

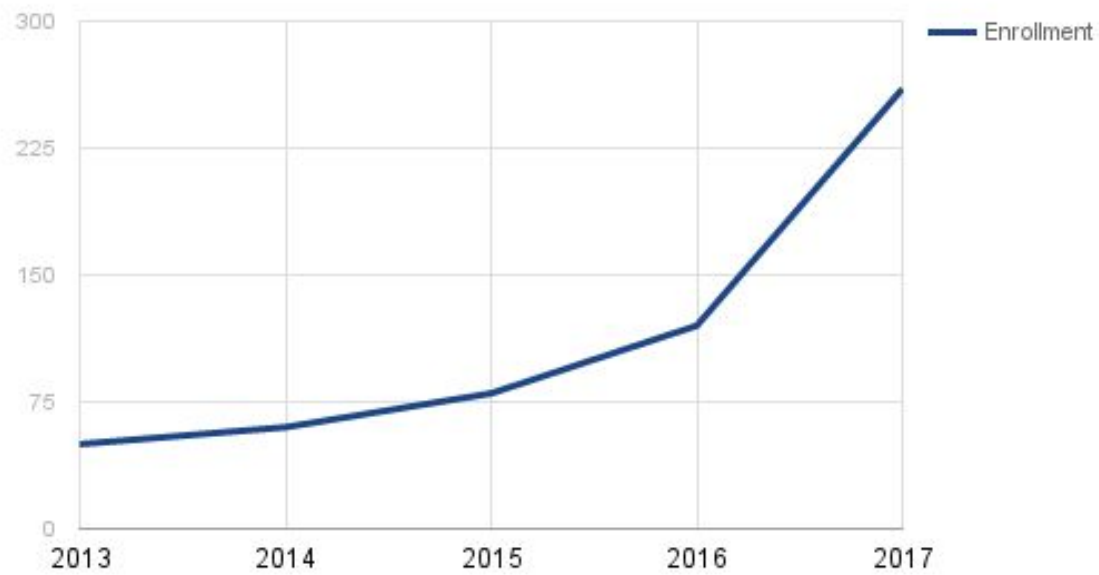
# ECE521 Lecture1

## Introduction



UNIVERSITY OF  
**TORONTO**

## ECE521 Enrollment



# Outline

- **History of machine learning**
- Types of machine learning problems

# What is machine learning?

- A scientific field is best defined by the central question it studies. The intellectual endeavour underlying the field of machine learning is:
    - “How can we build computer systems that automatically improve with experience, and what are the fundamental laws that govern all learning processes?”
- Tom Mitchell, Chair of the Machine Learning Department CMU, 2006

# What is machine learning?

- In other words, we are investigating the problems of how to get computers to program themselves.
  - ML has a strong computer science aspect: Which problems are inherently tractable? What architectures and algorithms are computational efficient?

# What is machine learning?

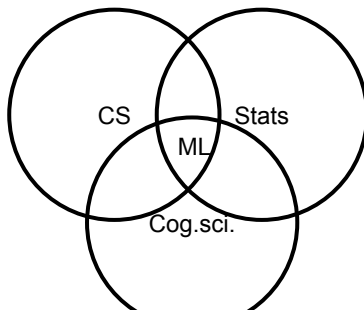
- In other words, we are investigating the problems of how to get computers to program themselves.
  - ML has a strong computer science aspect: Which problems are inherently tractable? What architectures and algorithms are computationally efficient?
  - ML borrows ideas from statistics: What can be inferred from data?

# What is machine learning?

- In other words, we are investigating the problems of how to get computers to program themselves.
  - ML has a strong computer science aspect: Which problems are inherently tractable? What architectures and algorithms are computationally efficient?
  - ML borrows ideas from statistics: What can be inferred from data?
  - ML tries to answer the same question asked in cognitive science / Psychology: How does human/machine intelligence emerge? Human/animal/machine learning are intertwined.

# What is machine learning?

- In other words, we are investigating the problems of how to get computers to program themselves.
  - ML has a strong computer science aspect: Which problems are inherently tractable? What architectures and algorithms are computational efficient?
  - ML borrows ideas from statistics: What can be inferred from data?
  - ML tries to answer the same question asked in cognitive science / Psychology: How does human/machine intelligence emerge? Human/animal/machine learning are intertwined.





# History of machine learning

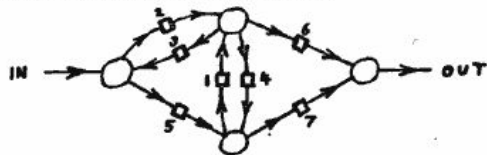
- At the beginning there is the “shallow” learning...

# History of machine learning

- Alan Turing wrote a little known paper in 1948 “Intelligent Machinery” that highlighted:
  - An unorganized machine that consists of randomly connected networks of NAND logic gates.
  - A general search algorithm that is similar to a “genetic algorithm” to organize the unorganized machine.
  - The unorganized machine resembles the cortex structure in the brain.

## 4. Organising unorganised machinery.

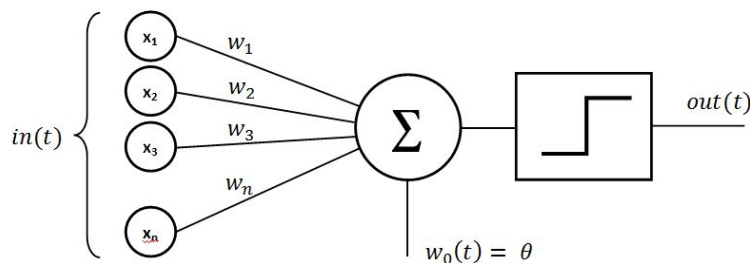
Many unorganised machines have configurations such that if once that configuration is reached, and if the interference thereafter is appropriately restricted, the machine behaves as one organised for some definite purpose. For instance the i-type machine shown below was chosen at random.



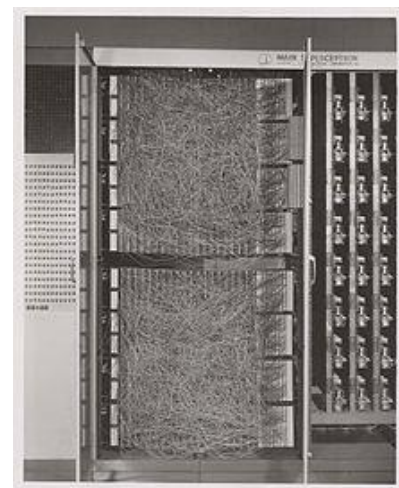
If the connections numbered 1, 3, 6, 4, are in condition ii) initially and connections

# History of machine learning

- Frank Rosenblatt in 1957 combined the ideas of the artificial neuron of McCulloch-Pitts and the Hebbian learning rule from Donald Hebb to develop the perceptron model:



$$Out = f\left(\sum_{i=1}^N w_i x_i + \theta\right)$$



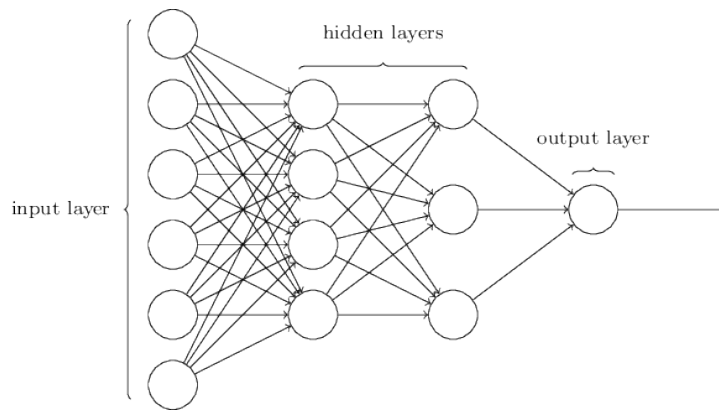
First implementation of perceptron [source](#)

# History of machine learning

- Then there was the first AI winter: 1970s
  - Machine translation did not make much progress from the breakthroughs of Chomsky's grammar
  - Perceptron was proven ineffective for non-linear classification problems

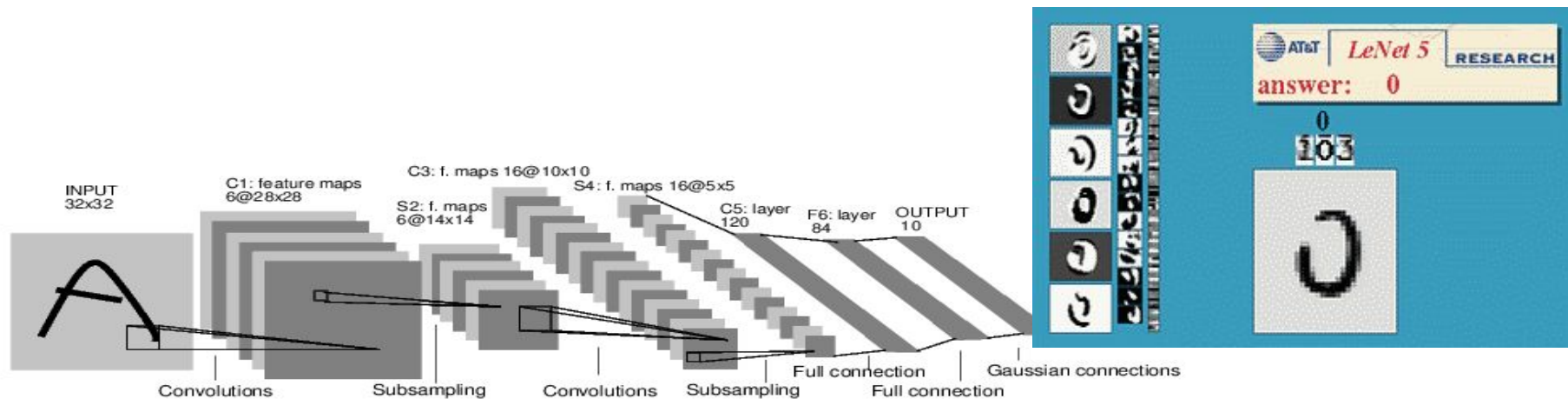
# History of machine learning

- The emergence of multi-layered perceptron and neural networks
  - Rumelhart, Hinton and Williams in 1986 highlighted a learning algorithm called “backpropagation” that can effectively train neural networks with multiple hidden layers.
  - Yann LeCun in 1989 proposed similar learning algorithm to train convolutional neural networks to recognize handwritten zip codes. Such a system has been used by USPS and bank ATMs saving hundreds of millions of dollars.



# History of machine learning

- The improved convolutional neural network LeNet that was deployed in 1997



LeNet-5 ([LeCun et al. 1998](#))

# History of machine learning

- Judea Pearl published Probabilistic Reasoning in Intelligent Systems in 1988 that changes the machine learning field to take statistical and probabilistic ideas seriously
  - Inspired statistical machine learning models for speech and language processing
  - It promotes the ideas of Hidden Markov Model, Kalman filter and particle filtering

























# History of machine learning

- Judea Pearl published Probabilistic Reasoning in Intelligent Systems in 1988 that changes the machine learning field to take statistical and probabilistic ideas seriously
  - Inspired statistical machine learning models for speech and language processing
  - It promotes the ideas of Hidden Markov Model, Kalman filter and particle filtering



# History of machine learning

- One interesting application of Bayesian inference is in matchmaking systems:

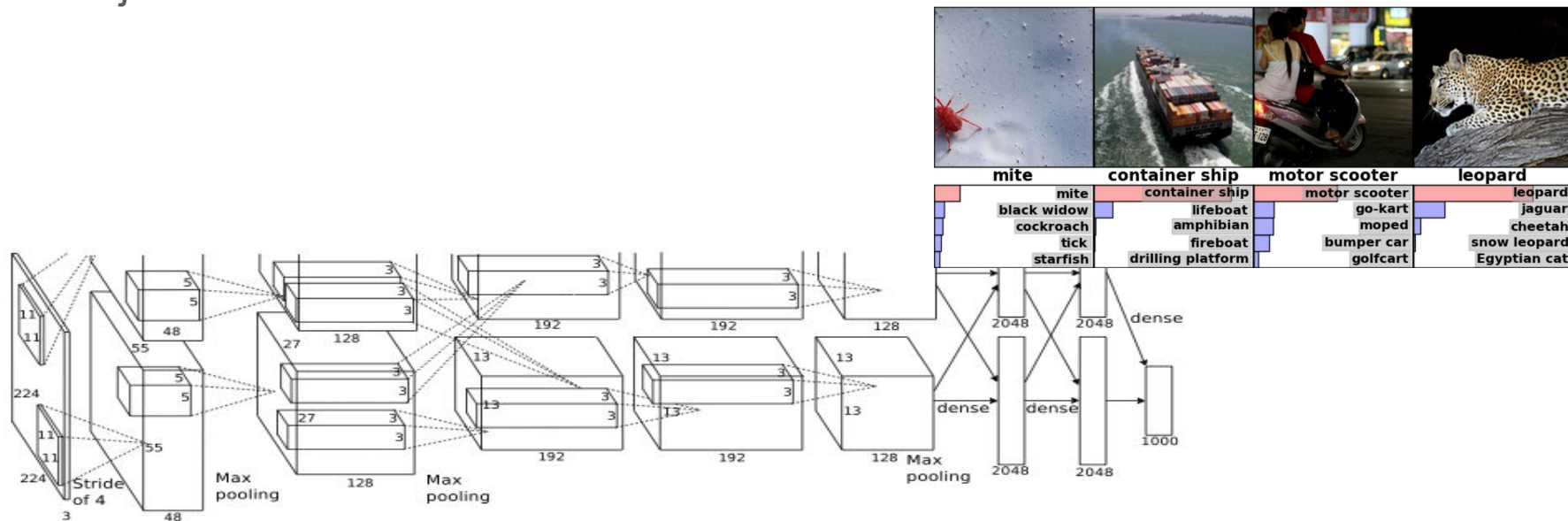
Blue Team								
	HorususVII		Unranked Division W: 0 L: 0	0% Wins	2 Lulu Games	2 / 5 / 19 Average Per Game	 0 / 0 / 30 Mastery	
	a68169356		Silver Division IV W: 409 L: 999	51% Wins	258 Fiora Games	7 / 7 / 8 Average Per Game	 9 / 21 / 0 Mastery	
	cheesus 198		Silver Division V W: 33 L: 30	52% Wins	6 Ashe Games	5 / 7 / 10 Average Per Game	 21 / 9 / 0 Mastery	
	Kriveless		Bronze Division I W: 17 L: 24	41% Wins	10 Lux Games	9 / 7 / 10 Average Per Game	 21 / 0 / 9 Mastery	
	NoAskill		Silver Division IV W: 221 L: 243	48% Wins	85 Xin Zhao Games	6 / 7 / 10 Average Per Game	 15 / 15 / 0 Mastery	
Champion	Summoner	Premade	Rank	Win%	Champion Games	Champion KDA	Runes	Mastery
Purple Team								
	KrustyKrust		Silver Division IV W: 209 L: 202	51% Wins	15 Sona Games	2 / 5 / 12 Average Per Game	 0 / 9 / 21 Mastery	
	megapaton		Silver Division IV W: 110 L: 100	52% Wins	6 Jarvan IV Games	4 / 5 / 10 Average Per Game	 9 / 21 / 0 Mastery	
	xYogiBear		Bronze Division III W: 114 L: 143	44% Wins	14 Tryndamere Games	9 / 5 / 6 Average Per Game	 21 / 9 / 0 Mastery	
	robinska		Gold Division IV W: 129 L: 112	54% Wins	0 Fizz Games	0 / 0 / 0 Average Per Game	 21 / 9 / 0 Mastery	
	doubleblade		Silver Division IV W: 468 L: 468	50% Wins	16 Ezreal Games	7 / 7 / 6 Average Per Game	 21 / 9 / 0 Mastery	

# History of machine learning

- Then the computers were too slow so we did not make much progress till 2012

# History of machine learning

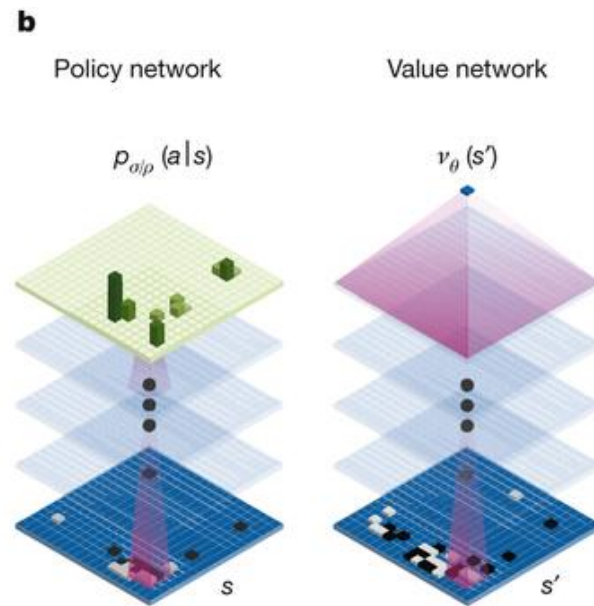
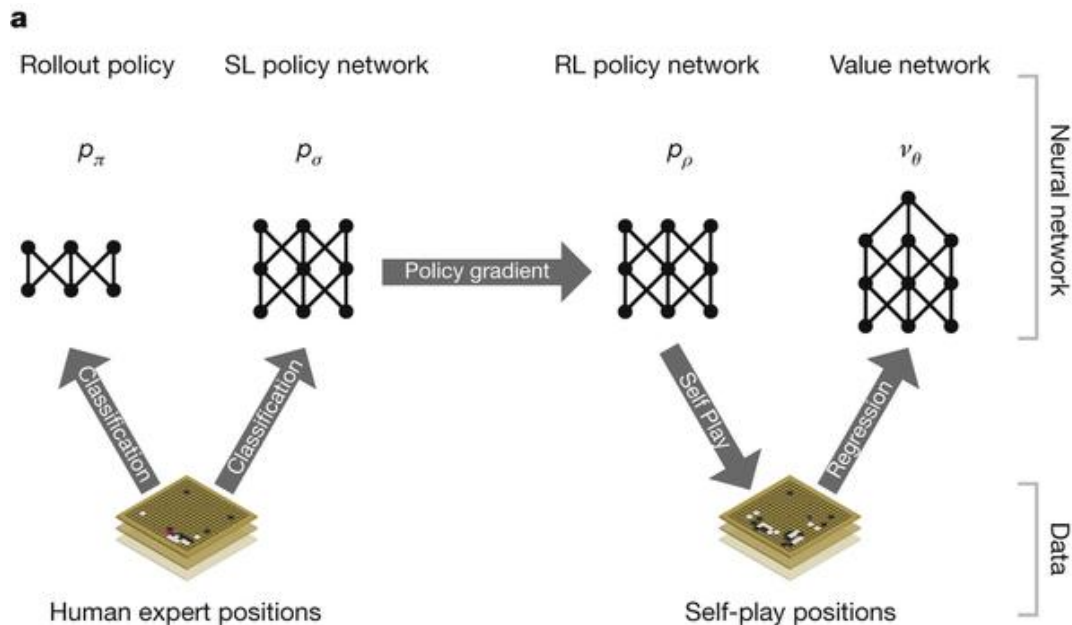
- A large-scale convolutional neural network that can recognize 1000s of objects



AlexNet([Krizhevsky et al. 2012](#))

# History of machine learning

- Mastering the game of Go with deep learning



AlphaGo([Silver et al. 2015](#))

# Current machine learning applications:

- Speech recognition



# Current machine learning applications:

- Computer vision





# Current machine learning applications:

- Natural language processing

EnglishSpanishFrenchSpanish - detected

EnglishSpanishArabic

Translate

Capítulo primero. Que trata de la condición y ejercicio del famoso hidalgo don Quijote de la Mancha

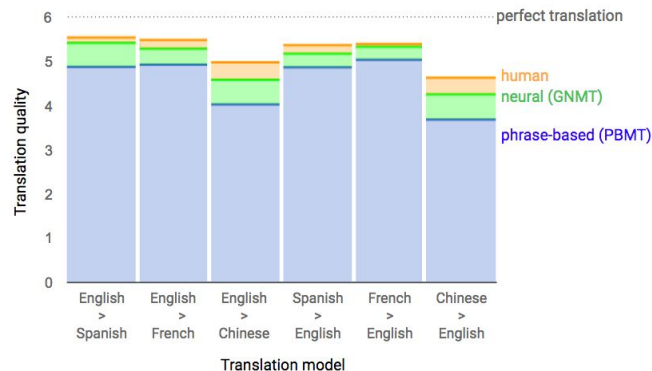
En un lugar de la Mancha, de cuyo nombre no quiero acordarme, no ha mucho tiempo que vivía un hidalgo de los de lanza en astillero, adarga antigua, rocín flaco y galgo corredor. Una olla de algo más vaca que carnero, salpicón las más noches, duelos y quebrantos los sábados, lantejas los viernes, algún palomino de añadidura los domingos, consumían las tres partes de su hacienda. El resto della concluían sayo de velarte, calzas de velludo para las fiestas, con sus pantuflos de lo mismo, y los días de entresemana se honraba con su vellorí de lo más fino. Tenía en su casa una ama que pasaba de los cuarenta, y una sobrina que no llegaba a los veinte, y un mozo de campo y plaza, que así ensillaba el rocín como tomaba la podadera. Frisaba la edad de nuestro hidalgo con los cincuenta años; era de complexión recia, seco de carnes, enjuto de rostro, gran madrugador y amigo de la caza. Quieren decir que tenía el sobrenombre de Quijada, o Quesada, que en esto hay alguna diferencia en los autores que deste caso escriben; aunque, por conjeturas verosímiles, se deja entender que se llamaba Quejana. Pero esto importa poco a nuestro cuento; basta que en la narración dél no se salga un punto de la verdad.

1307/5000

First chapter. Which deals with the condition and exercise of the famous nobleman Don Quixote de la Mancha

In a place of La Mancha, whose name I do not want to remember, there has not been a long time that lived a lord of the lance in shipyard, old pork, thin rocín and greyhound runner. A pot of something more cow than ram, spit most nights, duels and breaks on Saturdays, giblets on Fridays, some palomino in addition to Sundays, consumed the three parts of his estate. The rest of the party concluded a velvet dress, hairy tights for the parties, with their slippers of the same, and the days of midweek were honored with their vellorí of the finest. He had a housekeeper in his house who was in his forties, and a niece who was not in his twenties, and a boy in the country and square, who saddled the rocín as he took the pruning. He emphasized the age of our hidalgo at the age of fifty; Was of a hard complexion, dry of flesh, thin of face, great early bird and friend of the hunt. They mean that he had the nickname of Quijada, or Quesada, that in this there is some difference in the authors who in this case write; Although, by plausible conjectures, it is understood that it was called Quejana. But this matters little to our story; It is enough that in the narration of him a point of truth does not come out.

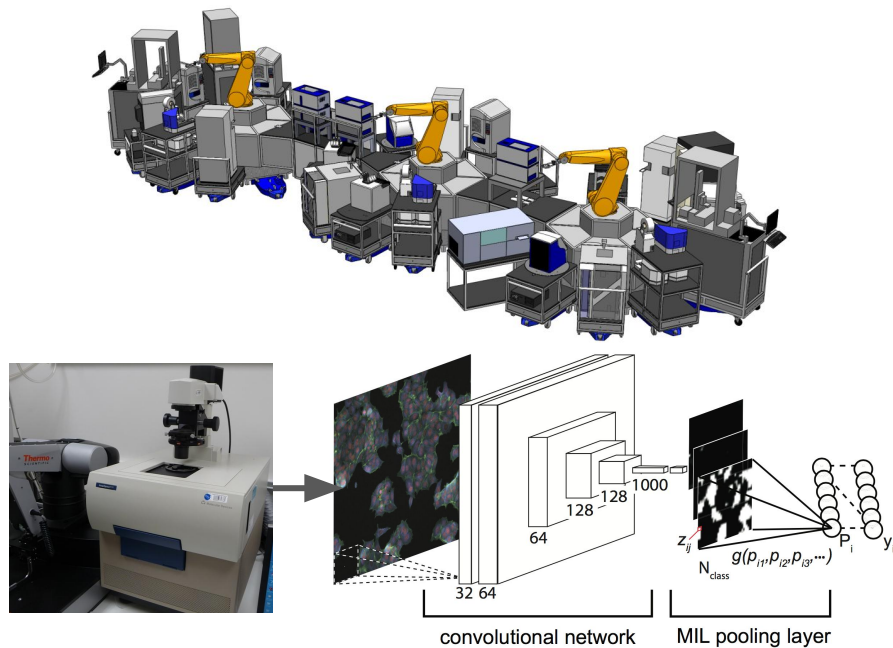
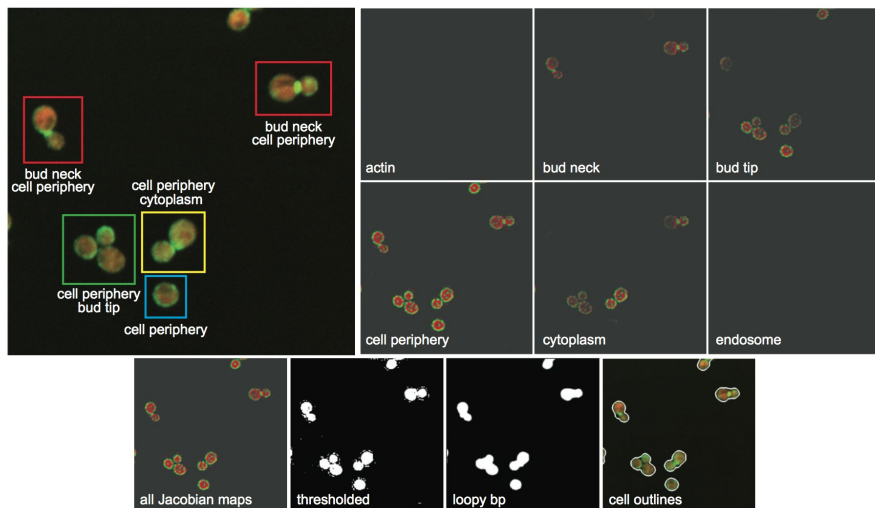
Suggest an edit



Google's Neural Machine Translation ([Wu et al. 2016](#))

# Current machine learning applications:

- Computational biology

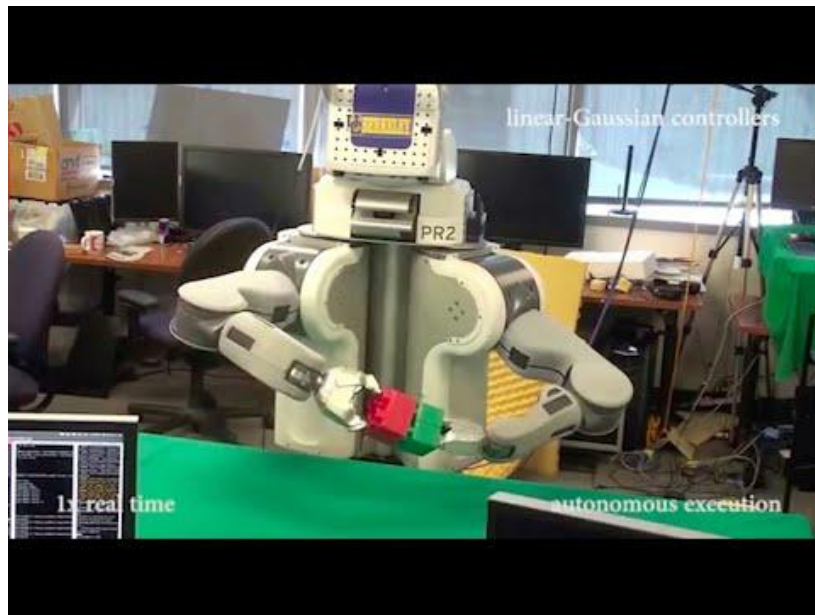
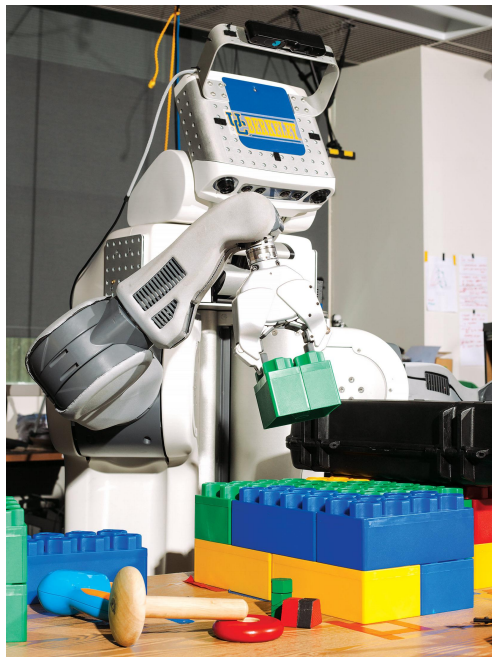


High-throughput microscopy  
of cellular data ([Oren et al. 2016](#))



# Current machine learning applications:

- Robotics



Berkeley's robot learnt using reinforcement learning([Levine et al. 2015](#))

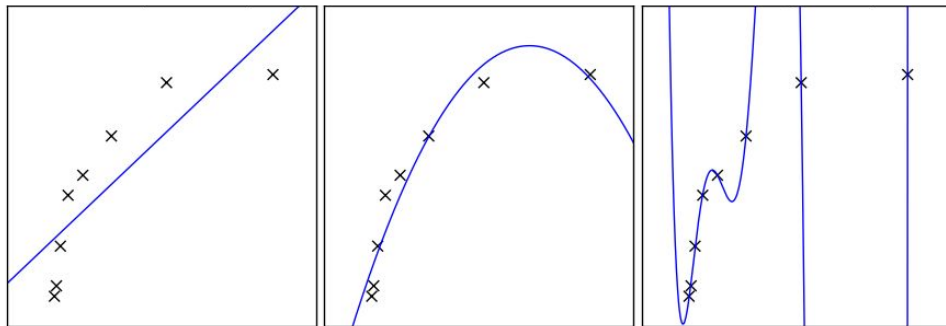
# Outline

- History of machine learning
- **Types of machine learning problems**

# Types of machine learning

- Supervised learning:

- Given a set of labeled training data points  $\{(x^i, y^i)\}$
- Space of input data and labels:  $x \in \mathcal{X}$ ,  $y \in \mathcal{Y}$
- The goal is to learn a function mapping  $f$ , that  $f : \mathcal{X} \rightarrow \mathcal{Y}$



# Types of machine learning

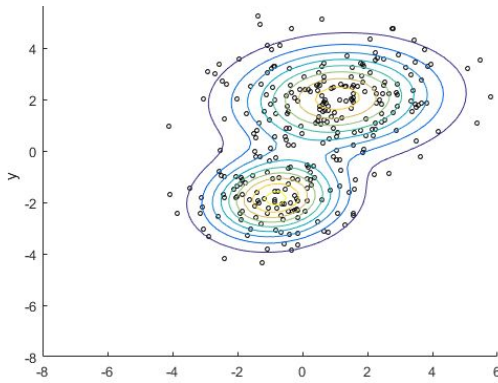
- “... Intelligence is not just about fitting some lines through bunch of points...”

# Types of machine learning

- Unsupervised learning: There is not label in the dataset. We would like to discover interesting patterns and structures within the input data.
  - Given a set of unlabelled training data points:  $\{x^i\}$
  - Space of input data:  $x \in \mathcal{X}$
  - One possible goal is to model the empirical distribution with a parametric distribution:  $p(x \mid \theta)$

# Types of machine learning

- Unsupervised learning: There is not label in the dataset. We would like to discover interesting patterns and structures within the input data.
  - Given a set of unlabelled training data points:  $\{x^i\}$
  - Space of input data:  $x \in \mathcal{X}$
  - One possible goal is to model the empirical distribution with a parametric distribution:  $p(x \mid \theta)$



# Types of machine learning

- Semi-supervised learning:
  - Given a dataset in terms of a mixture of labelled and unlabelled data

# Types of machine learning



How to grow a mind ([Tenenbaum, 2012](#))



# Types of machine learning

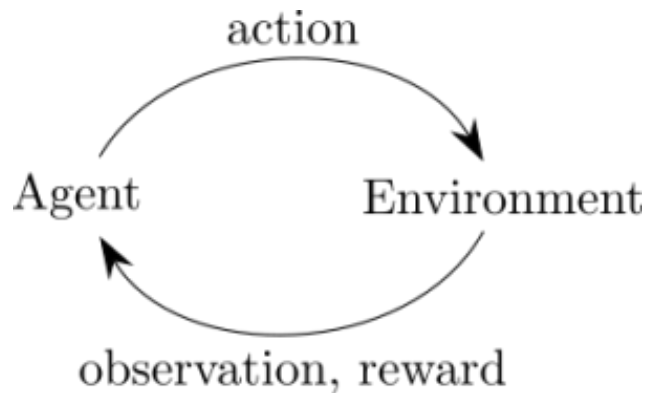
What are  
the other  
“Tufa”?



How to grow a mind ([Tenenbaum, 2012](#))

# Types of machine learning

- Reinforcement learning:



# What is this course all about?

- Concrete formulation of a learning problem in terms of a loss function
- Use gradient-based optimization algorithms to minimize the loss function
  - Learning: search for a set of parameters/weights that minimizes the loss function
  - Inference: search for a set of latent causes to explain the observed data

# What is this course all about?

- Learning algorithms
  - Back-propagation
  - Gradient descent
- Inference algorithms
  - Bayes rules
  - The sum-product algorithm

# What is this course all about?

- Supervised learning models
  - K-NN
  - Linear models: Linear regression, logistic regression
  - Neural networks
- Unsupervised learning models
  - K-means, Mixtures-of-Gaussians
  - PCA, Auto-encoder
  - Hidden Markov Models
  - Some acyclical graphs

# What is this course all about?

- Mechanical questions (easy free marks)
  - Carry out an algorithm on particular models and data
- Brain teasers
  - What happens when we do this?
  - Is it possible to have this scenario?

# Course topics:

covered in this class

machine learning

reinforcement learning

deep learning

graphical models

back-propagation

gradient-descent

neural nets

convolutional neural net

recurrent neural net

EM

mixture models

continuous latent  
variable models

Hidden Markov Models

Kalman filter

Markov random fields

Boltzmann machine

Bayesian inference

collaborative filtering

particle filtering

Monte Carlo  
methods

Bayesian non-parametrics

kernel methods

support vector  
machines

Gaussian  
processes