RNN

Practical NLP

Text Generation

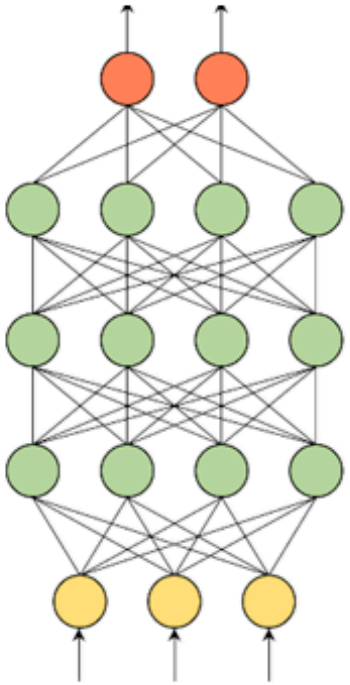
Data Trainers

How do we generate text?

Probabilities over char set

	a	b	c	d	e	f	g	...	z
0.01	0.02	0.36	0.25	0.02	0.001	0.22	0.001	...	0.06

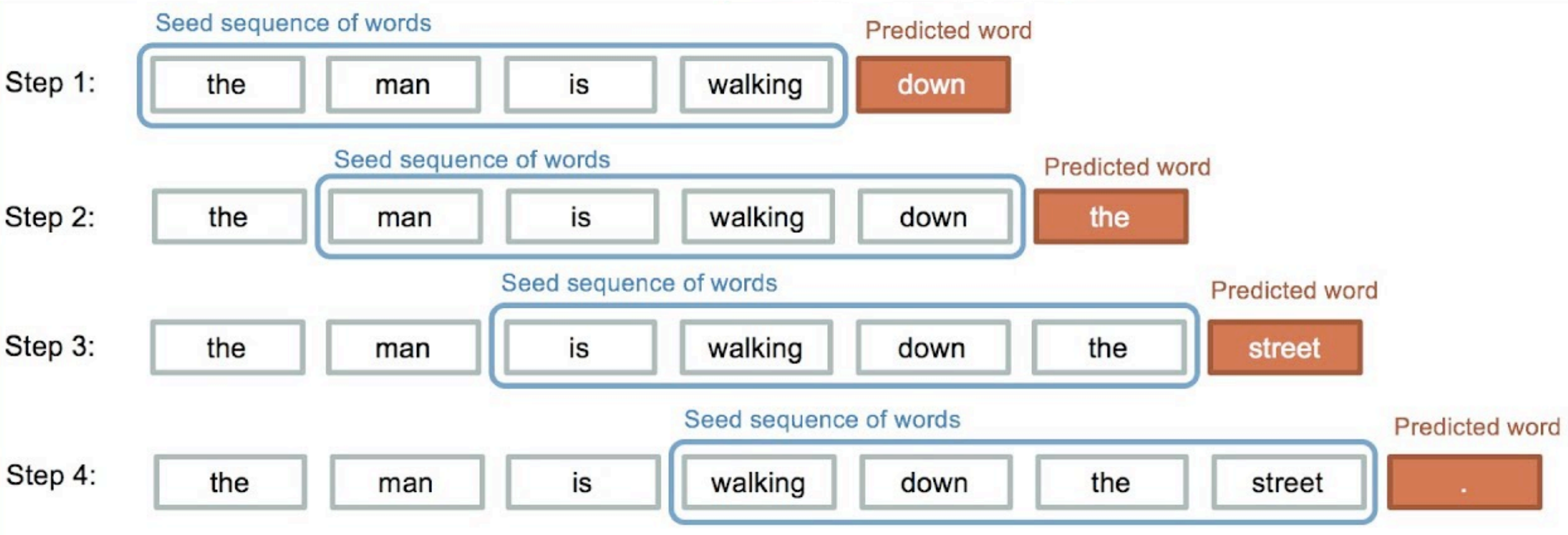
Language Model



Train Input from Corpus

I		w	a	n	t		t	o		b	a	k	e		

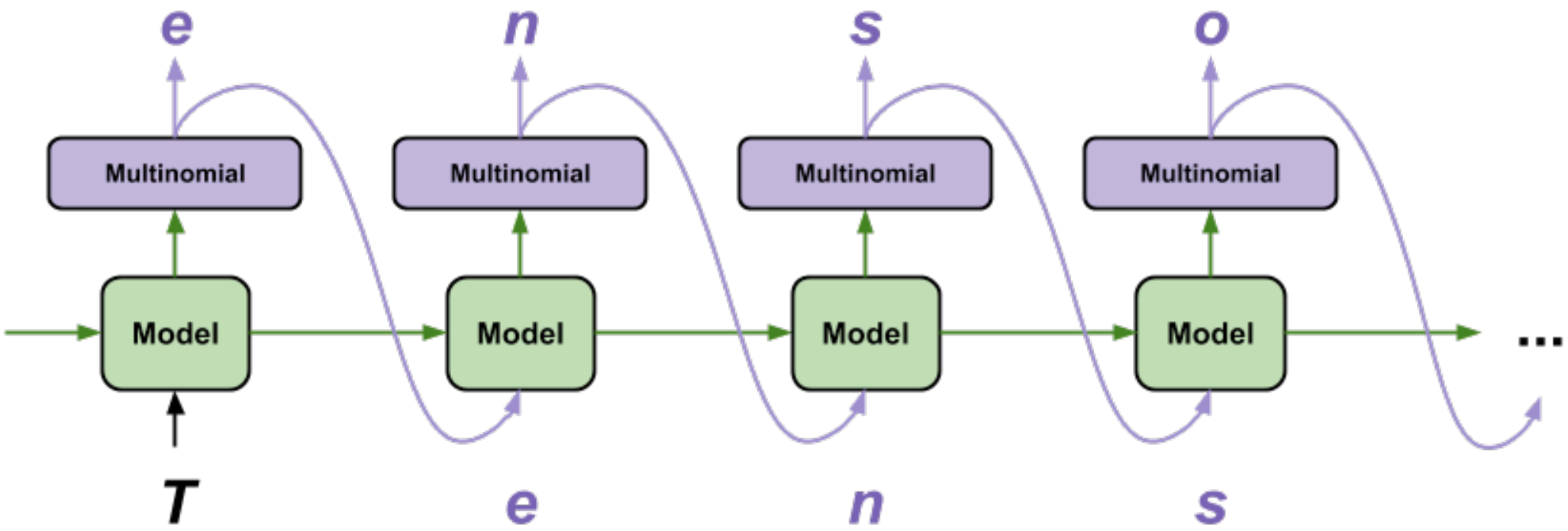
How do we generate text?



Problems

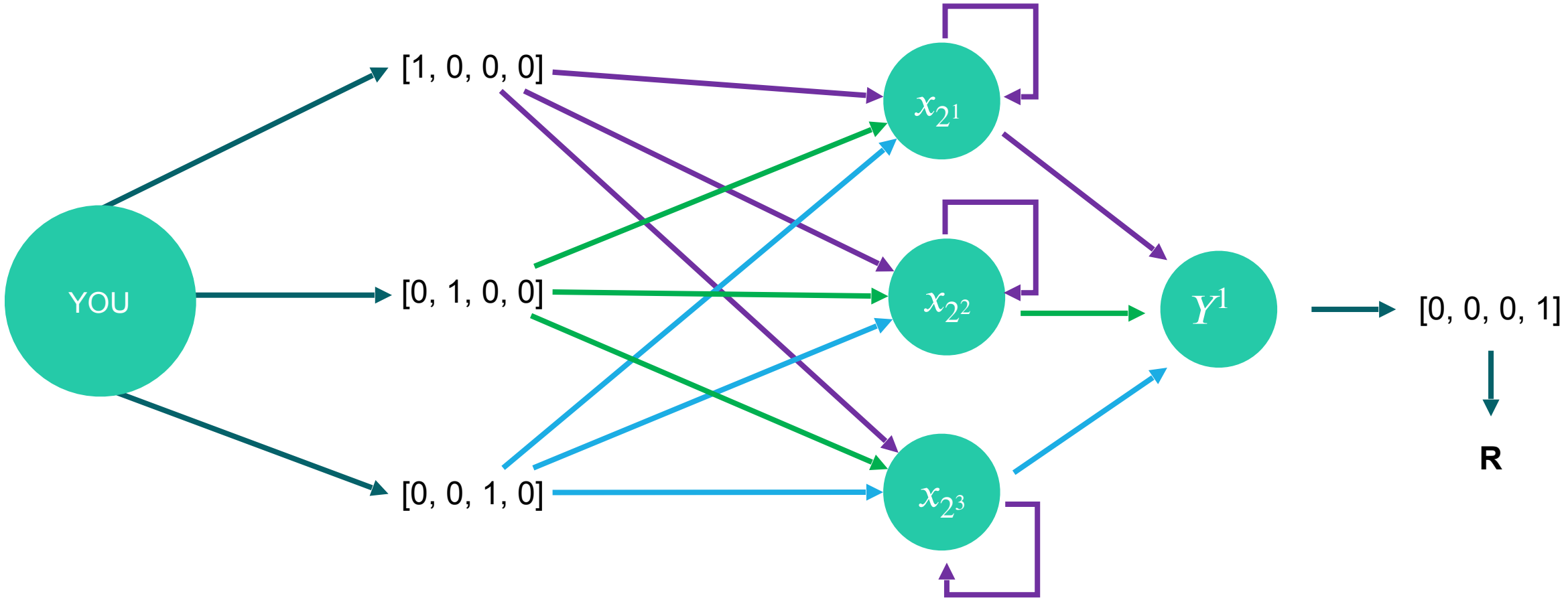
- If we used a normal NN, we would need a window of fixed size to predict the next character/word.
- But in text, sometimes the core meaning comes at the end, however long it is the sequence.
  - *“Hospitals are sued by 7 foot doctors”*
  - *Local high school dropouts cut in half*
- These problems are and for normal NN because they cannot remember!

Character level text generation “passing memory”

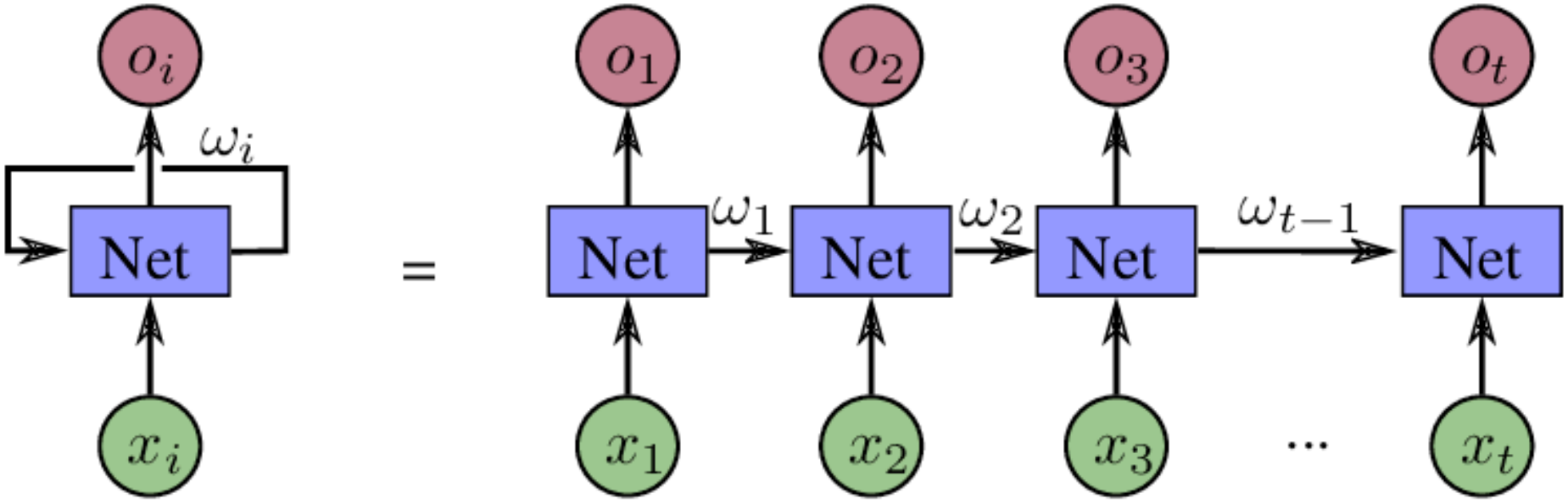


The “key” is to pass the state as input to remember

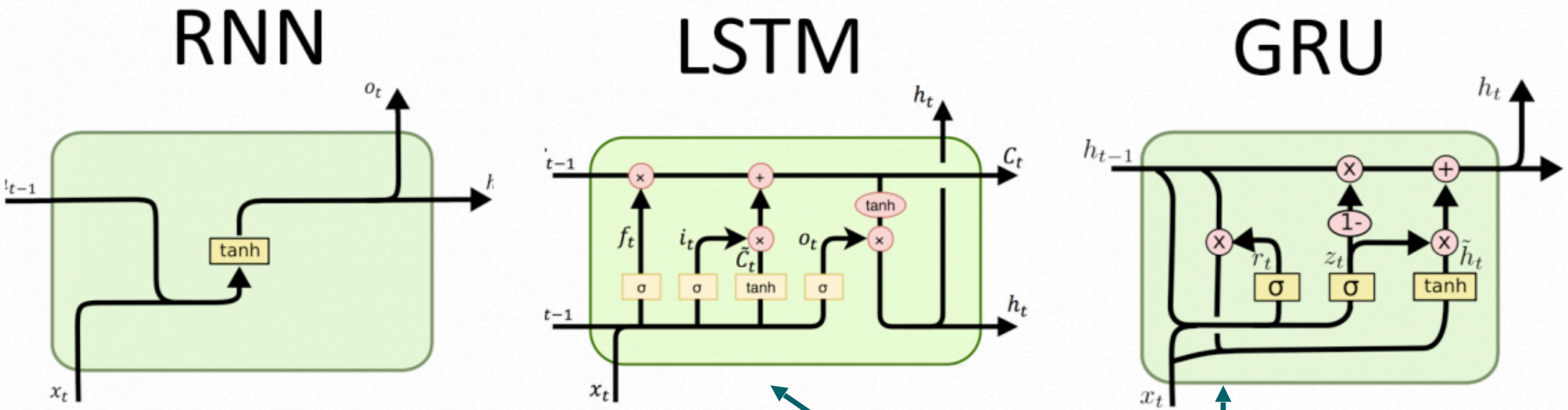
Recursive Neural Networks



Recursive Neural Networks



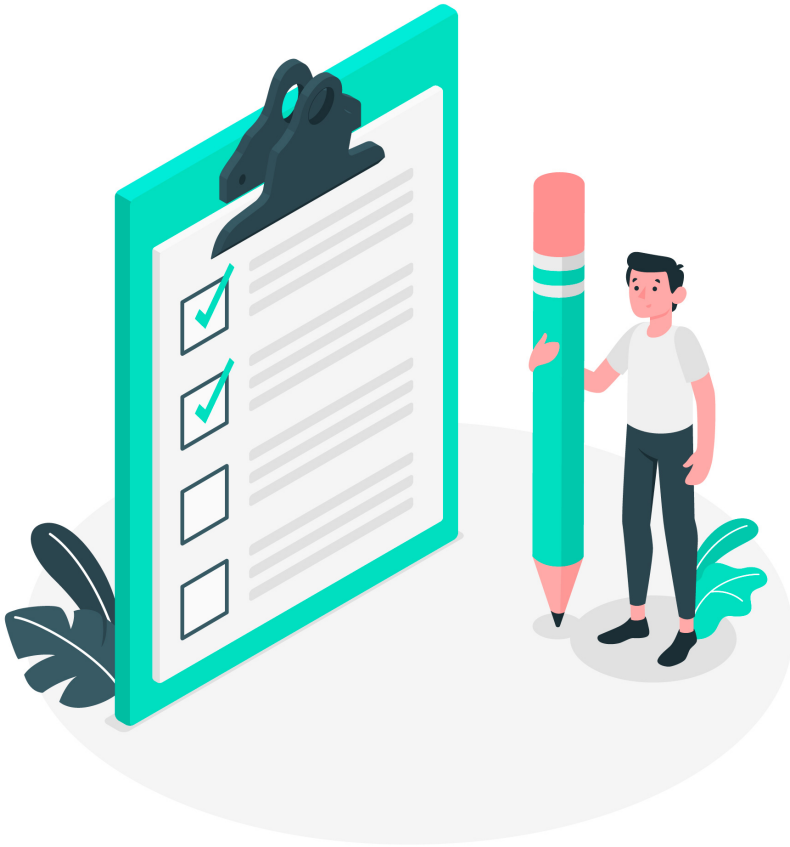
Variations of RNNs



LAB

Train a LSTM RNN

- Train an RNN from scratch with LSTM cells.
- Put attention to the dataset preparation since this will be a “word based” model.



Summary

- Using text generation we can generate rental ads easy.
- This way we can make the task easier for users by allowing them to be lazy.
- Under the hoods, we need a model with recursive structure to maintain memory of the state.
- If we have time at the end, we can discuss about distillation! (Please remind me)

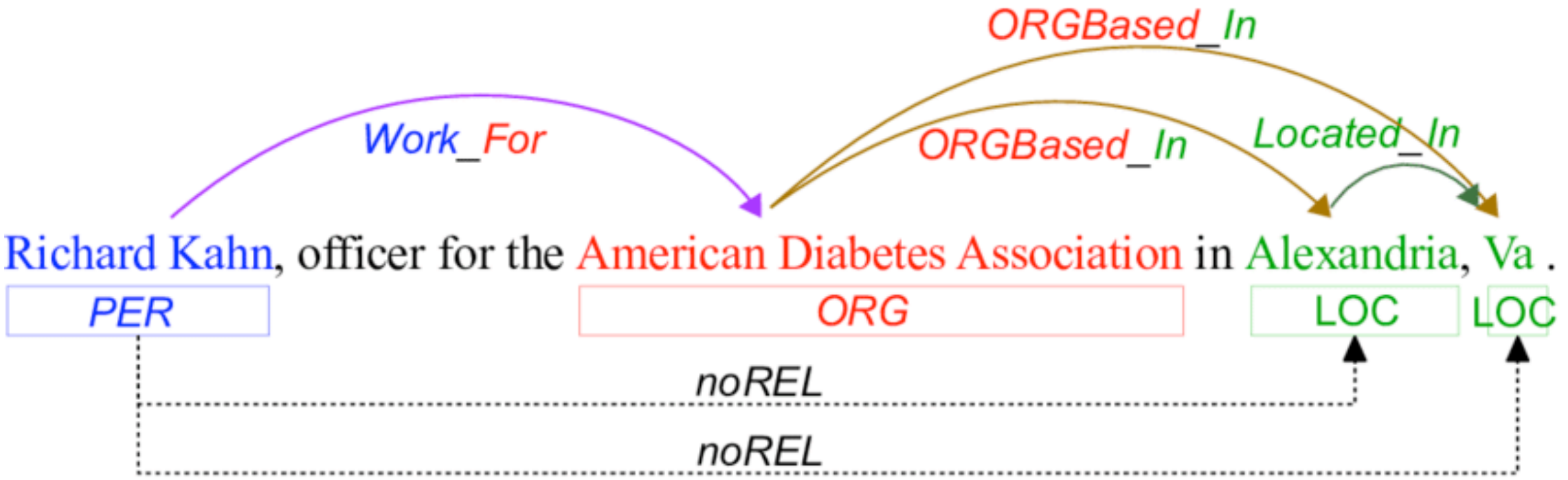
Named Entity Recognition

- All of these tags extract information from the text.

When Sebastian Thrun PERSON started at Google ORG in 2007 DATE, few people outside of the company took him seriously. “I can tell you very senior CEOs of major American NORP car companies would shake my hand and turn away because I wasn’t worth talking to”, said Thrun PERSON, now the co-founder and CEO of online higher education startup Udacity, in an interview with Recode ORG earlier this week DATE.

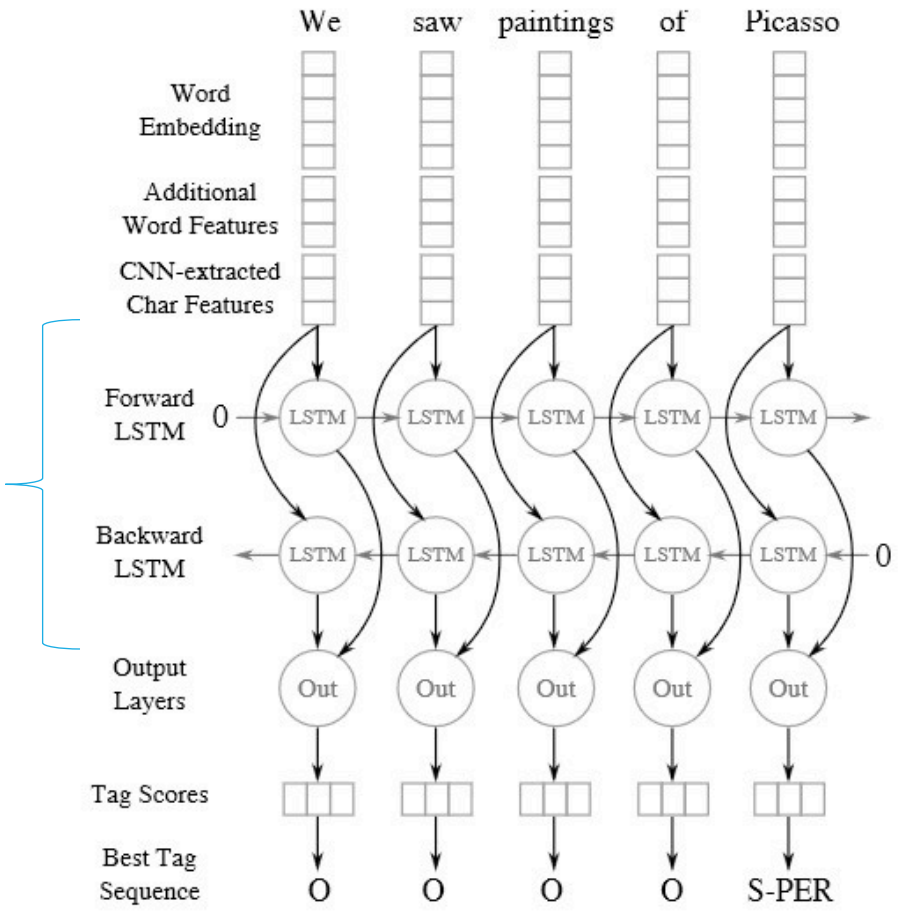
A little less than a decade later DATE, dozens of self-driving startups have cropped up while automakers around the world clamor, wallet in hand, to secure their place in the fast-moving world of fully automated transportation.

Named Entity Recognition



Bi-LSTM

- We need a forward and backward pass because some tags at the beginning only make sense after reading the whole sentence.



LAB

Perform NER with Bi-LSTM

- Create a Bidirectional LSTM trained on the NER dataset to predict entities.

- Put special attention to the **metric** used! Does accuracy make sense here?