COE206 – Principles of Artificial Intelligence

Mustafa MISIR

Istinye University, Department of Computer Engineering

mustafa.misir@istinye.edu.tr

http://mustafamisir.github.io
http://memoryrlab.github.io

L8: Learning Problem[1]

[1] the content is adapted from the RPI - CSCI 4100/6100 course slides on the Learning from Data book's first chapter, made by Malik Magdon-Ismail – https://www.cs.rpi.edu/~magdon/

Let's Define a Tree

Let's Define a Tree

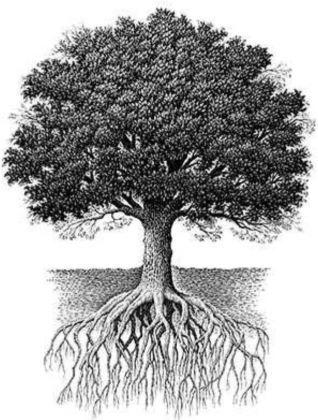
If you show a picture to a 3 year old kid and ask if there is a tree in it, you will probably get the right answer.

Let's Define a Tree

If you ask someone who is 30 years old the definition of the tree, you will probably get an inadequate answer

- We learn trees not from mathematical definitions, but by looking at them - that is, learn from data

Defining is Hard; Recognizing is Easy



Learning to Rate Movies

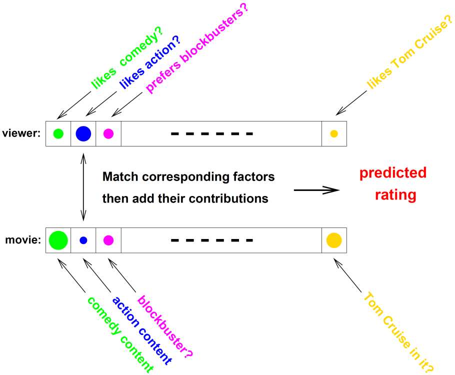
- Can we predict how a viewer would rate a movie?
- Why? So that Netflix can make better movie recommendations, and get more rentals
- Netflix Prize[2]:1 million prize for a mere 10% improvement in their recommendation system

NETFLIX

Rank	Team Name	Best Test Score	% Improvement	Best Submit Time
Grand Prize - RMSE = 0.8567 - Winning Team: BellKor's Pragmatic Chaos				
1	BellKor's Pragmatic Chaos	0.8567	10.06	2009-07-26 18:18:28
2	The Ensemble	0.8567	10.06	2009-07-26 18:38:22
3	Grand Prize Team	0.8582	9.90	2009-07-10 21:24:40
4	Opera Solutions and Vandelay United	0.8588	9.84	2009-07-10 01:12:31

[2] https://www.netflixprize.com/ – https://www.wired.com/2009/09/how-the-netflix-prize-was-won/ – https://www.thrillist.com/entertainment/nation/the-netflix-prize

Learning to Rate Movies



- Viewer taste and movie content imply viewer rating
- No magical formula to predict viewer rating
- Netflix has data. We can learn to identify movie categories as well as viewer preferences

A pattern exists even though we don't know it yet. We have data to learn it.

Learning to Rate Movies

- **Movies**: Let's define each movie by a series of different factors; e.g. how much comedy is in it, etc.
- **Viewers**: Let's define each viewer by a series of different factors; e.g. how much he likes comedy, prefers simple or complex graphics, how important the appearance of the lead actor, etc.

Based on **factors** that can explain the content of a movie and the taste of the viewer, the viewer's movie score can be estimated

Credit Approval

Will the loan application of a bank customer be approved?

Age	32
Gender	Male
Salary	40.000 TL / year
Debt	26.000 TL
Work Duration	1 Year

Credit Approval – Learning

- Related factors: salary, debt, years in residence etc. – input
  $\mathbf{x} \in \mathbb{R}^d = \mathcal{X}$
- Approve credit or not – output
  $y \in \{-1, +1\} = \mathcal{Y}$
- True relationship between x and y – target function
  $f : \mathcal{X} \to \mathcal{Y}$
- Data on customers – dataset
  $\mathcal{D} = (\mathbf{x}_1, y_1), ..., (\mathbf{x}_N, y_N)$ buradan $y_n = f(\mathbf{x}_n)$

$\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{D}$ are given, yet the target function f is unknown

Learn f from the data $\mathcal{D}$

Credit Approval – Learning

Start with a set of candidate **hypotheses** $\mathcal{H}$ which you think are likely to represent f

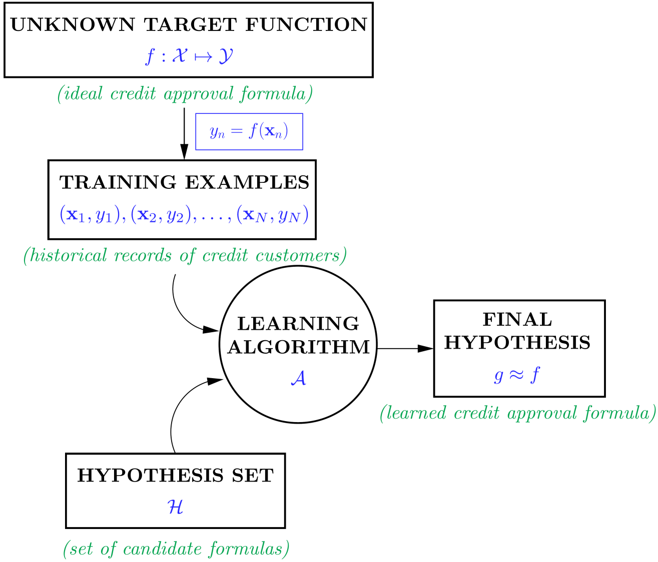
- Hypotheses set or Model : $\mathcal{H} = \{h_1, h_2, ..., h_m\}$

Select a **hypothesis** g from $\mathcal{H}$. The way we do this is called a **learning algorithm**.

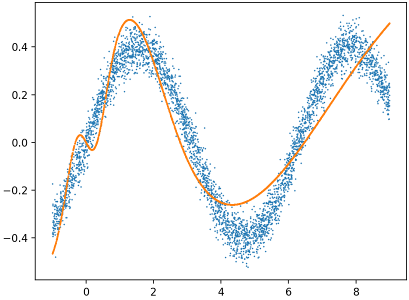
Use hypothesis g for new customers.

- The success of **hypothesis** g can be measured by the similarity between the selected **hypothesis** g and the actual function f: $g \approx f$

Summary of the Learning Setup



Learning – Function Approximation[3]



Learning can be seen as function approximation

[3] image source: https://stats.stackexchange.com/questions/320289/approximate-the-sine-function-with-shallow-neural-network

Learning – Compression[4]

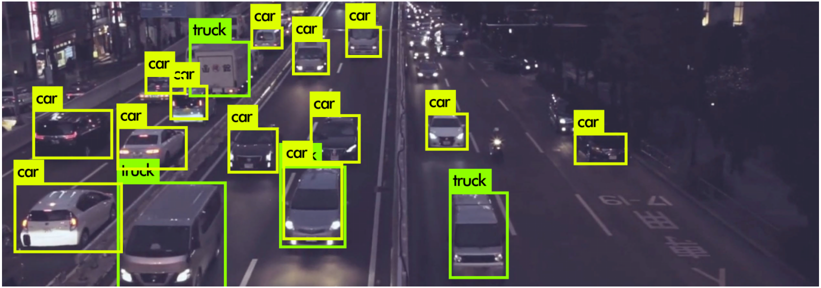
Learning performs compression by specifying a rule over the data, while getting a simpler explanation from the data, requiring less memory and computational resources

- e.g. when we learn the rules of addition, it is not necessary to remember the sum of every possible pair of numbers

[4] image source: https://www.vectorstock.com/royalty-free-vector/data-compression-vector-19122120

Learning – Labeling[5]

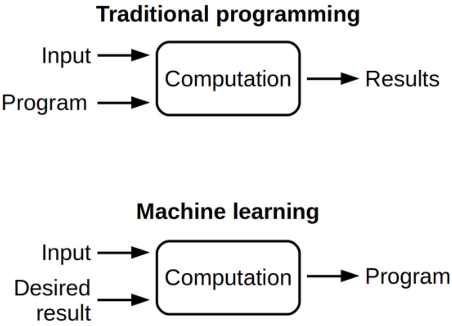
Learning can help top label new, unlabeled data



[5] image source: https://www.move-lab.com/blog/tracking-things-in-object-detection-videos

Machine Learning (ML)[7]

Learning without being explicity programmed[6]



[6] image source: https://www.futurice.com/blog/differences-between-machine-learning-and-software-engineering/

[7] https://en.wikipedi0.org/wiki/Machine_learning

Why Machine Learning?[8]

Tasks That Are Too Complex to Program: tasks that can be performed by humans or animals

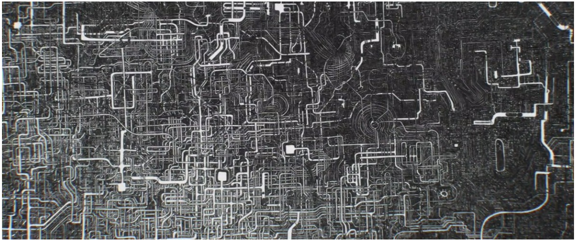
- As the tasks routinely performed by humans such as driving, speech recognition and image understanding, cannot be precisely described in detail, Machine Learning can be benefited to learn from their experiences (data)

[8] Chapter 1 – Understanding Machine Learning: From Theory to Algorithms, Shai Shalev-Shwartz and Shai Ben-David (1st Edition), 2014, Cambridge University Press – image source: https://www.ugent.be/ea/idlab/en/research/semantic-intelligence/speech-and-audio-processing.htm

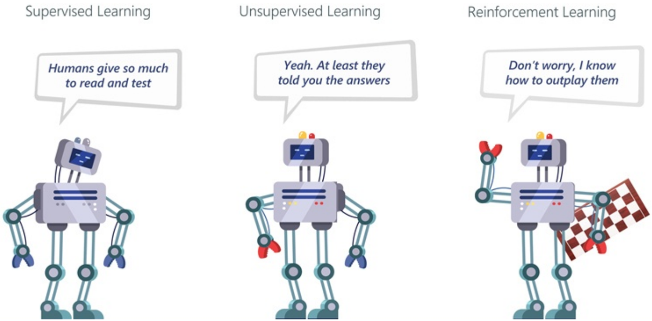
Why Machine Learning?

Tasks That Are Too Complex to Program: Tasks beyond human capabilities

- If the data is too much and extremely complex such that humans cannot make sense of them in practice, the increasing speeds and memory capabilities of the computing devices can be benefited with Machine Learning



Types of Learning[9]

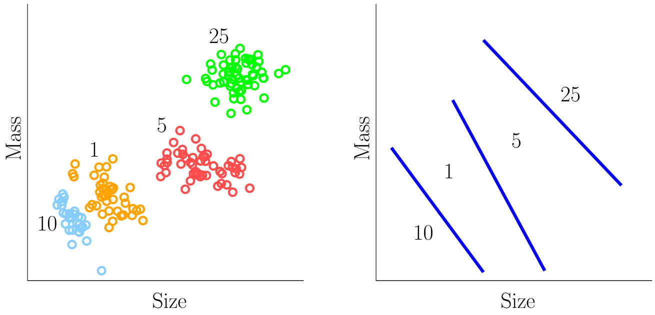


[9] image source: https://www.immuniweb.com/blog/machine-learning-AI-deep-learning-terms.html

Types of Learning: Supervised Learning (SL)

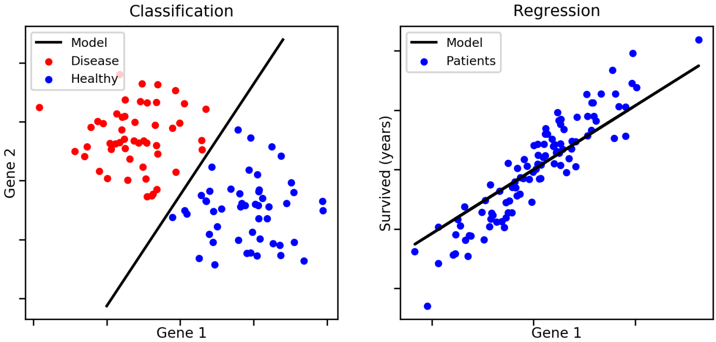
SL is concerned with the labeled datasets that can be used for training such that a resulting model or function can be applied to determine the label of any given, unseen data, e.g. coin classification

- The goal is to have a model that can generalize well on the unseen datasets



Types of Learning: Supervised Learning (SL)[10]

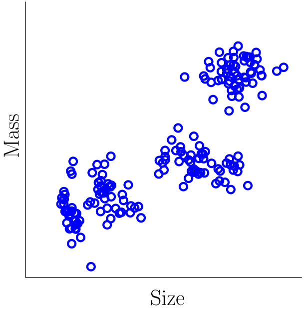
- **Classification** deals with predicting categorical labels
- **Regression** is a problem of predicting a real-valued label (target) given an unlabeled example

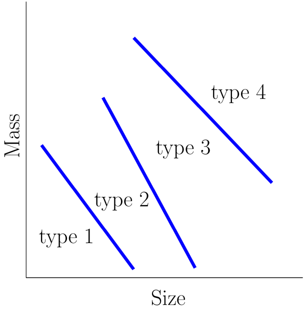


[10] image source: https://github.com/ramrathi/IECSE-ML-Winter18/wiki/Basic-Logistic-Regression

Types of Learning: Unsupervised Learning (UL)

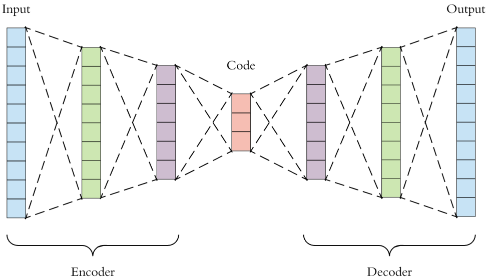
A type of learning aim to **infer** using unlabelled datasets, e.g. coin grouping





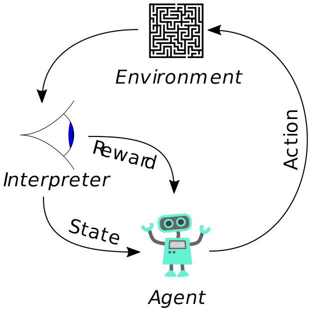
Types of Learning: Unsupervised Learning (UL)[11]

- **Clustering**: differentiate data samples w.r.t. their attributes / features
- **Dimensionality Reduction**: reducing the data sample size by decreasing the number of attributes (or determining more inclusive features) without undermining the overall structure of the data



[11] image source: https://ainews.spxbot.com/category/dimensionality-reduction/

Types of Learning: Reinforcement Learning[12]



A decision-making agent aims at solving a given problem, in a certain environment, by performig successive actions under a reward / penalty scheme

After a series of trial-and-error, a **policy** is expected to be learn, deciding on which action to apply in which order (relying on the present state) for maximizing the total (long-term) reward

[12] cart pole balancing : https://www.youtube.com/watch?v=Lt-KLtkDlh8 – image source: https://www.datahubbs.com/reinforcement-learning/