Natural Language Processing

Overview

1. Natural Language Processing

2. Word Representation

3. Sequence Models

1. Natural Language Processing

2. Word Representation

3. Sequence Models

- Introduction to NLP
- Tasks in NLP
  - Sentiment Classification
  - Language Modeling
  - Image Captioning
  - Named-Entity Recognition
  - Machine Translation
  - Text Summarization
  - Paraphrasing
  - Question-Answering
  - Chatbots

    .
    .
    .
- Deep Learning based NLP Models
  - RNN, GRU, LSTM
  - Attention + RNNs, GRUs, LSTMs
  - Transformer, BERT
  - GPT, GPT2, GPT3

1. Natural Language Processing

Introduction to NLP :

- give computers the ability to read, understand and interpret human language.

A corpus may contain texts in a single language (*monolingual corpus*) or text data in multiple languages (*multilingual corpus*).

In order to make the corpora more useful for doing linguistic research, they are often subjected to a process known as annotation. An example of annotating a corpus is part-of-speech tagging, or *POS-tagging*, in which information about each word's part of speech (verb, noun, adjective, etc.) is added to the corpus in the form of *tags*. Another example is indicating the lemma (base) form of each word. When the language of the corpus is not a working language of the researchers who use it, interlinear glossing is used to make the annotation bilingual.

Some corpora have further *structured* levels of analysis applied. In particular, a number of smaller corpora may be fully parsed. Such corpora are usually called Treebanks or Parsed Corpora. The difficulty of ensuring that the entire corpus is completely and consistently annotated means that these corpora are usually smaller, containing around one to three million words. Other levels of linguistic structured analysis are possible, including

1. Natural Language Processing

Tasks in NLP :

- Sentiment Classification

"Wow. This movie is great. Awesome acting by DiCaprio. The screenplay is just perfect."

"I didn't like that movie at all. Too much of violence. The story line was too slow."

1. Natural Language Processing

Tasks in NLP :

- Sentiment Classification

“Wow. This movie is great. Awesome acting by DiCaprio. The screenplay is just perfect.”

“I didn’t like that movie at all. Too much of violence. The story line was too slow.”

1. Natural Language Processing

Tasks in NLP :

- Named-Entity Recognition

“I am Max. I work on Machine Learning at Apple. I’m interested on Generative Modeling; experienced on Computer Vision, Time Series Analysis and NLP.”

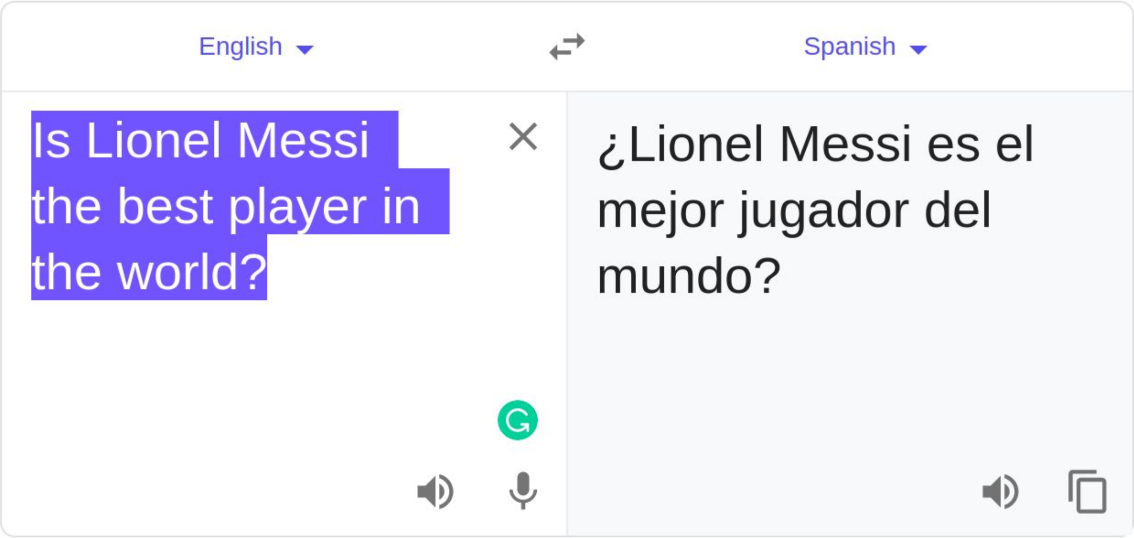
*“I am **Max**. I work on **Machine Learning** at **Apple**. I’m interested on **Generative Modeling**; experienced on **Computer Vision**, **Time Series Analysis** and **NLP**.”*

person names,
organizations,
locations,
professions, ….

1. Natural Language Processing

Tasks in NLP :

- Machine Translation



1. Natural Language Processing

Tasks in NLP :

- Image Captioning

“A person holding a flower.” [1]

“Green limes isolated on white background” [2]

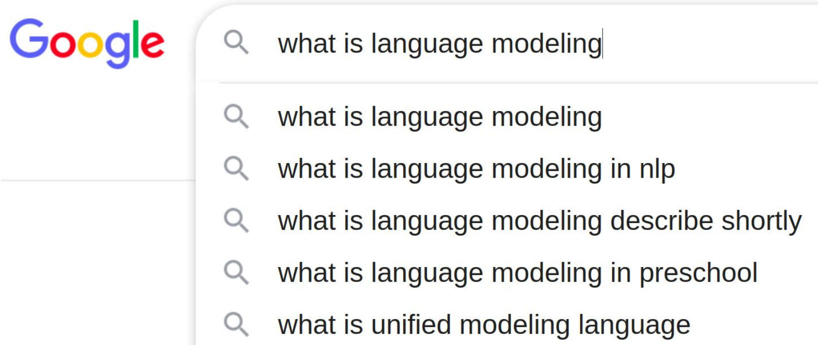
“A dog running in the green grass.” [3]

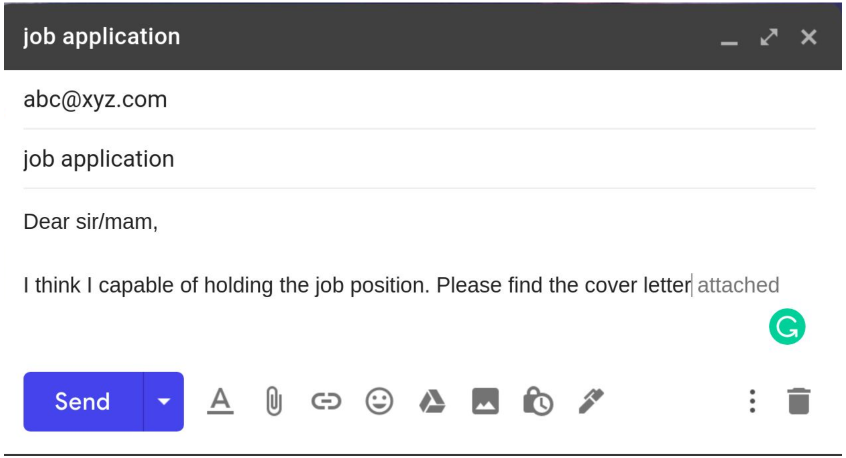
[1]https://unsplash.com/photos/ChPcC7satSU
[2]https://www.publicdomainpictures.net/en/view-image.php?image=3420&picture=green-limes
[3]https://www.publicdomainpictures.net/en/view-image.php?image=33633&picture=labrador-dog

1. Natural Language Processing

Tasks in NLP :

- Language Modeling

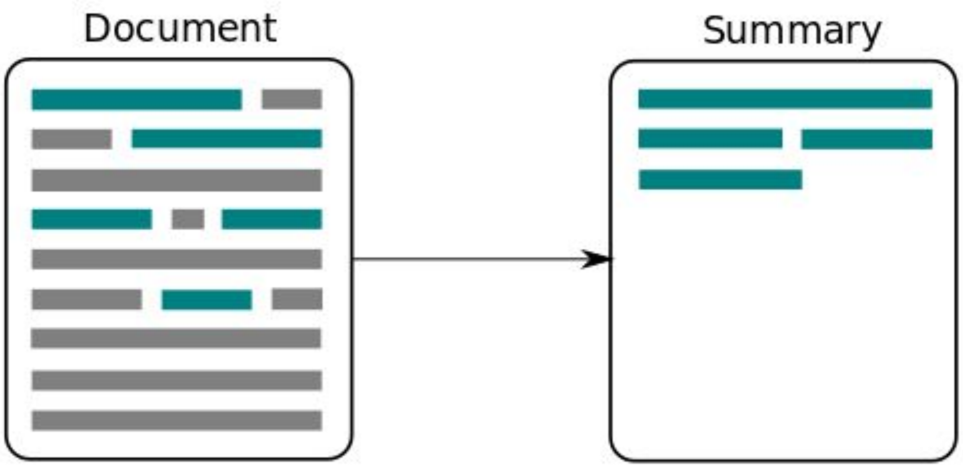




1. Natural Language Processing

Tasks in NLP :

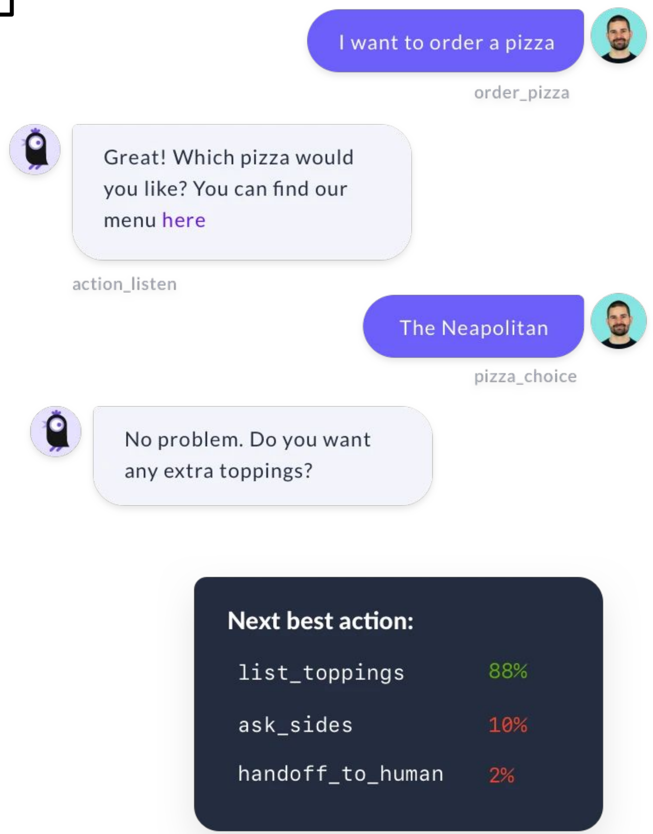
- Text Summarization



1. Natural Language Processing

Tasks in NLP :

- Contextual AI assistants or Chatbots



https://rasa.com/

1. Natural Language Processing

Tasks in NLP :

- Question-Answering

CBS broadcast Super Bowl 50 in the U.S., and charged an average of $5 million for a 30-second commercial during the game. The Super Bowl 50 halftime show was headlined by the British rock group Coldplay with special guest performers Beyoncé and Bruno Mars, who headlined the Super Bowl XLVII and Super Bowl XLVIII halftime shows, respectively. It was the third-most watched U.S. broadcast ever.

Which network broadcasted Super Bowl 50 in the U.S.?
Ground Truth Answers: CBS CBS CBS
Prediction: CBS

What was the average cost for a 30 second commercial during Super Bowl 50?
Ground Truth Answers: $5 million $5 million $5 million
Prediction: $5 million

1. Natural Language Processing

Deep Learning based NLP models :

- RNN, GRU, LSTM
- Attention based RNNs

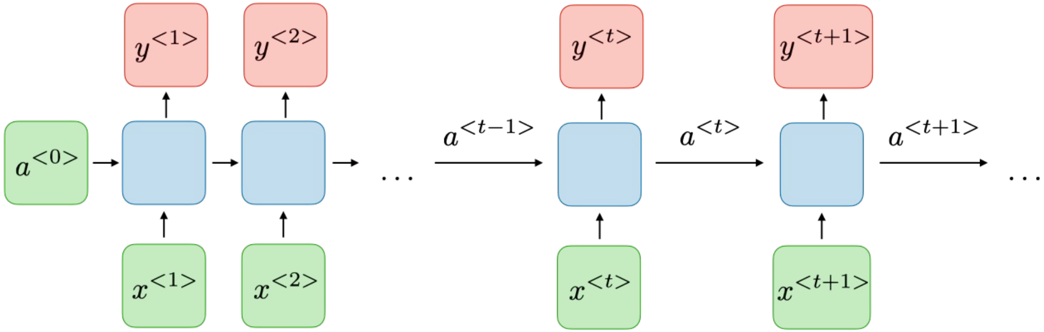
- Transformer, BERT
- GPT, GPT2, GPT3

1. Natural Language Processing

Deep Learning based NLP models :

- RNN, GRU, LSTM
- Attention based RNNs

- Transformer, BERT
- GPT, GPT2, GPT3



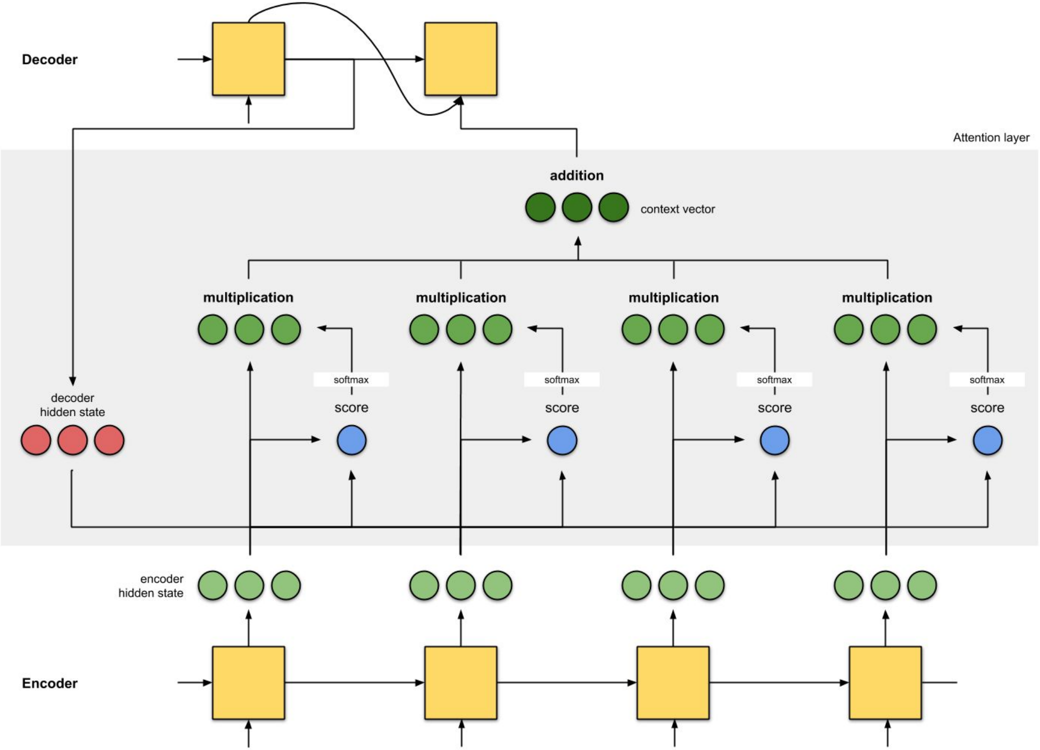
https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks

1. Natural Language Processing

Deep Learning based NLP model:

- RNN, GRU, LSTM
- Attention based RNNs

- Transformer, BERT
- GPT, GPT2, GPT3



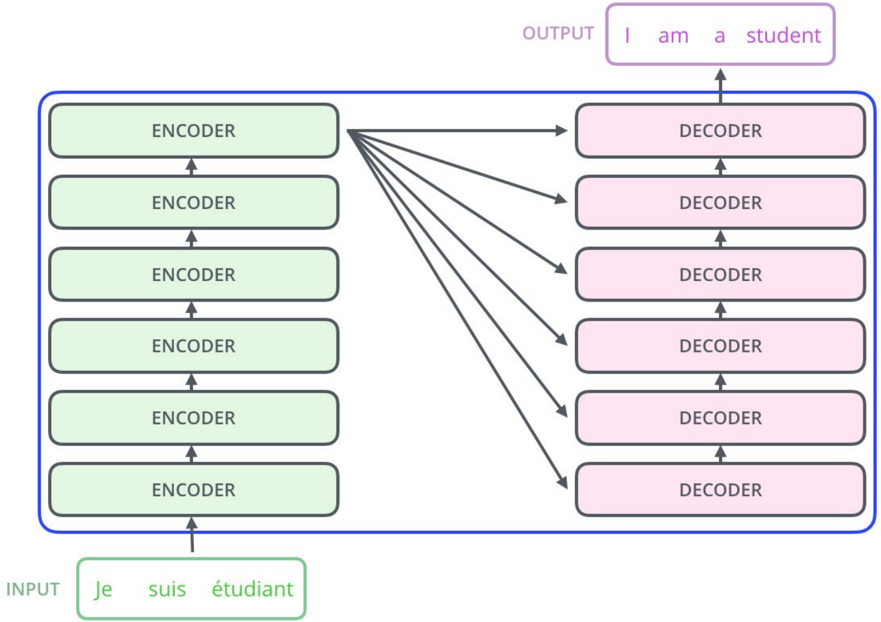
1. Natural Language Processing

Deep Learning based NLP models :

- RNN, GRU, LSTM
- Attention based RNNs

- Transformer, BERT
- GPT, GPT2, GPT3

Imagenet moment for NLP



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

1. Natural Language Processing
2. Word Representation
3. Sequence Models

1. One Hot Encoding
2. Featurized Representation
3. Word2Vec

2. Word Representation

Sentence: Corona has affected the world adversely.

2. Word Representation

Sentence: Corona has affected the world adversely.

Vocabulary: [adversely, affected, corona, has, the, world]

2. Word Representation

One-Hot Encoding

Vocabulary: [adversely, affected, corona, has, the, world]

2. Word Representation

One-Hot Encoding

Vocabulary: [adversely, affected, corona, has, the, world]

Corona	0	0	1	0	0	0
has	0	0	0	1	0	0
affected	0	1	0	0	0	0
the	0	0	0	0	1	0
world	0	0	0	0	0	1
adversely	1	0	0	0	0	0

2. Word Representation

One-Hot Encoding

Vocabulary: [adversely, affected, corona, has, the, world]

Corona	0	0	1	0	0	0
has	0	0	0	1	0	0
affected	0	1	0	0	0	0
the	0	0	0	0	1	0
world	0	0	0	0	0	1
adversely	1	0	0	0	0	0

$$e_{Corona} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

2. Word Representation

One-Hot Encoding

Vocabulary: [adversely, affected, corona, has, the, world]

Corona	0	0	1	0	0	0
has	0	0	0	1	0	0
affected	0	1	0	0	0	0
the	0	0	0	0	1	0
world	0	0	0	0	0	1
adversely	1	0	0	0	0	0

$$e_{has} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

2. Word Representation

One-Hot Encoding

Vocabulary: [adversely, affected, corona, has, the, world]

Corona	0	0	1	0	0	0
has	0	0	0	1	0	0
affected	0	1	0	0	0	0
the	0	0	0	0	1	0
world	0	0	0	0	0	1
adversely	1	0	0	0	0	0

$$e_{affected} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

2. Word Representation

One-Hot Encoding

Vocabulary: [adversely, affected, corona, has, the, world]

Corona	0	0	1	0	0	0
has	0	0	0	1	0	0
affected	0	1	0	0	0	0
the	0	0	0	0	1	0
world	0	0	0	0	0	1
adversely	1	0	0	0	0	0

$$e_{the} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

2. Word Representation

One-Hot Encoding

Vocabulary: [a, ..., adversely, ..., affected, ...,corona, ..., has, ..., the, ..., world, ...]

2. Word Representation

One-Hot Encoding

	0		120		363		4346		7232		23213		27316		
Vocabulary:	[a,	...,	adversely,	...,	affected,	...,	corona,	...,	has,	...,	the,	...,	world,	...]	

2. Word Representation

One-Hot Encoding

	0	120	363	4346	7232	23213	27316	
Vocabulary: [a, ...,		adversely, ...,	affected, ...,	corona, ...,	has, ...,	the, ...,	world,	...]

	120		363		4346		7232		23213		27316
Corona	0	...	0	...	1	...	0	...	0	...	0
has	0	...	0	...	0	...	1	...	0	...	0
affected	0	...	1	...	0	...	0	...	0	...	0
the	0	...	0	...	0	...	0	...	1	...	0
world	0	...	0	...	0	...	0	...	0	...	1
adversely	1	...	0	...	0	...	0	...	0	...	0

2. Word Representation

One-Hot Encoding

	0	120	363	4346	7232	23213	27316
Vocabulary:	[a, ...,	adversely, ...,	affected, ...,	corona, ...,	has, ...,	the, ...,	world, ...]

	120		363		4346		7232		23213		27316
Corona	0	...	0	...	1	...	0	...	0	...	0
has	0	...	0	...	0	...	1	...	0	...	0
affected	0	...	1	...	0	...	0	...	0	...	0
the	0	...	0	...	0	...	0	...	1	...	0
world	0	...	0	...	0	...	0	...	0	...	1
adversely	1	...	0	...	0	...	0	...	0	...	0

$$e_{Corona} = \begin{array}{r} \\ \\ \\ 4346 \\ \\ \\ \\ 29{,}999 \end{array} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

2. Word Representation

One-Hot Encoding

0	120	363	4346	7232	23213	27316

Vocabulary: [a, ..., adversely, ..., affected, ...,corona, ..., has, ..., the, ..., world, ...]

	120		363		4346		7232		23213		27316
Corona	0	...	0	...	1	...	0	...	0	...	0
has	0	...	0	...	0	...	1	...	0	...	0
affected	0	...	1	...	0	...	0	...	0	...	0
the	0	...	0	...	0	...	0	...	1	...	0
world	0	...	0	...	0	...	0	...	0	...	1
adversely	1	...	0	...	0	...	0	...	0	...	0

$$e_{has} = \begin{array}{r} \\ \\ \\ 7232 \\ \\ \\ \\ 29{,}999 \end{array} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

2. Word Representation

One-Hot Encoding

Pitfalls:

- Orthogonal Representation

$$e_{Corona} = \begin{array}{r} \\ \\ \\ 4346 \\ \\ \\ \\ 29{,}999 \end{array} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad e_{has} = \begin{array}{r} \\ \\ \\ 7232 \\ \\ \\ \\ 29{,}999 \end{array} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

2. Word Representation

Featurized Representation

					...	.
Man					...	
Woman					...	
King					...	
Queen					...	
Apple					...	
Orange					...	

2. Word Representation

Featurized Representation

	Gender				...	.
Man					...	
Woman					...	
King					...	
Queen					...	
Apple					...	
Orange					...	

2. Word Representation

Featurized Representation

	Gender				…	.
Man	-0.9				…	
Woman	0.9				…	
King	-0.9				…	
Queen	0.9				…	
Apple	0.01				…	
Orange	0.02				…	

2. Word Representation

Featurized Representation

	Gender	Royal			...	.
Man	-0.9				...	
Woman	0.9				...	
King	-0.9				...	
Queen	0.9				...	
Apple	0.01				...	
Orange	0.02				...	

2. Word Representation

Featurized Representation

	Gender	Royal			...	.
Man	-0.9	0.01			...	
Woman	0.9	0.001			...	
King	-0.9	1			...	
Queen	0.9	1			...	
Apple	0.01	0.02			...	
Orange	0.02	-0.01			...	

2. Word Representation

Featurized Representation

	Gender	Royal	Age		...	.
Man	-0.9	0.01			...	
Woman	0.9	0.001			...	
King	-0.9	1			...	
Queen	0.9	1			...	
Apple	0.01	0.02			...	
Orange	0.02	-0.01			...	

2. Word Representation

Featurized Representation

	Gender	Royal	Age		...	.
Man	-0.9	0.01	0.03		...	
Woman	0.9	0.001	0.02		...	
King	-0.9	1	0.85		...	
Queen	0.9	1	0.78		...	
Apple	0.01	0.02	0.03		...	
Orange	0.02	-0.01	0.02		...	

2. Word Representation

Featurized Representation

	Gender	Royal	Age	Food	...	.
Man	-0.9	0.01	0.03		...	
Woman	0.9	0.001	0.02		...	
King	-0.9	1	0.85		...	
Queen	0.9	1	0.78		...	
Apple	0.01	0.02	0.03		...	
Orange	0.02	-0.01	0.02		...	

2. Word Representation

Featurized Representation

	Gender	Royal	Age	Food	...	.
Man	-0.9	0.01	0.03	0.01	...	
Woman	0.9	0.001	0.02	0.001	...	
King	-0.9	1	0.85	-0.01	...	
Queen	0.9	1	0.78	0.0	...	
Apple	0.01	0.02	0.03	0.93	...	
Orange	0.02	-0.01	0.02	0.95	...	

2. Word Representation

Featurized Representation

	Gender	Royal	Age	Food	...	300th.
Man	-0.9	0.01	0.03	0.01	...	
Woman	0.9	0.001	0.02	0.001	...	
King	-0.9	1	0.85	-0.01	...	
Queen	0.9	1	0.78	0.0	...	
Apple	0.01	0.02	0.03	0.93	...	
Orange	0.02	-0.01	0.02	0.95	...	

2. Word Representation

Featurized Representation

	Gender	Royal	Age	Food	...	300th.
Man	-0.9	0.01	0.03	0.01	...	
Woman	0.9	0.001	0.02	0.001	...	
King	-0.9	1	0.85	-0.01	...	
Queen	0.9	1	0.78	0.0	...	
Apple	0.01	0.02	0.03	0.93	...	
Orange	0.02	-0.01	0.02	0.95	...	

2. Word Representation

Featurized Representation

	Gender	Royal	Age	Food
Man	-0.9	0.01	0.03	0.01
Woman	0.9	0.001	0.02	0.001
King	-0.9	1	0.85	-0.01
Queen	0.9	1	0.78	0.0
Apple	0.01	0.02	0.03	0.93
Orange	0.02	-0.01	0.02	0.95

2. Word Representation

Featurized Representation

Man → Woman as **King → ?**

2. Word Representation

Featurized Representation

Man → Woman as King → ?

$$e_{Man} - e_{Woman} \approx e_{King} - e_{?}$$

2. Word Representation

Featurized Representation

Man → Woman as King → ?

$$e_{Man} - e_{Woman} \approx e_{King} - e_{?}$$

$$e_{?} \approx e_{King} - e_{Man} + e_{Woman}$$

2. Word Representation

Featurized Representation

Man → Woman as King → ?

$e_{Man} - e_{Woman} \approx e_{King} - e_{?}$

$e_{?} \approx e_{King} - e_{Man} + e_{Woman}$

$$e_{?} = \arg\max_{w} \ \mathrm{sim}(e_{w}, e_{King} - e_{Man} + e_{Woman})$$

2. Word Representation

Featurized Representation

Man → Woman as King → ?

$$e_{Man} - e_{Woman} \approx e_{King} - e_{?}$$

$$e_{?} \approx e_{King} - e_{Man} + e_{Woman}$$

$$e_{?} = \arg\max_{w} \ sim(e_w, e_{King} - e_{Man} + e_{Woman})$$

	Gender	Royal	Age	Food
Man	-0.9	0.01	0.03	0.01
Woman	0.9	0.001	0.02	0.001
King	-0.9	1	0.85	-0.01
Queen	0.9	1	0.78	0.0
Apple	0.01	0.02	0.03	0.93
Orange	0.02	-0.01	0.02	0.95

2. Word Representation

Featurized Representation

Man → Woman as King → ?

$e_{Man} - e_{Woman} \approx e_{King} - e_{?}$

$e_{?} \approx e_{King} - e_{Man} + e_{Woman}$

$e_{?} = \arg\max_{w} \ \mathrm{sim}(e_{w}, e_{King} - e_{Man} + e_{Woman})$

$e_{King} - e_{Man} + e_{Woman} =$

	Gender	Royal	Age	Food
Man	-0.9	0.01	0.03	0.01
Woman	0.9	0.001	0.02	0.001
King	-0.9	1	0.85	-0.01
Queen	0.9	1	0.78	0.0
Apple	0.01	0.02	0.03	0.93
Orange	0.02	-0.01	0.02	0.95

2. Word Representation

Featurized Representation

Man → Woman as King → ?

$e_{Man} - e_{Woman} \approx e_{King} - e_?$

$e_? \approx e_{King} - e_{Man} + e_{Woman}$

$e_? = \underset{w}{\arg\max}\ sim(e_w, e_{King} - e_{Man} + e_{Woman})$

$\boxed{e_{King} - e_{Man} + e_{Woman} = [0.9, 0.991, 0.84, -0.019]}$

	Gender	Royal	Age	Food
Man	-0.9	0.01	0.03	0.01
Woman	0.9	0.001	0.02	0.001
King	-0.9	1	0.85	-0.01
Queen	0.9	1	0.78	0.0
Apple	0.01	0.02	0.03	0.93
Orange	0.02	-0.01	0.02	0.95

2. Word Representation

Featurized Representation

Man → Woman as King → ?

$e_{Man} - e_{Woman} \approx e_{King} - e_?$

$e_? \approx e_{King} - e_{Man} + e_{Woman}$

$e_? = \underset{w}{\arg\max}\ sim(e_w, e_{King} - e_{Man} + e_{Woman})$

$\boxed{e_{King} - e_{Man} + e_{Woman} = [0.9, 0.991, 0.84, -0.019]}$

	Gender	Royal	Age	Food
Man	-0.9	0.01	0.03	0.01
Woman	0.9	0.001	0.02	0.001
King	-0.9	1	0.85	-0.01
Queen	0.9	1	0.78	0.0
Apple	0.01	0.02	0.03	0.93
Orange	0.02	-0.01	0.02	0.95

2. Word Representation

Featurized Representation

Man → Woman as King --> ?

$e_{Man} - e_{Woman} \approx e_{King} - e_?$

$e_? \approx e_{King} - e_{Man} + e_{Woman}$

$e_? = \arg\max_{w} \; sim(e_w, e_{King} - e_{Man} + e_{Woman})$

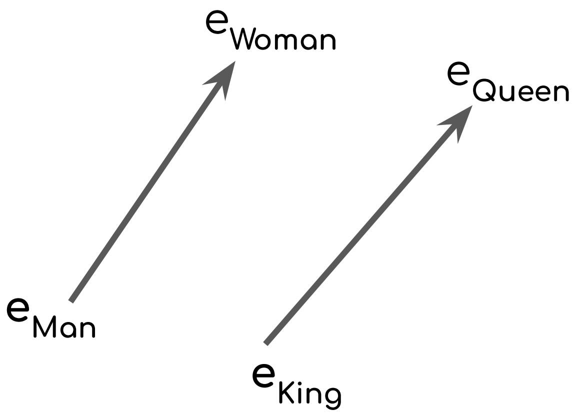
$\boxed{e_{King} - e_{Man} + e_{Woman} = [0.9, 0.991, 0.84, -0.019]}$

	Gender	Royal	Age	Food
Man	-0.9	0.01	0.03	0.01
Woman	0.9	0.001	0.02	0.001
King	-0.9	1	0.85	-0.01
Queen	0.9	1	0.78	0.0
Apple	0.01	0.02	0.03	0.93
Orange	0.02	-0.01	0.02	0.95

2. Word Representation

Featurized Representation

Man → Woman as **King --> Queen**

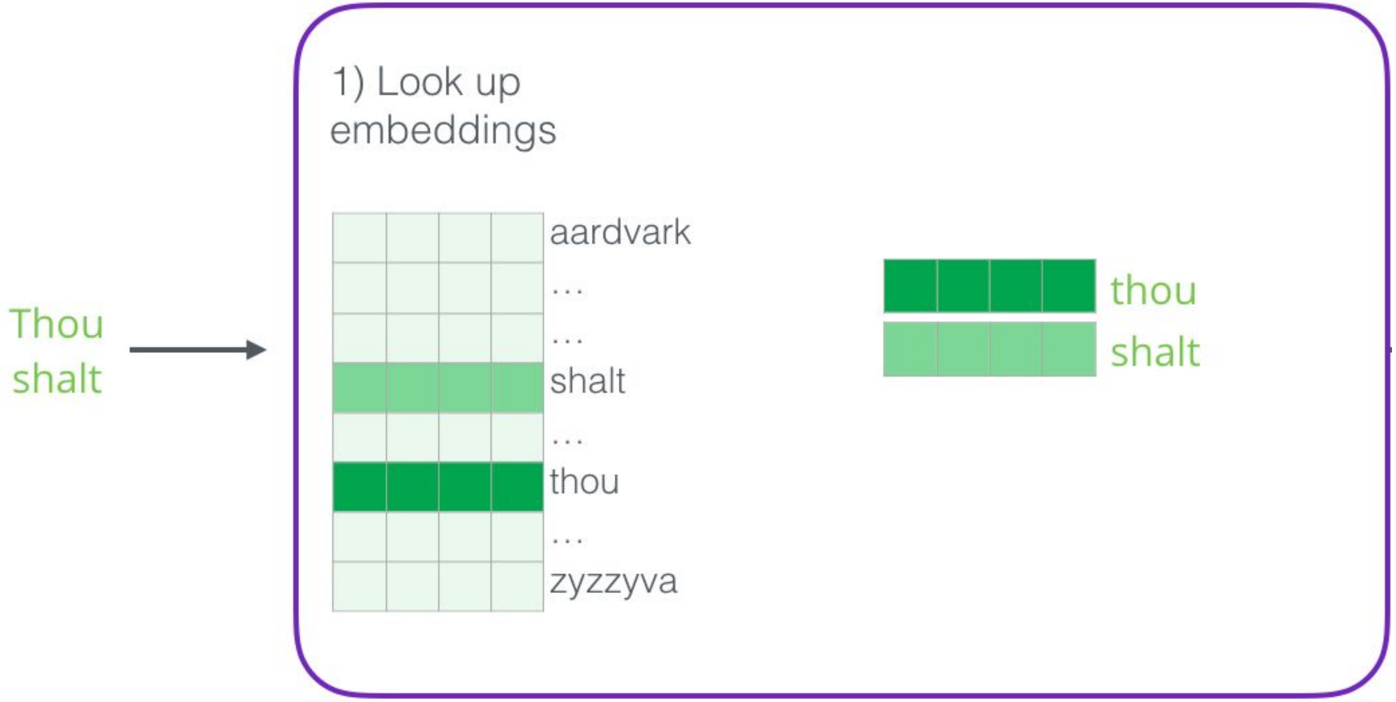


2. Word Representation

Word2Vec

2. Word Representation

Word2Vec

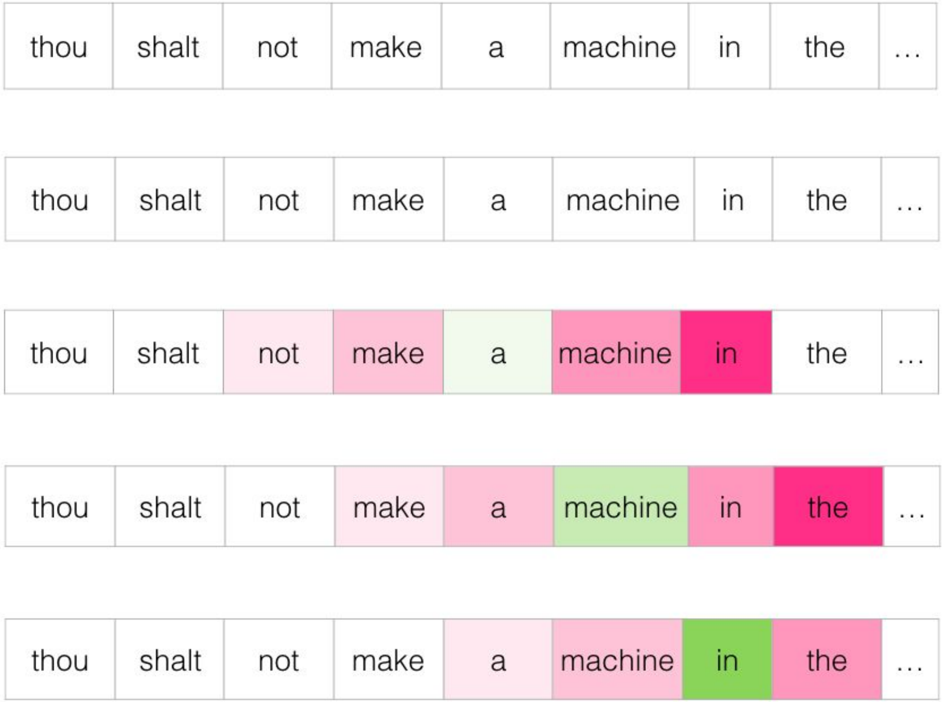


http://jalammar.github.io/illustrated-word2vec/

2. Word Representation

Word2Vec

Thou shalt not make a machine in the likeness of a human mind

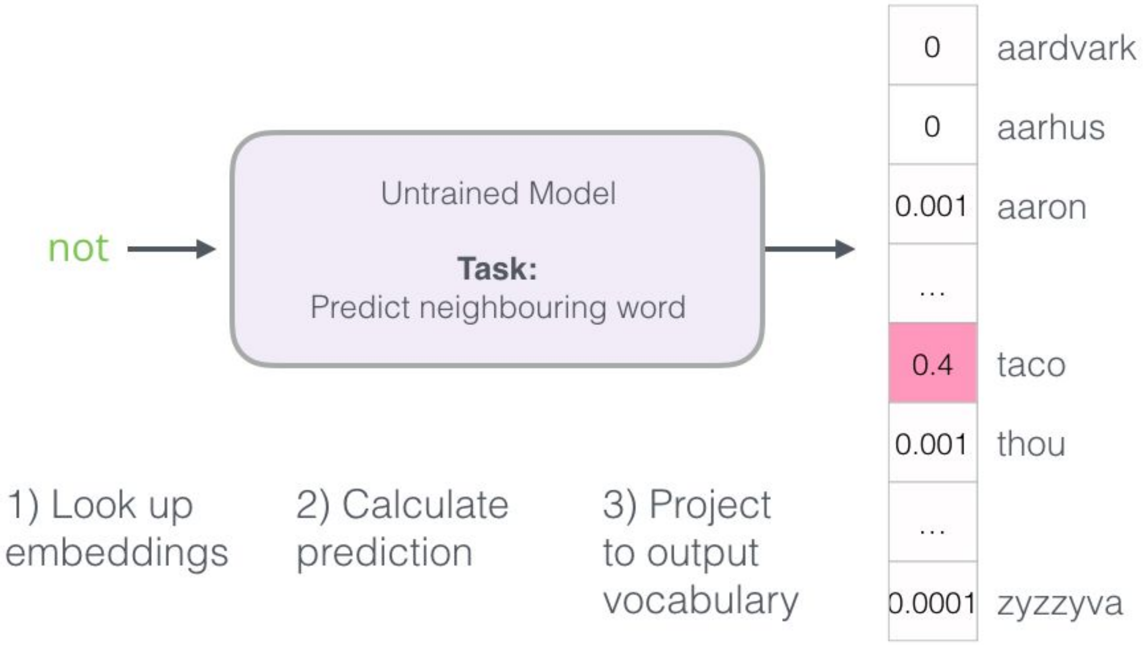


input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness

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

2. Word Representation

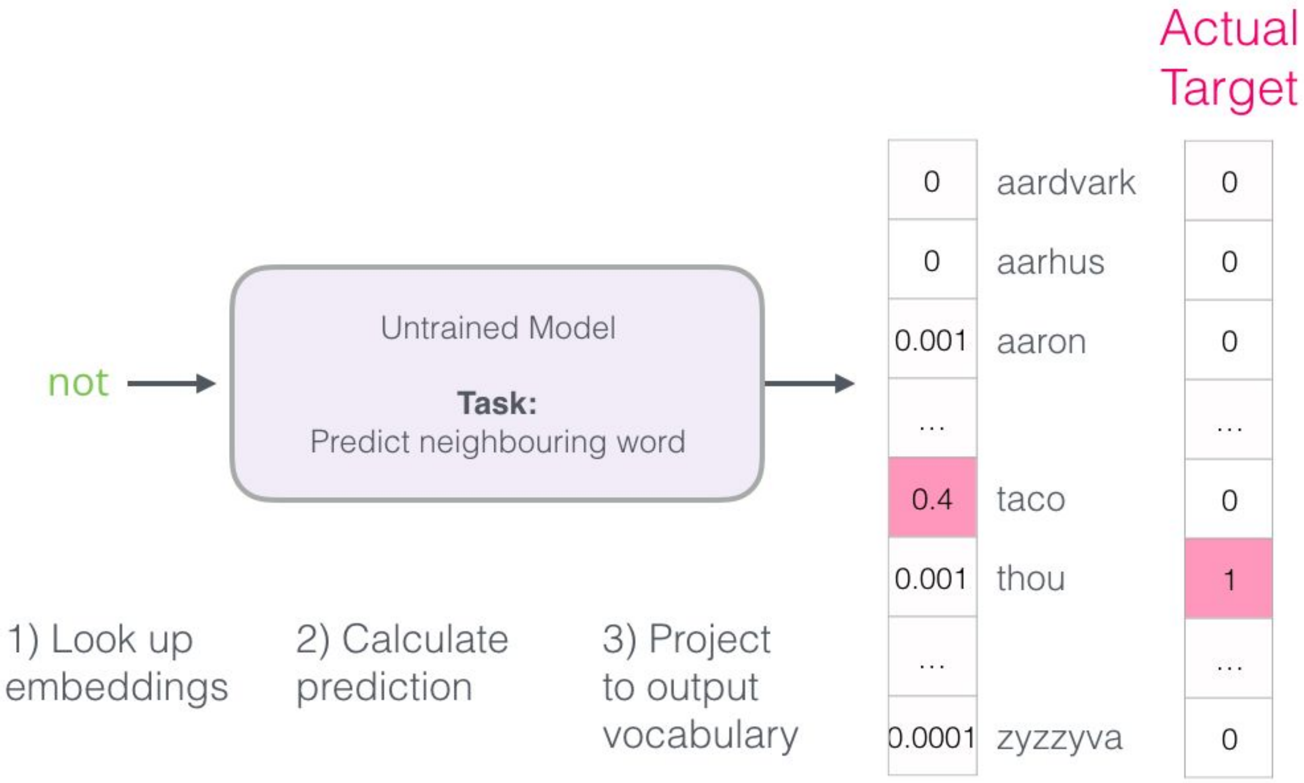
Word2Vec



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

2. Word Representation

Word2Vec



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

1. Natural Language Processing

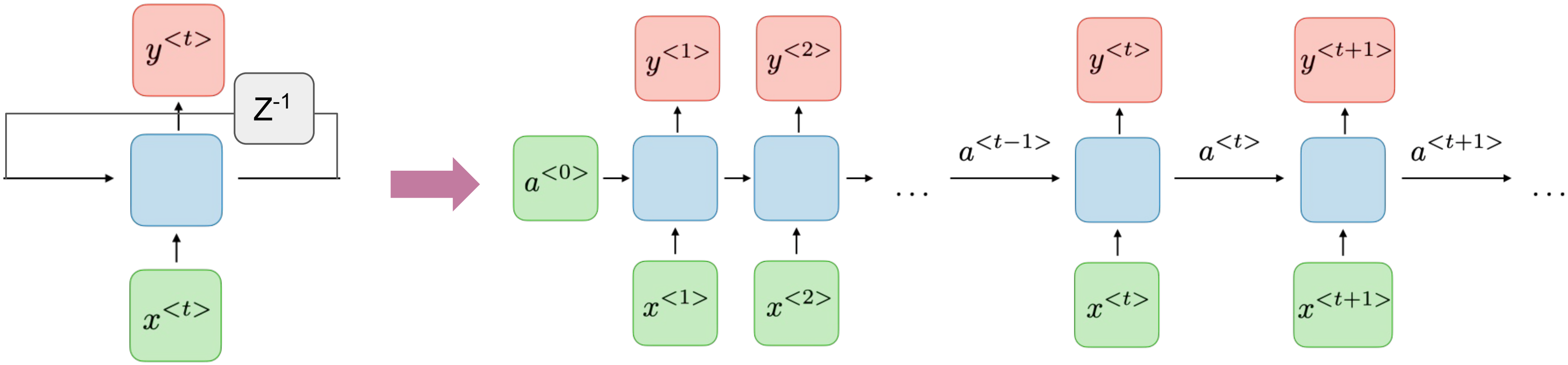
2. Word Representation

3. Sequence Models

1. RNN

3. Sequence Models

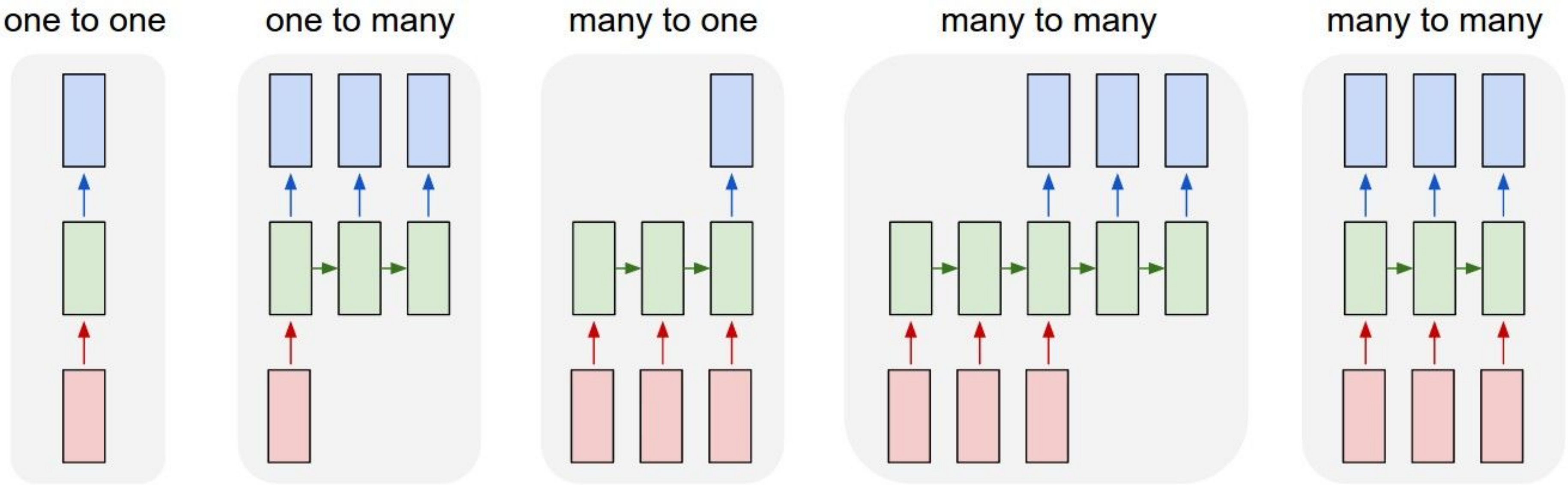
RNN



https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks

3. Sequence Models

RNN



http://karpathy.github.io/2015/05/21/rnn-effectiveness/

Thank you!