ANLY 590

Artificial Neural Networks and Deep Learning

Outline

- Brief History of “AI”
- Overview of Recent Excitement in Deep Learning
- Administrative Topics
  - Assignments, Exams, Projects, Grading
  - Course style
  - Prerequisite knowledge
- Major course topics

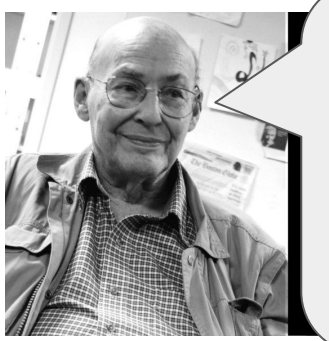
Brief History

SHORT-TERM HYPE

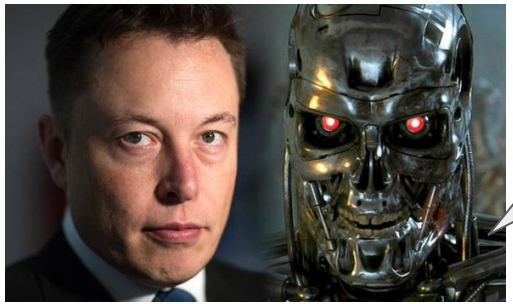
AI has gone through several hype periods fueled by early successes

AI failed to materialize and led to boom and bust cycles pushing many researchers to other fields

Two AI winters **1974-1980** and **1987-1993**, where funding dried up and governments shifted focus

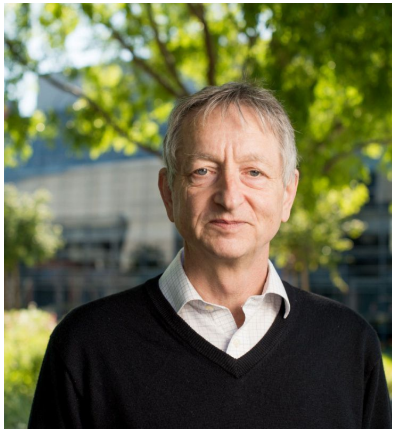


Within a generation... the problem of creating artificial intelligence will be substantially solved
- *Marvin Minsky 1967*



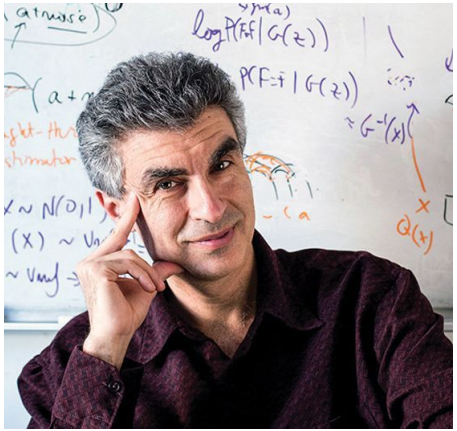
Disruption certainly. Deep AI is the real risk, though, not automation.
Elon Musk (@elonmusk)
June 9, 2017

Three pioneers kept AI alive

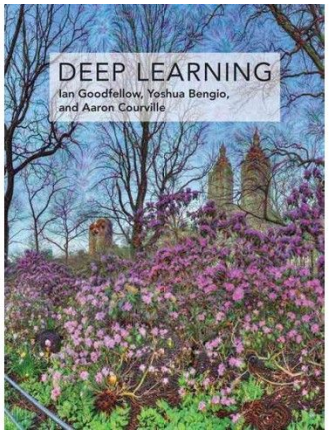


Geoffrey Hinton

Yann LeCunn



Yoshua Bengio



“Grandfathers of AI”

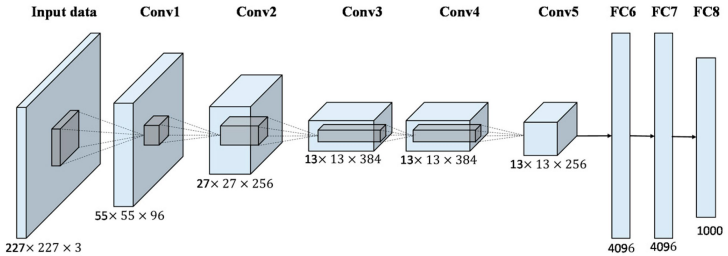
Why deep learning now?

Inflection point was in 2012 when convolutional net dominated the ImageNet Large Scale Visual Recognition Competition (ILSVRC)

Until 2010 the top methods predicted image classes with 75% accuracy

Incremental advances were fractions of a %

In 2012, researchers Alex Krizhevsky, Geoffrey Hinton, and Ilya Sutskever entered a ConvNet (now called Alexnet) into the challenge



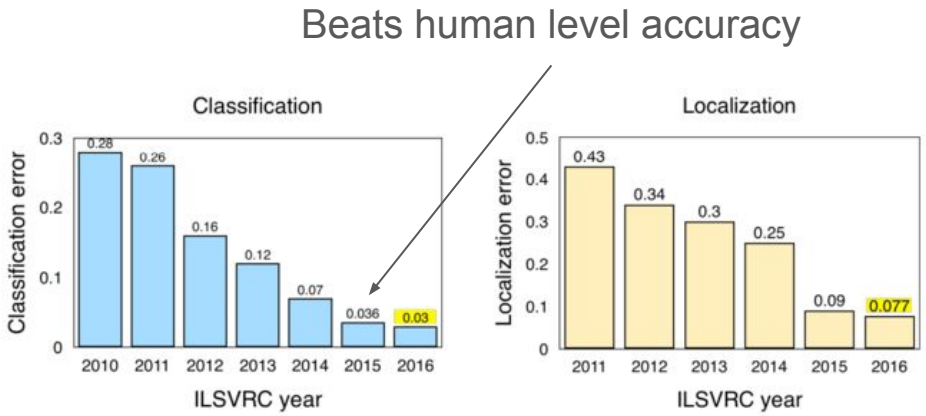
AlexNet shifted the stigma around Neural Networks

Before 2012 the best classification models were 75% accurate

AlexNet could predict classes with an incredible accuracy of ~85%

Human level accuracy is ~95%

Since AlexNet, all subsequent submissions used DL

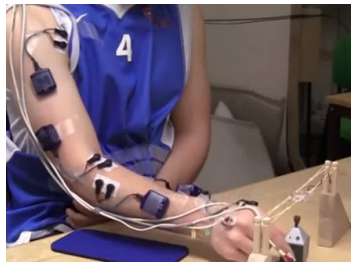


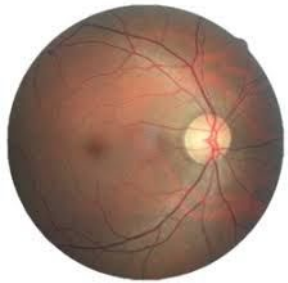
In 2016 the contest was considered solved as most of the entries were reaching near-human level accuracy

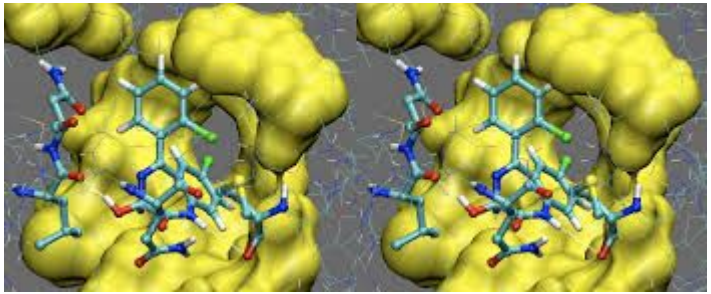
Kaggle has played an important in popularizing ML

In the early years Kaggle played an important role **connecting tangible business outcomes** with broader community adoption

Early adopters won $100Ks competitions







Examples include:

1. Grasp-and-Lift EEG Detection
2. Diabetic Retinopathy Detection
3. National Data Science Bowl
4. Right Whale Recognition
5. The Merck Molecular Activity Challenge
6. The Job Salary prediction challenge

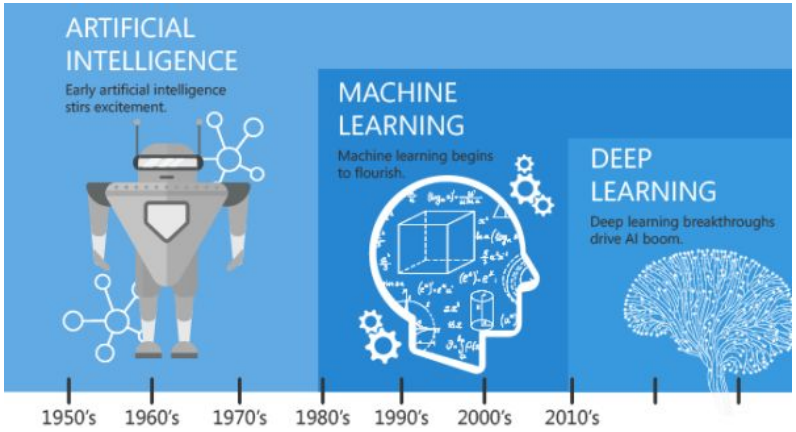
HOW DEEP LEARNING RELATES TO AI AND ML

AI was born in the 1950's

Deep learning is a subset of artificial intelligence (AI) and machine learning (ML)

AI encompasses ML and other approaches that don't use learning

Early chess programs use hard-coded rules, are a subset of the field of AI
This approach was known as symbolic AI



Reference to recent advances in AI they are to the advances deep learning and not canonic machine learning

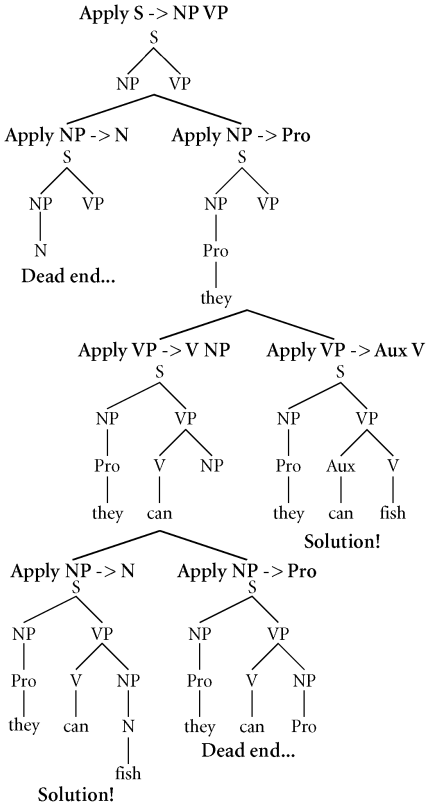
Symbolic AI was what people thought of as AI

Researchers postulated that for a sufficiently large rule-set human-level artificial could be achieved

Dominant paradigm in AI from the 1950s to the 1980s - lots of hand crafted features

Reached its peak with expert systems in the 1980s

We still use variants of it, we'll see in our NLP section with recurrent networks that feature engineering can be abstracted



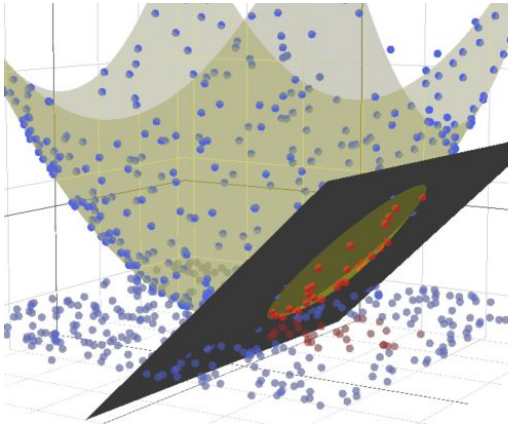
MACHINE LEARNING BEFORE DEEP LEARNING

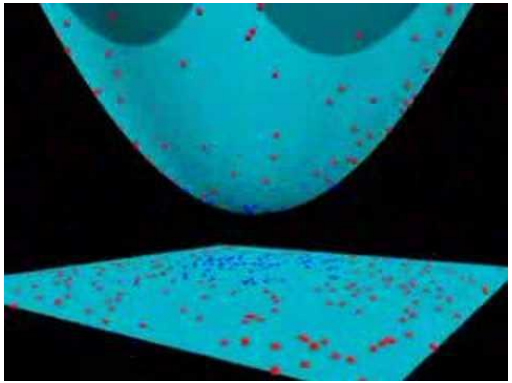
Faster hardware and larger data sets have been drivers for ML flourishing since the 1990s

Towards the mid 2000s most improvements on problems and models were incremental and faced **scaling**, **training** and **parallelization**

Two popular approaches during this time were **kernel** and **Bayesian** approaches

SVMs have O(n^2) training complexity and do not scale well



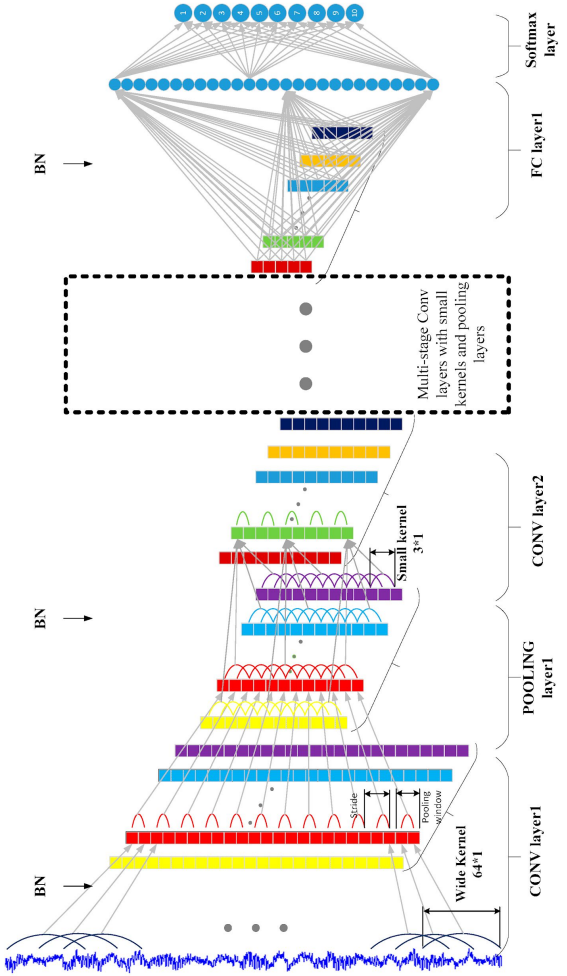


Shallow learning vs deep learning

Deep learning models stack simple models on top of one another generating hierarchical structures

They therefore have effect greater **representational capacity**

Whereas shallow models tend to make fairly complicated assumptions about the nature of the data, deep models make assumptions about network connectivity



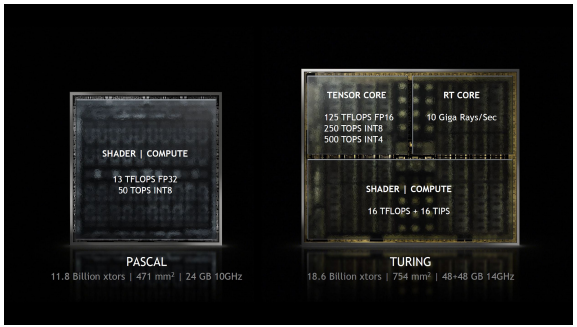
3 FACTORS CONTRIBUTING TO DL'S SUCCESSES

Hardware

Increase in speed

Increase in capacity

Gaming market subsidized supercomputing for the next gen of artificial intelligence



Data

Speed and affordability
- NVMe drives are 25x quicker than conventional HDDs
- Cost per GB decreases yearly

Availability
- Easy access to data via APIs
- Sites collect and share data

Amount
- The amount of data required to train a deep learning models is enormous
- In the early 2000s a standard hard drive had between 10-40Gb Most production-level corpora well exceed the 50Gb mark.

Algorithms

Specialized algorithms
- Developed to leverage general purpose hardware for DL

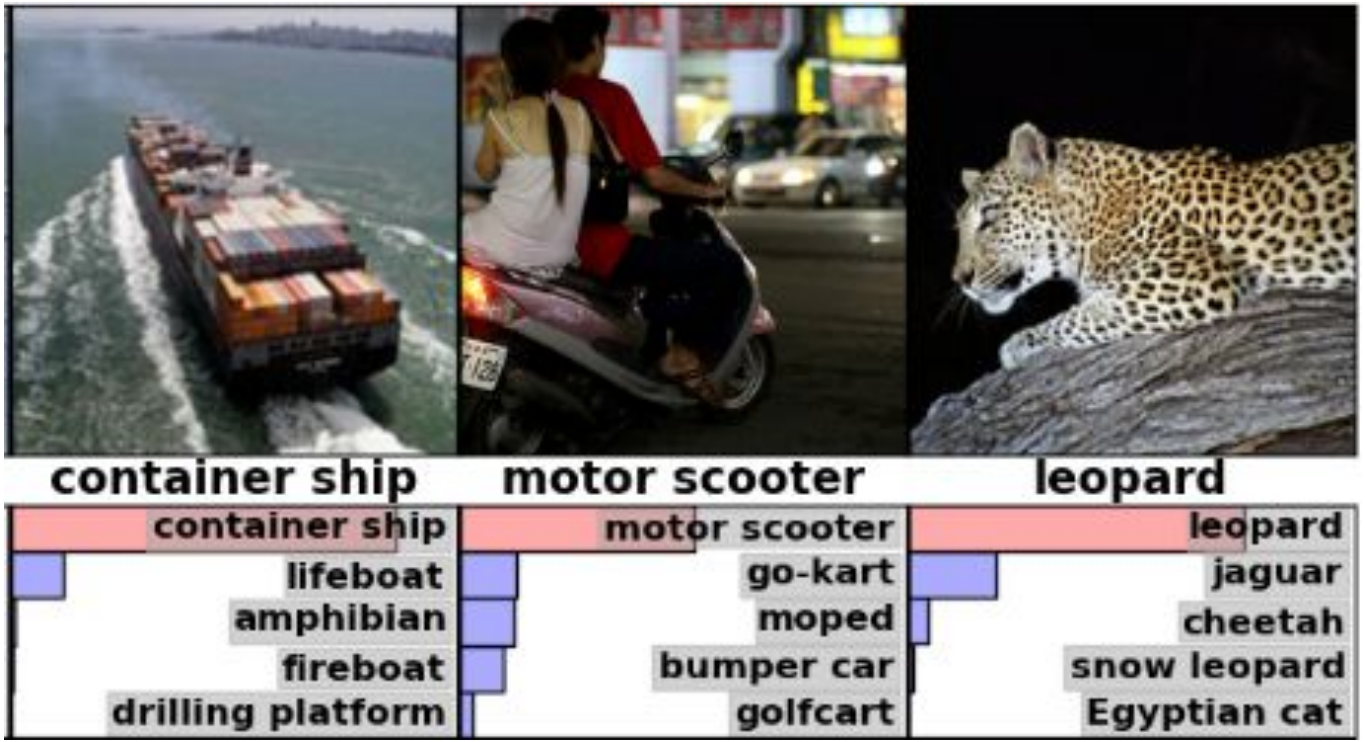
Algorithmic & research advances
- Logistic vs ReLu networks
- Embeddings
- Batch norm
- Residual networks
- Attention

DL IN INDUSTRY

- A less discussed aspect of DL are the **lack of use-cases in industry**
- A good portion of research simply hasn't trickled into business
- While several reasons may exist, three resonate
  a. Large established firms that have the resources simply **haven't made the investment** to get up to speed ts who are growing quickly and well funded
  b. The **field is evolving so quickly** that staying on top of the research isn't possible so the most salient use-cases fall by the way side in favor of more traditional implementations
  c. Connecting **business and DL models is hard**

Deep Learning Successes

Image Analysis



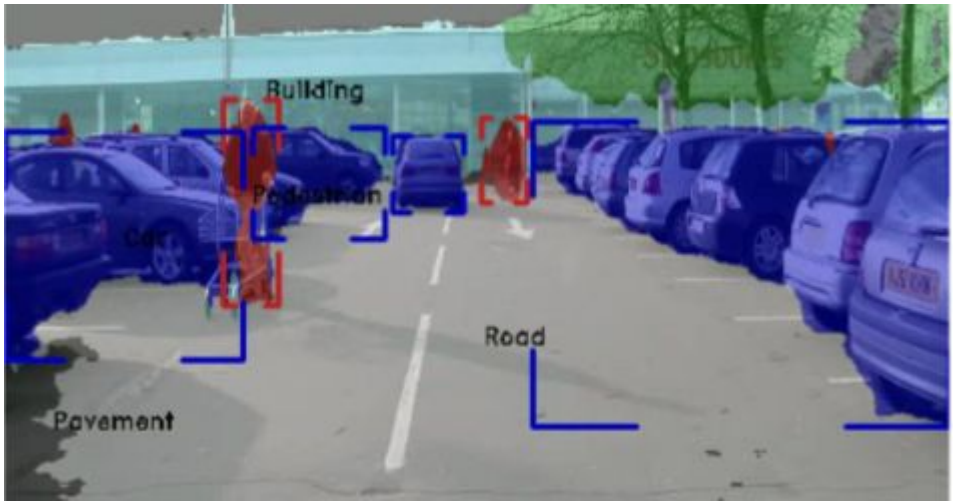
https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

Image Analysis

https://arxiv.org/pdf/1703.06870.pdf

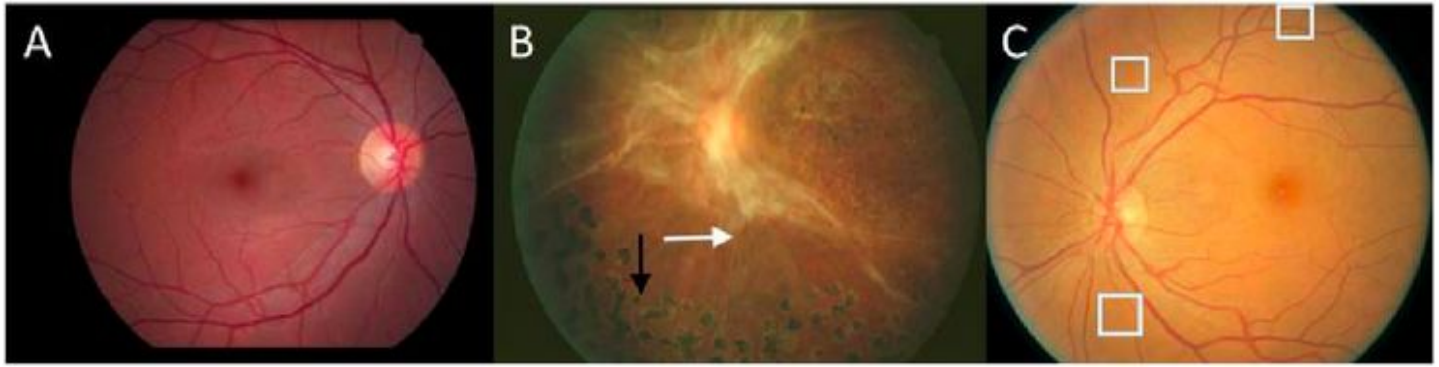
Image Analysis

https://arxiv.org/pdf/1611.08303.pdf



https://arxiv.org/pdf/1707.02432.pdf

Image Analysis

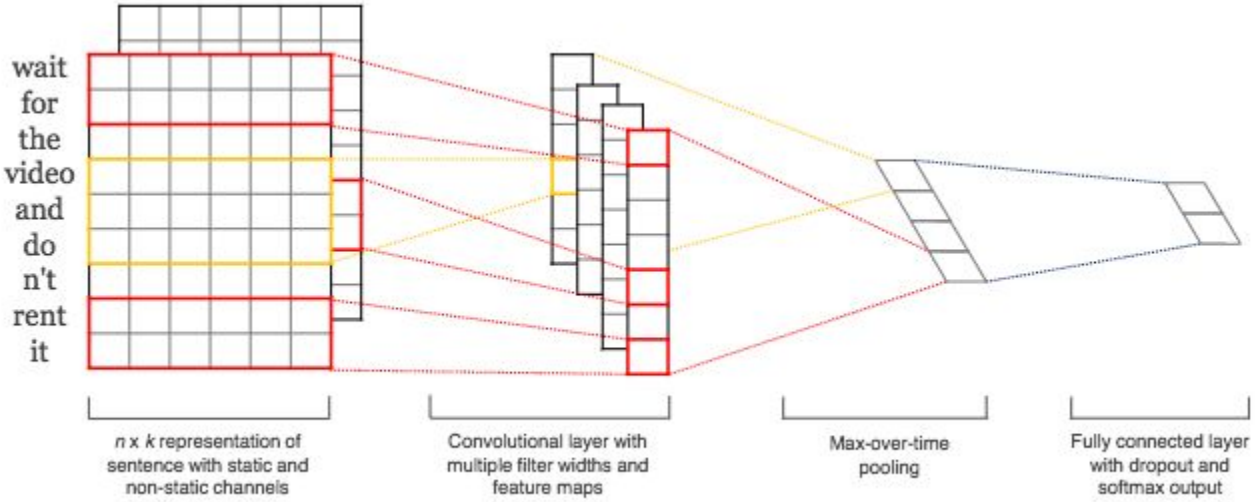


https://www.healthcarefinancenews.com/news/fda-approves-first-ai-tool-detecting-retinopathy-nih-shows-machine-learning-success-imaging

https://jamanetwork.com/journals/jama/fullarticle/2588763

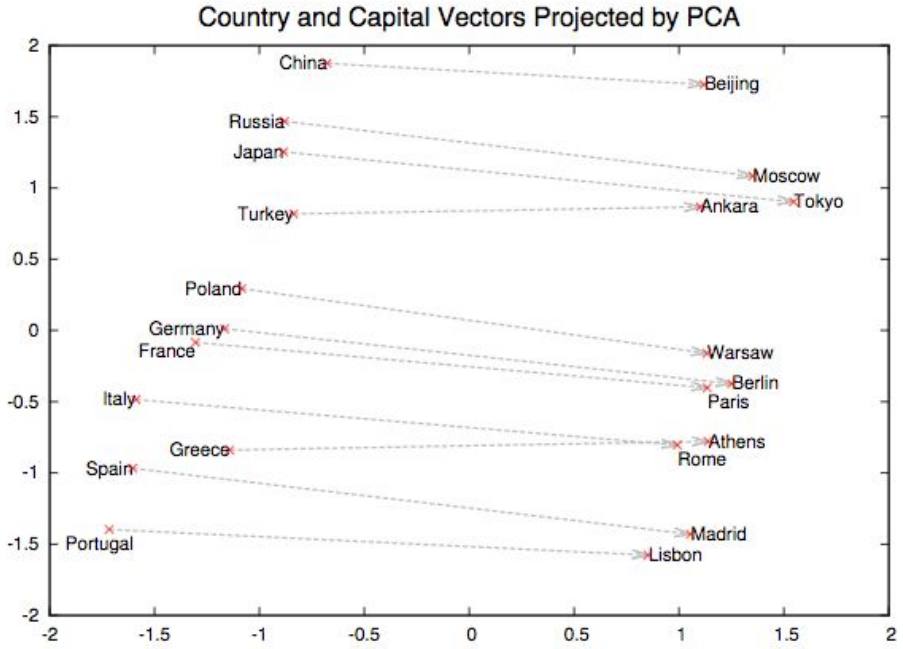
https://www.semanticscholar.org/paper/Automated-Detection-of-Diabetic-Retinopathy-using-Lam-Yi/e9fbe155954c01560d8e2504741692a40e3ffa31

NLP - Sentence Classification



http://www.aclweb.org/anthology/D14-1181

NLP - Word Embeddings



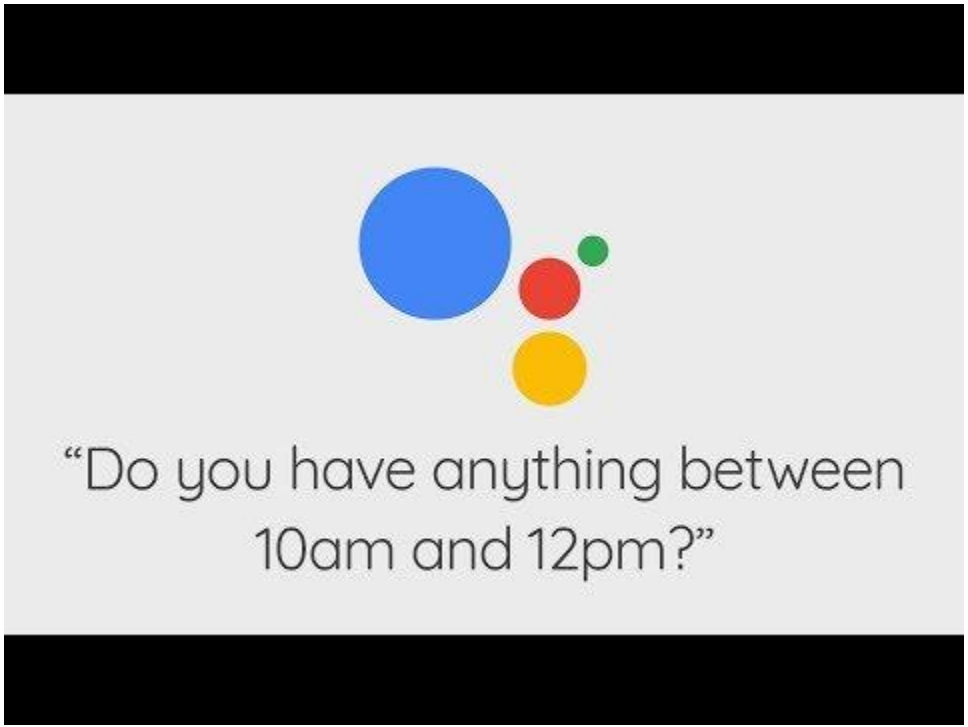
German + airlines	Russian + river	French + actress
airline Lufthansa	Moscow	Juliette Binoche
carrier Lufthansa	Volga River	Vanessa Paradis
flag carrier Lufthansa	upriver	Charlotte Gainsbourg
Lufthansa	Russia	Cecile De

https://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf

NLP - Machine Translation

NLP - Machine Translation

NLP - Natural Language Understanding



Reinforcement Learning - Atari Games

Reinforcement Learning - AlphaGo

Reinforcement Learning - StarCraft and Dota

https://deepmind.com/blog/deepmind-and-blizzard-open-starcraft-ii-ai-research-environment/

https://github.com/Blizzard/s2client-proto

https://blog.openai.com/openai-five/

Course Logistics

Course Work

- Assignments
  - There will be 3 graded assignments
  - These will be a combination of theory and coding projects
  - Assignments will be turned in electronically
- Quiz
  - There be at least one in-class quiz before the Midterm
- Midterm
  - There will be an in-class Midterm Exam. You will be expected to be able to derive and interpret relevant equations, reason about core concepts, interpret model outputs etc
- Final Project
  - See next slide
- There will be no final exam

Final Project

- The culmination of the course will be a group project
  - You will work in groups of 2-4 people
  - You will devise a project of your own choosing to apply course concepts to any task or dataset that is of interest to you
  - Your major deliverable will be a scientific poster/paper for our course **Poster Session** at the end of the semester
  - Your group will also complete a write-up of your work and results

Github

- Each of you will use a Github repo to track and submit your coursework
- Learning git and github can be a little tricky, but has important benefits

  1. The tech world uses github, it's time to learn

  2. Version controlling (and collaborating) your coding work is the **smart thing to do**

  3. At the end of our course, you will have a **public portfolio** of machine learning projects you have completed. This will aid in job search and interviews. More generally, this helps to inform the world about what you're interested in and knowledgeable about.

- Git tutorials
  - https://www.dataschool.io/git-and-github-videos-for-beginners/
  - https://s3.amazonaws.com/anly-590/git-for-anly590.mov

Course Topics

Feedforward Networks

Convolutional Networks

Generative Adversarial Networks

Backprop and Inference

Recurrent Networks

Reinforcement Learning

Deep Learning Toolkits

Unsupervised Deep Learning

Guest Topics