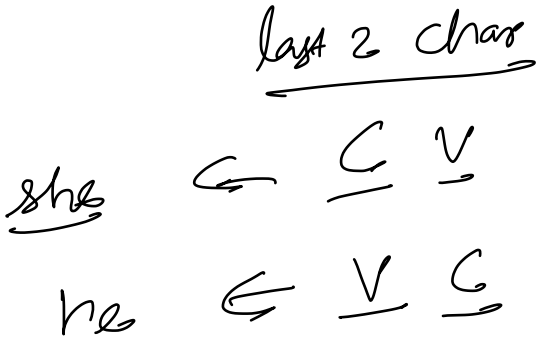
NLP Interaction Hour

January 19th, 2021

Some fun with NLP



Hebrew

diber	'he spoke'
hasav	'he thought'
sagra	'she shut'
hasva	'she thought'
kalat	'he received'
kalta	'she received'
milmel	'he muttered'
kimet	'he wrinkled'

CV

VC

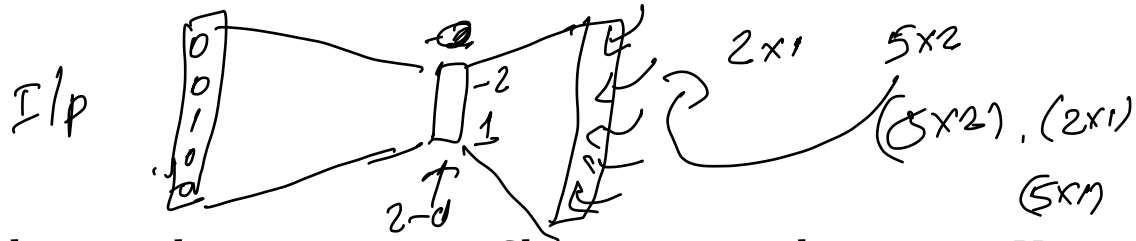
CV

Translate the following -

dibra	she spoke
kimta	
milmla	
sagar	

Questions from the Lecture Videos?

Solve this Problem



Suppose you are computing the word vectors using Skip-gram architecture. You have 5 words in your vocabulary, {passed, through, relu, activation, function} in that order and suppose you have the window, '*through relu activation*' in your corpora. You use this window with `relu' as the center word and one word before and after the center word as your context.

Also, suppose that for each word, you have 2-dim in and out vectors, which have the same value at this point given by [1,-1],[1,1],[-2,1],[0,1],[1,0] for the 5 words, respectively. As per the Skip-gram architecture, the loss corresponding to the target word "activation" would be –log(x). What is the value of x?

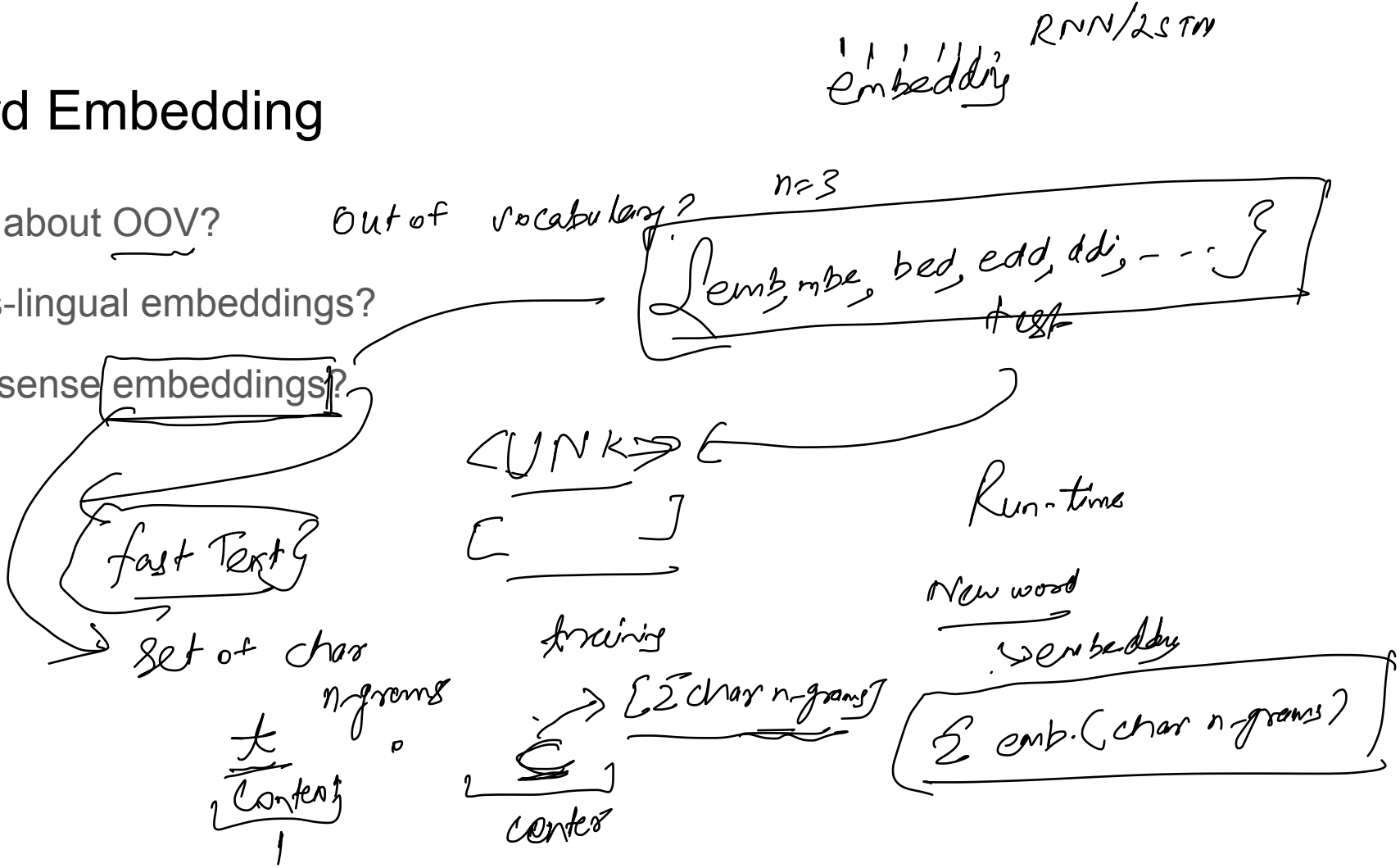
Let’s try a quick Demo

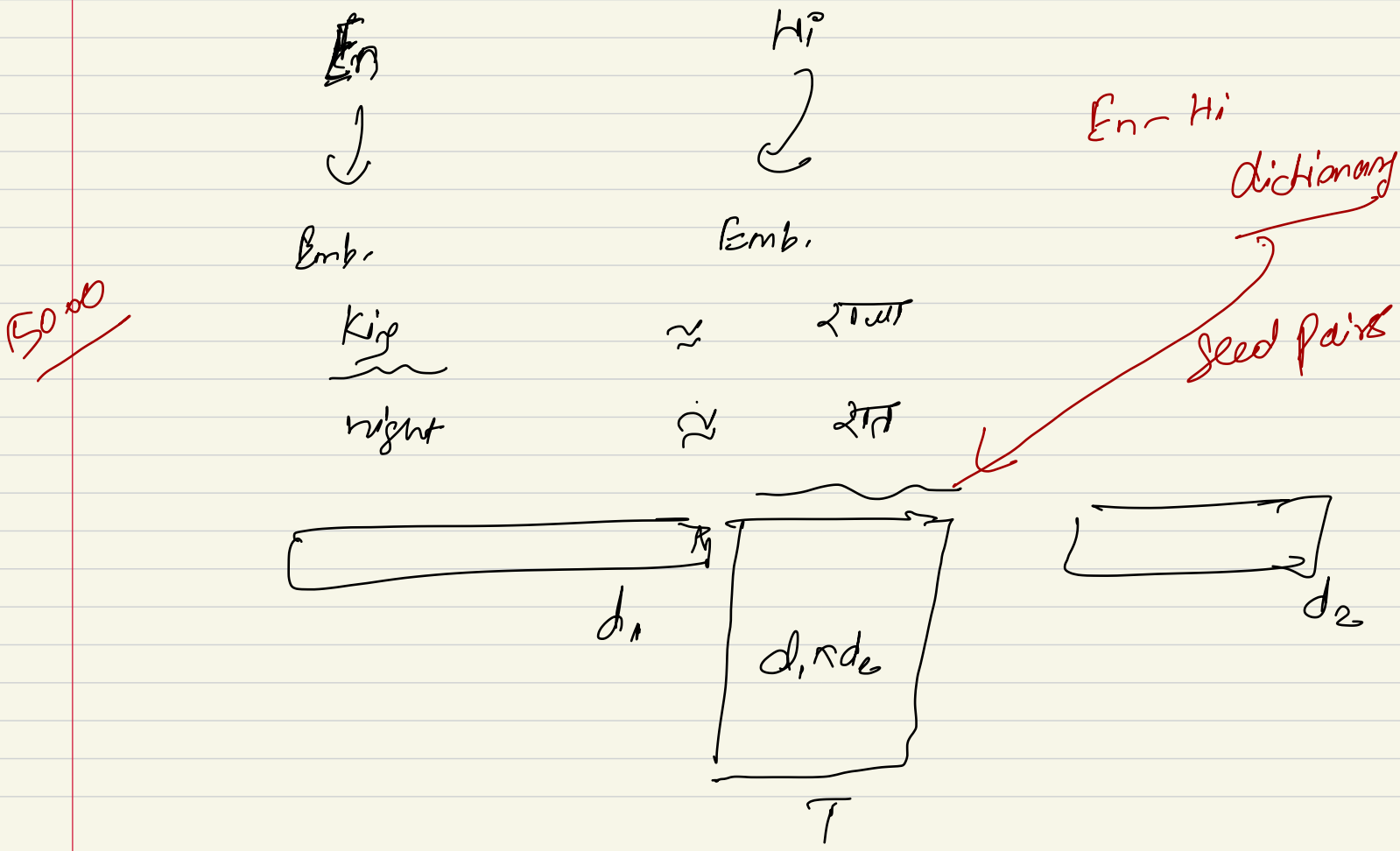
Word Embedding

What about OOV?

Cross-lingual embeddings?

Multi-sense embeddings?




En
Hi
En - Hi
dictionary
Emb.
Emb.
Kiss
≈
राजा
right
≈
सही
seed pairs
d_1
$d_1 \times d_6$
d_2
T

Word Embeddings for various paradigms

Text Classification

Sequence Labeling

Text Generation?