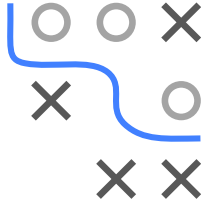
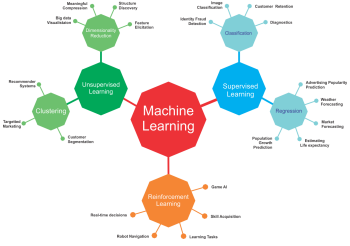
Introduction to Machine Learning



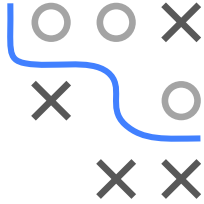
ML-Basics
What is Machine Learning?



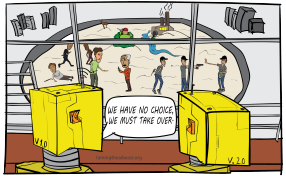
Learning goals

- Understand basic terminology of and connections between ML, AI, DL and statistics
- Know the main directions of ML: Supervised, Unsupervised and Reinforcement Learning

ML IS CHANGING OUR WORLD

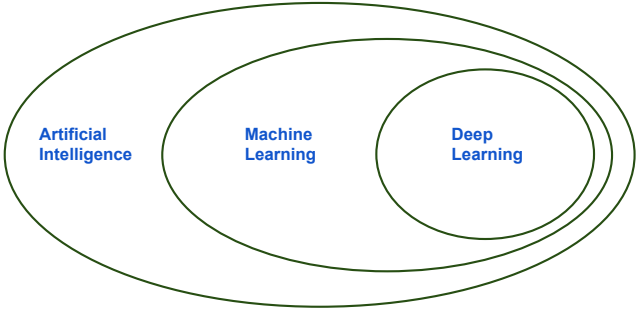


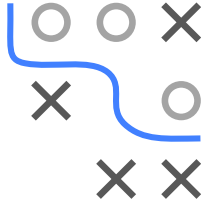
- Search engines learn your search preferences
- Recommender systems learn your taste in books, music, movies,...
- Algorithms do automatic stock trading
- Tools can accurately translate between many different languages
- DeepMind beats humans at Go
- Physicians are supported by personalized medicine
- LLMs revolutionize many fields (currently especially coding)
- Data-driven discoveries are made in physics, biology, genetics, astronomy, chemistry, neurology,...
- ...



AI End-Scenario: Necessary Rescue

AI, ML AND DL





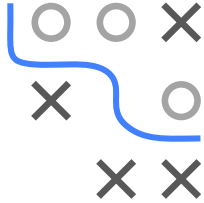
Many people are confused what these terms actually mean

And what does all this have to do with statistics?

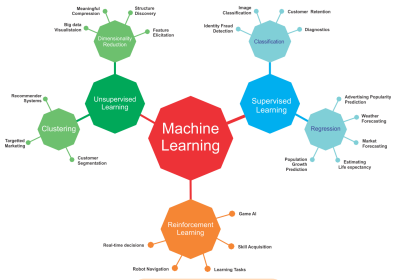
ARTIFICIAL INTELLIGENCE

- General term for very large and rapidly developing field
- No strict definition, but often used when machines perform tasks that could only be solved by humans or are very difficult and assumed to require “intelligence”.
- Started in the 1940s – when the computer was invented. Turing and von Neumann immediately asked: If we can formalize computation, can we use that to formalize “thinking”?
- Includes ML, NLP, computer vision, robotics, planning, search, intelligent agents, ...
- Sometimes misused as a “hype” term for ML or ... basic data analysis
- Or people refer to the fascinating developments in the area of foundation models

▸ Click for source

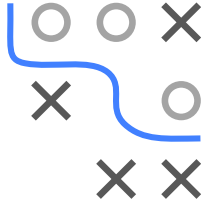


MACHINE LEARNING



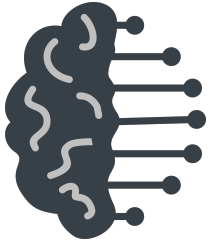
▸ Click for source

- Mathematically well-defined and solves reasonably narrow tasks
- Usually construct predictive models from data, instead of explicitly programming them
- “A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.”
  Tom Mitchell, Carnegie Mellon University, 1998

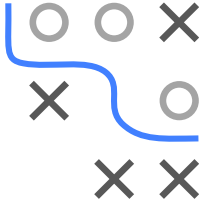


DEEP LEARNING

- Subfield of ML which studies neural networks
- Artificial neural networks are roughly inspired by the human brain, but we treat them as useful mathematical models
- Studied for decades (start in the 1940/50s). Uses more layers, might use specific neurons, e.g., for images, many computational improvements to train on large data.
- Can be used on tabular data but typical applications are images, texts or signals
- Last 15-20 years have produced remarkable results and imitations of human ability where the result looked intelligent

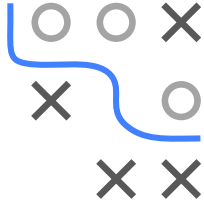


▸ Click for source

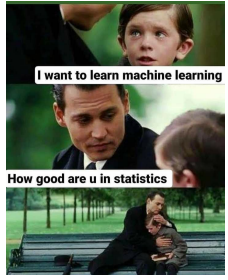


"Any sufficiently advanced technology is indistinguishable from magic." *Arthur C. Clarke's 3rd law*

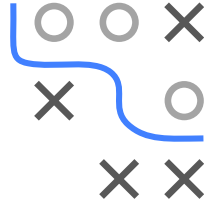
ML VS. STATS



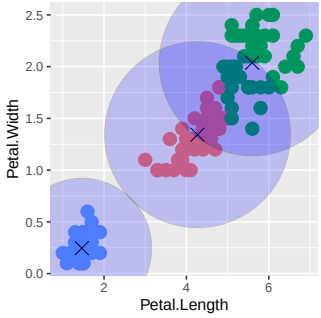
- Historically developed as different fields, but many methods and concepts are pretty much the same
- ML: Rather accurate predictions with more complex models
- Stats: More interpreting relationships and sound inference
- Now: Both basically work on same problems with same tools
- Communities are still divided
- Often different terminology for the same concepts
- Most parts of ML we could also call:
  Nonparametric statistics plus efficient numerical optimization
- Personal opinion: Nowadays few practical differences, seeing differences instead of commonalities mainly holds you back



UNSUPERVISED LEARNING

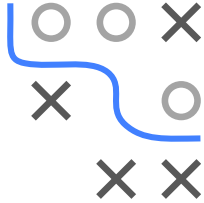


- Data without labels *y*
- Search for patterns within the inputs **x**
- *Unsupervised* as there is no “true” output we can optimize against

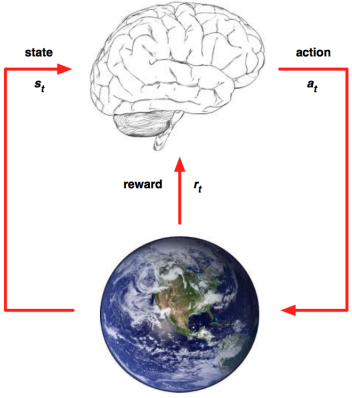


- Dimensionality reduction (PCA, autoencoders ...):
  compress information in $\mathcal{X}$
- Clustering: group similar observations
- Outlier detection, anomaly detection
- Association rules

REINFORCEMENT LEARNING

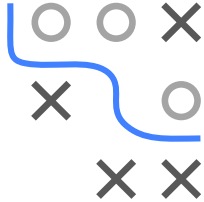


- General-purpose framework: At each time step *agent* interacts with *environment*: observes state, receives reward, selects action



- Goal: Select actions to maximize future reward
- Reward signals may be sparse, noisy and delayed

WHAT COMES NEXT



- **Supervised learning** for regression and classification: predict labels y through features $\mathbf{x}$ based on training data
- First we will go through fundamental concepts in supervised ML:
  - What kind of “data” do we learn from?
  - What is a “prediction model”?
  - How can we quantify “predictive performance”?
  - What is a “learning algorithm”
  - How can we operationalize learning?
- We will also introduce first concrete learning algorithms: Linear models, trees and forests
- More complex stuff comes later