Machine Learning: Overview

September 25, 2017

How (I guess) many people view Machine Learning

Machine learning

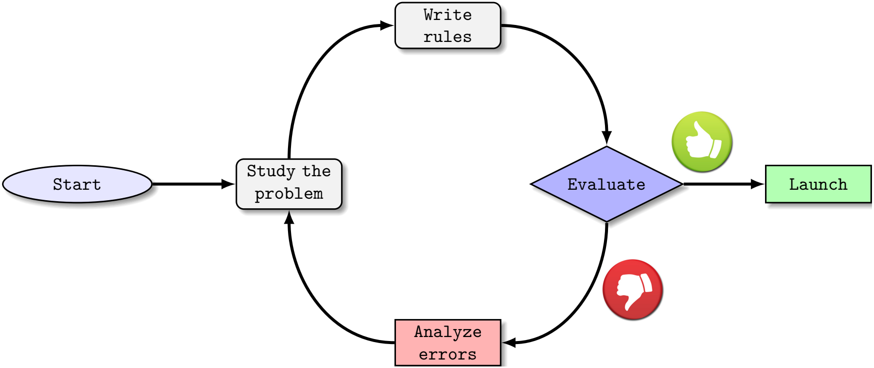
[Machine Learning is the] field of study that gives computers the ability to learn without being explicitly programmed.

- **Arthur Lee Samuel, 1959**

A computer program is said to learn from experience E with respect to some task T and some performance P, if its performance on T, as measured by P, improves with experience E.

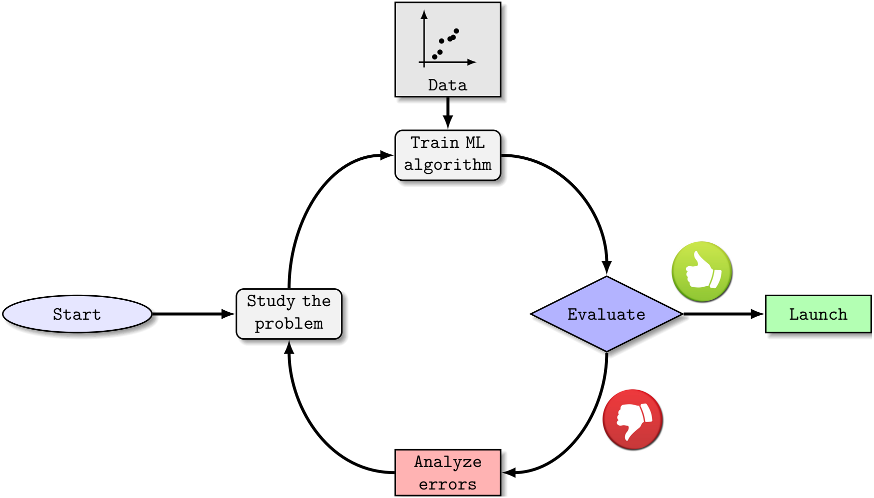
- **Tom M. Mitchell, 1997**

Why do Machine Learning: Traditional approach



Writing general rules that cover every possibility can be hard!

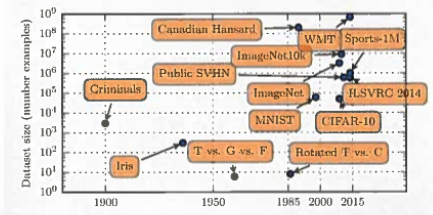
Why do Machine Learning: Machine Learning approach

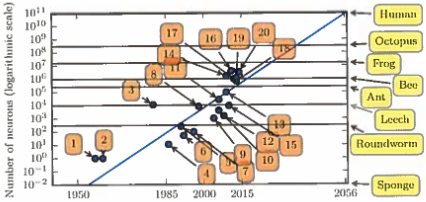


Do not write rules, instead let the machine figure it it out!

I've heard about ML since the 50's. Why now?

The **most** significant factor is increase in computing power!





Rule of thumb: A Machine Learning algorithm will achieve **acceptable performance** at 5000 examples, and **super-human performance** with at least 10 million examples.

Big Data

The amount of data generated in the world increases each year.

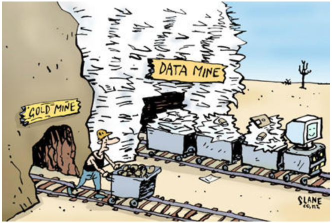
"90% of the worlds data has been generated in the last two years." (2013)

Data mining (Facebook, Google, etc.).

Machine learning is ideal for exploiting opportunities hidden in big data.

We are drowning in information and starving for knowledge.

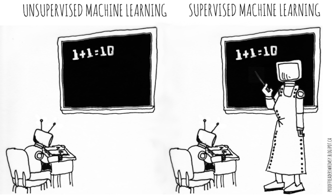
- **Rutherford D. Rogers**



Types of Machine Learning problems

Unsupervised Learning

Learning from *unlabeled* data.



Supervised Learning

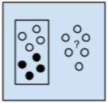
Learning from *labeled* data (examples).

Semi-supervised Learning

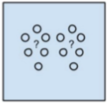
Combination of unsupervised and supervised learning.

Reinforcement Learning

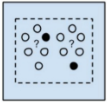
Learning from experience.



Supervised Learning



Unsupervised Learning



Semi-Supervised Learning

Reinforcement Learning

Small Machine Learning (ML) dictionary

Attribute: A data type. For car data it might be, *e.g.* "Mileage".

Feature: Depends somewhat on context, but usually means *attribute* plus *value*, *e.g.* "Mileage = 15.000".

Instance: A collection of all *features* for a single data point.

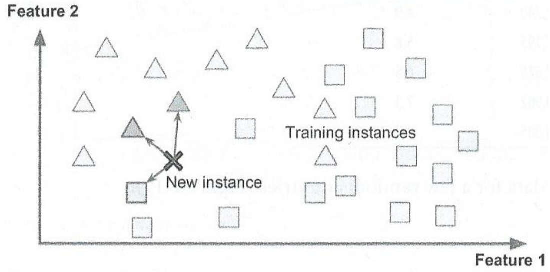
Predictors: *Features* that you want use for predictions on new *instances*.

Label: Desired output given input.

Instance-based learning

Instance- or **memory** based learning compares new problem instances with instances seen in training.

Learning by heart.

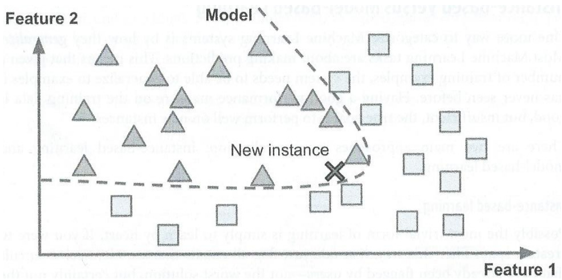


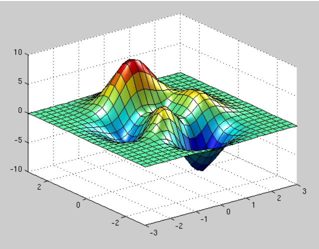
Model based learning

Build a model from examples, and use the model to make predictions.

Specific instances in the training are not used to compare against!

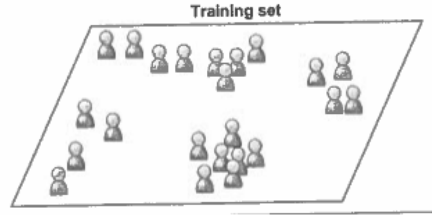
Depending on the model training might be a multidimensional **optimization** problem!

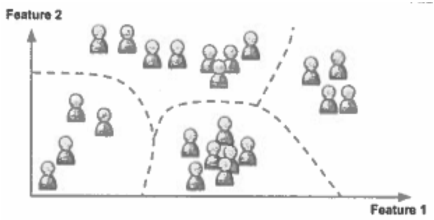




Unsupervised learning

Unsupervised learning works with **unlabeled data**.



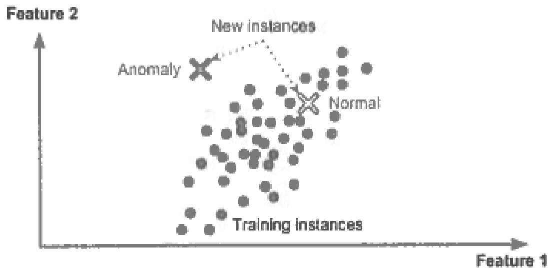


Typical use is **Clustering**.

Example: Users of a homepage, who is visiting?

Unsupervised learning: Continued

Also used for **Anomaly detection**.

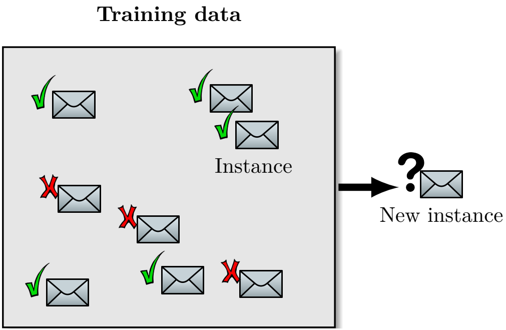


Some unsupervised learning methods:

- k-Means (Clustering)
- Principal Component Analysis (Visualization, Dimensionality reduction, Feature extraction)

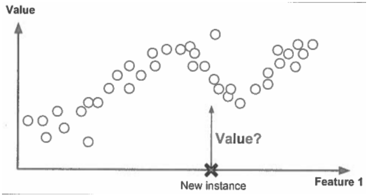
Supervised learning

Supervised learning works with **labeled data**, *i.e.* some input accompanied by desired output.



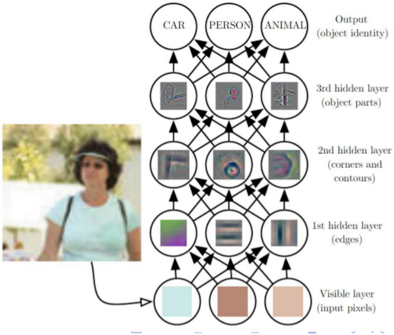
Typical task is **classification**, *e.g.* classify an email as spam or ham.

Predict a **target** numeric value, given a set of **features** called **predictors** (Regression).

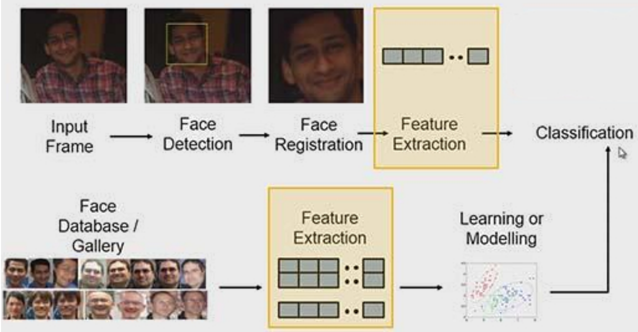


Some supervised learning algorithms:

- k-Nearest Neighbours
- Linear Regression
- Support Vector Machines (SVMs)
- GPR
- Neural Networks
- Etc., etc., etc, ...



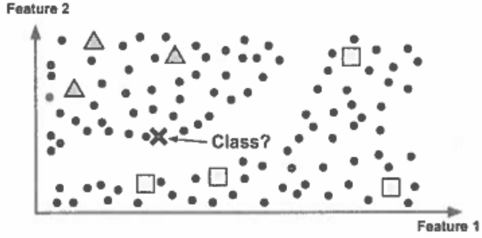
Example: Facial recognition



Semi-supervised learning

Semi-supervised learning deals with labeled and unlabeled data.

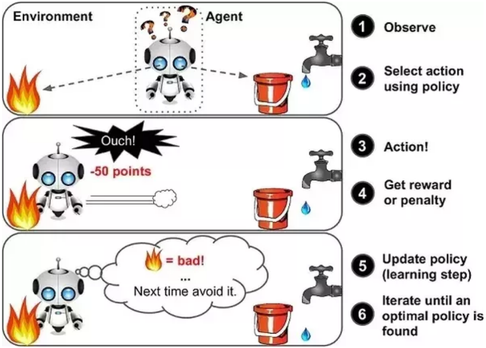
Most algorithms are combinations of supervised and unsupervised learning algorithms.



A semi-supervised algorithm could be a *Deep belief network*.

Reinforcement learning

Reinforcement learning is learning by **experience**.

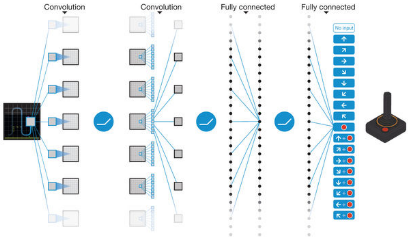


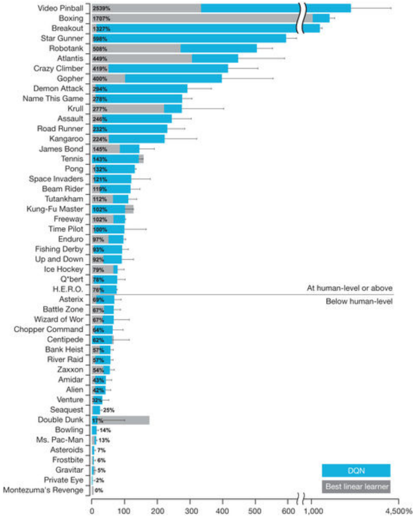
Some reinforcement learning algorithms:

- Q-learning (Q-networks)
- Genetic algorithms

Yo dawg, I heard you like computers...

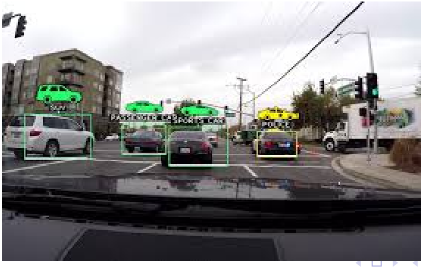
... so I taught your computer to use the computer! (Q-Learning)





Lets have a look: Link

Example: Self-driving cars



Correlation doesn't mean causation!

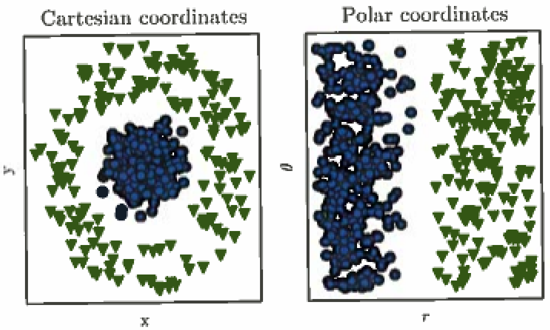
"The machine learning algorithm wants to know if we'd like a dozen wireless mice to feed the Python book we just bought."

"You can't keep adjusting the data to prove that you would be the best Valentine's date for Scarlett Johansson."

Representation of Data

Tasks may be **impossible** in one representation, but **easy** in another.

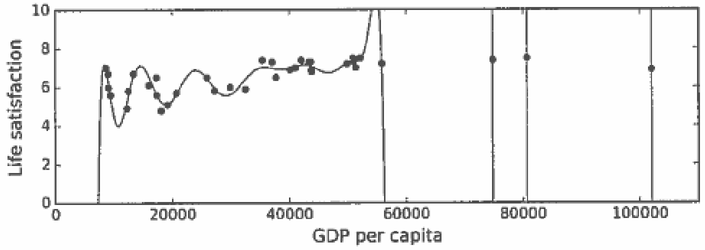
Can you draw a straight line separating the two categories?



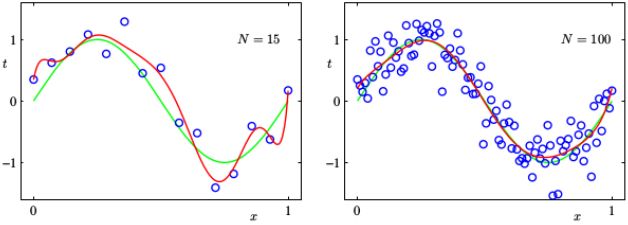
It is crucially important which **features** to use in the model.

Overfitting

If the model has to much freedom, overfitting can occur.



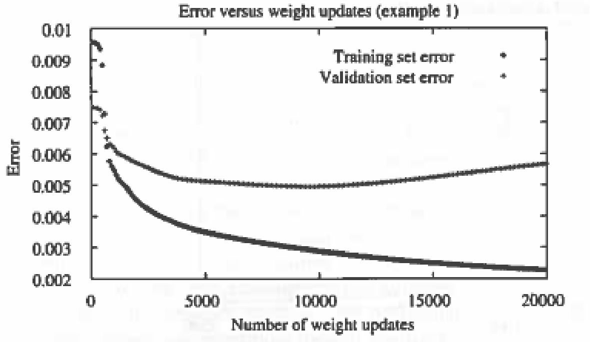
This can be mitigated somewhat by large data sets.



Bias towards training data

Be careful to have a broad spectrum of training data.

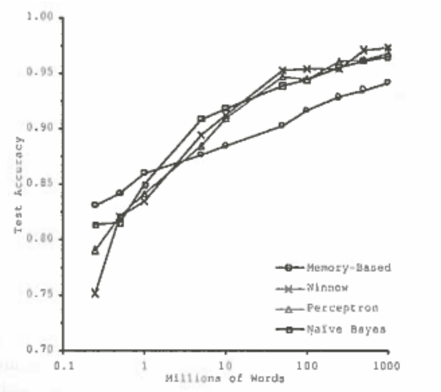
Some algorithms can be overtrained to idiosyncrasies of the training data.



Be sure to validate your model!

Needs lots of data

Actually, having the right (amount of) data is often more important that having the right model.

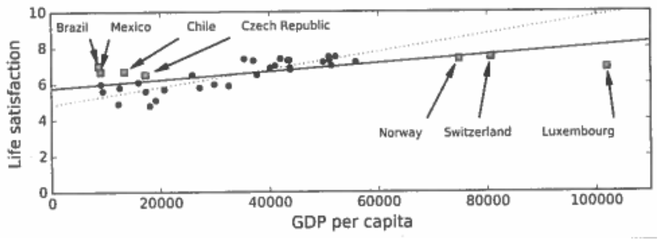


"Data is ten times more powerful than algorithms"

- **Peter Norvig**

Quality of data

Use data that is **representative** of the problem you want to model.



Poor data quality: clean out **noise** and **outliers**.

Apply **feature engineering**:

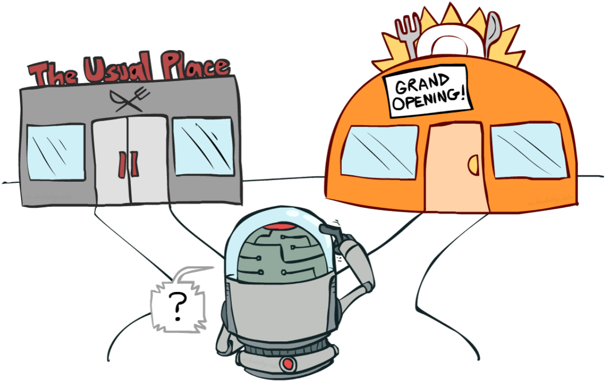
Feature selection: Select the most important features.

Feature extraction: Combine features to produce more useful ones.

Reinforcement learning: Exploration VS Exploitation

Exploitation: Make *best descision* given current information.

Exploration: gather *more information*.



No Free Lunch

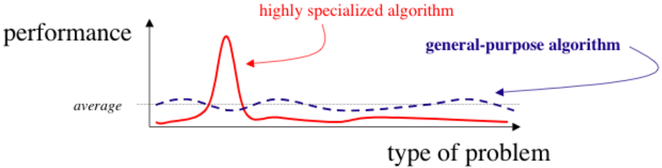
"No Free Lunch"

\- **David H. Wolpert**

A model is a simplification of reality.

Simplification is based on assumption.

Assumptions **always** fail in certain situations.



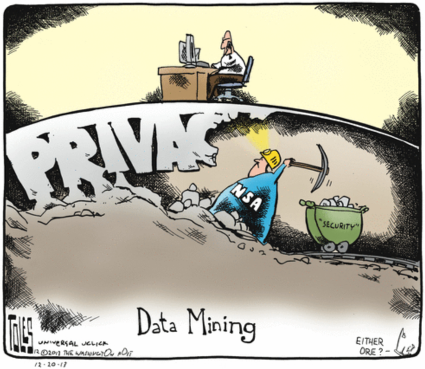
No one model works best in all possible situations.

Social and ethical aspects

Cathy O'Neil: The era of blind faith in big data must end

Algorithms decide who gets a loan, who gets a job interview, who gets insurance and much more -- but they don't automatically make things fair. Mathematician and data scientist Cathy O'Neil coined a term for algorithms that are secret, important and harmful: "weapons of math destruction." Learn more about the hidden agendas behind the formulas.

http://www.ted.com/talks/cathy_o_neil_the_era_of_blind_faith_in_big_data_must_end



MAGIC?!

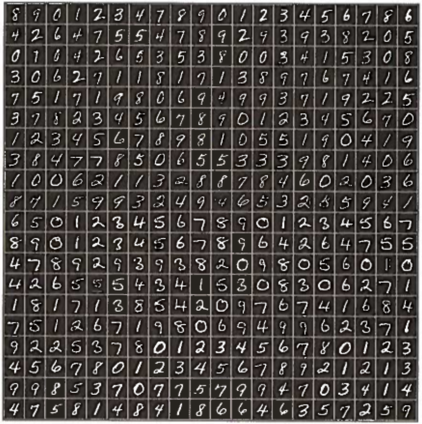
Just feed in the data and watch the computer magically figure out the equation that fits the data!

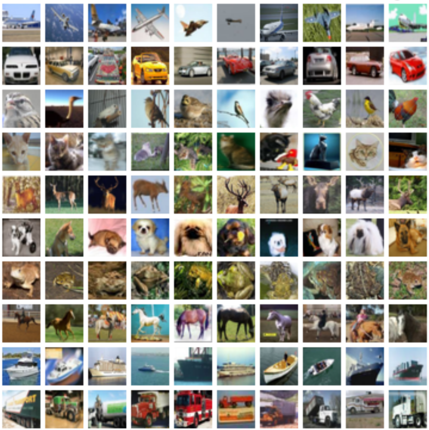
But, it's important to remember that machine learning only works if the problem is actually solvable with **the data that you have**.

Remember, if a human expert couldn't use the data to solve the problem manually, a computer probably won't be able to either.

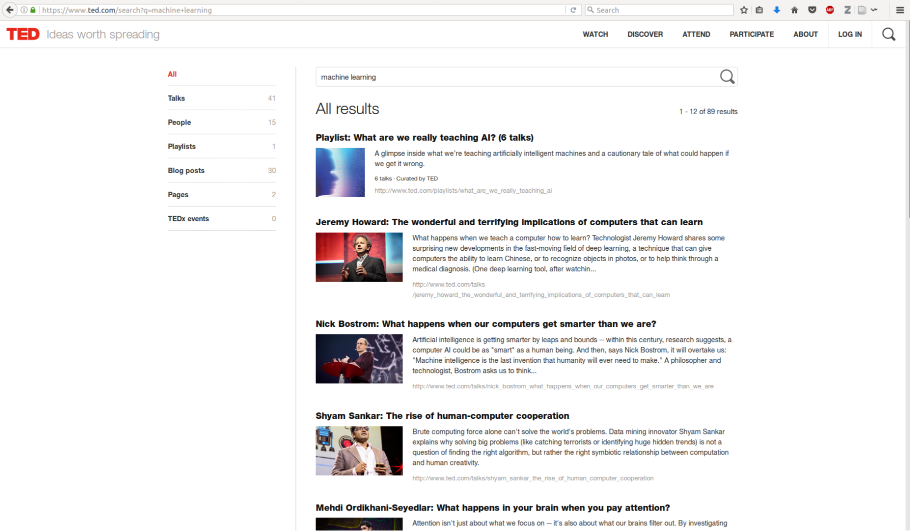
Focus on problems where a human could solve the problem, but where it would be great if a computer could solve it much more quickly.

Test data bases: MNIST and CIFAR-10





Other resources: TED



Final note

"I propose to consider the question 'Can machines think?'".

- Alan Turing, 1950