Deep Learning Intro

2018. 5. 11.

Lee, Gyeongbok

DI-LAB
Data Intelligence Laboratory

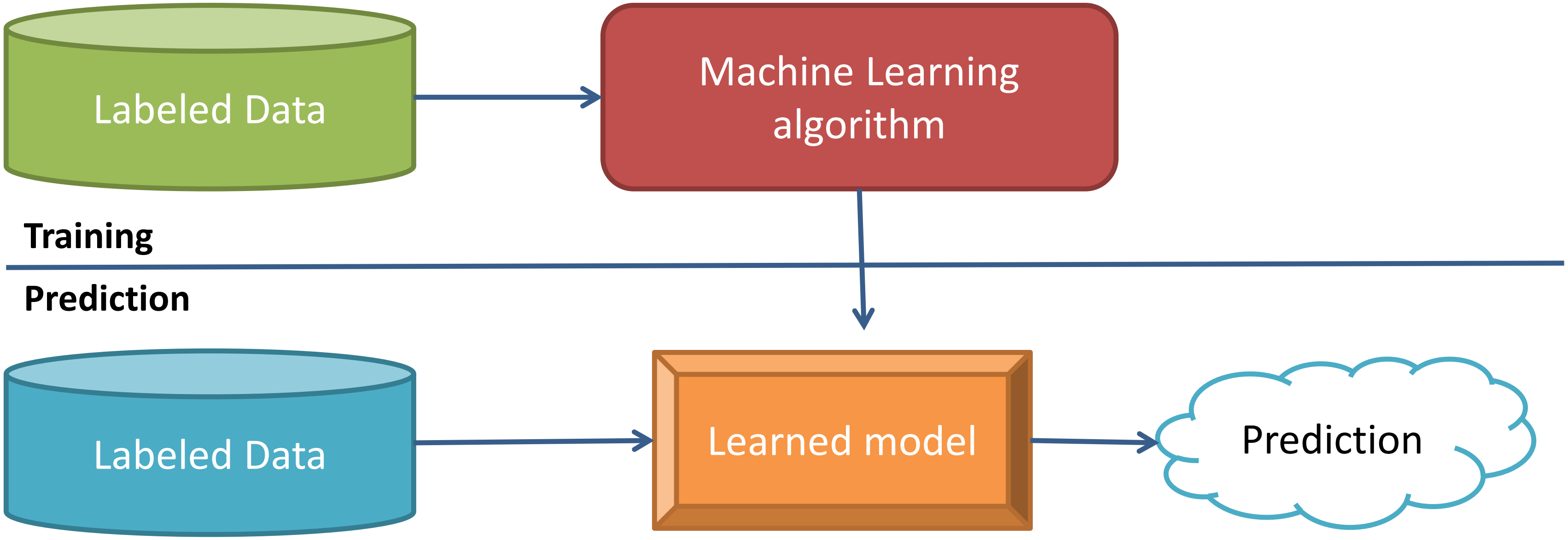
Contents

- Machine Learning and Deep Learning
- Neural Network Architectures
  - Convolutional Neural Network (CNN)
  - Recurrent Neural Network (RNN)
- …and some practices later (with `pytorch`)

※ Most of the material is from [1], [2], [3] in References slide.

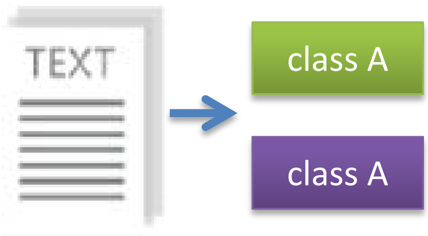
Machine Learning?

- Machine learning: a field of computer science that gives computers the ability to learn without being explicitly programmed
  - Can learn from and make predictions on data



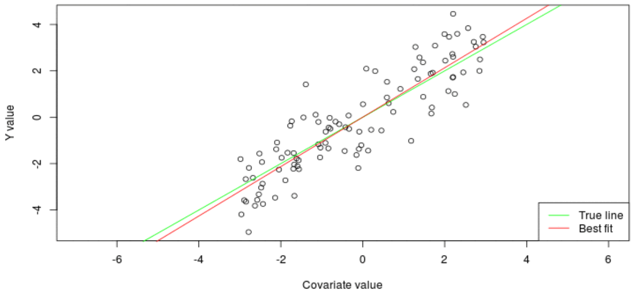
Types of Learning

- **Supervised**: Learning with a labeled training set
  - email *classification* with already labeled emails
- **Unsupervised**: Discover patterns in unlabeled data
  - *cluster* similar documents based on text
- **Reinforcement learning**: learn to act based on feedback/reward
  - Go agent (alphaGo) - reward: *win or lose*

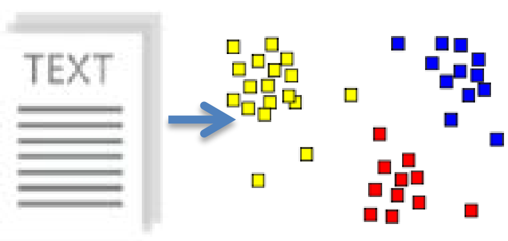


Classification

Anomaly Detection
Sequence labeling
...



Regression

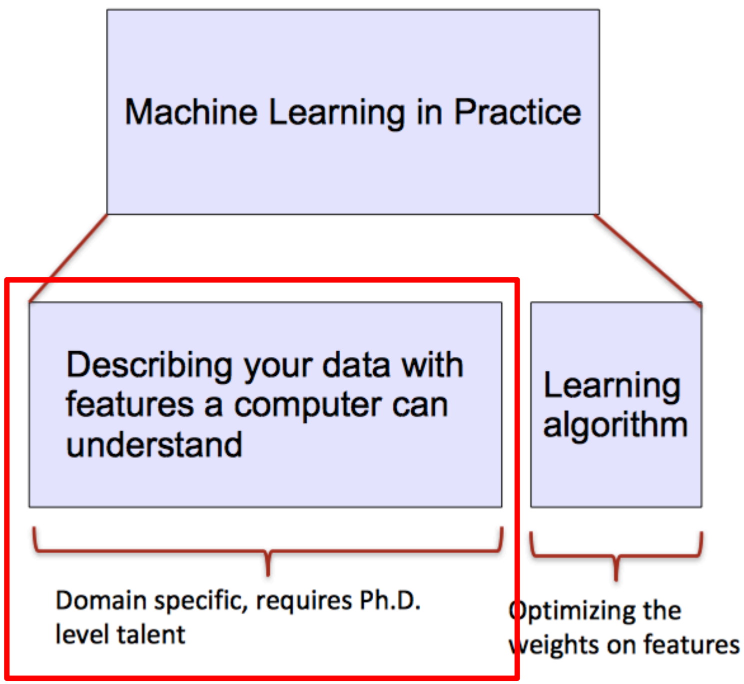


Clustering

http://mbjoseph.github.io/2013/11/27/measure.html

ML vs. Deep Learning

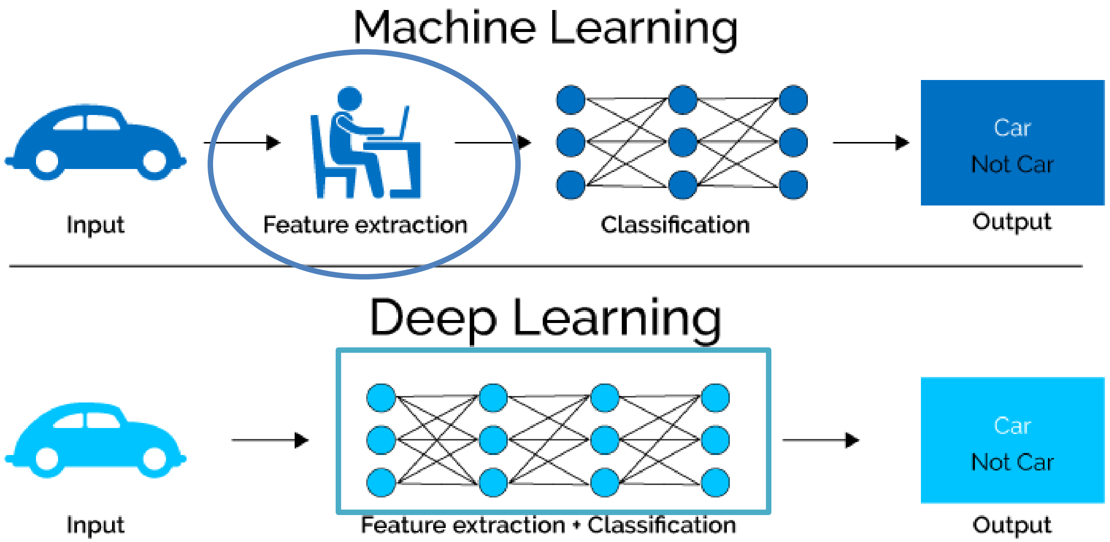
- Still needs human works
  - Most machine learning methods work well because of human-designed representations and input features
  - ML becomes just optimizing weights to best make a final prediction (tuning)



Feature	NER
Current Word	✓
Previous Word	✓
Next Word	✓
Current Word Character n-gram	all
Current POS Tag	✓
Surrounding POS Tag Sequence	✓
Current Word Shape	✓
Surrounding Word Shape Sequence	✓
Presence of Word in Left Window	size 4
Presence of Word in Right Window	size 4

Deep Learning?

- Subfield of ML: learning representations of data.
  - Attempt to learn (multiple levels of) representation by using a hierarchy of multiple layers
  - If you provide the system tons of information, it begins to understand it and respond in useful ways.
  - Exceptional effective at learning patterns!

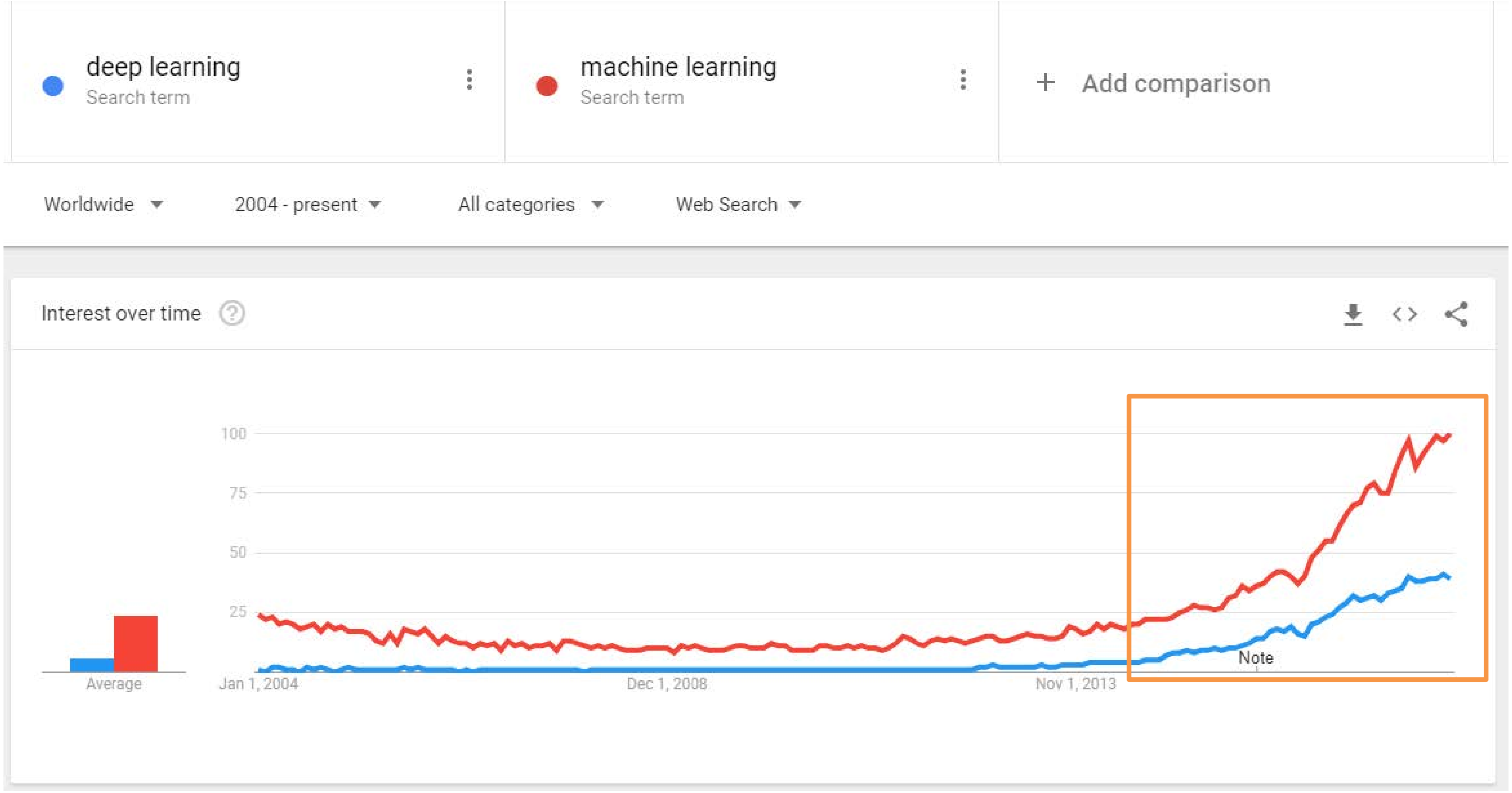


https://www.xenonstack.com/blog/static/public/uploads/media/machine-learning-vs-deep-learning.png

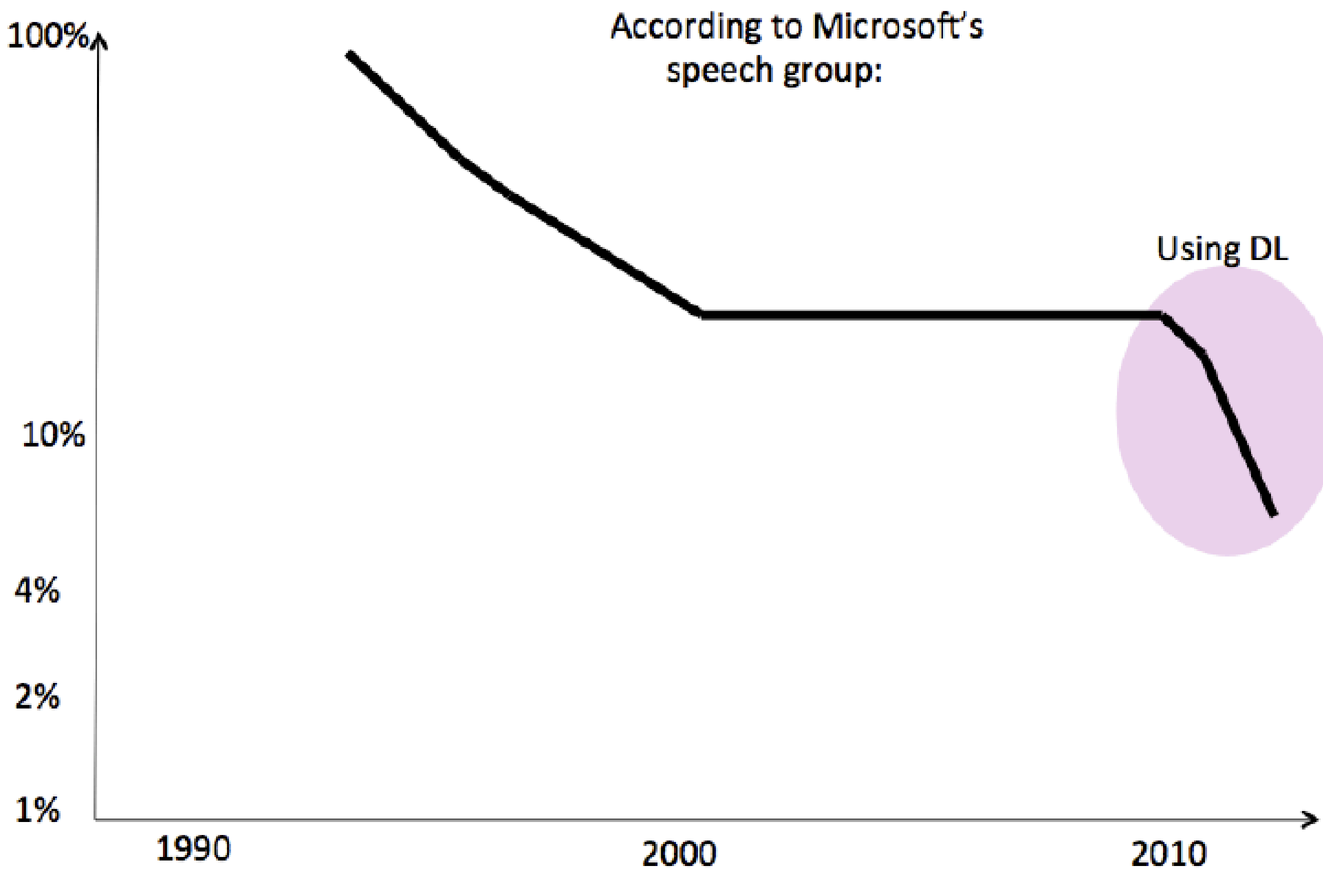
Why is DL useful?

- Existing ML uses manually designed features
  - often over-specified and incomplete
  - take a long time to design and validate
- Learned Features are easy to adapt, fast to learn
- Deep learning provides a very flexible, (almost?) universal, learnable framework for representing world, visual and linguistic information.
  - For both unsupervised and supervised
- Effective end-to-end joint system learning
- Utilize large amounts of training data

In Google Trend…

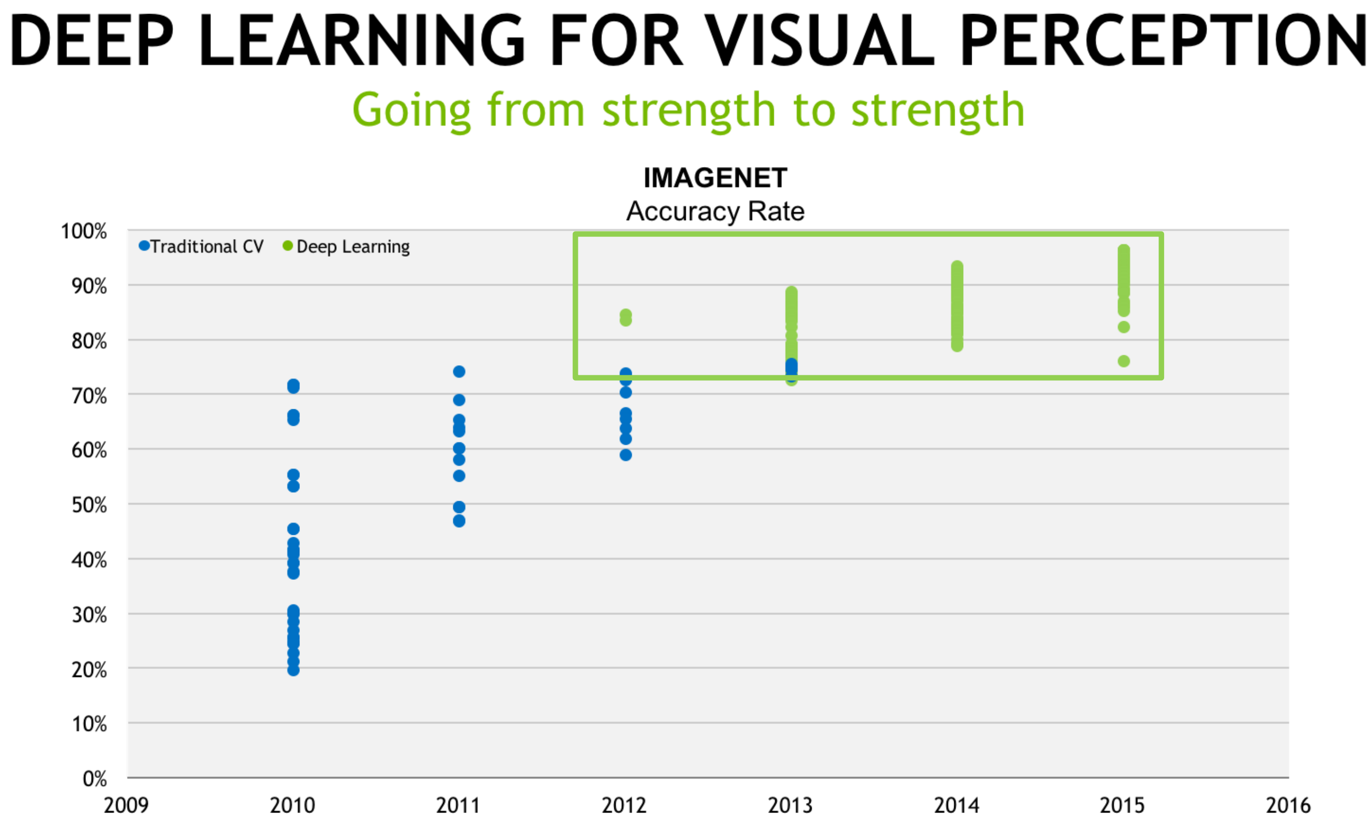


State of the art in …



Deep Learning in Speech Recognition

State of the art in ...

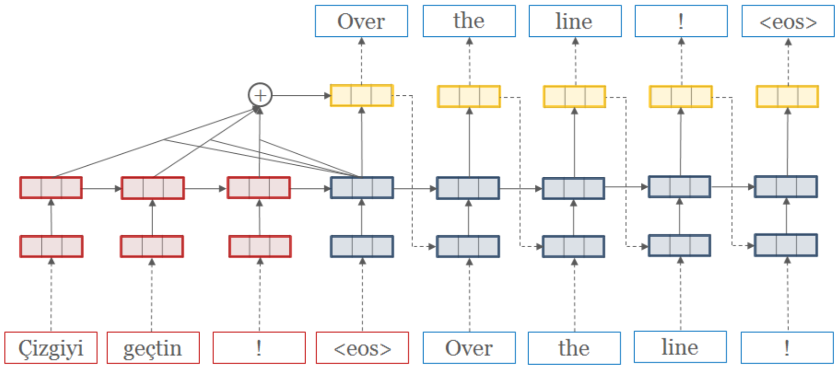


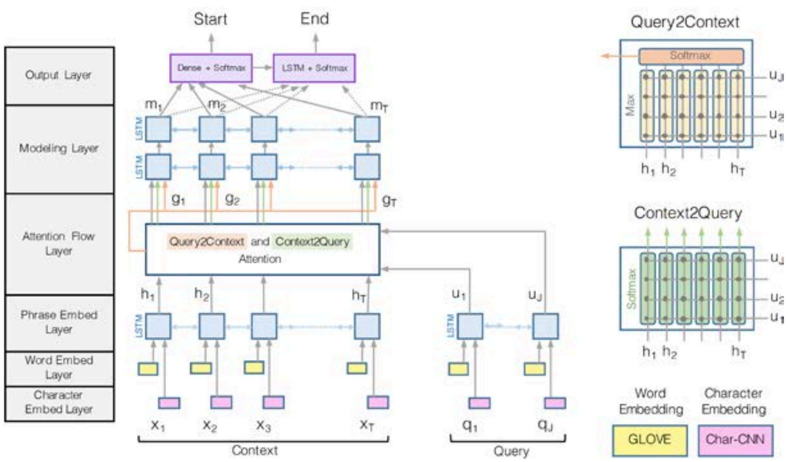
Ends in 2017

Image from https://blogs.nvidia.com/blog/2016/06/29/deep-learning-6/

State of the art in …

- Several big improvements in recent years in NLP
  - Machine Translation
  - Sentiment Analysis
  - Dialogue Agents
  - Question Answering
  - Text Classification





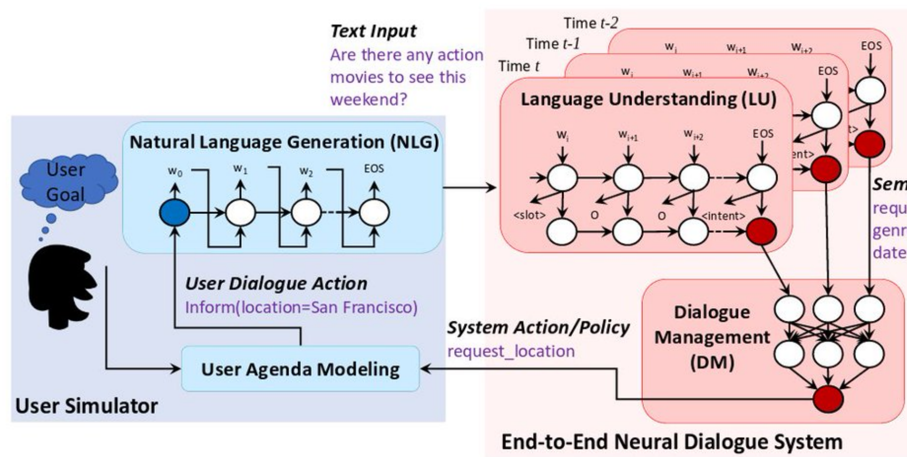
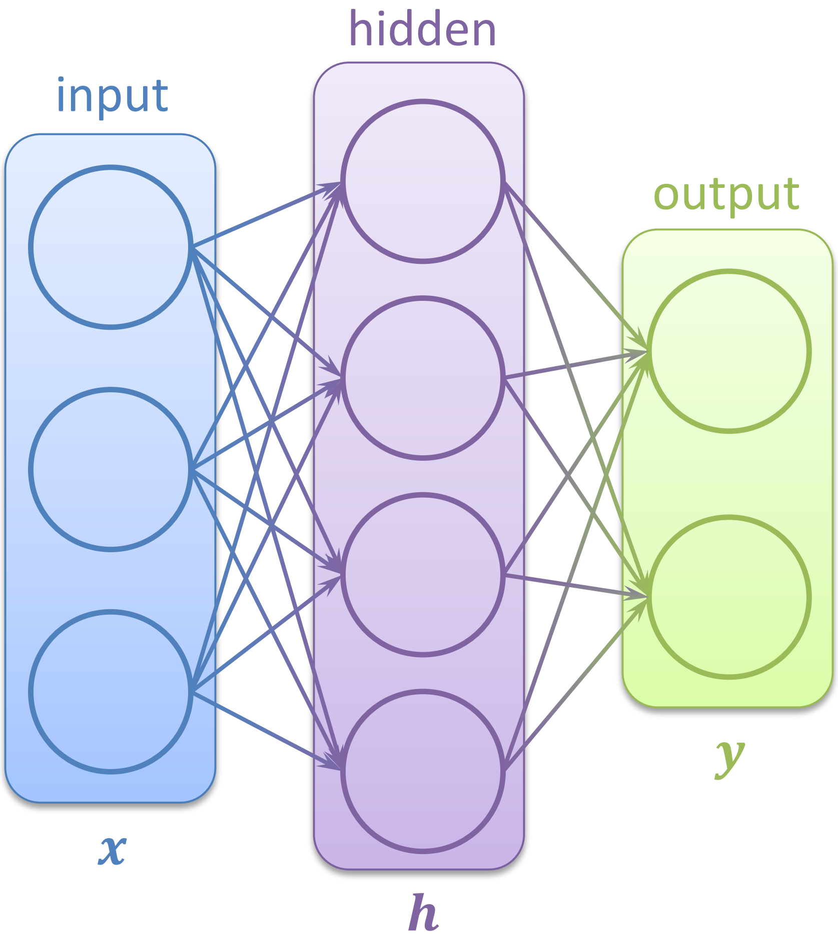


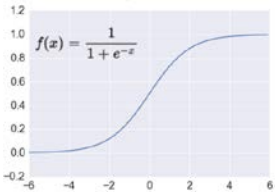
Figure 1: Illustration of the end-to-end neural dialogue system: given user utteranc
learning is used to train all components in an end-to-end fashion.

Neural Network Basis



$$h = \sigma(W_1 x + b_1)$$
$$y = \sigma(W_2 h + b_2)$$

Weights & Activation Functions



4 + 2 = 6 neurons (not counting inputs)
[3 x 4] + [4 x 2] = 20 weights
4 + 2 = 6 biases

26 learnable **parameters**

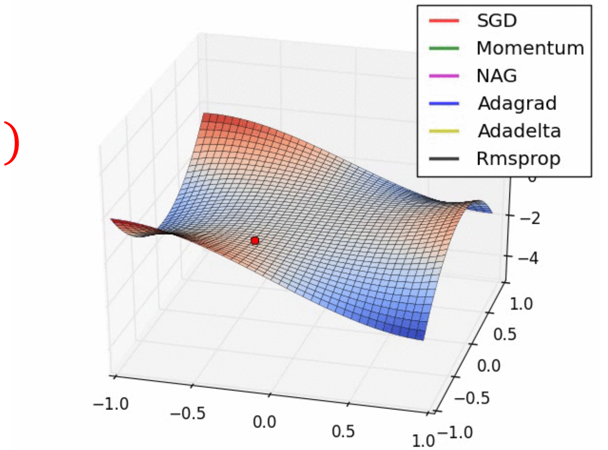
[Demo (tensorflow playground)]

How do we train?

Training Process

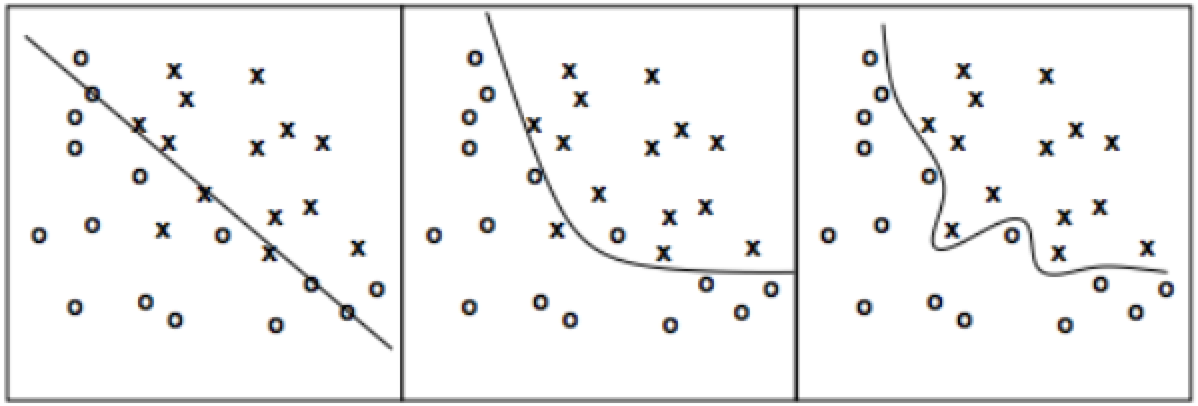
1. Sample labeled data (batch input)
2. Forward it through the network, get predictions
3. Back-propagate the errors
4. Update the network weights

Optimize (min. or max.) objective/cost function $J(\theta)$
Generate error signal that measures difference between predictions and target values



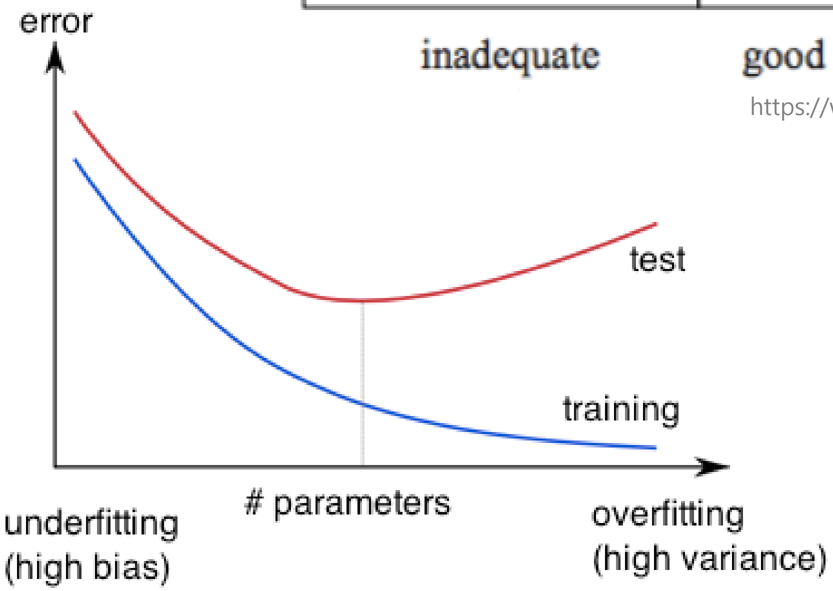
https://medium.com/@ramrajchandradevan/the-evolution-of-gradient-descend-optimization-algorithm-4106a6702d39

Problems?



inadequate good compromise over-fitting

https://www.neuraldesigner.com/images/learning/selection_error.svg



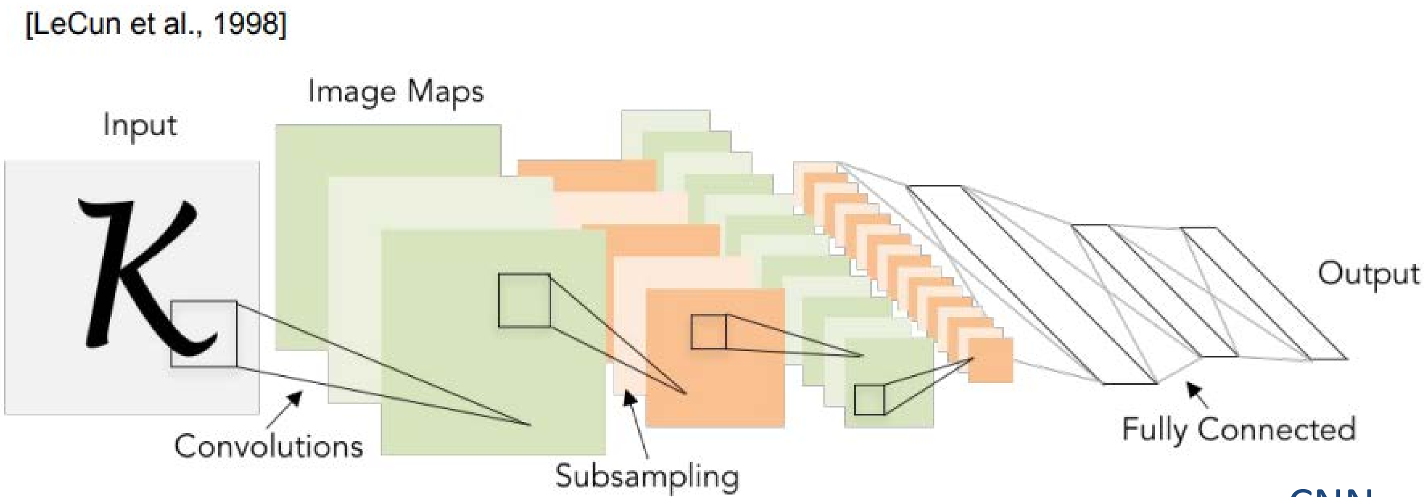
https://www.neuraldesigner.com/images/learning/selection_error.svg

Over-fitting

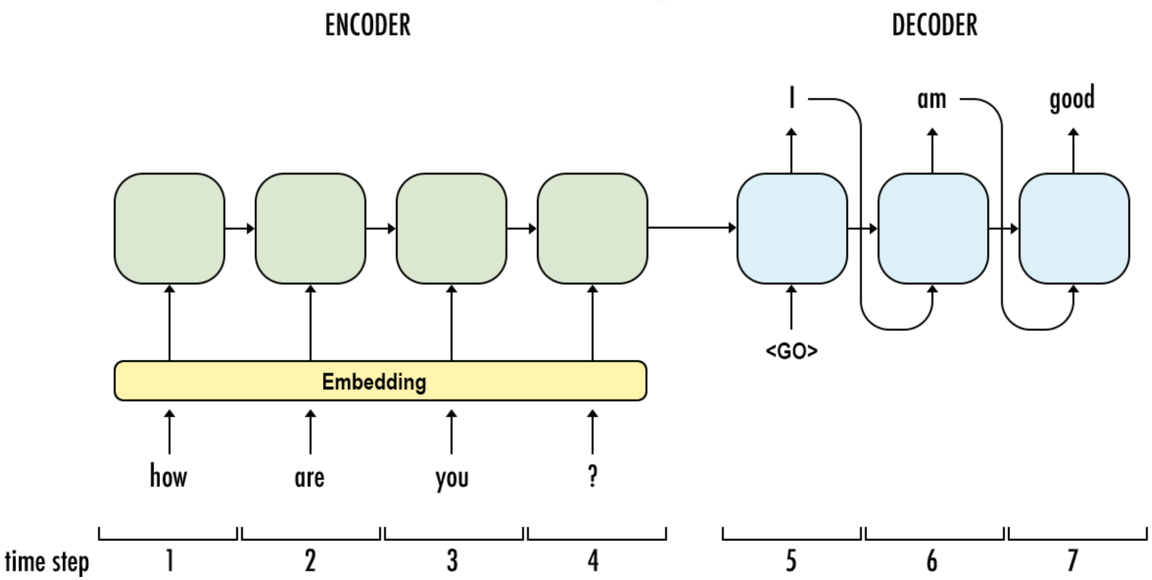
Learned hypothesis may fit the training data very well, even for outliers (noise) but fail to generalize to new examples (test data)

Solution: regularization, etc

Neural Network Architectures



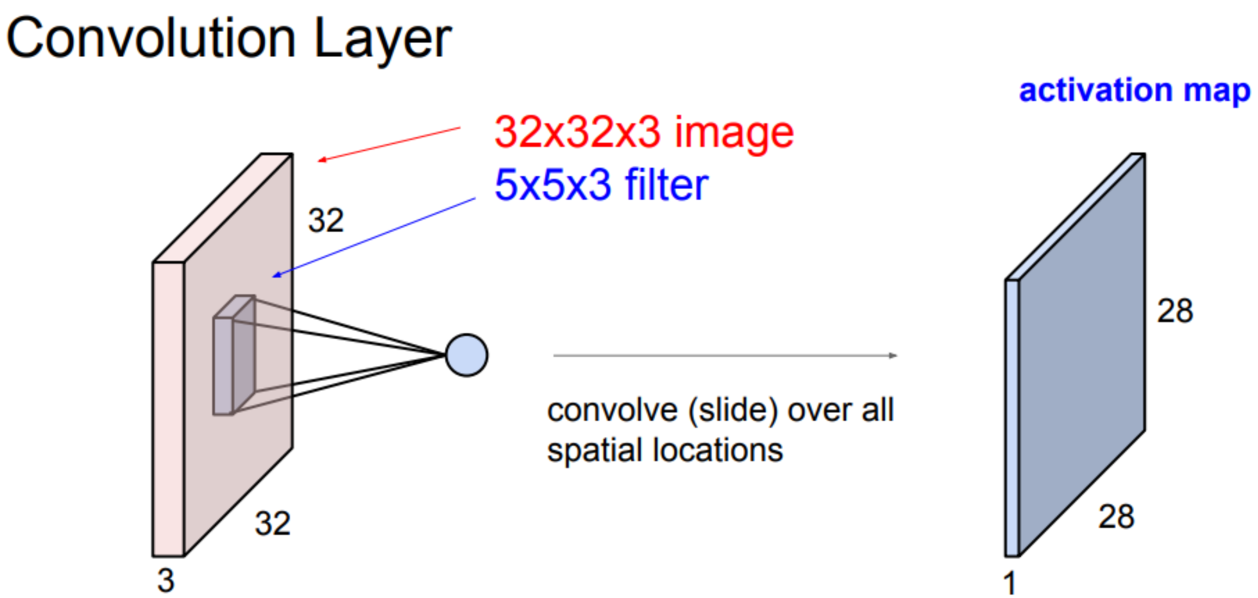
CNN

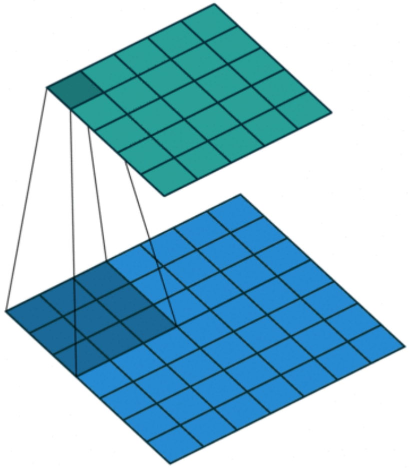


RNN

Images from CS231n lecture slides

Convolution Neural Network (CNN)

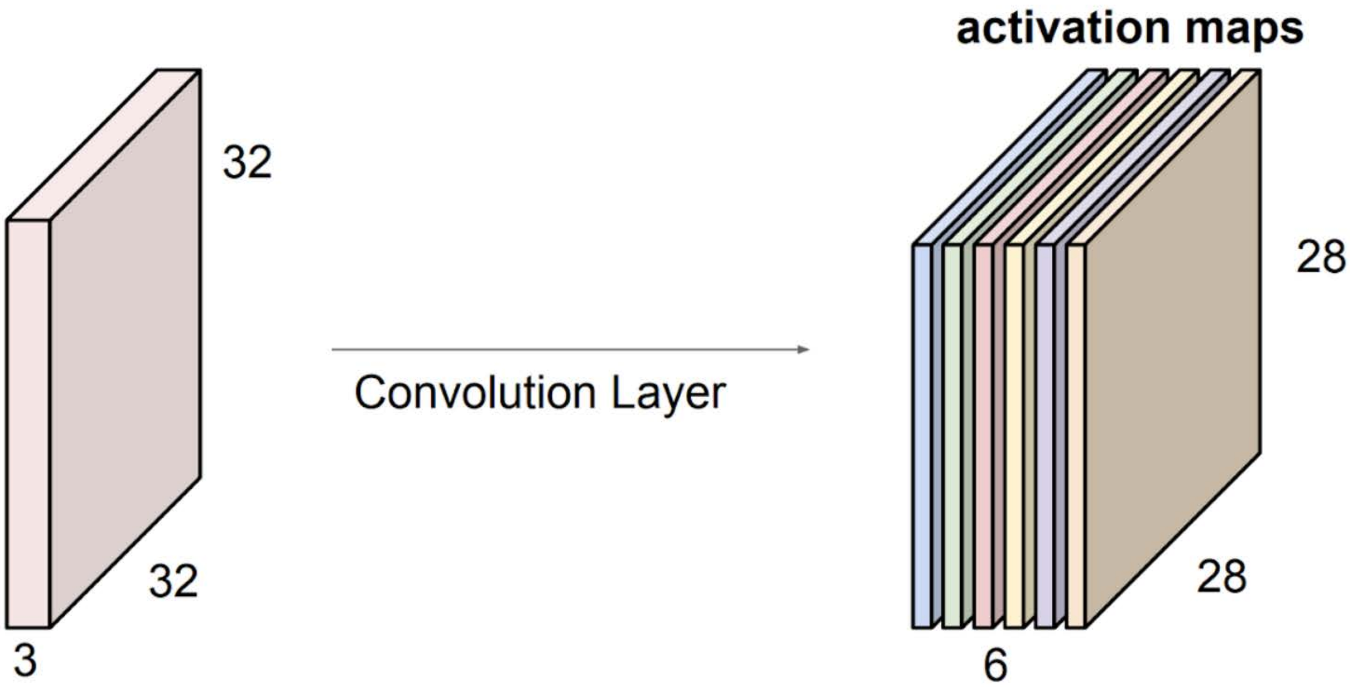




Images from CS231n lecture slides
Animation from https://github.com/vdumoulin/conv_arithmetic

Convolution Neural Network (CNN)

For example, if we had 6 5x5 filters, we'll get 6 separate activation maps:

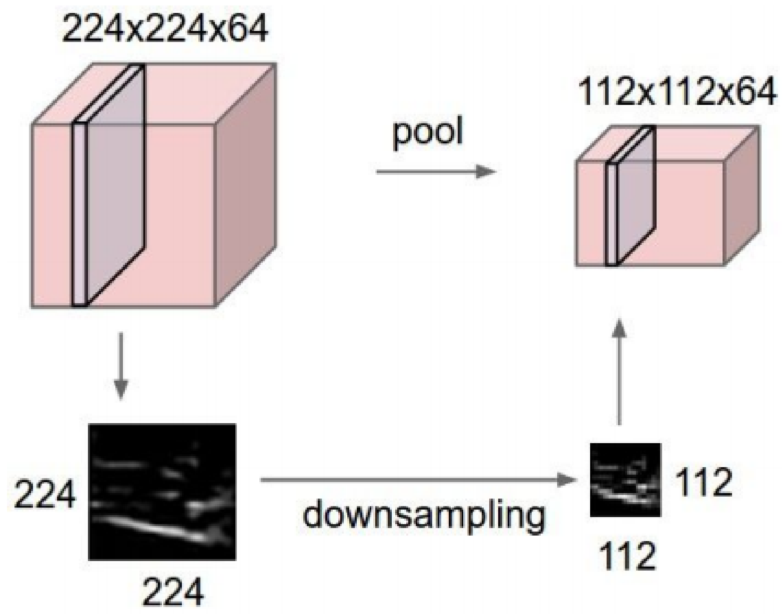


We stack these up to get a "new image" of size 28x28x6!

Pooling layer

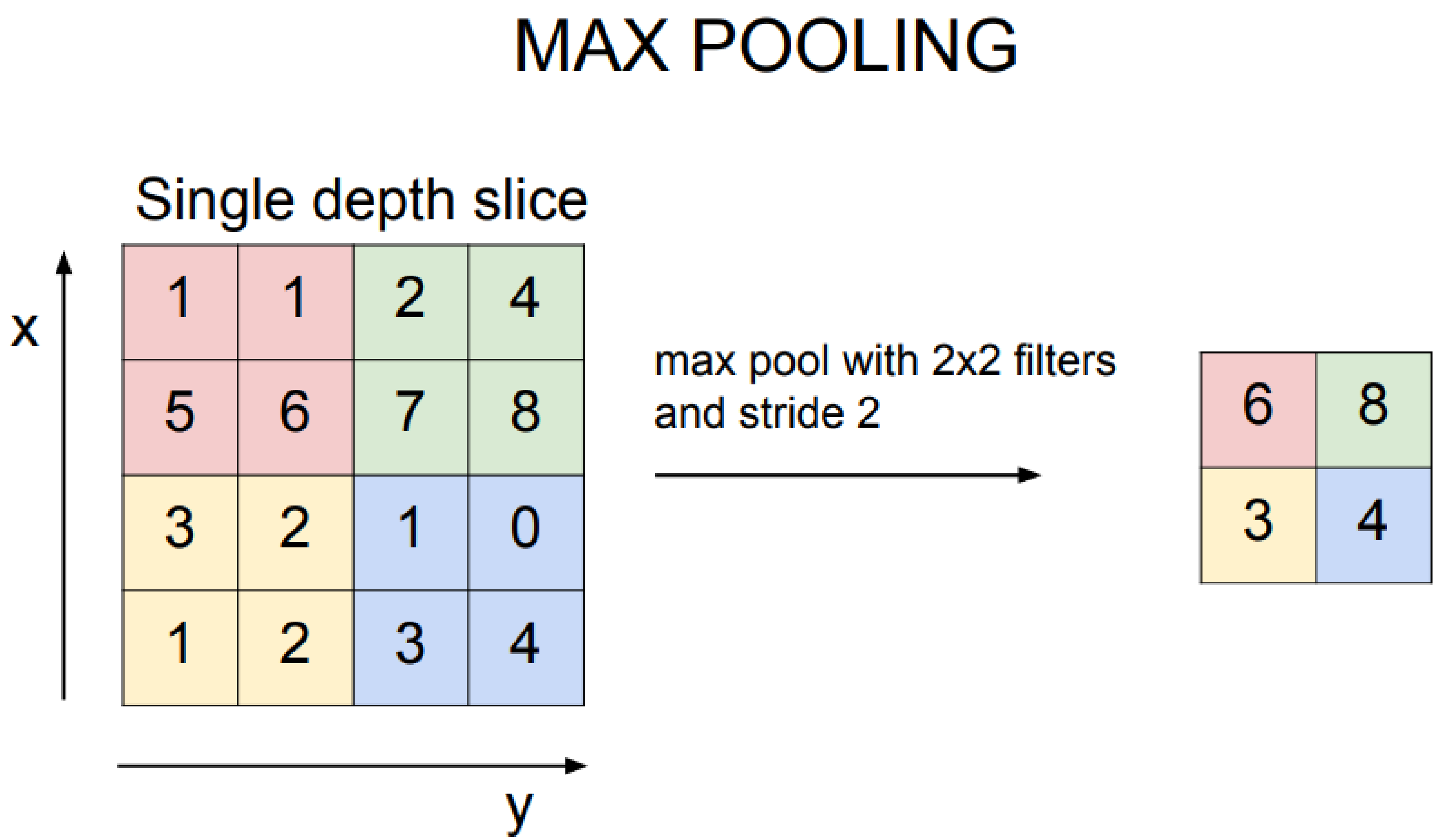
Pooling layer

- makes the representations smaller and more manageable
- operates over each activation map independently:

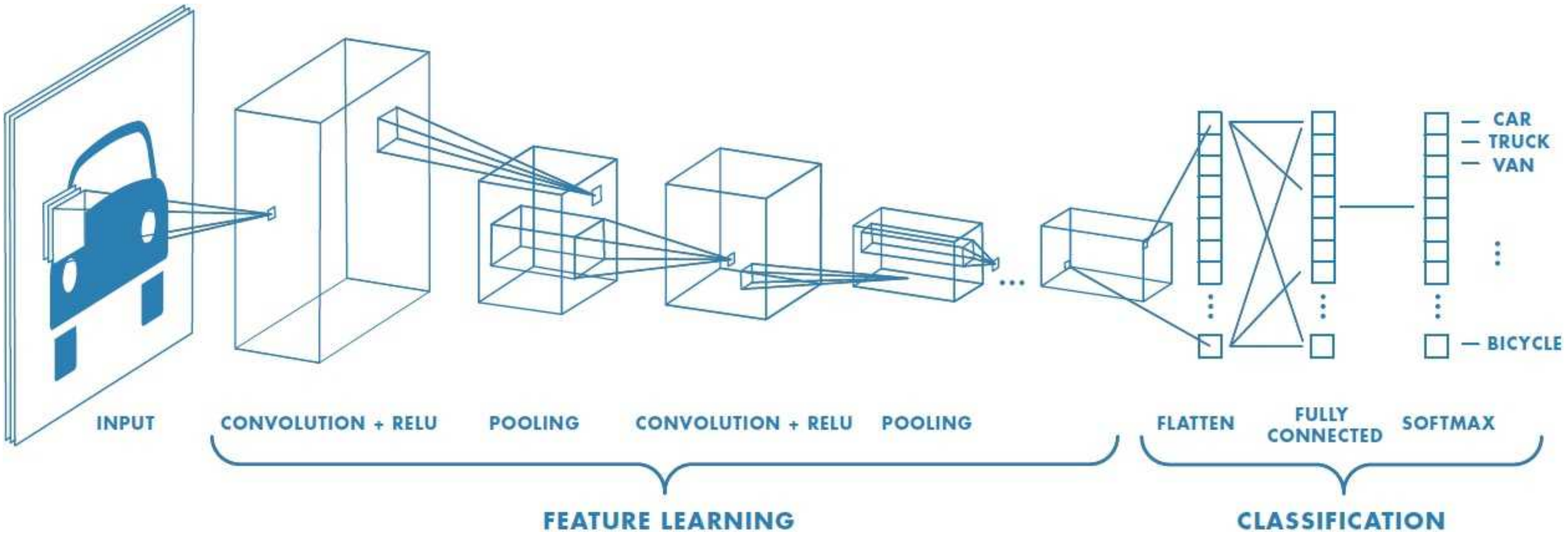


Images from CS231n lecture slides

Max Pooling


MAX POOLING
Single depth slice
x
y
1 1 2 4
5 6 7 8
3 2 1 0
1 2 3 4
max pool with 2x2 filters
and stride 2
6 8
3 4

ConvNet



Images from https://blog.floydhub.com/building-your-first-convnet/

CNN Applications

Fast-forward to today: ConvNets are everywhere

Classification

Retrieval

Figures copyright Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, 2012. Reproduced with permission.

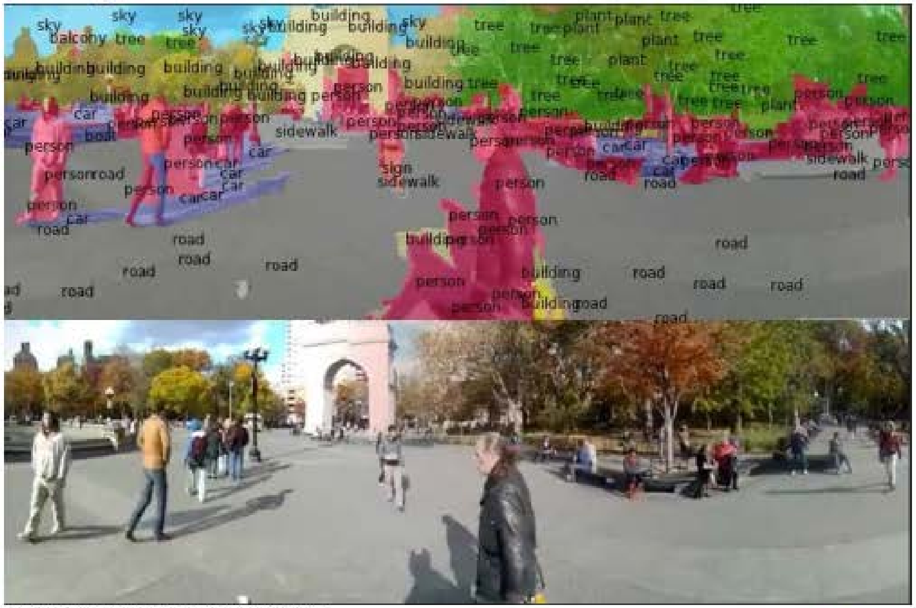
CNN Applications

Fast-forward to today: ConvNets are everywhere

Detection

[Faster R-CNN: Ren, He, Girshick, Sun 2015]

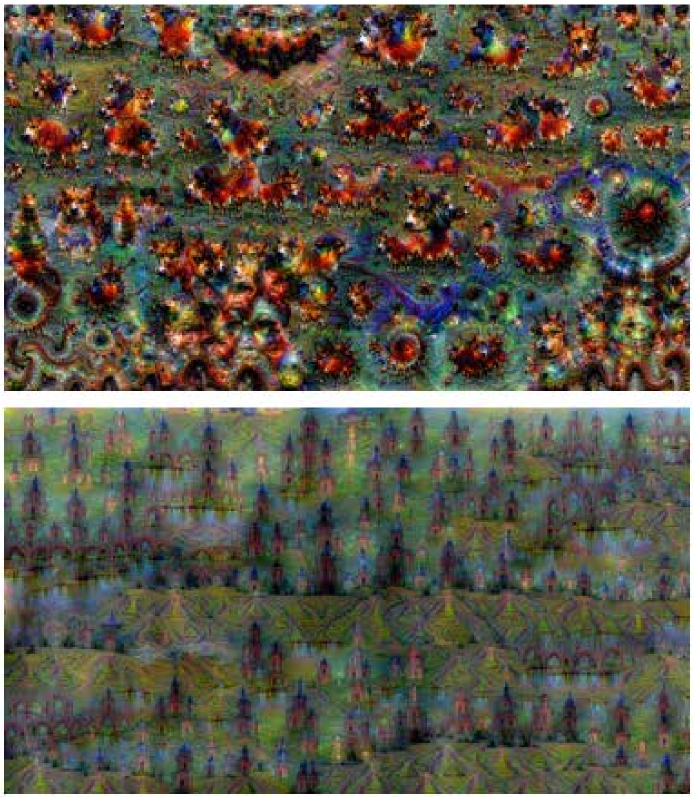
Segmentation



[Farabet et al., 2012]

Images from CS231n lecture slides

CNN Applications



Figures copyright Justin Johnson, 2015. Reproduced with permission. Generated using the Inceptionism approach from a blog post by Google Research.

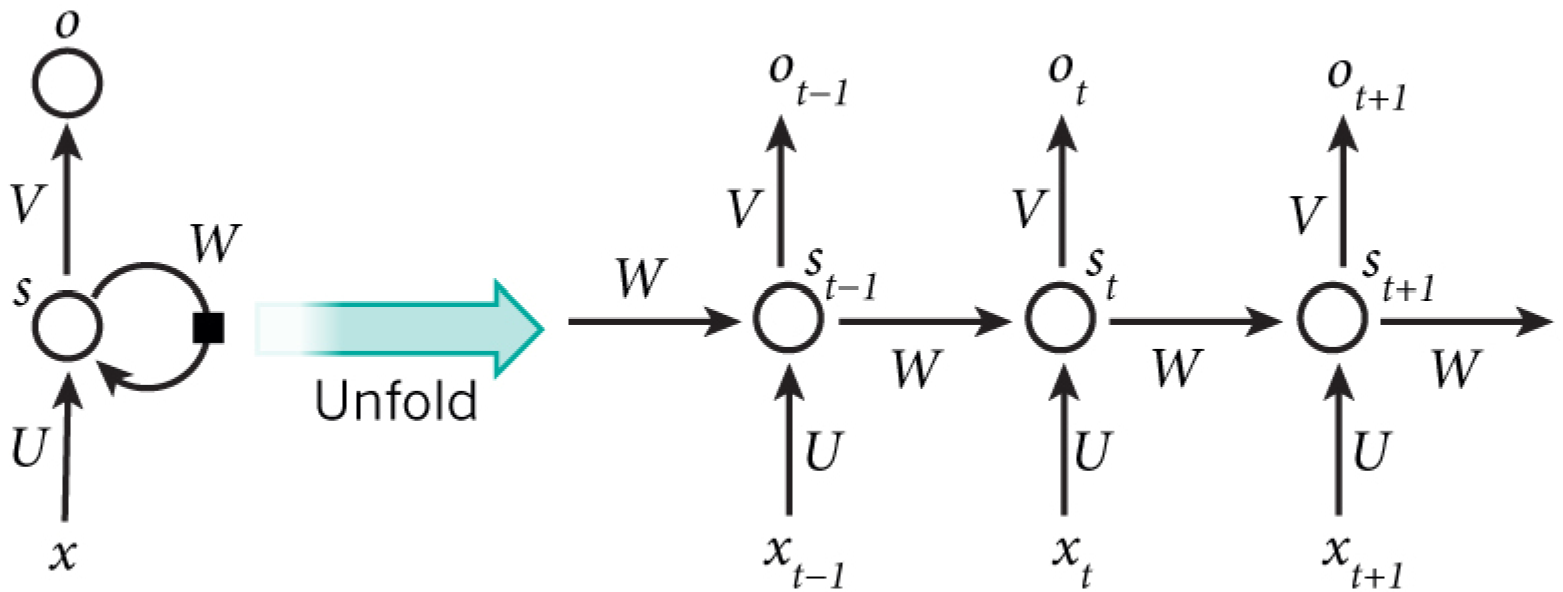
Original image is CC0 public domain
Starry Night and Tree Roots by Van Gogh are in the public domain
Bokeh image is in the public domain
Stylized images copyright Justin Johnson, 2017; reproduced with permission

Gatys et al, "Image Style Transfer using Convolutional Neural Networks", CVPR 2016
Gatys et al, "Controlling Perceptual Factors in Neural Style Transfer", CVPR 2017

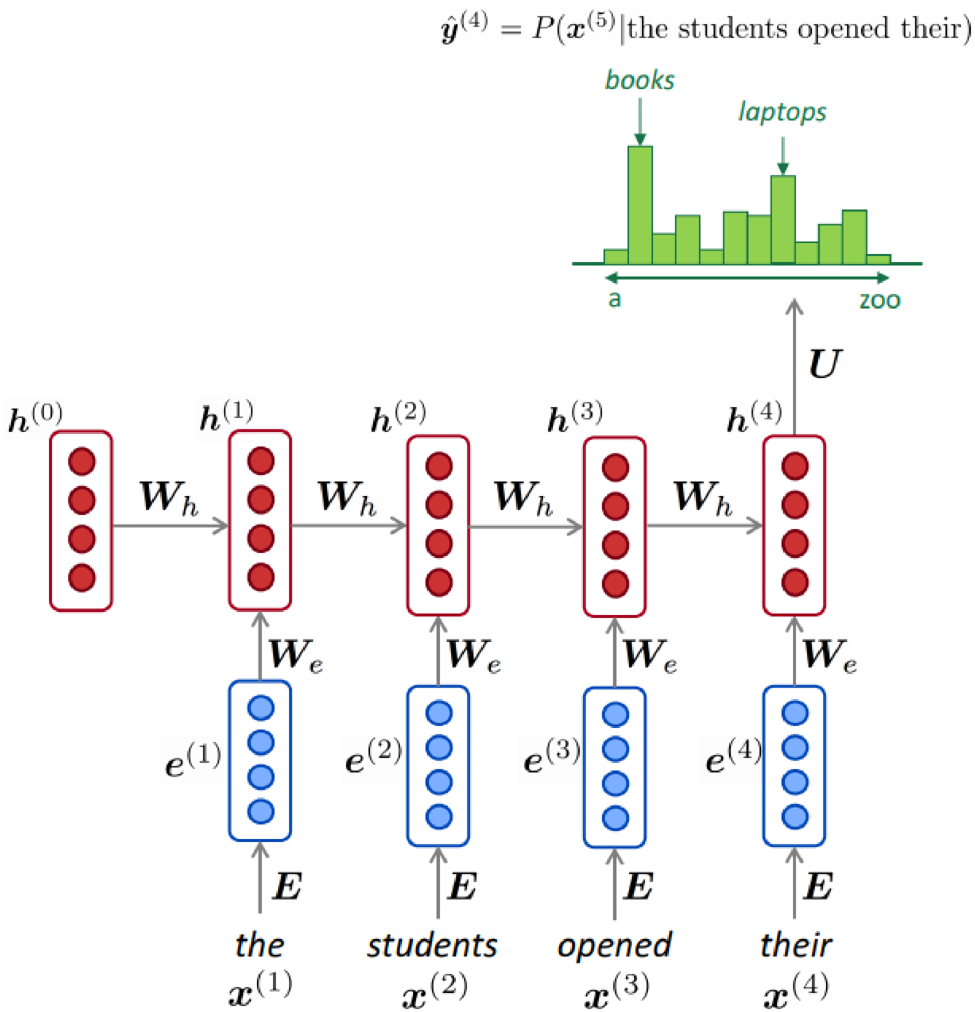
Style Transfer

Images from CS231n lecture slides

Recurrent Neural Network (RNN)



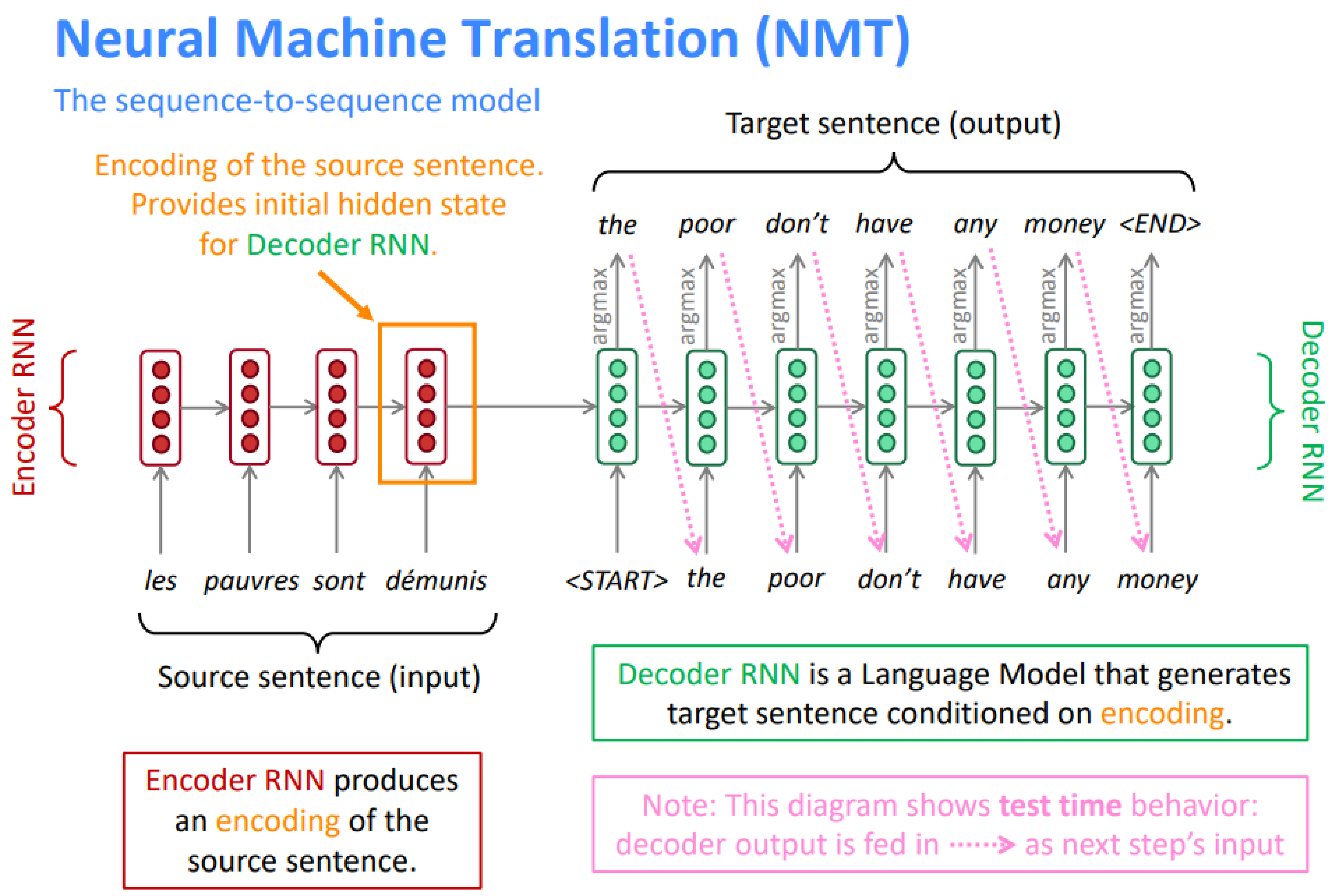
RNN Applications



Language Model

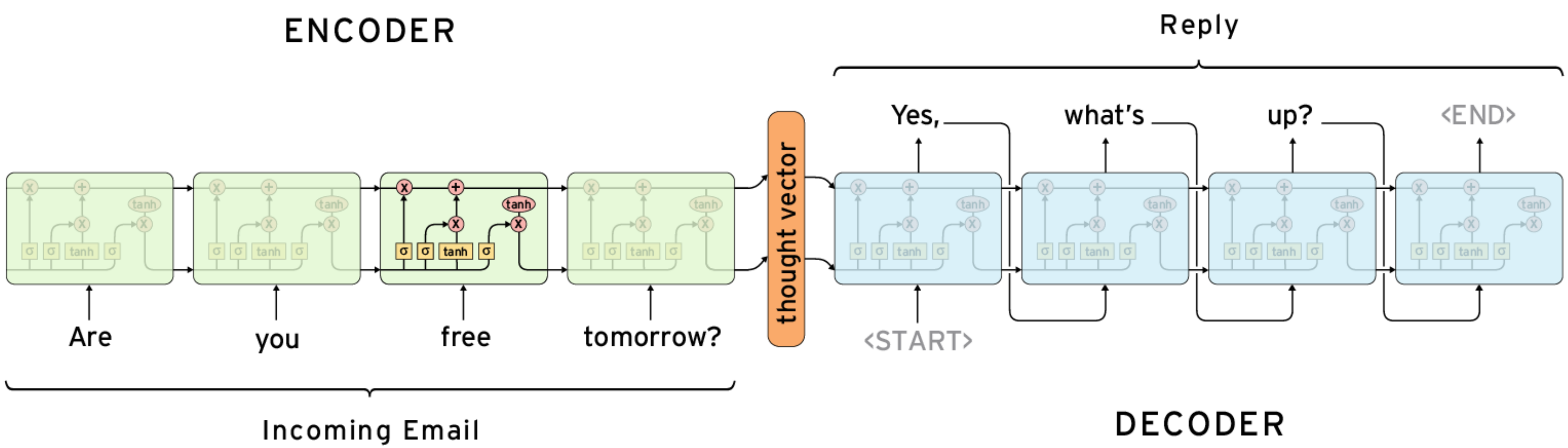
Images from CS224n lecture slides

RNN Applications



RNN Applications

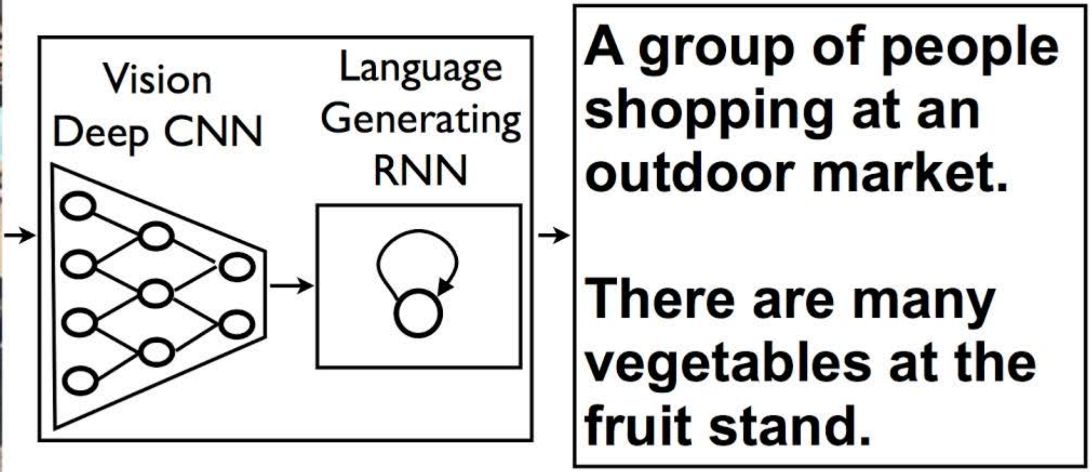
message	Where do you live now?
response	I live in Los Angeles.
message	In which city do you live now?
response	I live in Madrid.
message	In which country do you live now?
response	England, you?



Question Answering, Conversation (Chatbot)

Images from:
http://www.wildml.com/2016/04/deep-learning-for-chatbots-part-1-introduction/
https://medium.com/botsupply/generative-model-chatbots-e422ab08461e

RNN Applications



Image/Video Caption

Images from https://research.googleblog.com/2014/11/a-picture-is-worth-thousand-coherent.html

References

[1] UIUC CS 510 Course Material made by Ismini Lourentzou

– http://times.cs.uiuc.edu/course/510f17/ppt/deep-learning.pptx

[2] Stanford CS231n lecture slides (CNN/Visual Recognition)

– http://cs231n.stanford.edu/syllabus.html
– https://www.youtube.com/playlist?list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv (2017 Lecture Videos)

[3] Stanford CS224n lecture slides (RNN/Language)

– http://web.stanford.edu/class/cs224n/syllabus.html
– https://www.youtube.com/playlist?list=PLqdrfNEc5QnuV9RwUAhoJcoQvu4Q46Lja (2017 Lecture Videos)