Deep Learning

Jian Tang

tangjianpku@gmail.com

HEC MONTRÉAL

Institut des algorithmes d'apprentissage de Montréal MILA

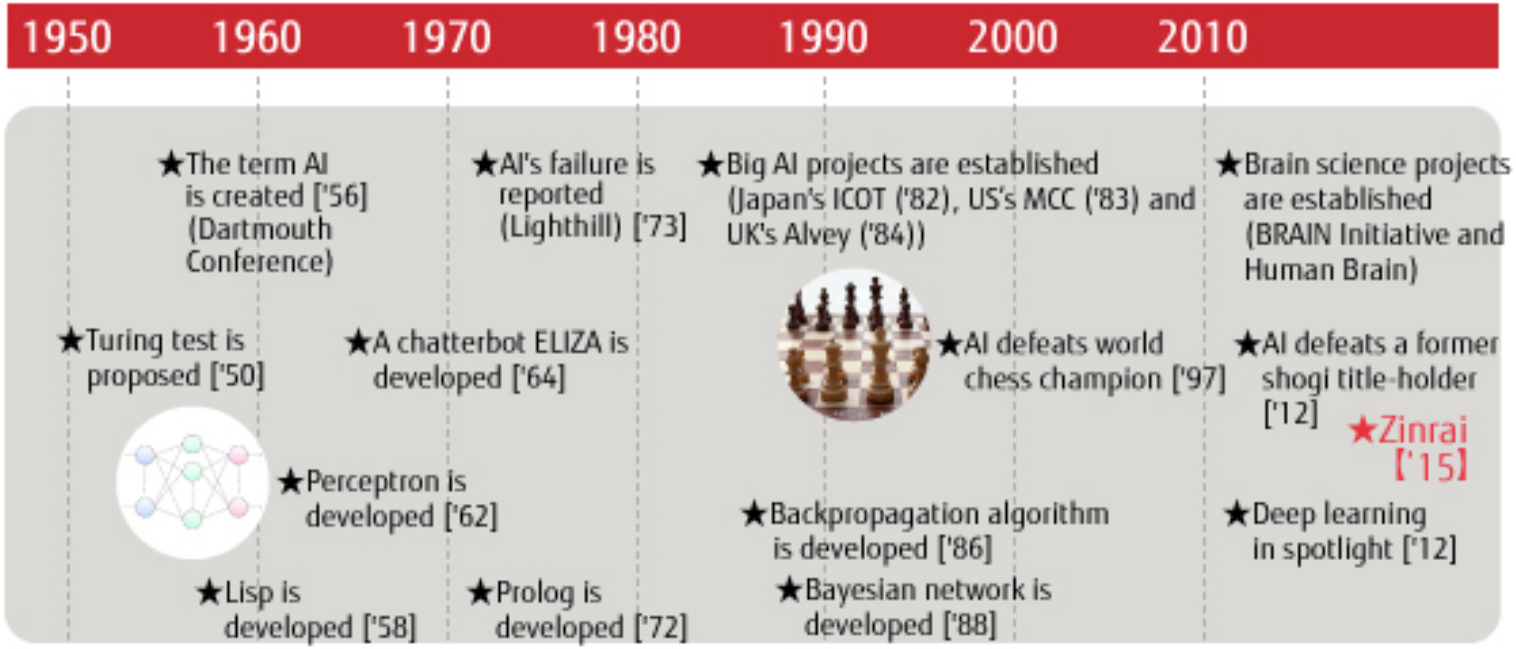
Artificial Intelligence:
Fourth Industrial Revolution

What is Artificial Intelligence?

*“**Artificial intelligence** (AI, also machine intelligence, MI) is intelligence demonstrated by machines, in contrast to the natural intelligence (NI) displayed by humans and other animals. In computer science AI research is defined as the study of "**intelligent agents**": any device that **perceives its environment and takes actions that maximize its chance of successfully achieving its goals**. Colloquially, the term "artificial intelligence" is applied when a machine mimics "cognitive" functions that humans associate with other human minds, such as "learning" and "problem solving".”*

-Wikipedia: https://en.wikipedia.org/wiki/Artificial_intelligence

The History of Artificial Intelligence



The age of search and deductive reasonin

The first AI boom

AI programs are able to solve only trivial "toy" problems but hardly any practical ones. Into "AI winter"

The age of knowledge acquisition

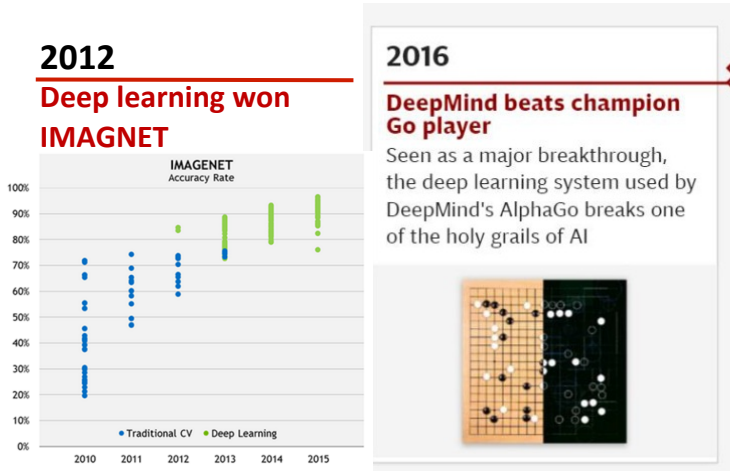
The second AI boom

Researchers tried to teach machines the knowledge of experts, which turned out to be extremely difficult. Into "AI winter" again

The age of machine learning

The third AI boom

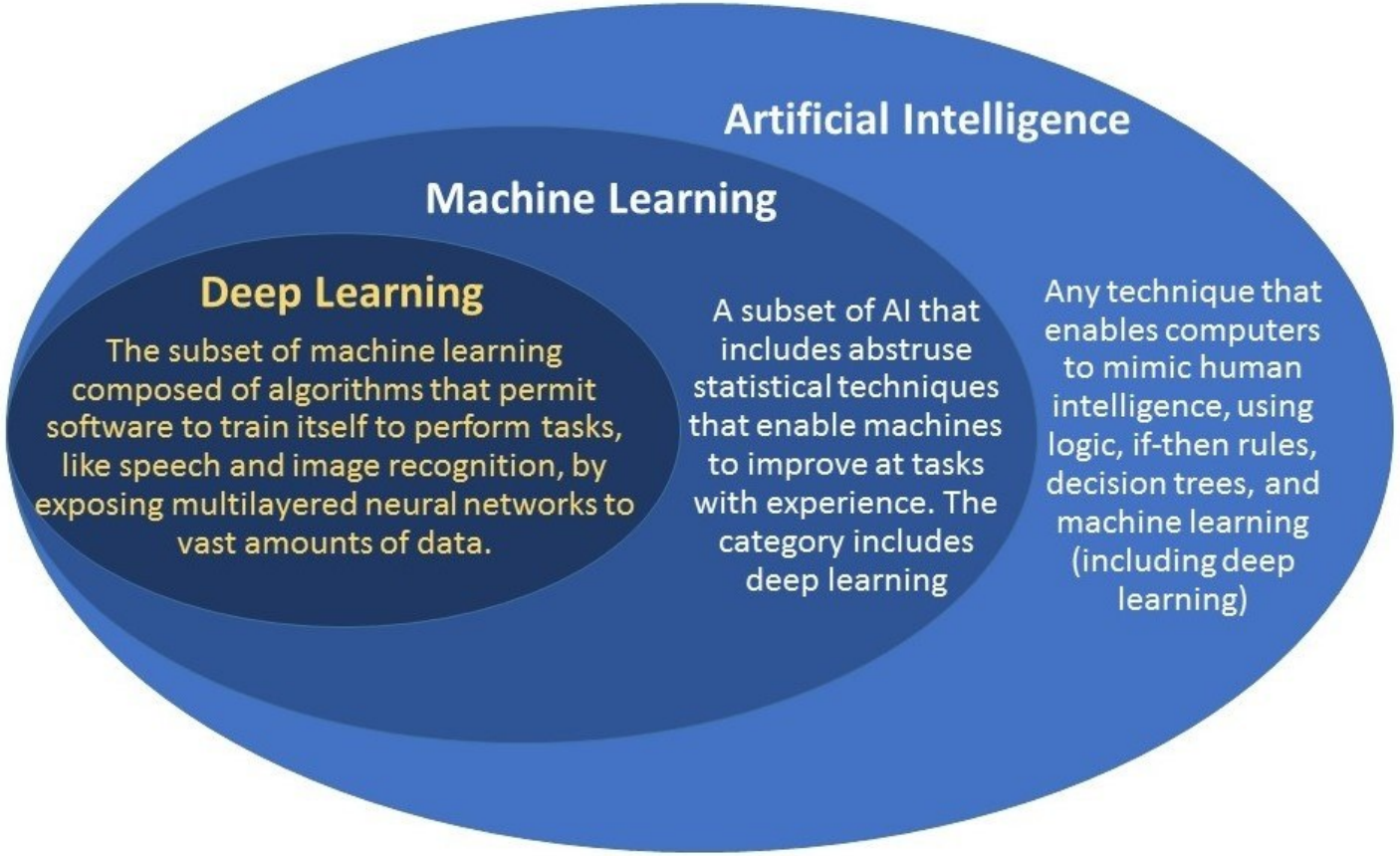
Into the age where computers acquire knowledge from data, not from humans by machine learning



Deep neural networks are widely adopted, achieving or outperforming human performance in a variety of applications

Images from internet

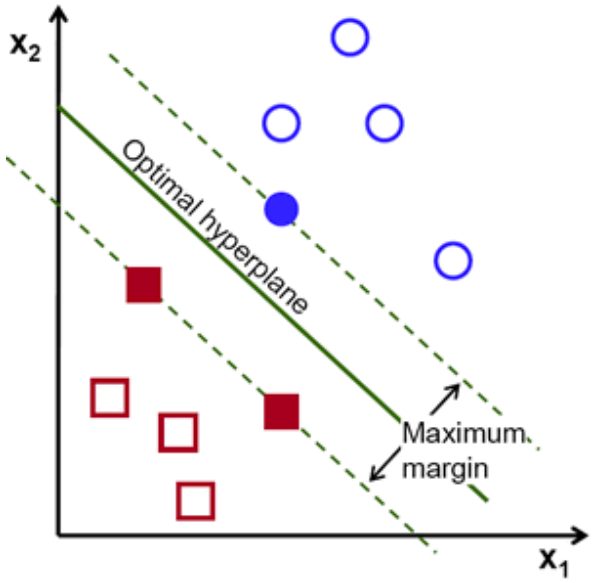
Artificial Intelligence v.s. Machine Learning v.s. Deep Learning



Machine Learning

- "**Machine learning** is a field of computer science that uses statistical techniques to give computer systems the ability to "learn" (i.e., progressively improve performance on a specific task) with data, without being explicitly programmed."

-Wikipedia



Support vector machines

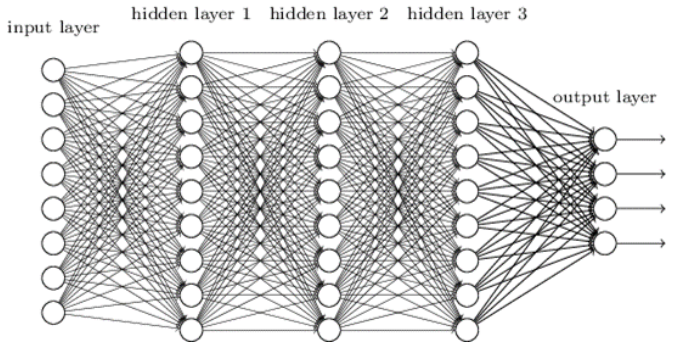
Hand-crafted
Feature Extractor

Simple Trainable Classifier
e.g., SVM, LR

Domain experts

Deep Learning

- Algorithms that allow to learn from features from data (a.k.a, End-to-end learning)



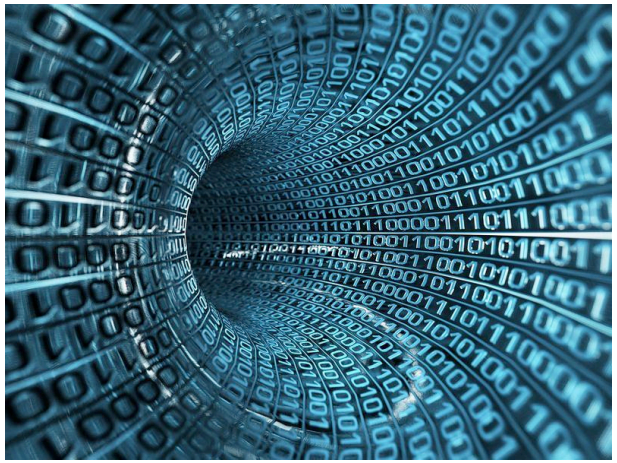
Deep Neural Networks

Trainable
Feature Extractor

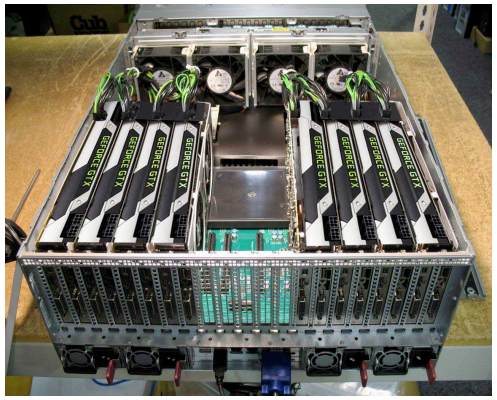
Simple Trainable Classifier
e.g., SVM, LR

Domain experts

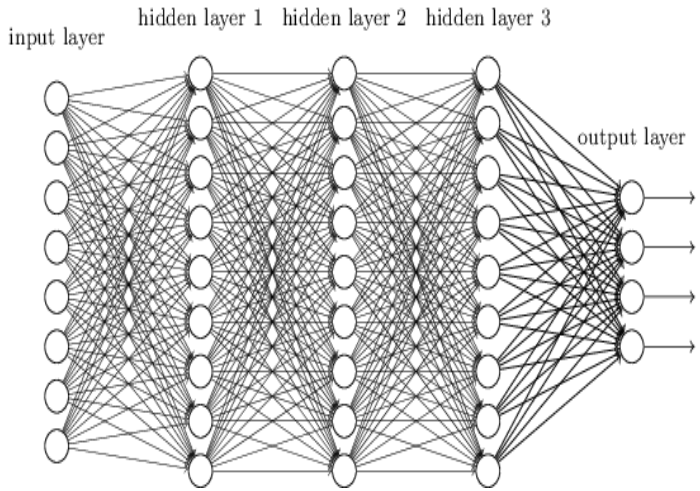
Why Deep Learning Now?



Big Data



Big Computation

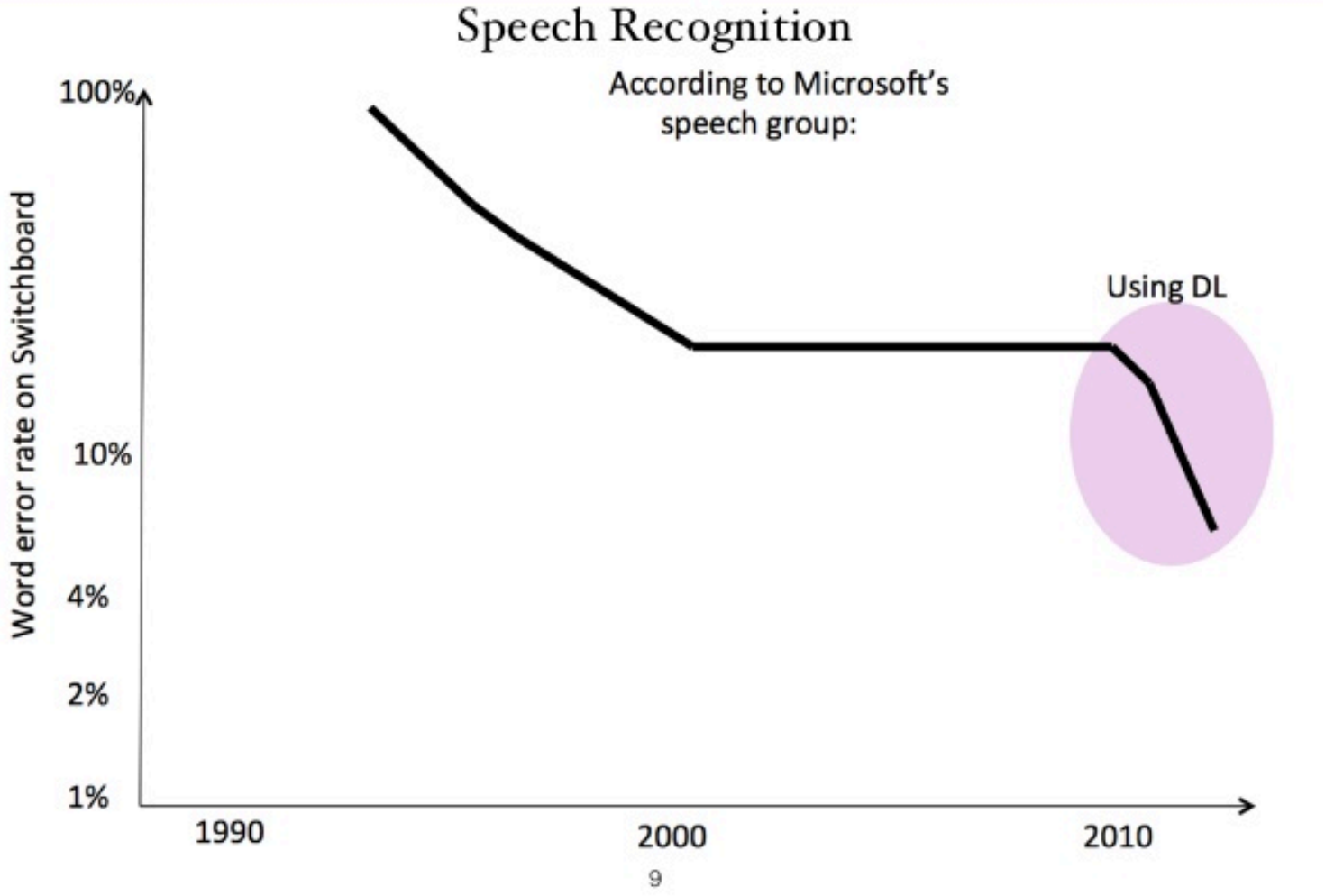


Big Model

Speech Recognition

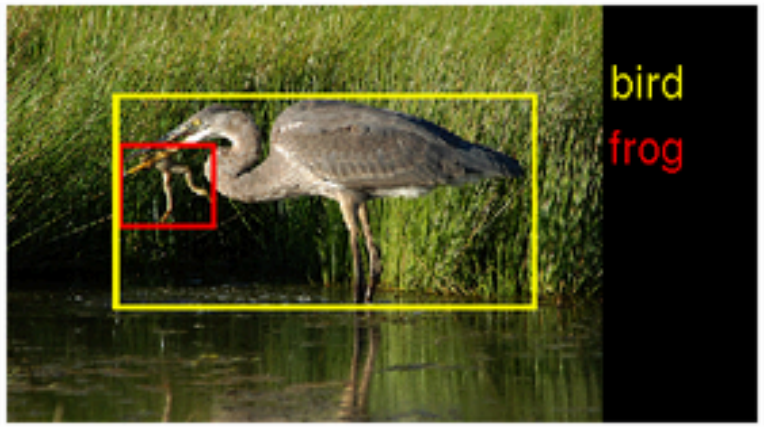
Skype to get 'real-time' translator

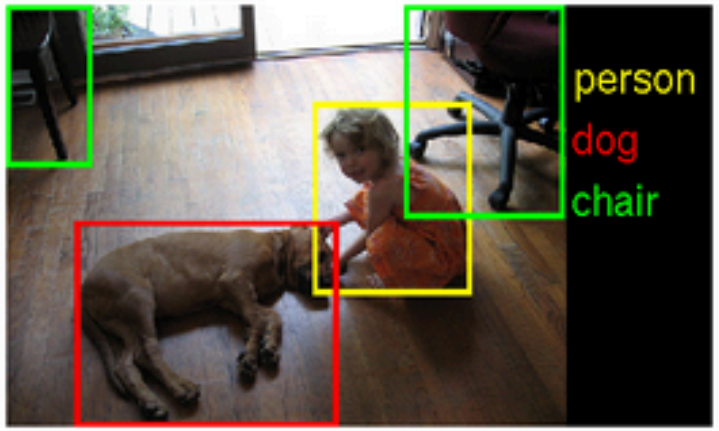
Speech Recognition Results

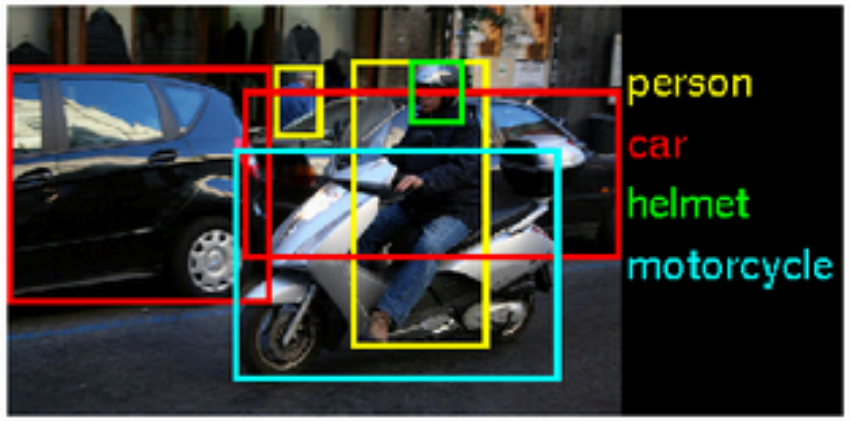


(Figure from Microsoft's speech Group)

Image Recognition







Results on ImageNet

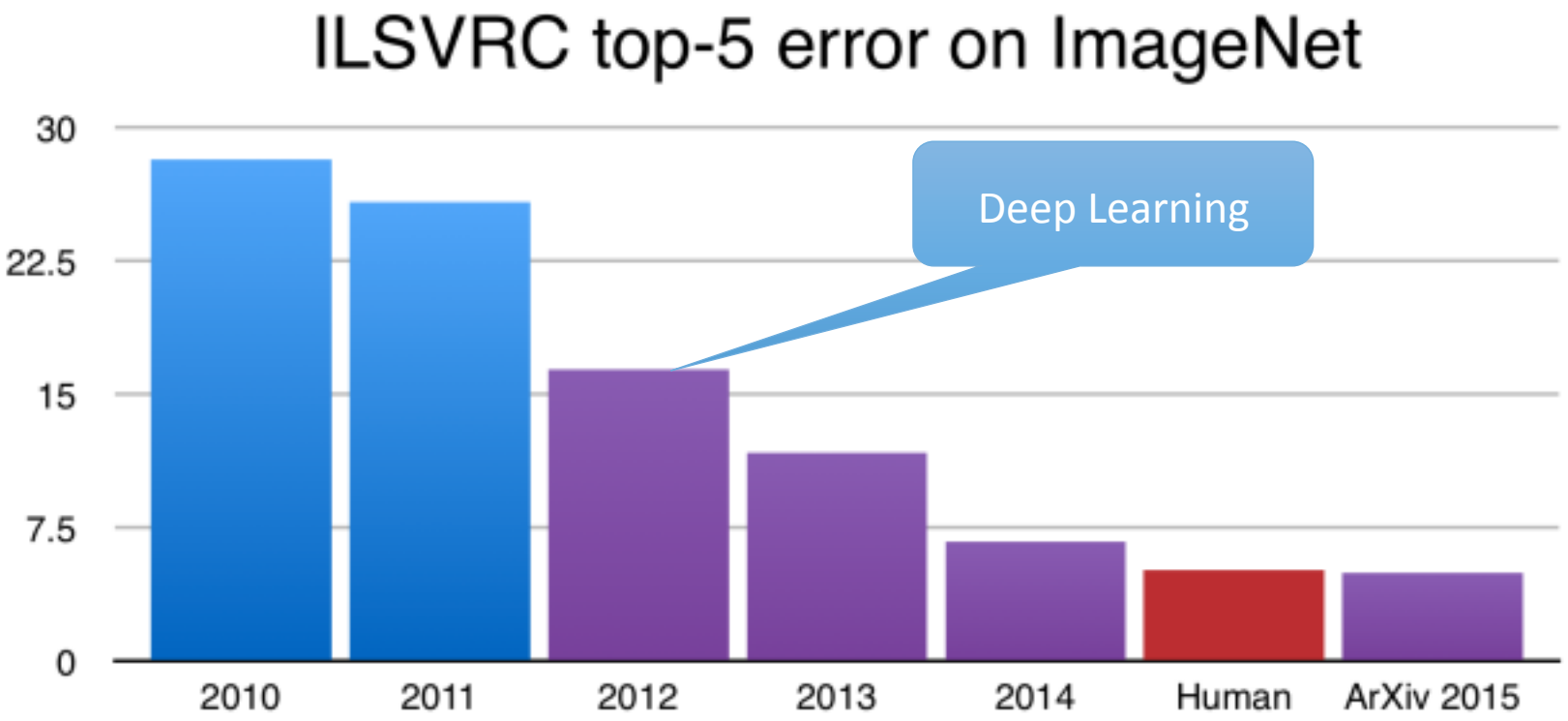


Image Generation

Volcano

Volcano

(Figure from Nguyen et al. 2016)

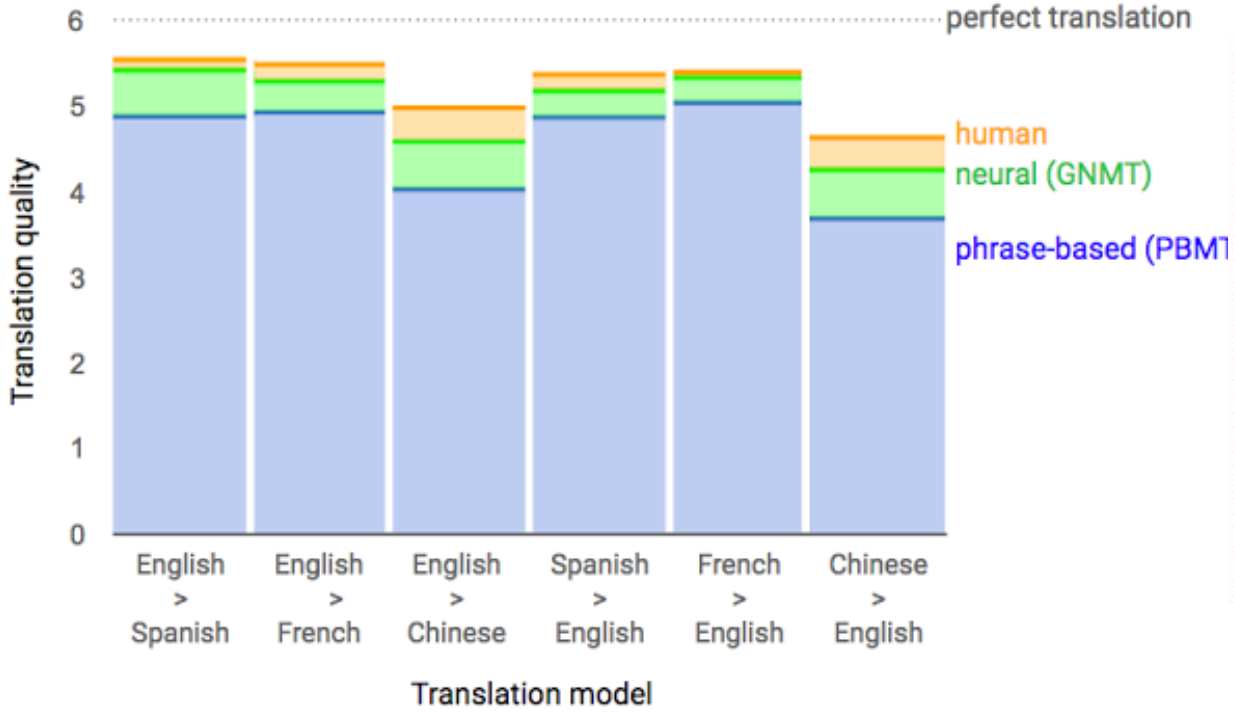
AlphaGo

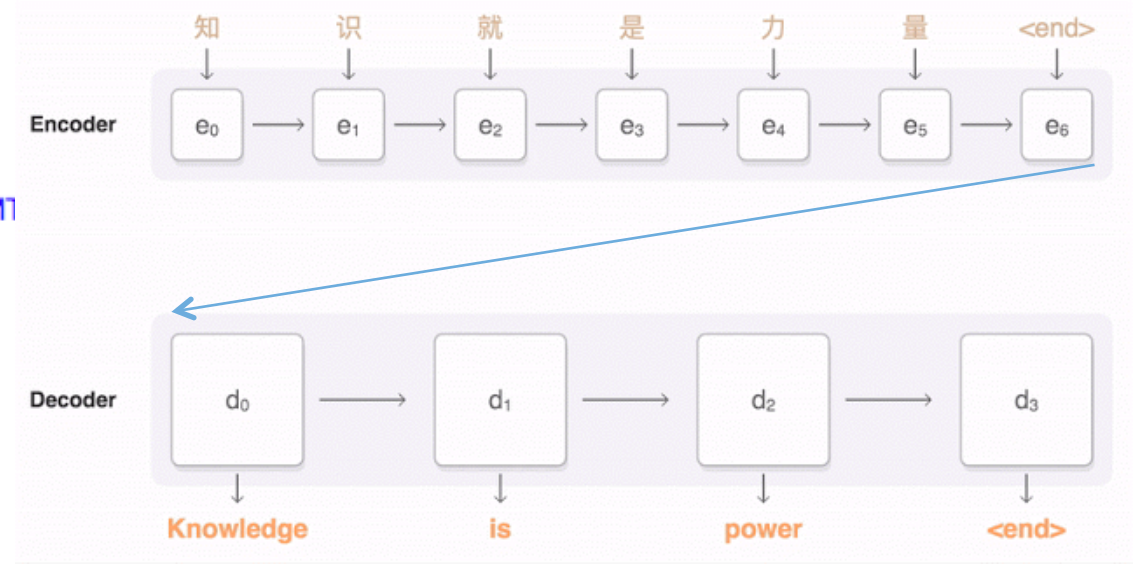
4:1
vs Sedol Lee
(2016.3)

3:0
vs Ke Jie
(2017.5)

Machine Translation

- 2016.9, Google announce its *neural machine translation* system.
- 2018.3 , Microsoft claimed its NMT achieved “human parity” on automatic Chinese to English news translation.





(Seq2Seq, Sutskever et al. 2014)

Machine Reading Comprehension

Passage: Tesla later approached Morgan to ask for more funds to build a more powerful transmitter. **When asked where all the money had gone, Tesla responded by saying that he was affected by the Panic of 1901**, which he (Morgan) had caused. Morgan was shocked by the reminder of his part in the stock market crash and by Tesla's breach of contract by asking for more funds. Tesla wrote another plea to Morgan, but it was also fruitless. Morgan still owed Tesla money on the original agreement, and Tesla had been facing foreclosure even before construction of the tower began.

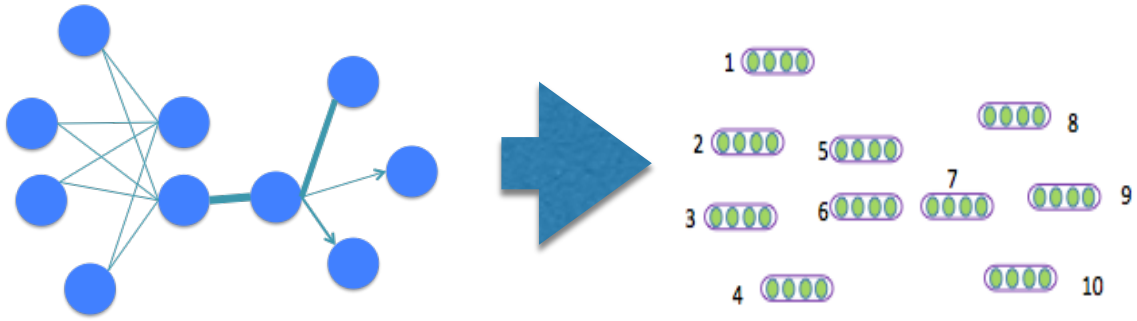
Question: On what did Tesla blame for the loss of the initial money?

Answer: Panic of 1901

Rank	Model	EM	F1
	Human Performance *Stanford University* (Rajpurkar et al. '16)	82.304	91.221
1 Mar 19, 2018	QANet (ensemble) *Google Brain & CMU*	**83.877**	**89.737**
2 Jan 22, 2018	Hybrid AoA Reader (ensemble) *Joint Laboratory of HIT and iFLYTEK Research*	82.482	89.281

Analyzing Graphs

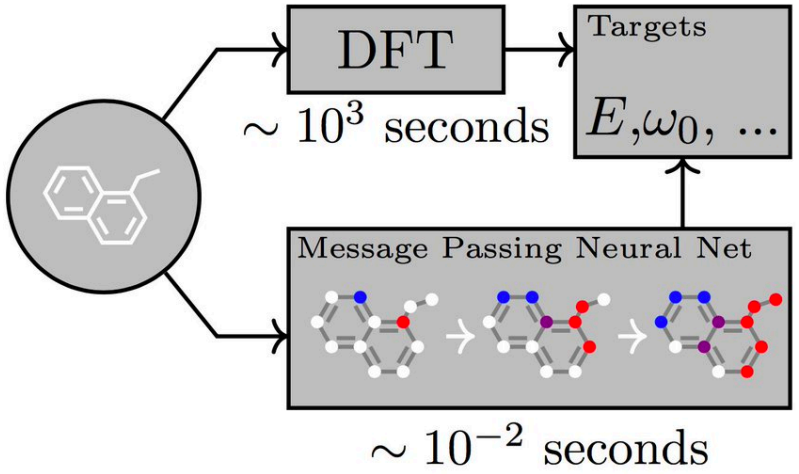
- Representing graphs in low-dimensional spaces
  - Node representation, graph representation



Graph

Node representations

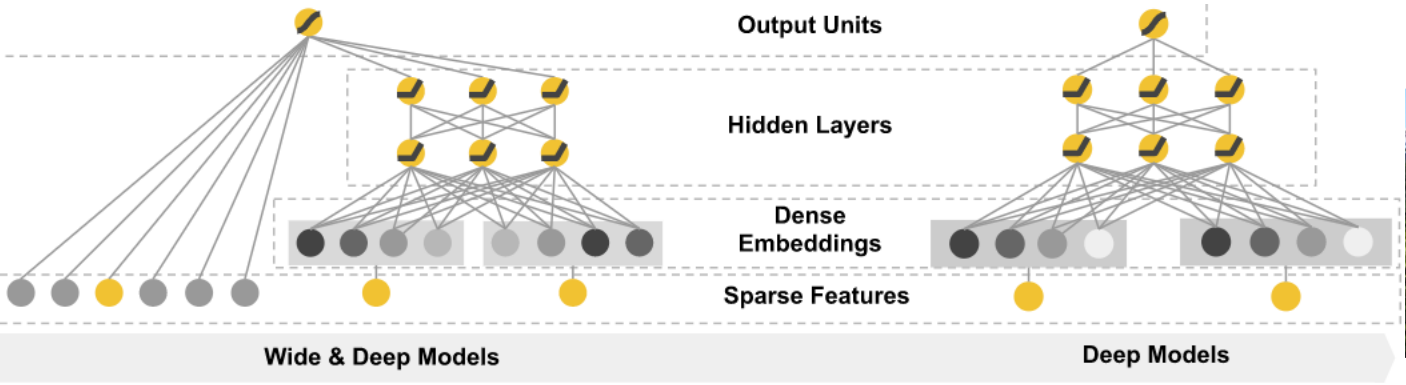
(LINE, Tang et al. 2015)



Molecule properties prediction

(Gilmer et al. 2016)

Recommender Systems



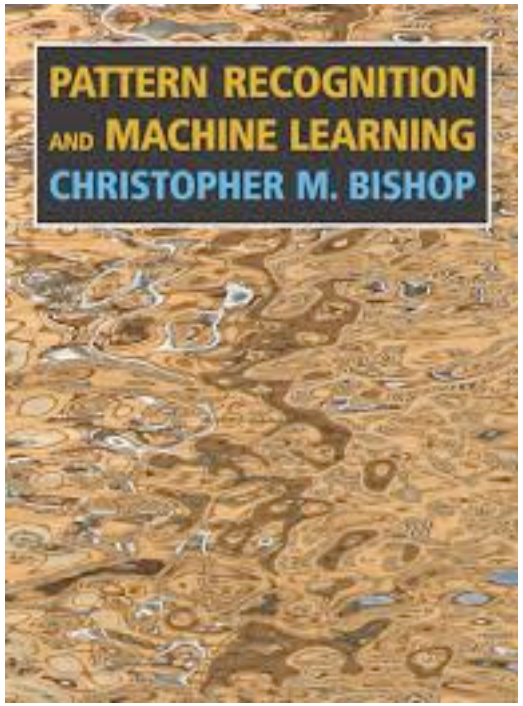
Wide & deep learning for recommender systems (Google 2016)

Workshops on Deep Learning for Recommender Systems

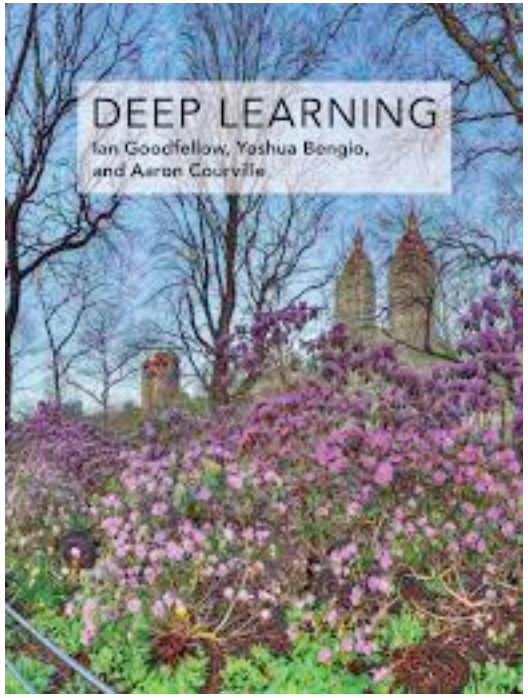
This Course

- **Objectives**
  - Understand the basic techniques of machine learning and deep learning
  - Learn advanced topics/latest progress of deep learning (selected topics)
  - Know how to apply deep learning techniques to real-world applications
- **Prerequisite**
  - Some basics of probability, statistics, and linear algebra
  - No programming is required

Textbooks



Christopher Bishop. “Pattern Recognition and Machine Learning”. Springer, 2006.



Ian Goodfellow, Yoshua Bengio and Aaron Courville. “Deep Learning”. MIT, 2016.

Online Resources

- Stanford course: “CS224d: Deep Learning for Natural Language Processing”. http://cs224d.stanford.edu/index.html
- CMU course: “Topics in Deep Learning” http://www.cs.cmu.edu/~rsalakhu/10807_2016/
- Hugo Larochelle Neural Network Course: http://info.usherbrooke.ca/hlarochelle/neural_networks/description.html
- Deep learning summer school in Montreal: https://sites.google.com/site/deeplearningsummerschool2016/home
- Many of the slides and materials are borrowed from the resources and books

Evaluation

- **Course Projects**:
  - Students should work on course projects in teams (at most 4 students).
  - At the end of this course, each team should make a poster (30%) and also hand in a project report (70%, due in two weeks after the course is finished).
- **Course report**
  - Should give a clear definition of the problem (10%)
  - A detailed survey of the problem (25%)
  - A proposal (35%)
  - Some preliminary results (not required, + 10 %)
  - Five pages in total (NIPS format, English)

Course Outline

- Introduction & Mathematics (Day 1)
- Machine Learning Basics (Day 2)
- Feedforward Neural Networks & Optimization Tricks (Day 3)
- Convolutional Neural Networks (Day 4)
- Recurrent Neural Networks (Day 4)
- Deep Learning for Natural Language Understanding (Day 5)
- Graph Representation Learning & Recommendation (Day 6)
- Poster Session (Day 6)

Thanks!