Deep Learning 101

Word Embedding

Schedule

week	Date	Topic
9	10.27	Environment setup, python, Jupyter, PyCharm, TensorFlow, & regression
10	11.03	Training and testing
11	11.11	CNN
12	11.18	RNN
13	11.24	Word embedding & confusion matrix
14	12.01	Autoencoder & GAN

Today’s Class

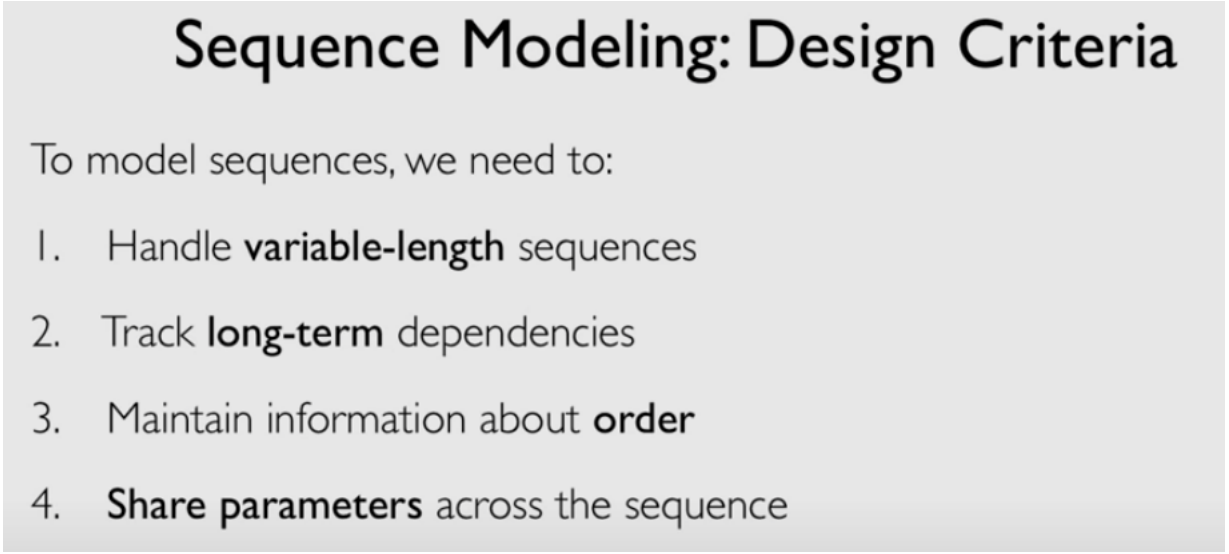
- Recap
- Word embedding
- Confusion matrix
- Lab time

How to model sequence?

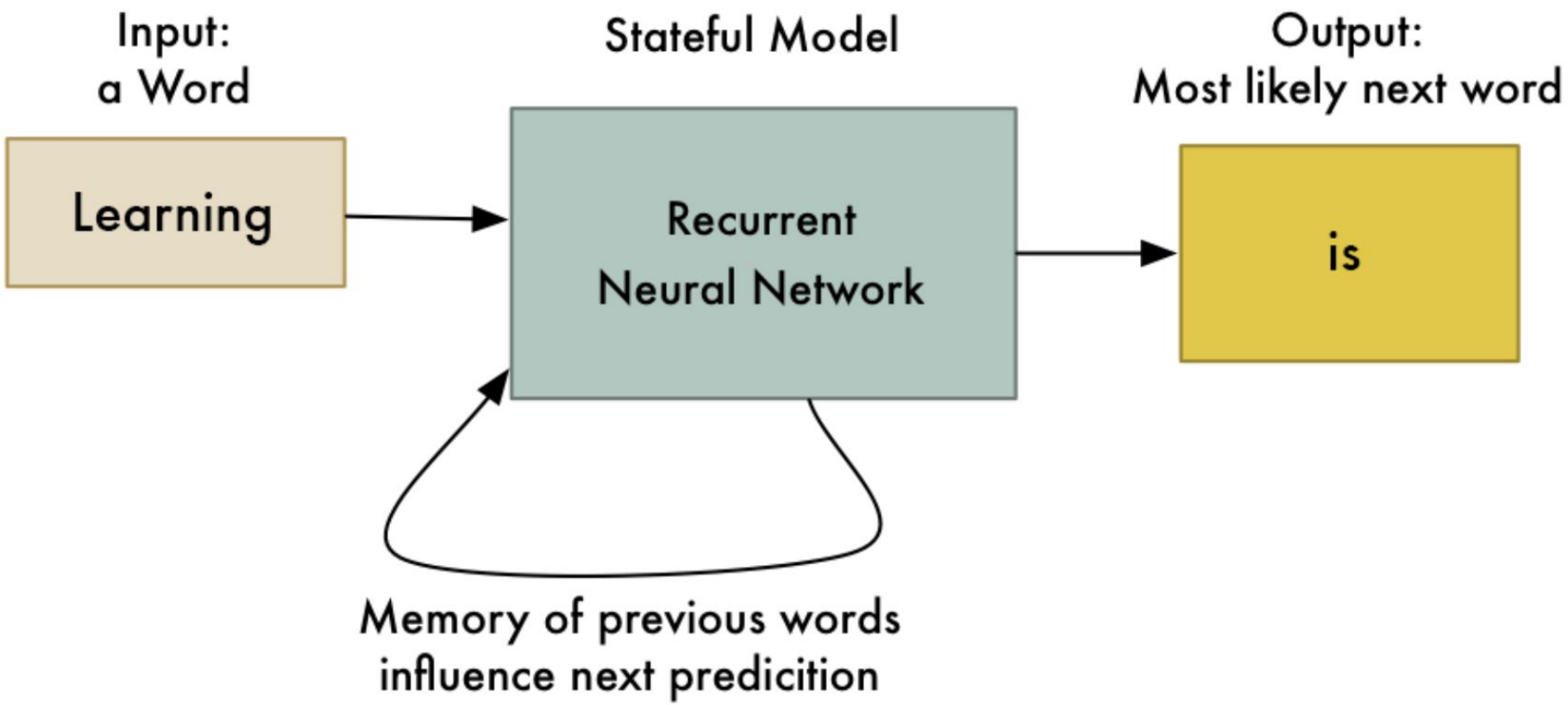
Predicting the next word

“The quick brown fox jumped over the lazy ___”

1. Fixed window
- 2. Bag of words
- 3. Big fixed window?
  - Long term dependency

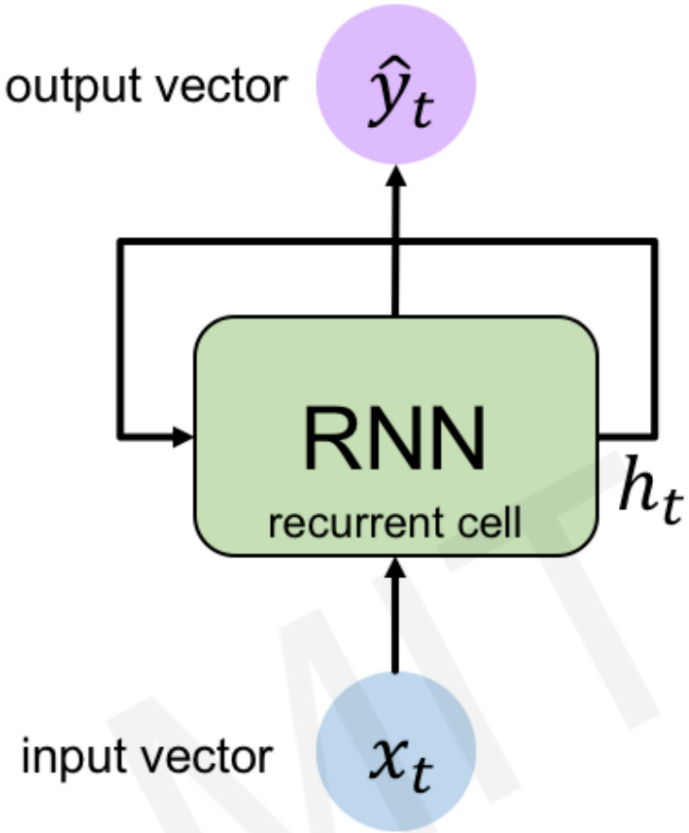


RNN (Intuition)



https://wikidocs.net/35011

RNN in a nutshell



Apply a **recurrence relation** at every time step to process a sequence:

$$h_t = f_W(h_{t-1}, x_t)$$

cell state

function parameterized by W

old state

input vector at time step *t*

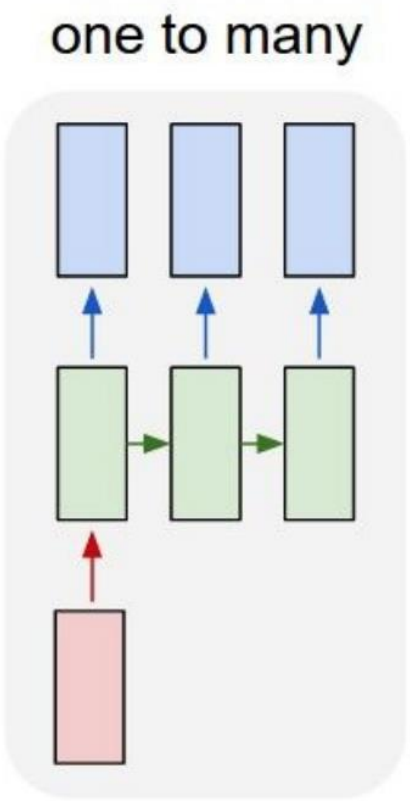
Note: the same function and set of parameters are used at every time step

http://introtodeeplearning.com

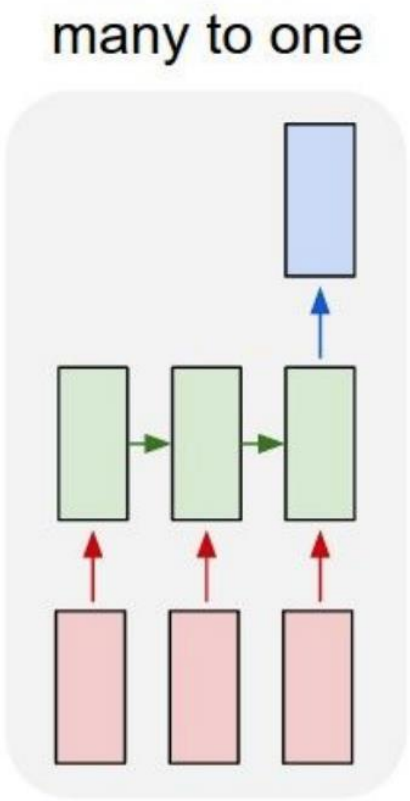
Recap - RNN

- Question: How to model a sequence?
  - e.g. next word prediction
- RNN intuition: keeping the memory of previous words
- $h_t = f_w(h_{t-1} x_t)$
- Problems with Vanilla RNN: vanishing gradients
- ➔ LSTM – gated cells

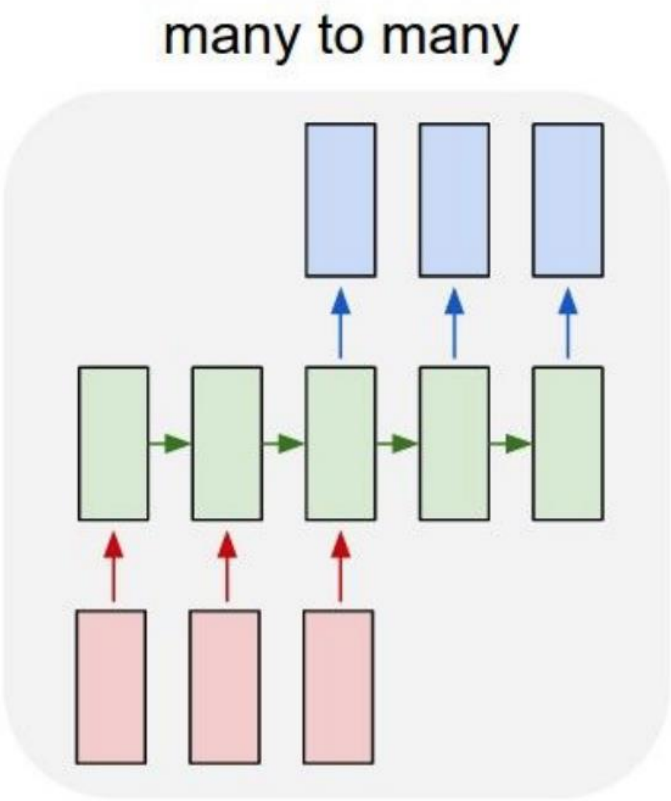
RNN Applications



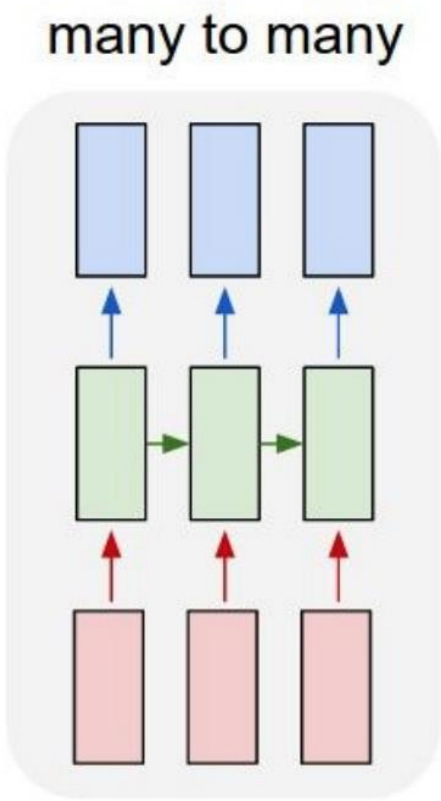
Captioning



Sentiment



Translation



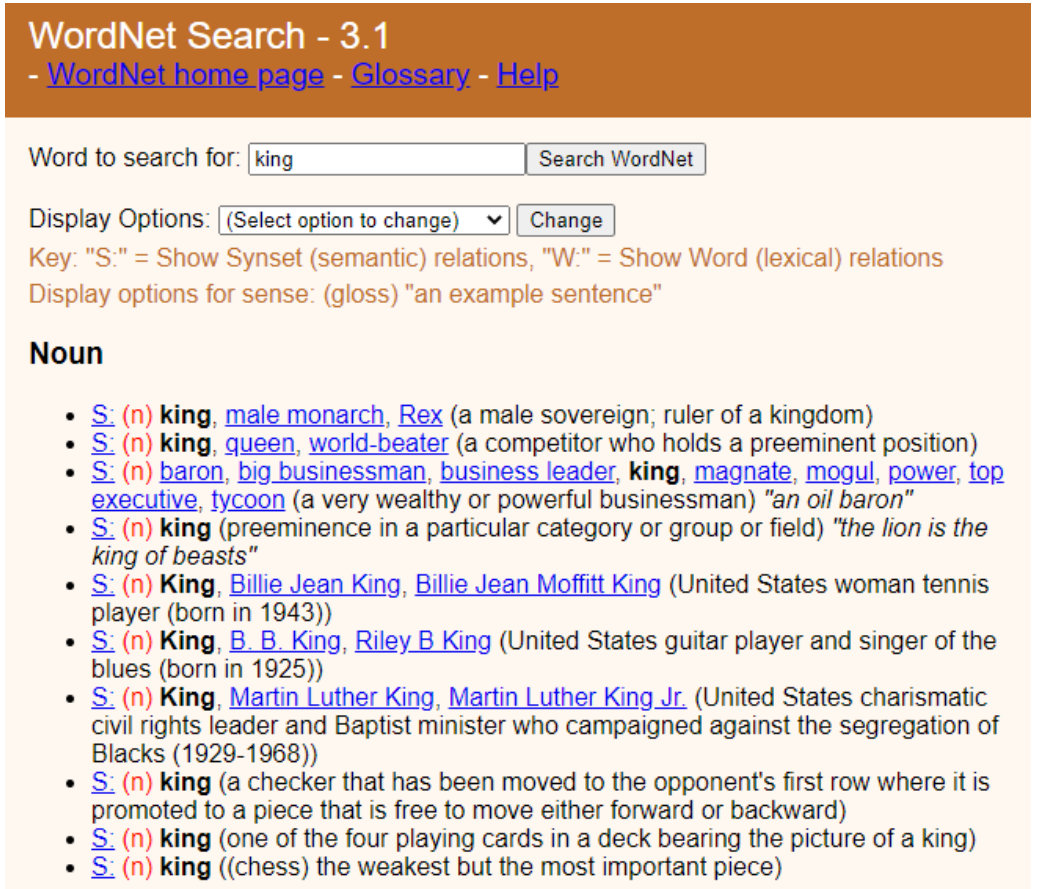
Music Generation

What is the meaning of a word?

- Question: How do you represent words?

How to represent the meaning of a word?

- WordNet? Explicit handmade definitions for each word (aka, discrete semantics)



WordNet Search - 3.1
- WordNet home page - Glossary - Help

Word to search for: king Search WordNet

Display Options: (Select option to change) Change

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations

Display options for sense: (gloss) "an example sentence"

Noun

- S: (n) **king**, male monarch, Rex (a male sovereign; ruler of a kingdom)
- S: (n) **king**, queen, world-beater (a competitor who holds a preeminent position)
- S: (n) baron, big businessman, business leader, **king**, magnate, mogul, power, top executive, tycoon (a very wealthy or powerful businessman) *"an oil baron"*
- S: (n) **king** (preeminence in a particular category or group or field) *"the lion is the king of beasts"*
- S: (n) **King**, Billie Jean King, Billie Jean Moffitt King (United States woman tennis player (born in 1943))
- S: (n) **King**, B. B. King, Riley B King (United States guitar player and singer of the blues (born in 1925))
- S: (n) **King**, Martin Luther King, Martin Luther King Jr. (United States charismatic civil rights leader and Baptist minister who campaigned against the segregation of Blacks (1929-1968))
- S: (n) **king** (a checker that has been moved to the opponent's first row where it is promoted to a piece that is free to move either forward or backward)
- S: (n) **king** (one of the four playing cards in a deck bearing the picture of a king)
- S: (n) **king** ((chess) the weakest but the most important piece)

Problems with discrete semantics

- Time consuming
- Requires human labor (huge)
- Incomplete: new words, etc.
- Subjective
- Missing nuances
- Can’t compute accurate word similarity

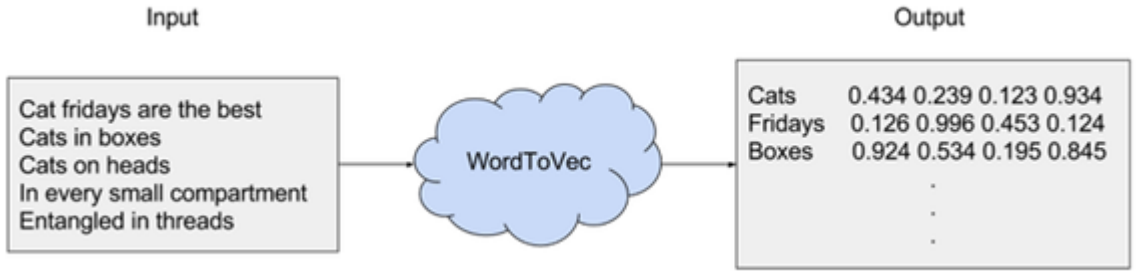
What if you represent the wordings in “features”?

Word	Male	Female	Royalty	Age
King	Yes	No	Y	50
Queen	No	Yes	Y	50
Man	Yes	No	N	45
Woman	No	Yes	N	40
Prince	Yes	No	Y	20
Princess	No	Yes	Y	18

What if you turn the features into numbers?

Word	Male	Female	Royalty	Age
King	0.9	0.2	0.9	0.9
Queen	0.1	0.8	0.8	0.8
Man	0.8	0.1	0.2	0.7
Woman	0.1	0.7	0.3	0.8
Prince	0.8	0.3	0.6	0.5
Princess	0.2	0.8	0.7	0.6

Word embedding intuition

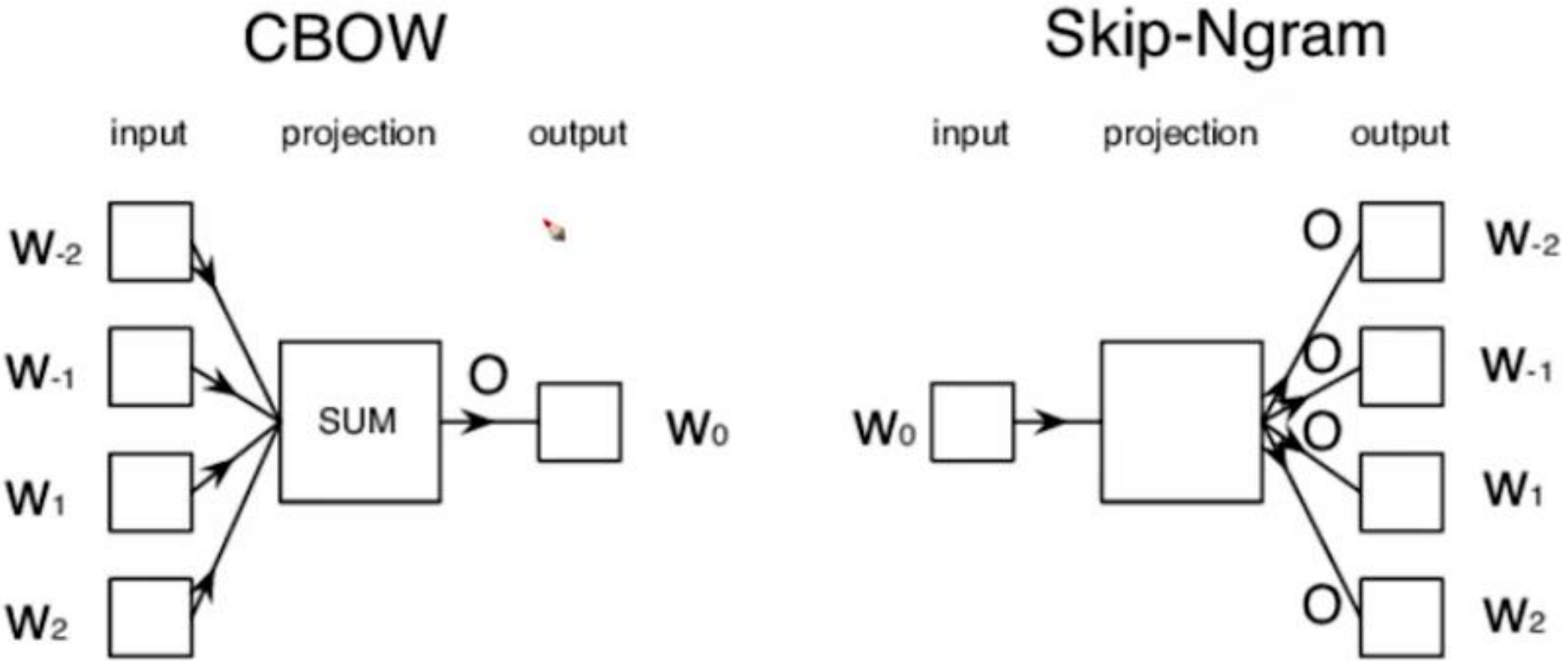


Word Embedding Intuition

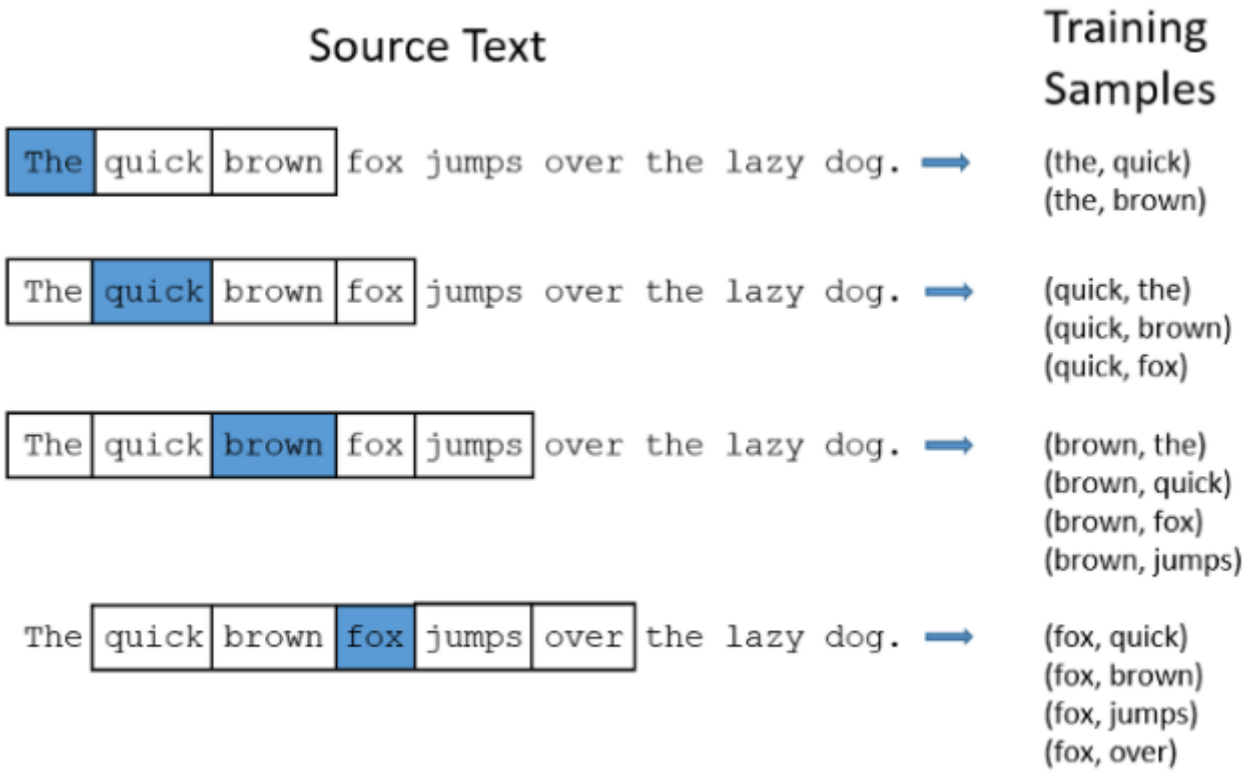
- “You shall know a word by the company it keeps.”
  - J. R. Firth (1957)

- Define a model that aims to predict between a center word w_t and context words in terms of word vectors. (Skip Gram)
  - $p(\text{context}|w_t)$
- Skip grams: Predict context words given target (position independent)
- Continuous Bag of Words (CBOW): Predict target word from bag-of-words context

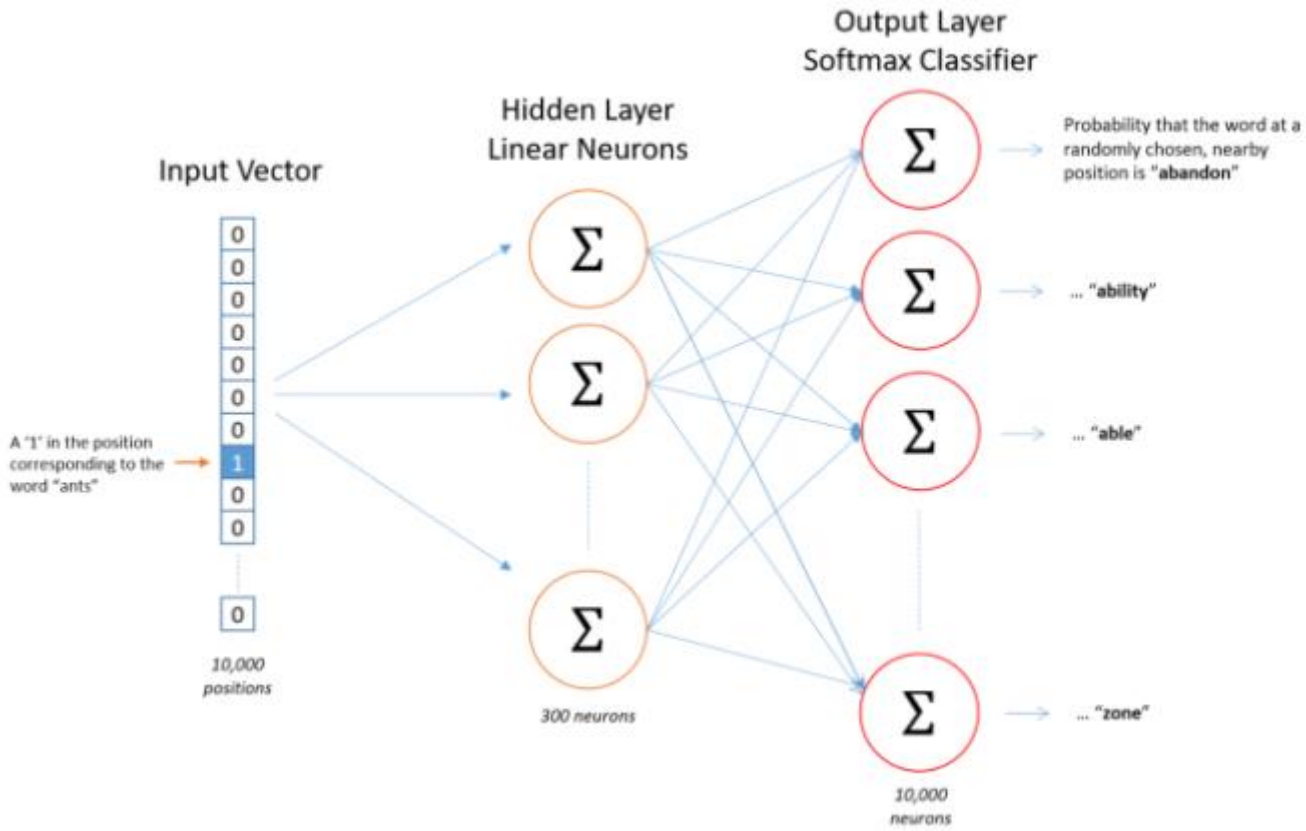
Two models of word embedding



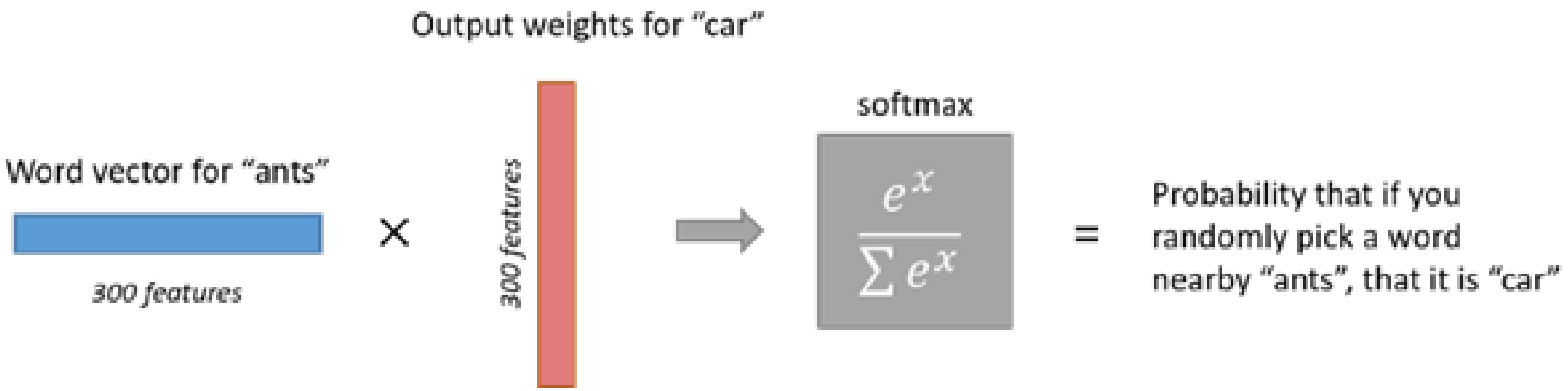
Word embedding – training data



Word embedding



https://towardsdatascience.com/word2vec-skip-gram-model-part-1-intuition-78614e4d6e0b


Output weights for "car"
Word vector for "ants"
300 features
300 features
softmax
$\frac{e^x}{\sum e^x}$
=
Probability that if you randomly pick a word nearby "ants", that it is "car"

Word embeddings

king − man + woman ~= queen

https://jalammar.github.io/illustrated-word2vec/

Visualization of word embedding

- Korean:
  - http://word2vec.kr
- TensorFlow word embeddings:
  - https://projector.tensorflow.org/
- TensorFlow pre-trained models:
  - https://tfhub.dev/s?module-type=text-embedding

Lab time

- To clone: from your terminal
  - >git clone https://github.com/changsin/DeepLearning-101.git
- Or use google colab to point to the git hub repository
- Git is an open source version control system
  - Github is a host service using git.