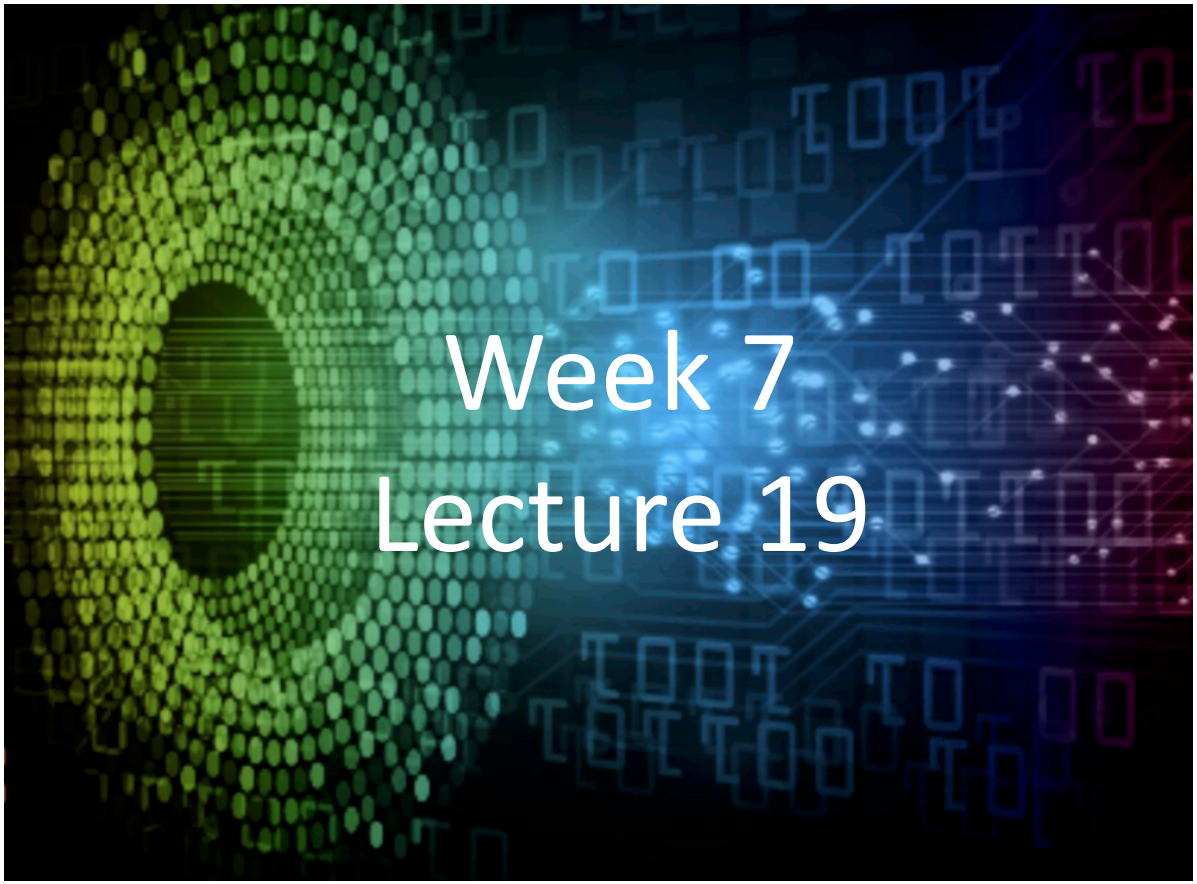
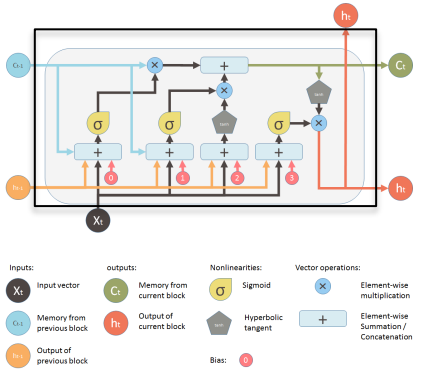
Introduction to Deep Learning
Applications and Theory

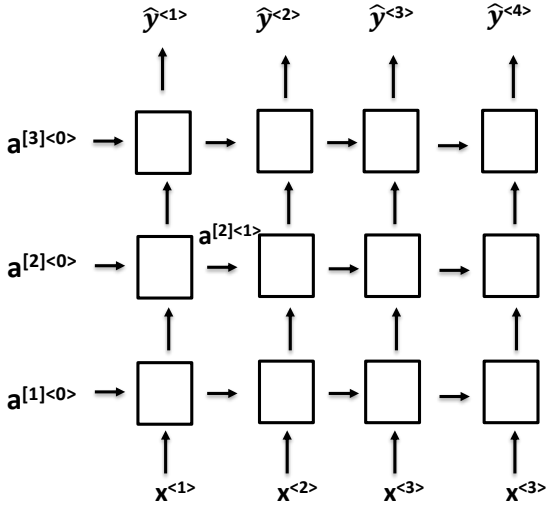


ECE 596 / AMATH 563

Previous Week: Recurrent Neural Networks (RNNs) Setup

- Diminishing/ Exploding Gradients
- Gated RNNs
  - GRU
  - LSTM
- Architectures of RNN Blocks

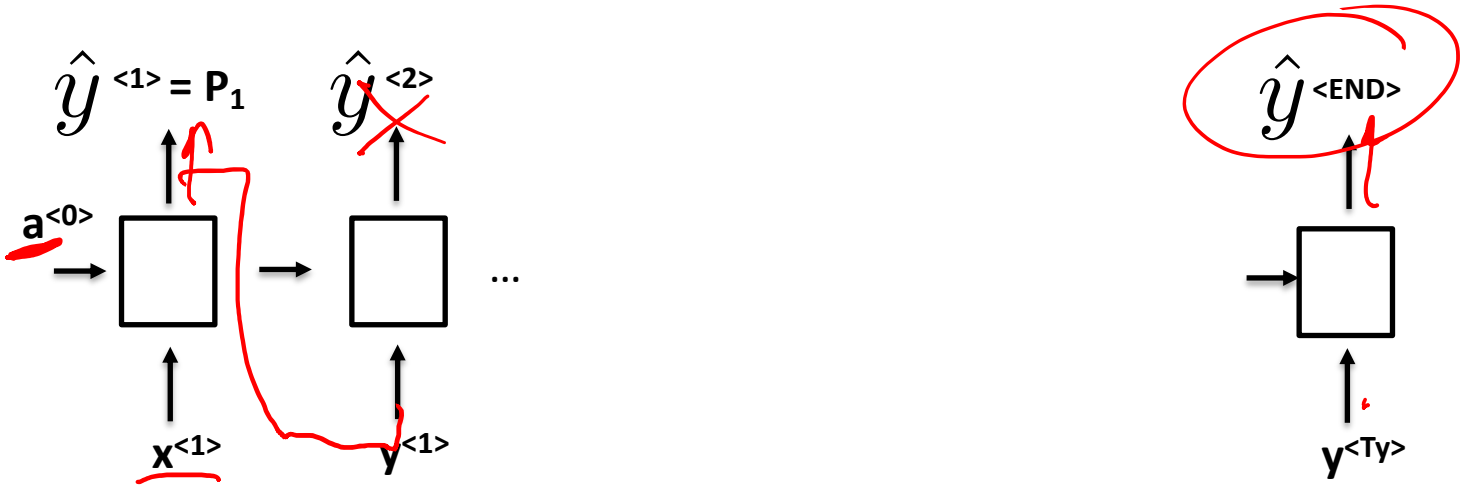




This Week:
Advanced Applications of RNNs

- Today: NLP Applications
  - Word embeddings
  - Word2vec approaches
  - Negative Sampling
  - Sentiment Classification

Named Entity Recognition Example

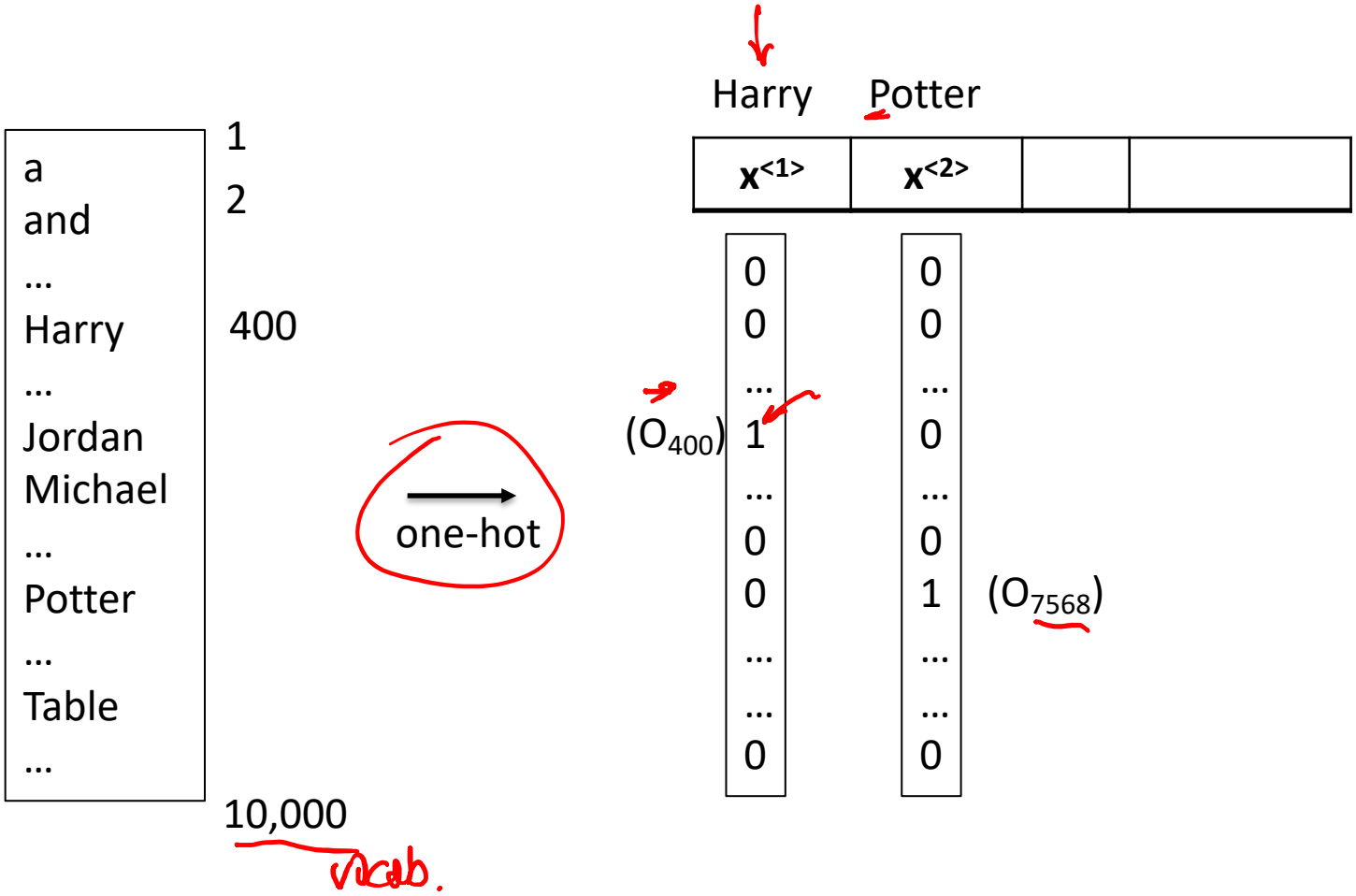


Sally Johnson is a Porsche driver.

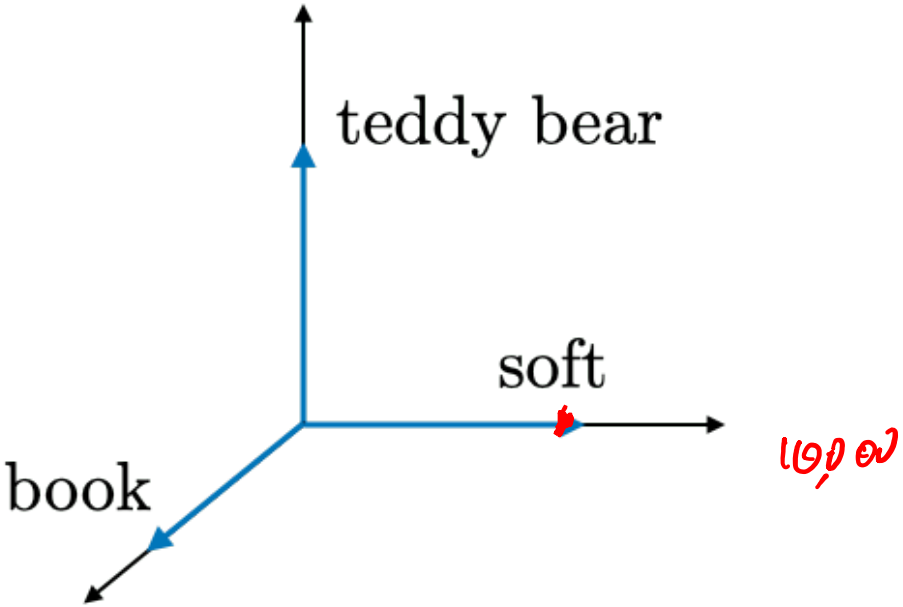
Magic Jordan is a Tesla motorist.

Natural Language Processing

Vocabulary/Dictionary



Representation Techniques



1-hot representation O_w

Word Embeddings

Vocab 10,000

Embedding vector

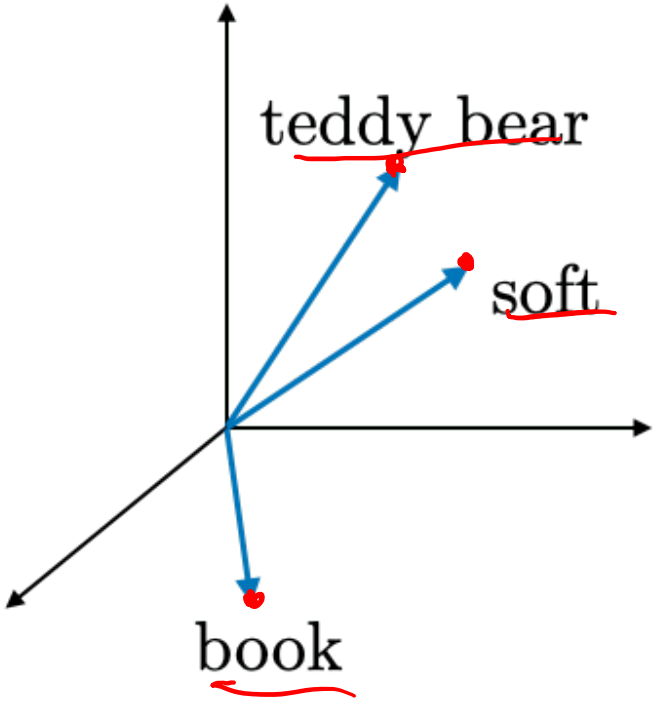
	Man	Woman	King	Queen	Orange	Apple	Tesla	Porsche
Gender	-1	1	-0.95	0.97	0.01	0	-0.5	0.03
Royal	0.01	0.02	0.94	0.95	-0.02	0.04	-0.01	0.1
Food								
Size								
Engine	0.07	-0.01	0.03	0.02	0.001	0	0.95	0.98

<<10,000

300

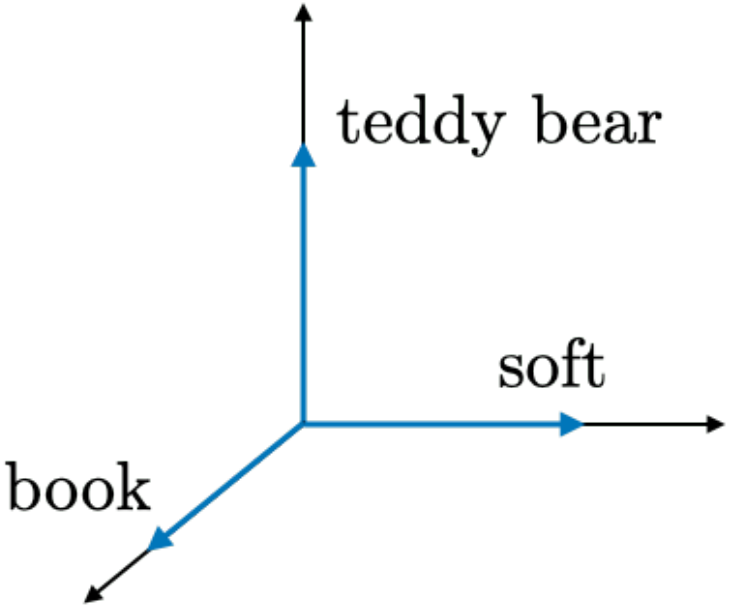
Visualization 300 -> 2D: t-SNE

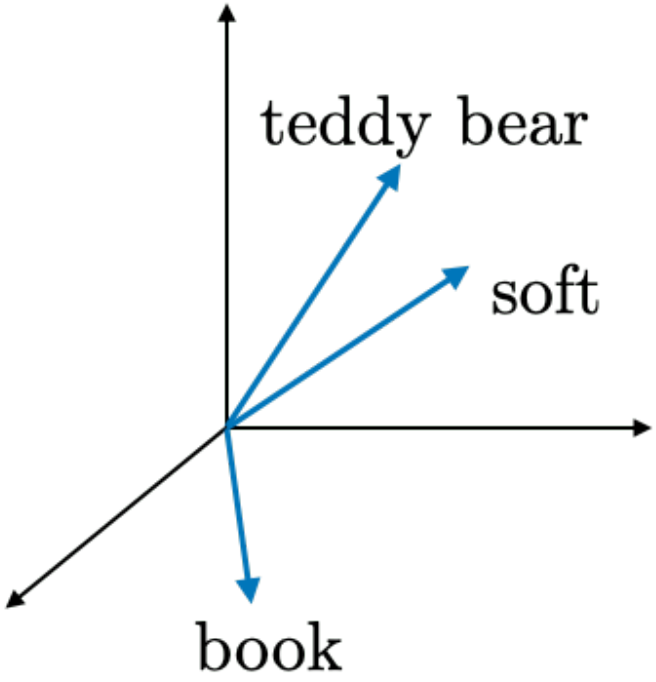
Representation Techniques



Word embedding representation e_w

Embedding



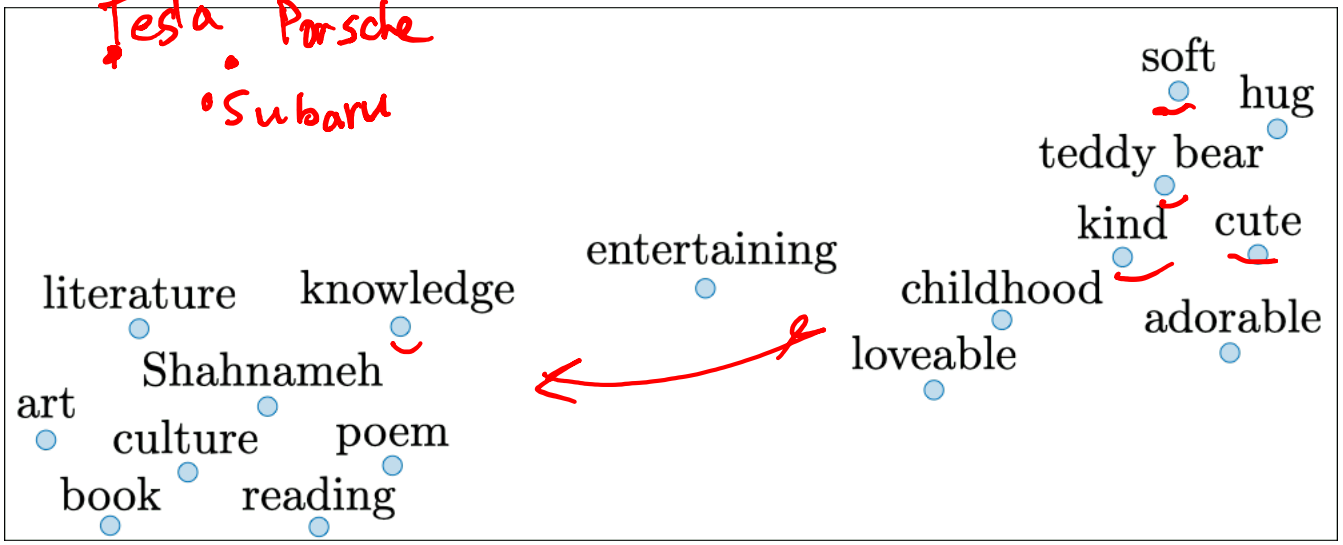


E

$o_w \rightarrow e_w$ (not E^{-1})

$o_w \leftarrow e_w$

Embedding Organization



t-SNE (t-distributed Stochastic Neighbor Embedding):
Reducing high-dimensional embeddings into a lower dimensional space.

Properties of Embedding Vectors

	Man	Woman	King	Queen	Orange	Apple	Tesla	Porsche
Gender	-1	1	-0.95	0.97	0.01	0	-0.5	0.03
Royal	0.01	0.02	0.94	0.95	-0.02	0.04	-0.01	0.1
Food								
Size								
Engine	0.07	-0.01	0.03	0.02	0.001	0	0.95	0.98

Can define distances (similarity):

$e_{Man} - e_{Woman} \sim = [-2;0;0;..;0]$

$e_{King} - e_{Queen} \sim = [-2;0;0;..;0]$

$e_{Man} - e_{Woman} \sim = e_{King} - e_{?}$

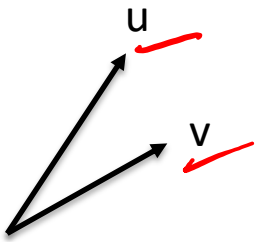
Similarity

We can define similarity in the space of embedding vectors (Full space)

$$\text{argmax}_w \ \text{sim}(e_w, v)$$

Similarity

We can define similarity in the space of embedding vectors (Full space)



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

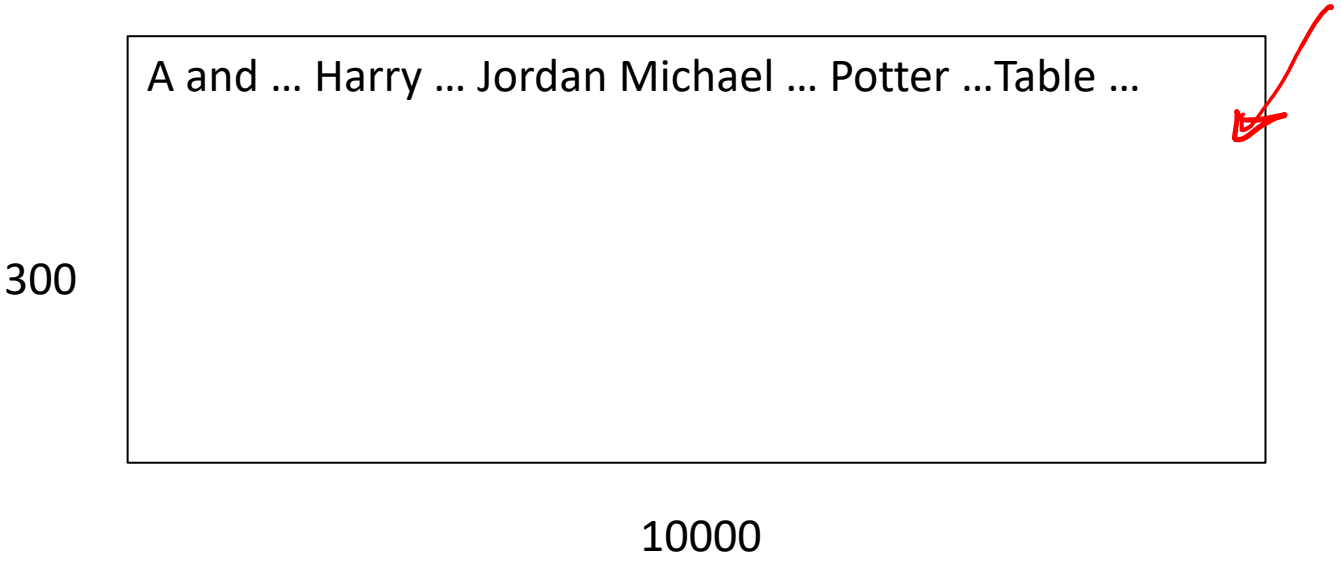
cosyne

$$sim(u, v) = ||u - v||^2$$

$\text{argmax}_w \text{ sim}(e_w, v)$

Embedding Matrix

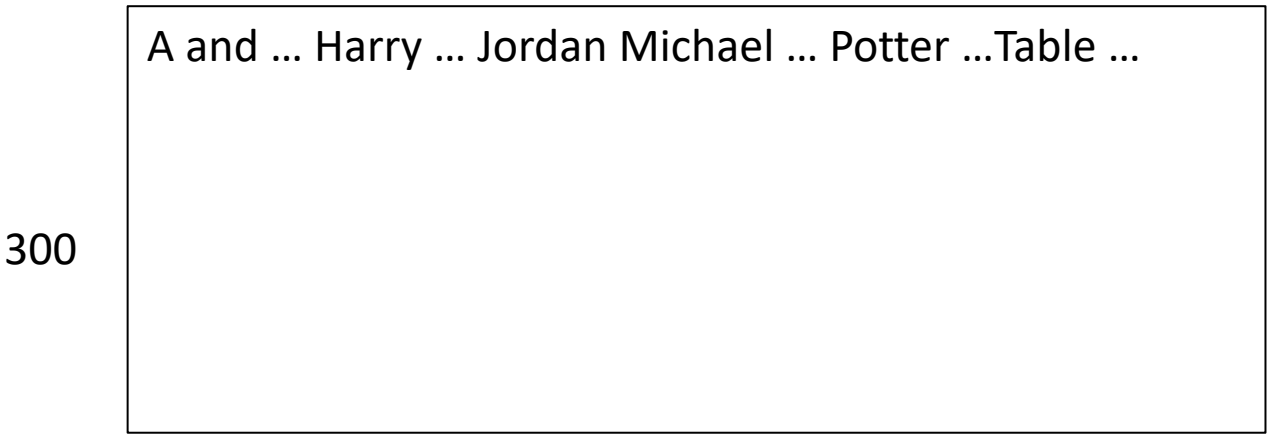
E=



$$\overrightarrow{e_j} = E \cdot \overrightarrow{o_j}$$

Embedding Matrix

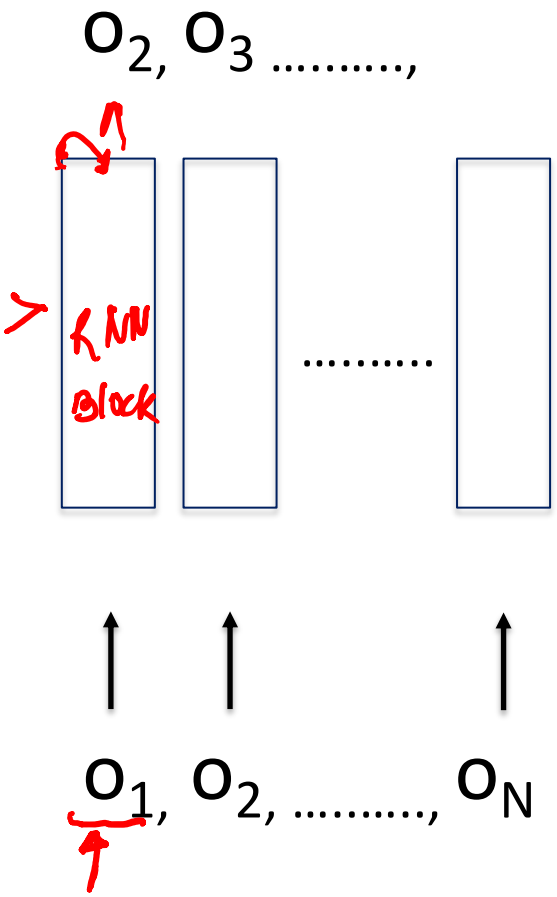
E=



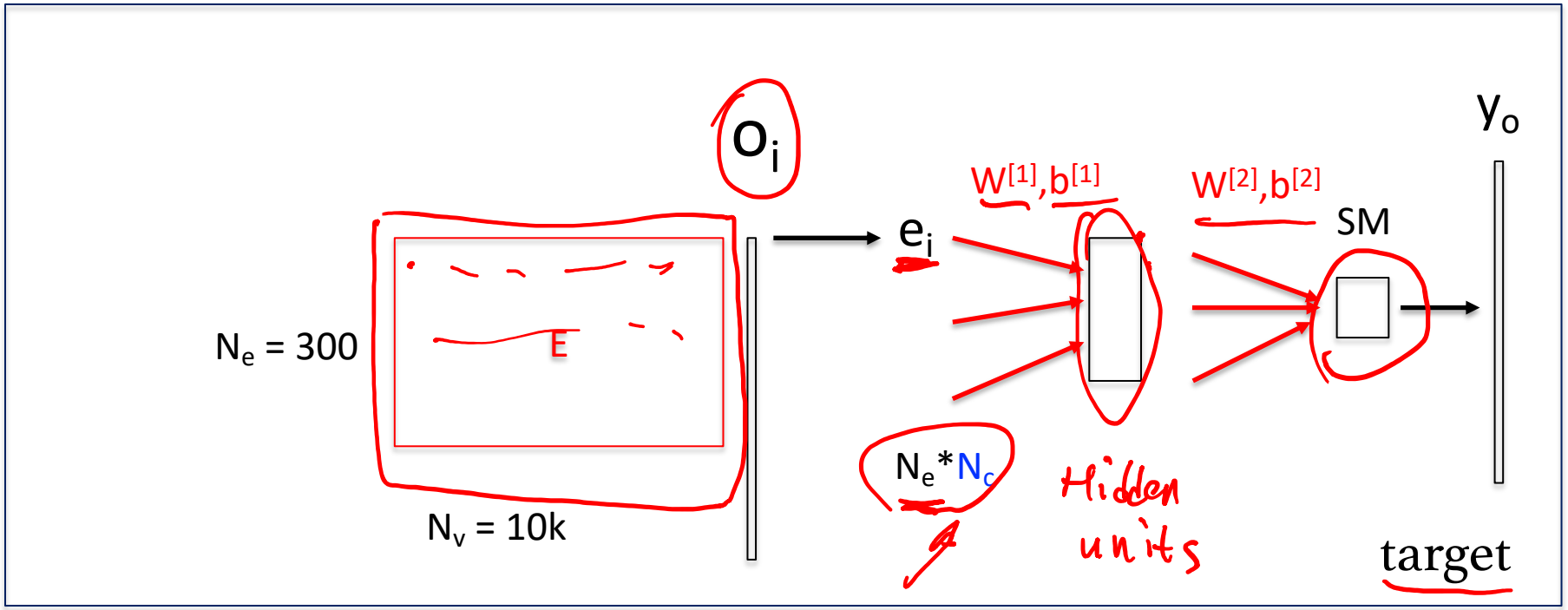
$$\overrightarrow{\boldsymbol{e_j}} = \boldsymbol{E} \cdot \overrightarrow{\boldsymbol{o_j}}$$

Neural Language Model

When an object is detected in your blind spot or close to the side of Model S (such as a vehicle, guard rail, etc.), colored lines radiate from the image of Model S on the instrument panel. The location of the lines correspond to the location of the detected object. The color of the lines (white, yellow, orange, or red) represents the object's proximity to Model S, with white being the farthest and red being very close and requiring your immediate attention. These colored lines only display when driving between approximately 7 and 85 mph (12 and 140 km/h). When Autosteer is active, these colored lines also display if driving slower than 7 mph (12 km/h). However, the colored lines do not display if Model S is at a standstill (for example, in heavy traffic).



Neural Language Model



N_c = # context words

Context and Target

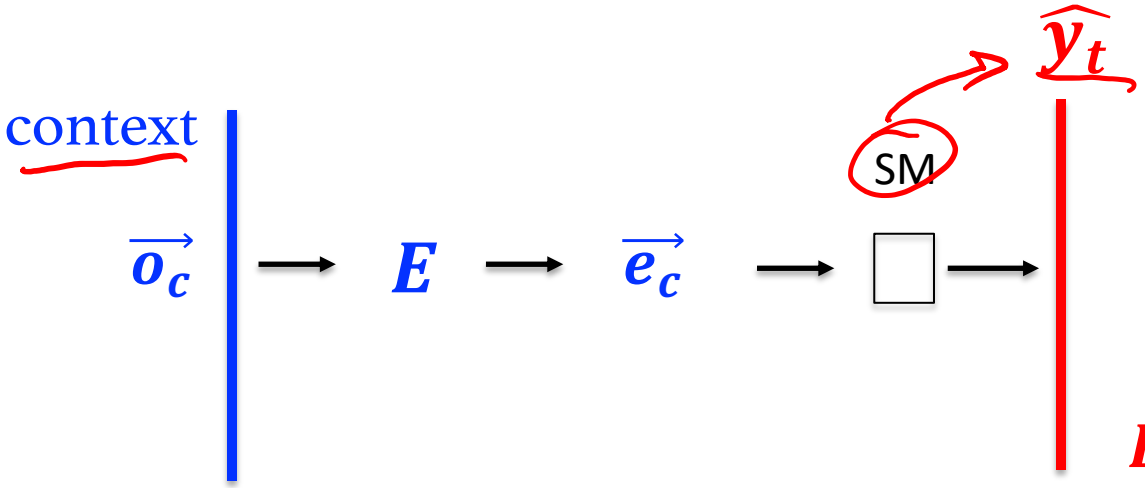
I want a glass of orange juice to go along with my cereal

target

a glass of orange	context	
orange	context	
orange	context	to
glass	context	

Word2Vec (skip gram)

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

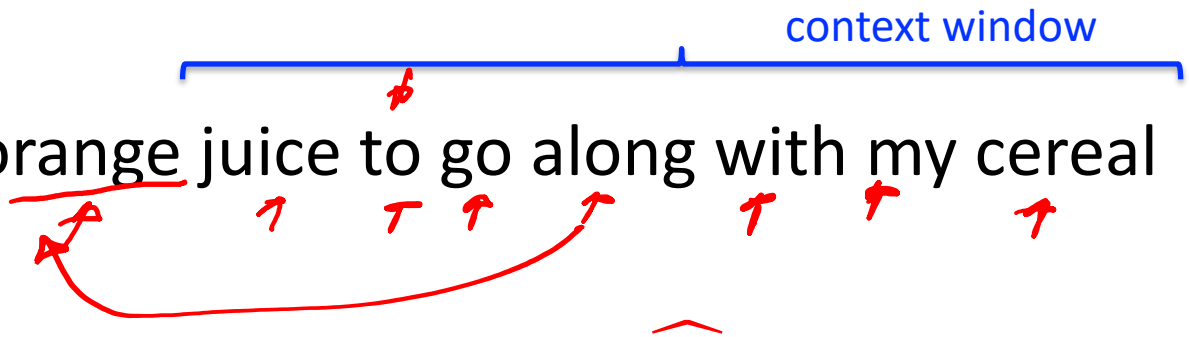


$$L(\hat{y}, y) = \sum_{i=1}^{N_v} y_i \log \hat{y}_i$$

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

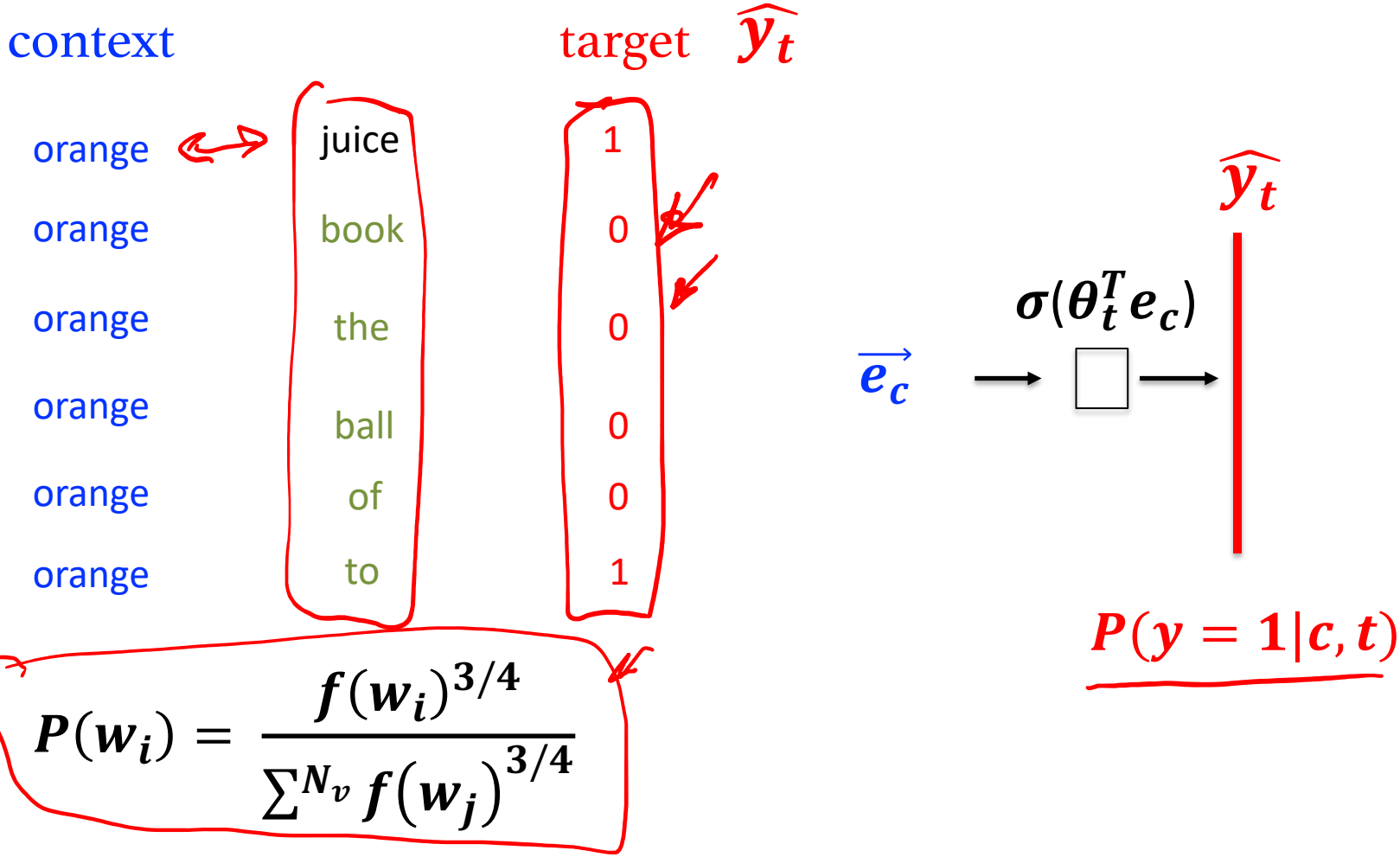
θ_t^T
Params associated with t

Negative Sampling



context	target	$\hat{y}_t$
orange	juice	1
k	book	0
	the	0
	ball	0
	of	0
	to	1

Negative Sampling



Sentiment Classification

“ this film was just brilliant”

“I loved the wildly uneven far more than I should have, but that doesn't mean it is perfect.”

“Despite a compelling lead performance by Tom Hanks and a great soundtrack, “

“with this movie you're better off saving your money, your popcorn, and time.”

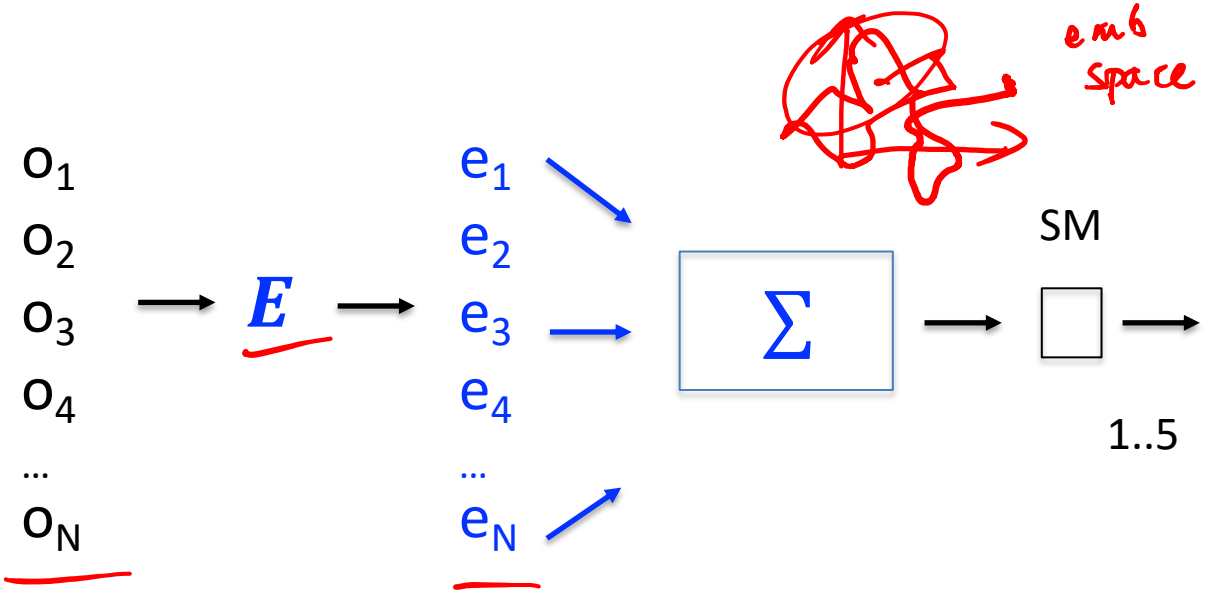
Sentiment Classification

Input: X

" this film was just brilliant"

Output: Y

Simple approach:



RNN Sentiment Classification