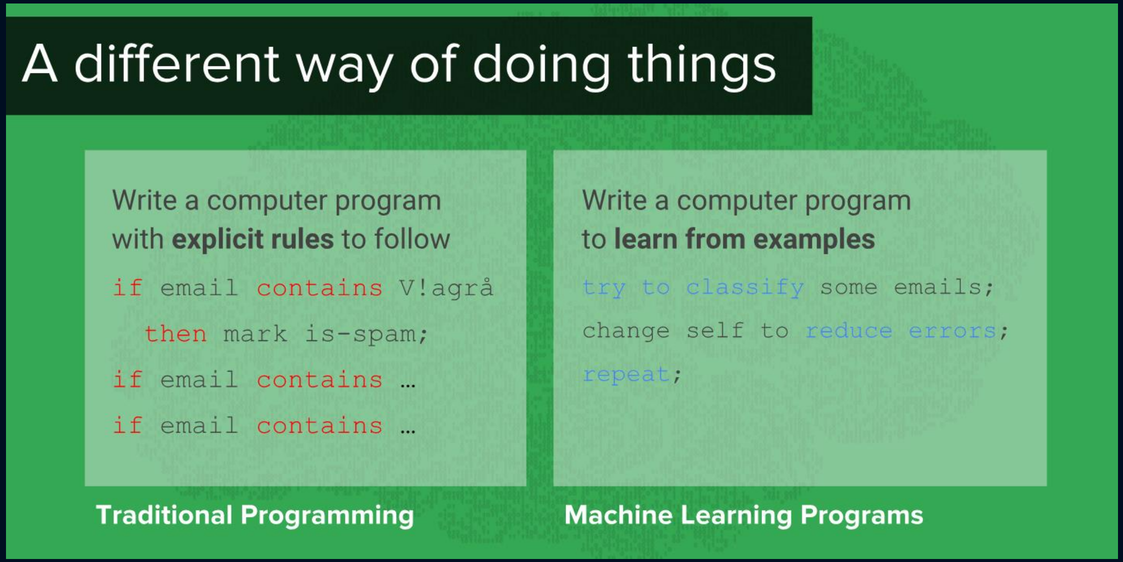
Session 7. Big Data, Artificial Intelligence, Machine Learning

NATALIA MILOVANTSEVA, PHD
NRU HSE, FACULTY OF WORLD ECONOMY AND INTERNATIONAL AFFAIRS
MARCH 16, 23 2019

Learning outcomes

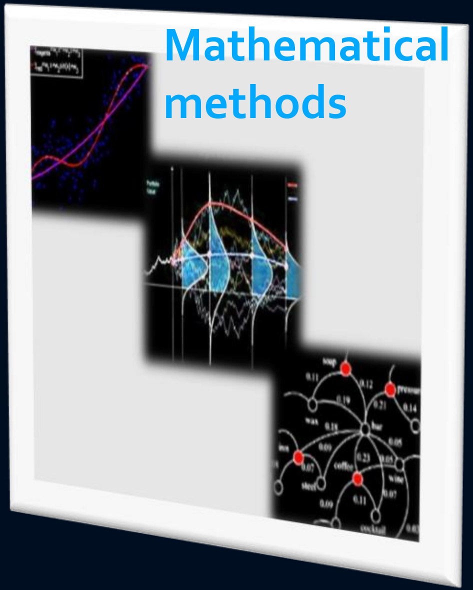
- LO1: Understand the basics of machine learning technique.
- LO2: Identify increasing areas or functions where machines are simply better performers than humans.
- LO3: Recognize the role of human guidance in digital transformation.
- LO4: Investigate how machine learning could be used in a business context.

Categories of AI

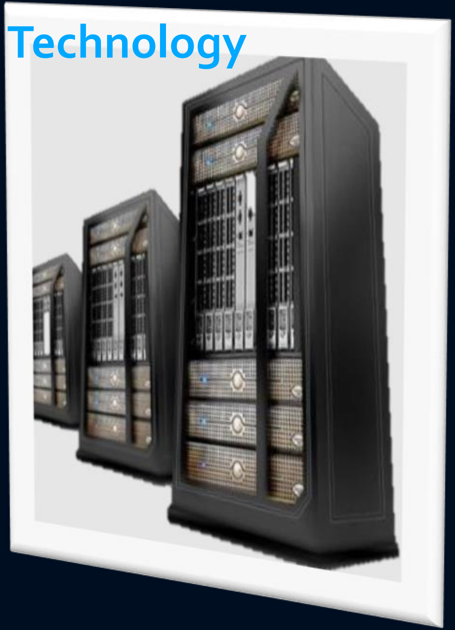


1) Symbolists - write rule-based systems and try to symbolically solve problems
2) Analogizers - make analogies from one to another
3) Evolutionists - game theory and games
4) Bayesians - statistical methods
5) Connectionists - mathematical functions used to show connections between different relationships

Machine learning concepts



+



- Artificial Intelligence
  - Coded rules are required

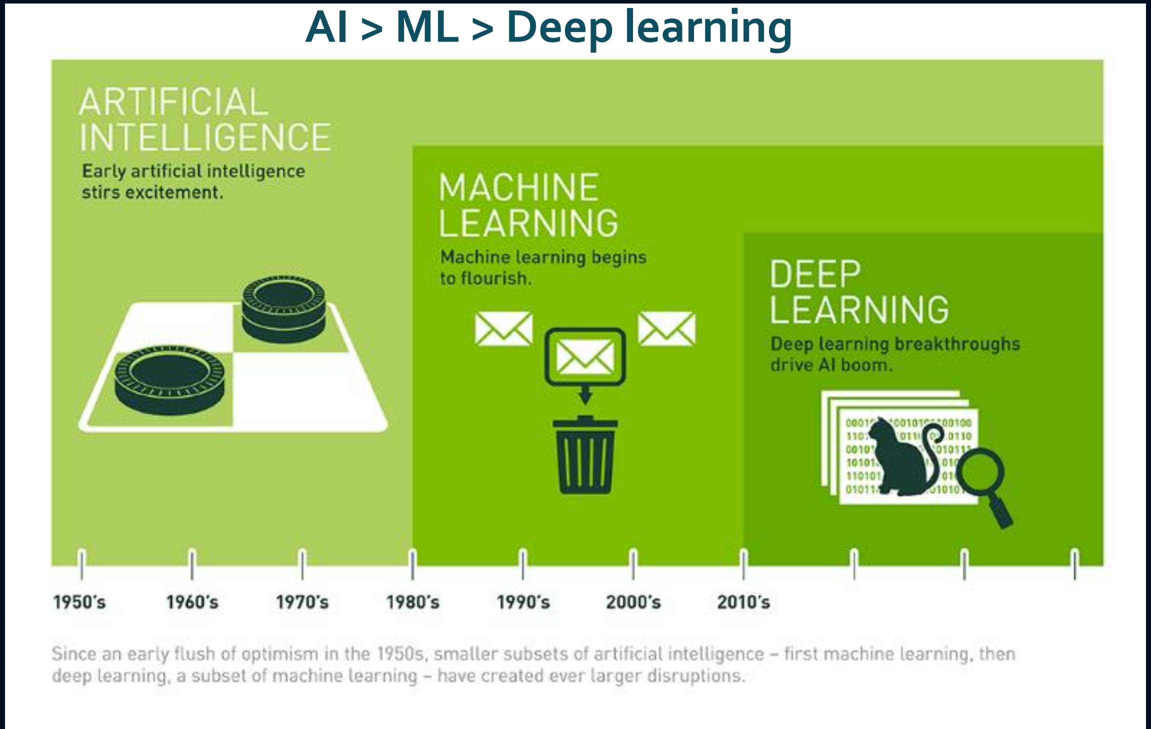
- Machine learning is an AI technique that enables a machine to learn using large data sets
  - ML allows computers to learn by themselves taking advantage of the processing power of modern computers
  - To solve a problem ML needs big data and computing power

ML explained by Yufeng Guo from Google (10:35)

- https://www.youtube.com/watch?v=HcqpanDadyQ&list=PLIivdWyY5sqJxnwJhe3etaK7utrBiPBQ2

Deep learning

- Deep learning is a technique of ML
- Deep learning uses neural nets (or networks) (NNs, or deep neural nets DNNs)
- Neurons' importance is dictated by weights
- Iterating through data set reduces AI's error

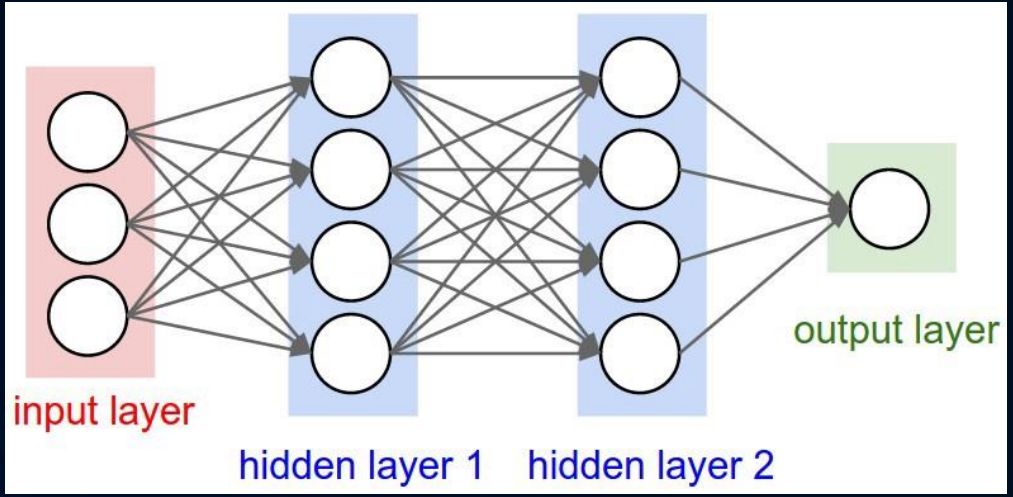


Neural nets

Business implication:
"*Deep neural nets*" (7 to 15 layers deep):

- the deeper the net, the more sophisticated the decisions it can make
- Examples:
  - decision regarding image recognition
  - decisions about lending credit

- Modeled after neural networks in the brain
- Based on matching inputs to outputs
- Example:
  - Input - files of human speech
  - Output - written text of the words that correspond to those sound waves



Regression:

$Y = f(X, \beta)$

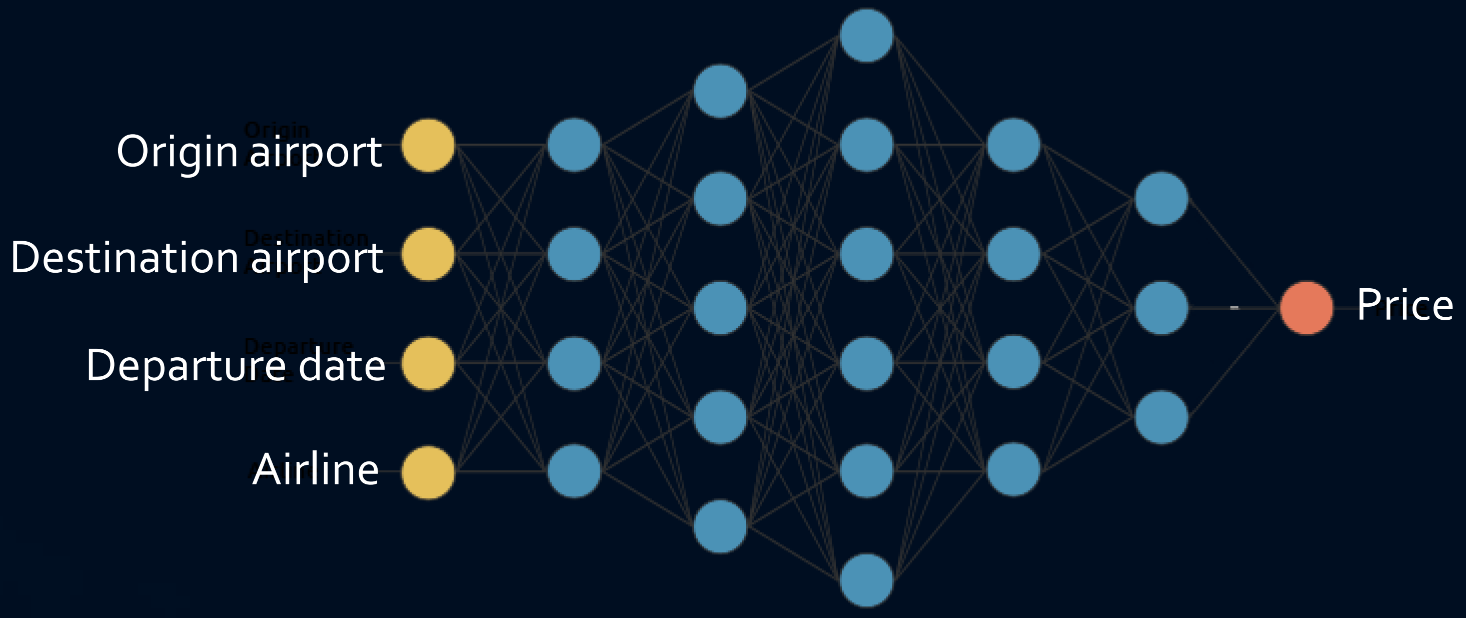
ML problem formulation

- **Data** - collection of objects
- **Objects** - are described by a set of observables and target variables
- **Observed variables** - can easily be measured for an arbitrary object
- **Target variables** - are known for a limited number of objects (the so-called training sample)
- **Task** - to predict the value of target object variables from the observed variables

Compute price prediction by deep learning

Simple task

- To build a service that will estimate price of airline tickets

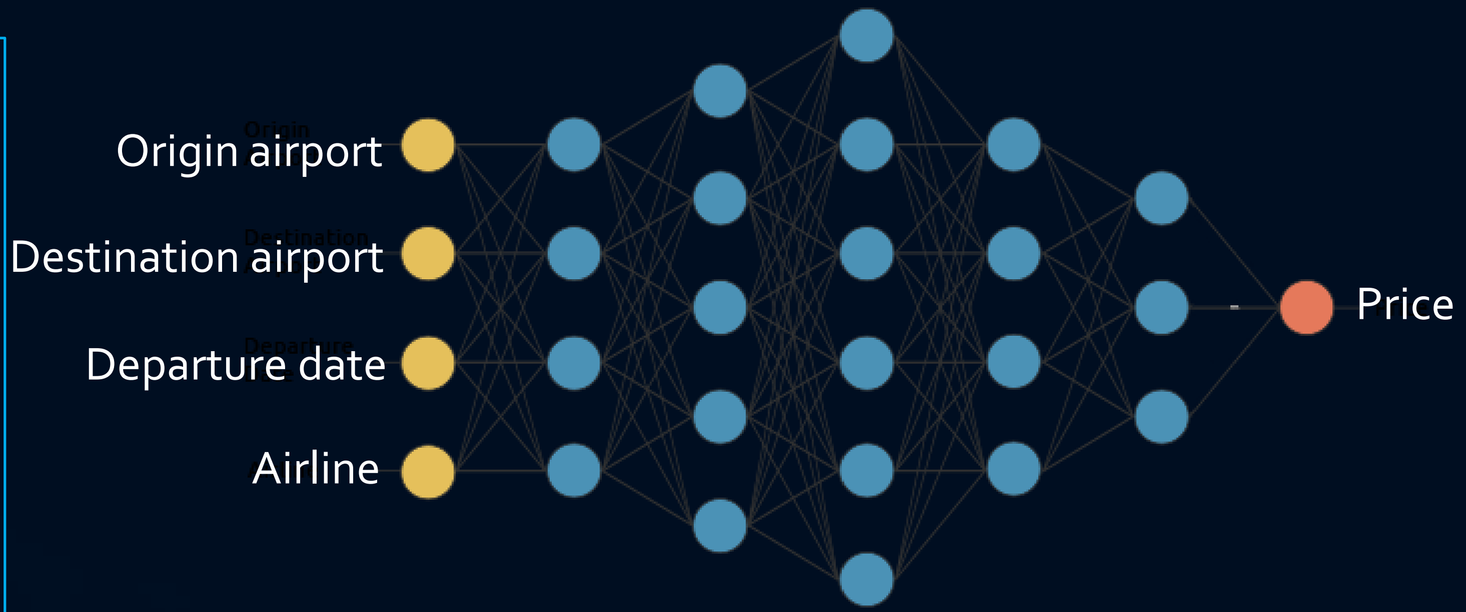


Input: origin airport, destination airport, departure date, airline

Output: predict the price of a one-way ticket

Training AI for price prediction by deep learning

- Our big data: historical data of ticket prices. Need a very large list of ticket prices because of the large amount of possible airports and departure date combinations
- Untrained AI goes through entire data set
- Compare its outputs with data set
- Create a cost function
- When cost function=0, AI is trained (AI's outputs=data set outputs)



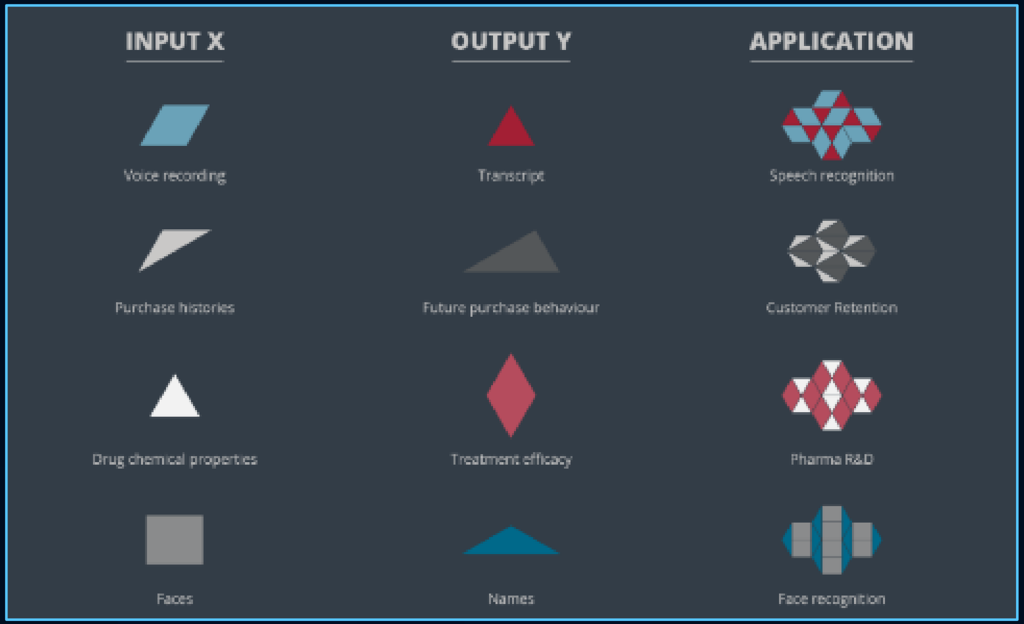
Task: to build a service that will estimate price of airline tickets
Input: origin airport, destination airport, departure date, airline
Output: predict the price of a one-way ticket

Supervised and unsupervised learning

SUPERVISED LEARNING

- Many samples of good, labeled data
- AI is given them as inputs and told of the expected outputs
- If AI gives wrong output, it readjusts its calculations iteratively until no mistakes are made
- Example: predicting weather. AI is trained on historical data. Inputs: pressure, humidity, wind speed. Outputs: temperature.

BUSINESS IMPLICATION



Supervised and unsupervised learning

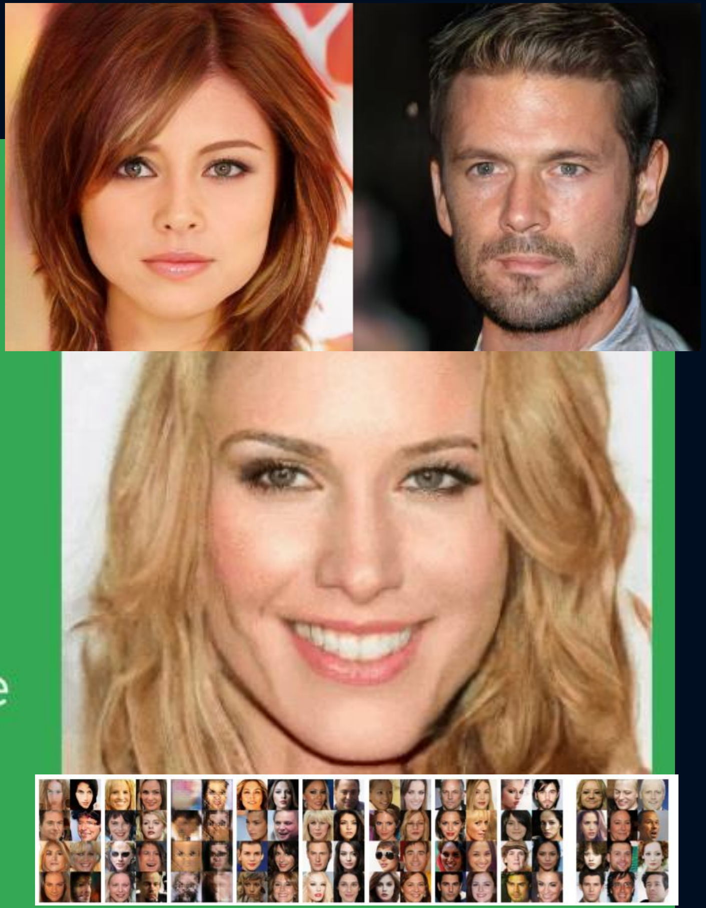
SUPERVISED LEARNING

- Labelled data sets
- AI is given them as inputs and told of the expected outputs
- If AI gives wrong output, it readjusts its calculations iteratively until no mistakes are made
- Example: predicting weather. AI is trained on historical data. Inputs: pressure, humidity, wind speed. Outputs: temperature.

UNSUPERVISED LEARNING

- Data sets have specified structure
- AI is allowed to make logical classifications of data
- Example. Predicting behavior for an e-commerce website:
  - **Instead of learning by using labelled data set of inputs and outputs, AI classifies the input data. Based on this classification, it will tell which kind of users are most likely to buy certain products.**

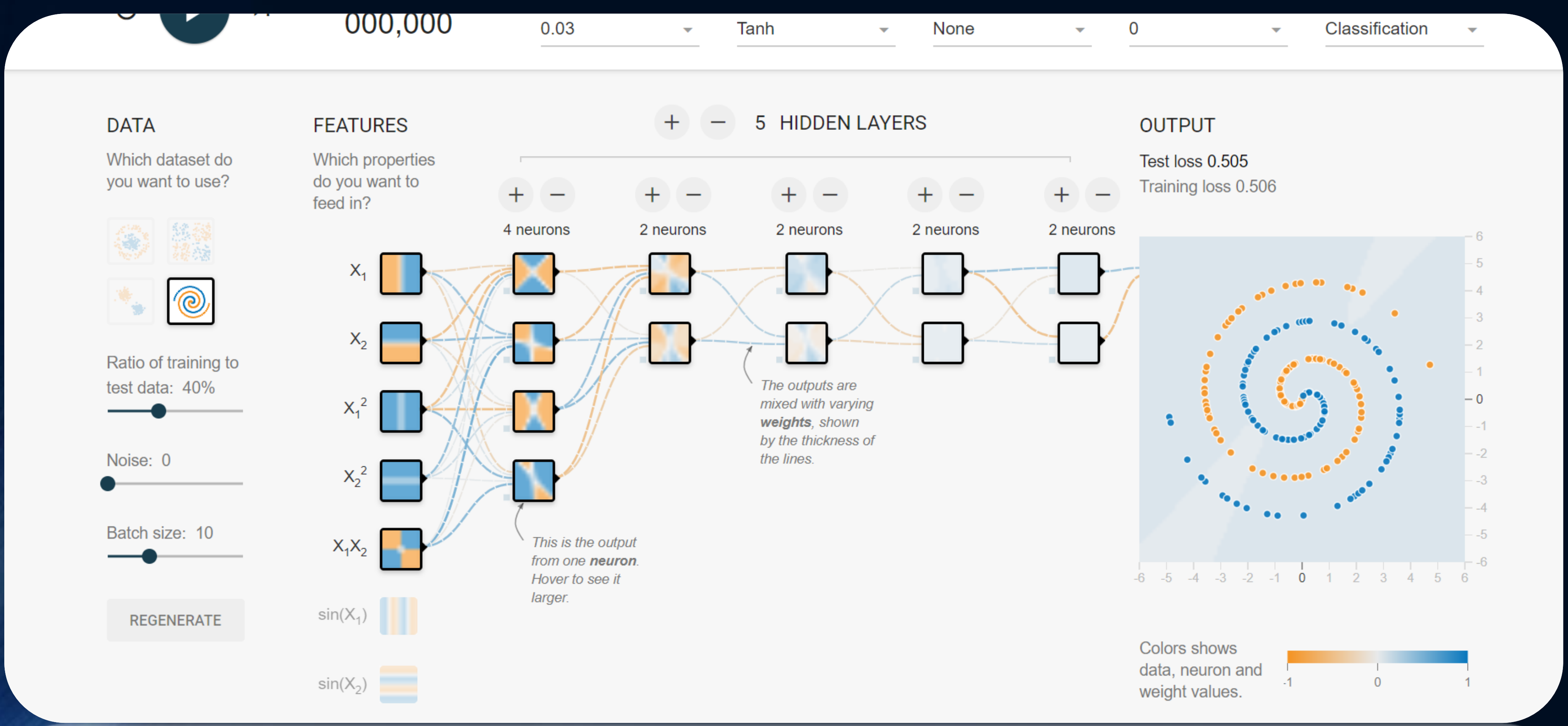
AI: Common Use Cases



- **Object recognition**
- **Speech recognition / sound detection**
- **Natural Language Processing** / Sentiment analysis
- **Creative** (e.g. Style Transfer - learning to draw an image in the style of an artist)
- **Prediction** - given some inputs, what is the expected output for unseen examples
- **Translation** between languages
- **Restoration / Transformation** - eg taking an image and using ML to figure out what should be there, or generating faces based on what it knows a face to be.

Experience a Neural Network

https://playground.tensorflow.org

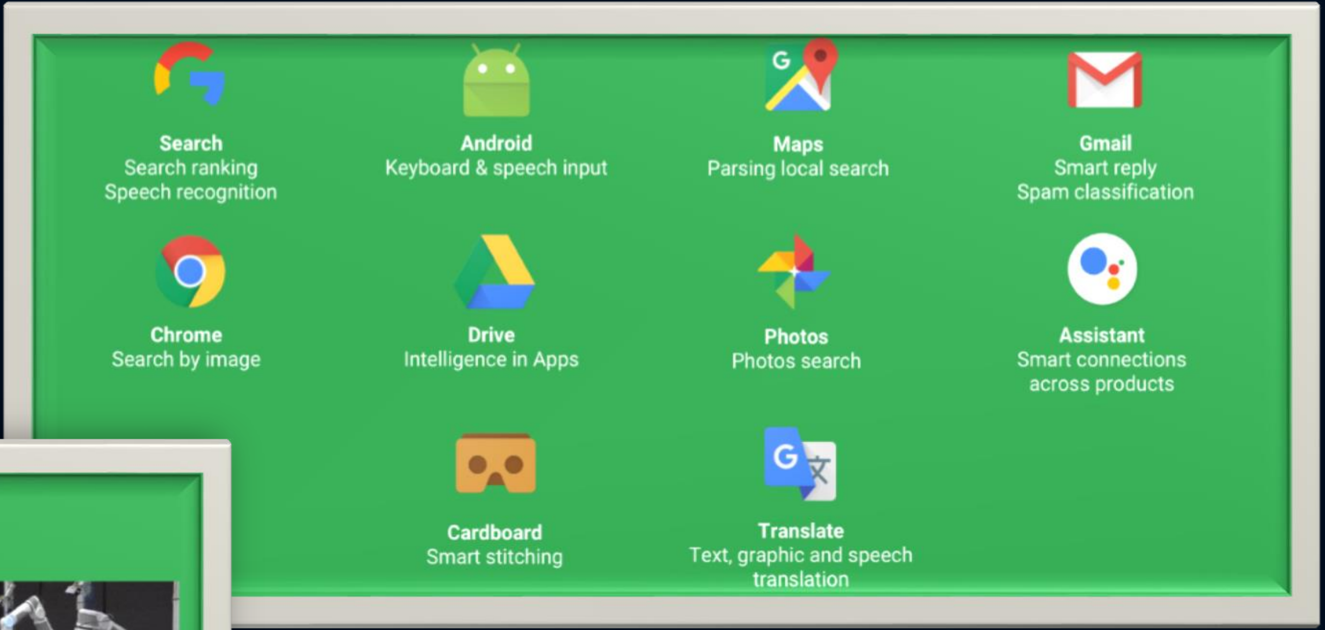


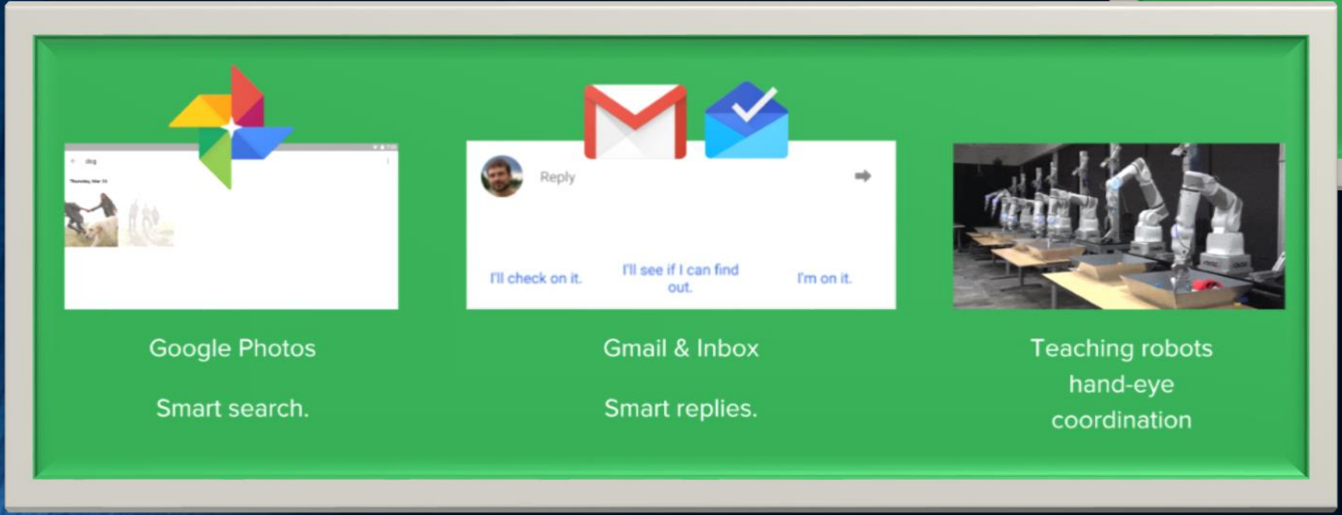
Limits to ML today (open to discussion!)

- Rapid advancements in computing underlies advancements in ML
- Computing power will level off due to physical limitations of computers
- ML as well be limited/level off soon (how soon?), unless the QC takes off

https://www.youtube.com/watch?v=JhHMJCUmq28

Examples of ML: Google and others





In next Session we'll review a case of Google's use of reinforcement learning to reduce the cost of cooling Google Data Centers.

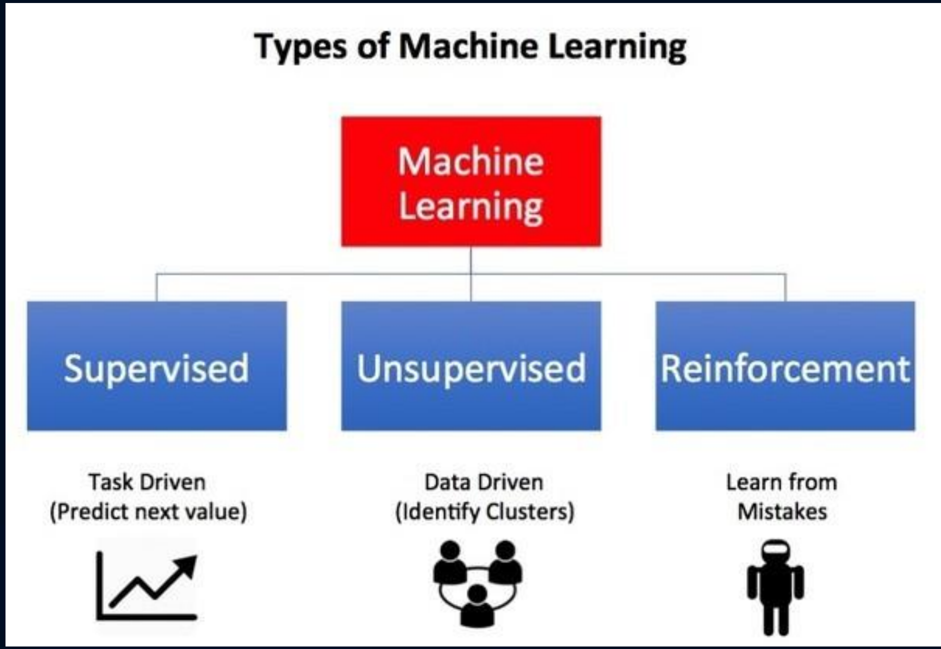
Experience autodraw.com

The role of human guidance. What's the big picture?

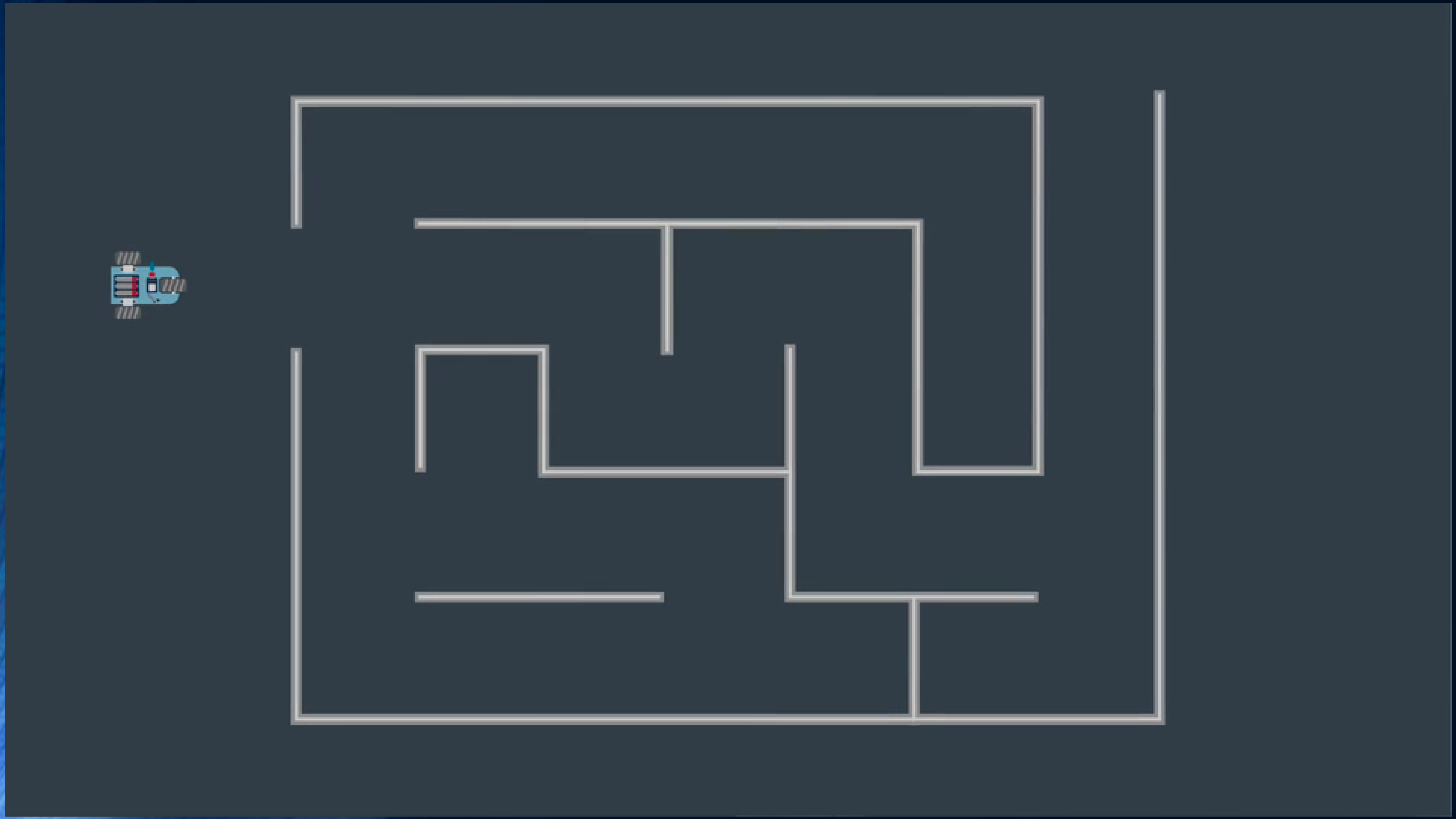
- Digitizing manual tasks (execution) is NOT TRANSFORMATIONAL
- People will be left to GUIDE STRATEGY, rather than execution
- For that, we need to understand the execution
  - We want machines to do data processing, analysis and pattern discovery to intelligent
  - We want them to produce insights and to act on them

ML concepts, cont.:

- A digital agent put in a physical or simulated environment and given a goal, such as “find a way out of the maze”
- A company doesn’t know the correct answer but has some way to score the machine’s decisions as better or worse
- Advantage: no need for a large set of data for a machine to learn



Reinforcement learning



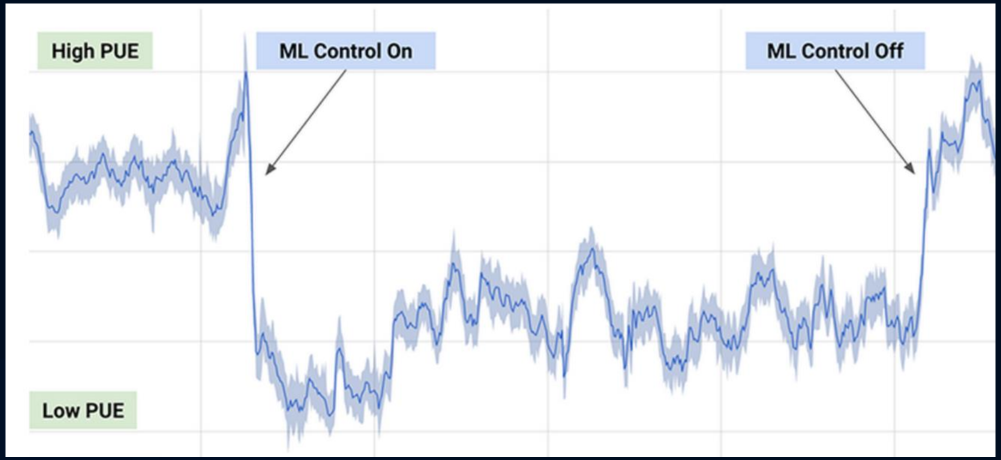
- Agent receives positive feedback for doing the right things
- Agent makes progress toward its ultimate goal

- Agent - robot
- Environment - maze
- Goal – to get out of the other end of the maze

Case study: Google's use of reinforcement learning

How reinforcement learning was used to reduce the cost of cooling Data Centers:

- Agent – algorithm
- Environment - data center treated like a video game
- Goal - to get a higher score (better energy efficiency)
- Gave the algorithms historical data on centers' fluctuating computing loads, sensor readings, and environmental variations



Results:

- total energy consumed for cooling fell by 40%
- facilities' overhead improved by about 15%.

Factors of AI progress in recent years

1. Moore's law
2. Big data
3. Working with existing technology (tinkering)
4. Improvements in computer architecture

Examples of AI application

- AI-based systems spread quickly after surpassing human performance at a given task
- Using improved vision systems to automate much of the work of security guards:
  - Aptonomy - makers of drones and robots
  - Sanbot – makers of robots
  - Affectiva – recognition of emotions such as joy, surprise, and anger
  - Enlitic - deep-learning start-up, scans medical images to help diagnose cancer

Using Natural Language Processing for legal research, etc.

References

- Contact instructor for references used in this presentation