

# Applied Machine Learning

Introduction

**Siamak Ravanbakhsh, Reihaneh Rabbany**

**COMP 551 (winter 2020)**

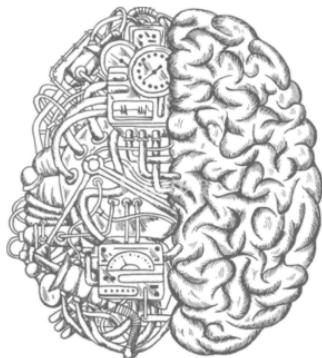
# Objectives

- understanding the scope of machine learning
  - relation to other areas
- understanding types of machine learning

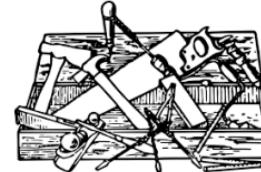
# What is Machine Learning?

ML is the set of "algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions"

## an inadequate history of ML



- 1950: Turing test
- 1956: checker player that learned as it played (Arthur Samuel)
  - coined the term Machine Learning
- 1958: first artificial neural networks called Perceptron (Frank Rosenblatt),
  - 1959: ADELINe (Widrow and Hoff)
- 1963: support vector machines (Vapnick & Ya)
- 1969: Minsky and Papert show the limitations of single-layer neural networks
- 1970-80s rule-based and symbolic AI dominates (two AI winters)
- 1980's Bayesian networks (Judea Pearl)
- 1986 Backpropagation rediscovered (Rumelhart, Hinton & Williams)
- 1991 Kernel trick for SVM
- 2012 AlexNet wins Imagenet by a large margin
- 2012 - now deep learning explosion...
- next? AI winter? AGI?



# Placing ML: overlapping fields

*subset (learn from data)*

- **Artificial Intelligence:** its a broader domain (includes search, planning, multiagent systems, robotics, etc.)
- **Statistics:** historically precedes ML. ML is more focused on algorithmic, practical and powerful models (e.g., neural networks) and is built around AI
- **Vision & Natural Language Processing:** use many ML algorithms and ideas
- **Optimization:** extensively used in ML
- **Data mining:** scalability, and performance comes before having strong theoretical foundations, more space for using heuristics, exploratory analysis, and unsupervised algorithms
- **Data science:** an umbrella term for the above mostly used in industry when the output is knowledge/information to be used for decision making

# Placing ML: main venues

top computer science conferences

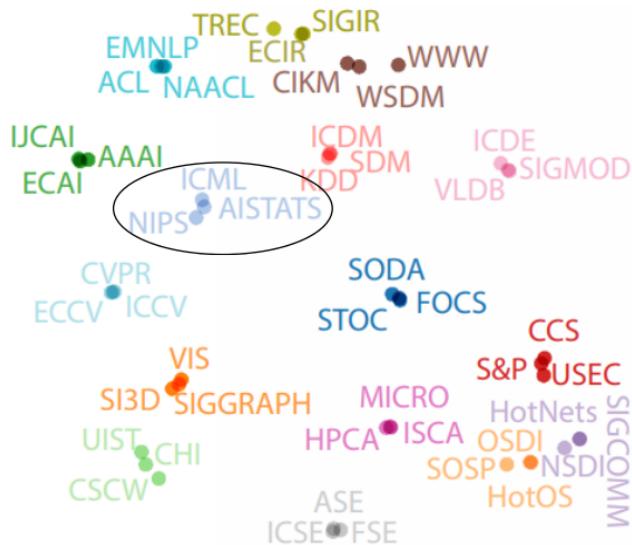
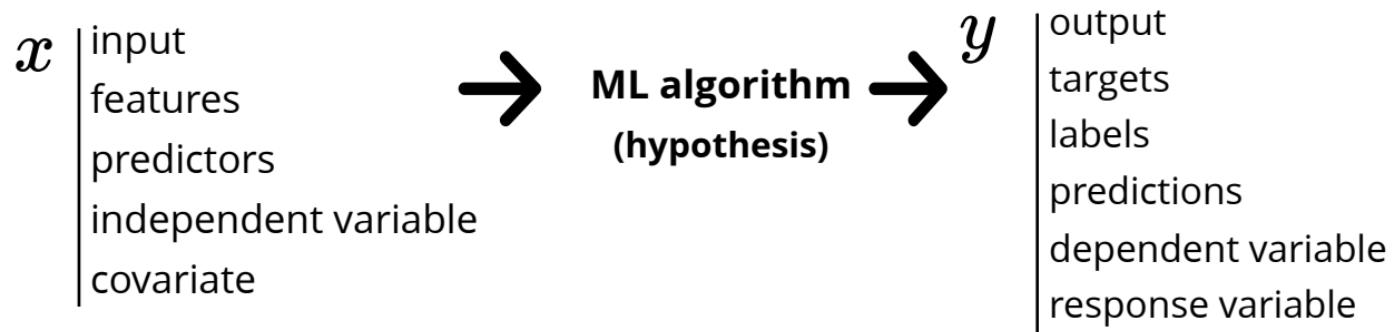


Figure from Dong et al.

# Some terminology



example

$\langle$ tumorsize, texture, perimeter $\rangle = \langle 18.2, 27.6, 117.5 \rangle \rightarrow$  cancer = No

# Some terminology

(labelled) **datasets**: consist of many training examples or **instances**

<tumorsize, texture, perimeter> , <cancer, size change>

<18.2, 27.6, 117.5> , <No, +2>

$x^{(1)}$

<17.9, 10.3, 122.8> , <No, -4>

$x^{(2)}$

one instance

<20.2, 14.3, 111.2> , <Yes, +3>

$x^{(3)}$

⋮

⋮

<15.5, 15.2, 135.5> , <No, 0>

$x^{(N)}$

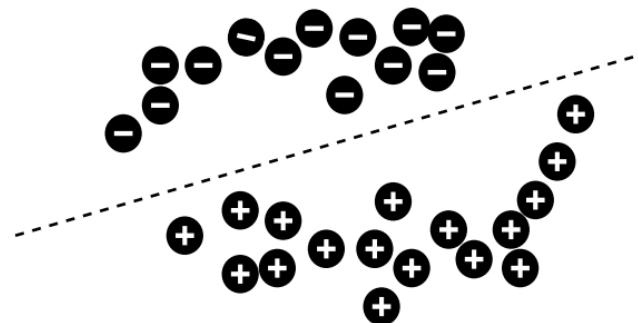
# "Types" of Machine Learning



**Supervised Learning:** we have **labeled** data

- classification
- regression
- structured prediction

most of this course!



# Supervised Learning

**Classification:** categorical/discrete output

<tumorsize, texture, perimeter>	,	<cancer>
<18.2,	27.6,	117.5>
<17.9,	10.3,	122.8>
<20.2,	14.3,	111.2>
<15.5,	15.2,	135.5>

target

**Regression:** continuous output

<tumorsize, texture, perimeter>	,	<size change>
<18.2,	27.6,	117.5>
<17.9,	10.3,	122.8>
<20.2,	14.3,	111.2>
<15.5,	15.2,	135.5>

target

# Supervised Learning: Example

MIT  
Technology  
Review

Topics+



(a)

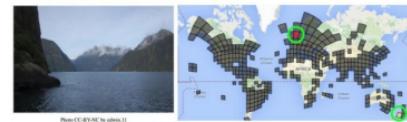
---

Intelligent Machines

## Google Unveils Neural Network with “Superhuman” Ability to Determine the Location of Almost Any Image

Guessing the location of a randomly chosen Street View image is hard, even for well-traveled humans. But Google’s latest artificial-intelligence machine manages it with relative ease.

by Emerging Technology from the arXiv February 24, 2016



# Supervised Learning: Example

## Machine Translation:

data consists of input-output sentence pairs ( $x, y$ )

**DeepL schools other online translators with clever  
machine learning**

Devin Coldewey, Frederic Lardinois / 1:57 pm EDT • August 29, 2017

Comment



Image Credits: H. Armstrong Roberts/Getty Images

# Supervised Learning: Example

## Image captioning

**input:** image

**output:** text



The man at bat readies to swing at the pitch while the umpire looks on.



A large bus sitting next to a very tall building.

image: COCO dataset

# Supervised Learning: Example

## Object detection

**input:** image

**output:** a set of bounding box coordinates

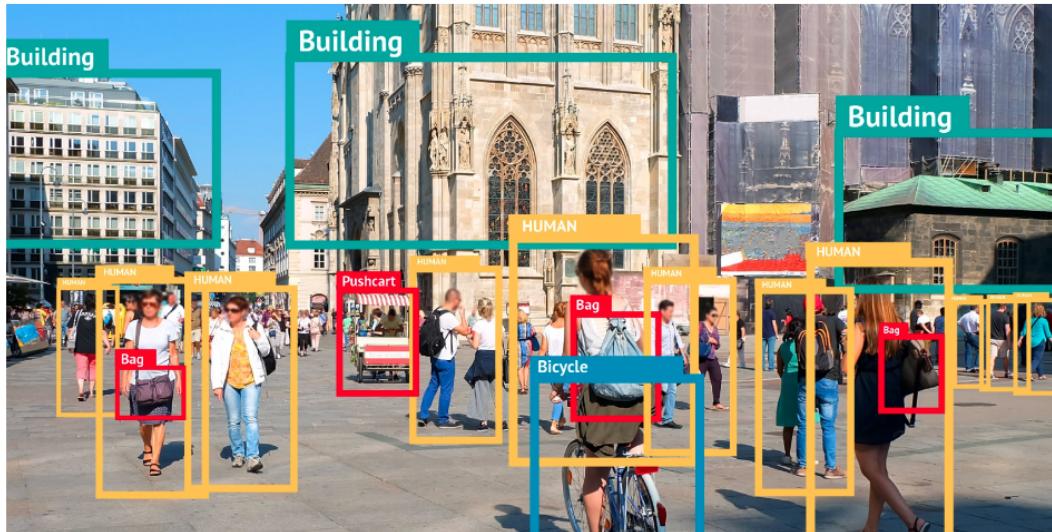


image: <https://bitmovin.com/object-detection/>

# "Types" of Machine Learning



**Unsupervised Learning:** only unlabeled data

- clustering
- dimensionality reduction
- density estimation / generative modeling
- anomaly detection
- discovering latent factors and structures

helps explore and understand the data  
closer to data mining  
we have much more unlabeled data  
more open challenges

# Unsupervised Learning: Example

## clustering

similar to classification but labels/classes should be inferred and are not given to the algorithm

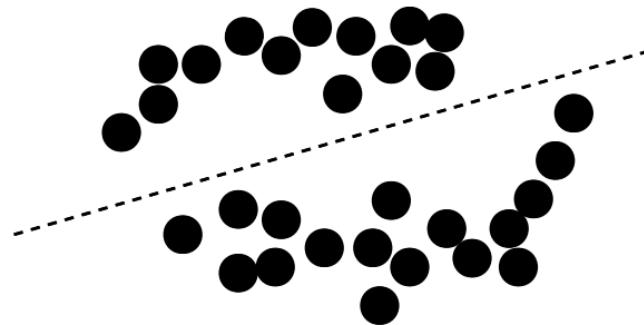
<tumorsize, texture, perimeter>

<18.2, 27.6, 117.5>

<17.9, 10.3, 122.8>

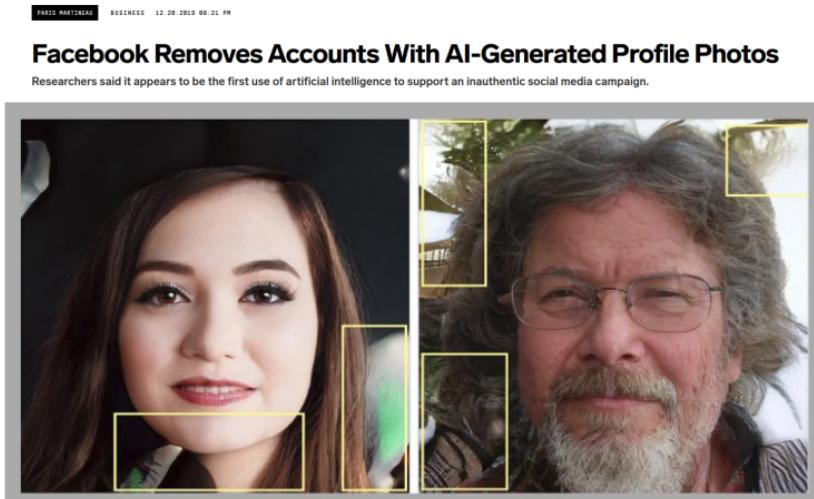
<20.2, 14.3, 111.2>

<15.5, 15.2, 135.5>



# Unsupervised Learning: Example

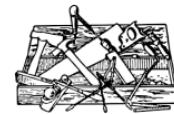
**Generative modeling** (density estimation):  
learn the data distribution  $p(x)$



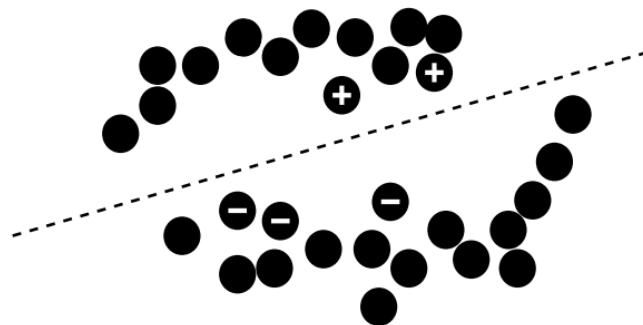
Profile pictures for Facebook accounts "Mary Keen" and "Jacobs Guillermo," admins on groups associated with The BL highlighted by Graphika. COURTESY OF GRAPHIKA

# "Types" of Machine Learning

**Semisupervised learning:** a few labeled examples



- we can include structured problems such as
  - matrix completion (a few entries are observed)
  - link prediction



# Netflix Awards \$1 Million Prize and Starts a New Contest

BY STEVE LOHR SEPTEMBER 21, 2009 10:15 AM



on Kempin/Getty Images Netflix prize winners, f  
Chabbert, Martin Piotte, Michael Jahrer, Andreas To:

**Update | 1:45 p.m.** Adding details announced  
close finish to the contest.



# "Types" of Machine Learning

## Reinforcement Learning:

- weak supervision through the reward signal
- sequential decision making
- biologically motivated

also related:

**imitation learning:** learning from demonstrations

- behavior cloning (is supervised learning!) *predict the action given the state*
- inverse reinforcement learning (learning the reward function)



# Reinforcement Learning: Example



## Human Level Control Through Deep Reinforcement Learning

### Abstract

The theory of reinforcement learning provides a normative account deeply rooted in psychological and neuroscientific perspectives on animal behaviour, of how agents may optimize their control of an environment. To use reinforcement learning successfully in situations approaching real-world complexity, however, agents are confronted with a difficult task: they must derive efficient representations of the environment from high-dimensional sensory inputs, and use these to generalize

# Reinforcement Learning: Example

## *Google's AlphaGo Defeats Chinese Go Master in Win for A.I.*

[点击查看本文中文版](#)

By PAUL MOZUR MAY 23, 2017



Ke Jie, the world's top Go player, reacting during his match on Tuesday against AlphaGo, artificial intelligence software developed by a Google affiliate. China Stringer Network, via Reuters

### RELATED COVERAGE



A.I. Is Doing Legal Work. But It Won't Replace Lawyers, Yet. MARCH 19, 2017



China's Intelligent Weaponry Gets Smarter FEB. 3, 2017



THE FUTURE OF WORK  
The Future of Not Working FEB. 23, 2017



Master of Go Board Game Is Walloped by Google Computer Program MARCH 9, 2016

# Summary

- **Supervised Learning:** we have labeled data
  - classification
  - regression
  - structured prediction
- **Unsupervised Learning:** only unlabeled data
  - clustering
  - dimensionality reduction
  - density estimation / generative modeling
  - anomaly detection
  - discovering latent factors and structures
- **Semisupervised learning:** a few labeled examples
- **Reinforcement Learning:** reward signal