Introduction to Machine Learning

Shahane Arushanyan
Machine Learning. Summer 2020

My Background (in Data Science and ML)

- Ucom LLC (2017-2019)
- Philip Morris Armenia (2020-present)

Teaching:

- AUA Machine Learning, Algorithms (2017-2018)
- ISTC (2018)
- ACA (2019-present)

Course Outline

- Intro to Python (optional)
- Numpy, Pandas, Matplotlib
- Regression Algorithms (Linear r., Polynomial r., Ridge r.)
- Classification algorithms (Logistic r., SVM, Decision tree, Random forest)
- Introduction to unsupervised learning
- Introduction to Neural networks

Course Structure

- Lectures
- Practical sessions
- Homeworks (practical + homework)
- 3 Midterm exams (at the end of each month)
- Slack

Lecture Outline

- What is Machine Learning?
- Set ups
- Github

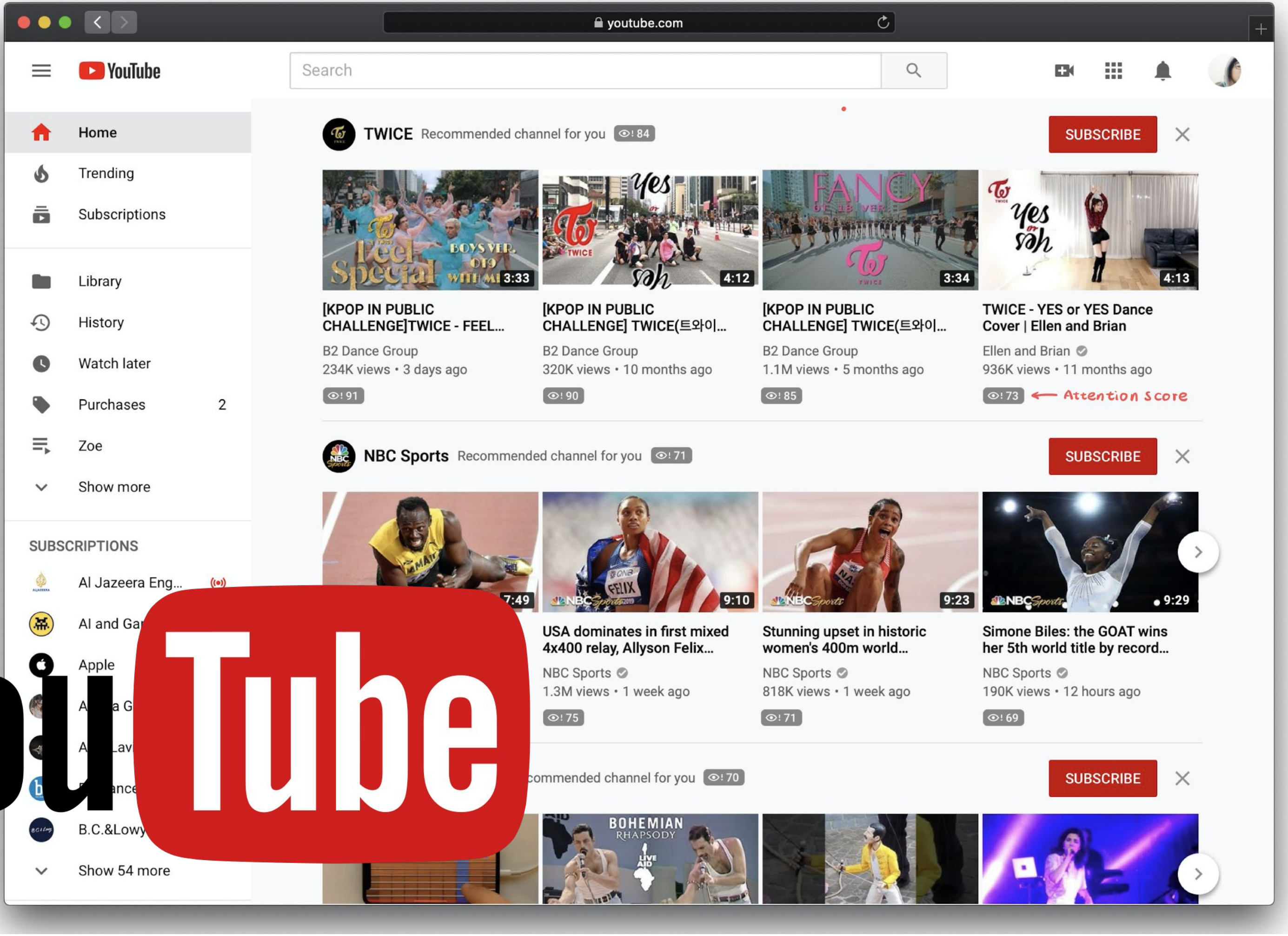
Data Science vs Machine Learning

DS -> using scientific approach to extract meaning and insights from data

ML -> a group of techniques that allow computers to learn from data

What is Machine Learning?

- Data is everywhere and in every field
- Huge amounts of data collected and stored
- Machine learning techniques in our everyday lives


youtube.com
YouTube
Search
Home
Trending
Subscriptions
Library
History
Watch later
Purchases 2
Zoe
Show more
SUBSCRIPTIONS
Al Jazeera Eng...
Al and Ga
Apple
B.C.&Lowy
Show 54 more
TWICE Recommended channel for you 84
SUBSCRIBE
[KPOP IN PUBLIC CHALLENGE]TWICE - FEEL...
B2 Dance Group
234K views • 3 days ago
91
[KPOP IN PUBLIC CHALLENGE] TWICE(트와이...
B2 Dance Group
320K views • 10 months ago
90
[KPOP IN PUBLIC CHALLENGE] TWICE(트와이...
B2 Dance Group
1.1M views • 5 months ago
85
TWICE - YES or YES Dance Cover | Ellen and Brian
Ellen and Brian
936K views • 11 months ago
73
NBC Sports Recommended channel for you 71
SUBSCRIBE
7:49
USA dominates in first mixed 4x400 relay, Allyson Felix...
NBC Sports
1.3M views • 1 week ago
75
Stunning upset in historic women's 400m world...
NBC Sports
818K views • 1 week ago
71
Simone Biles: the GOAT wins her 5th world title by record...
NBC Sports
190K views • 12 hours ago
69
commended channel for you 70
SUBSCRIBE

← Attention Score

YouTube

Face recognition

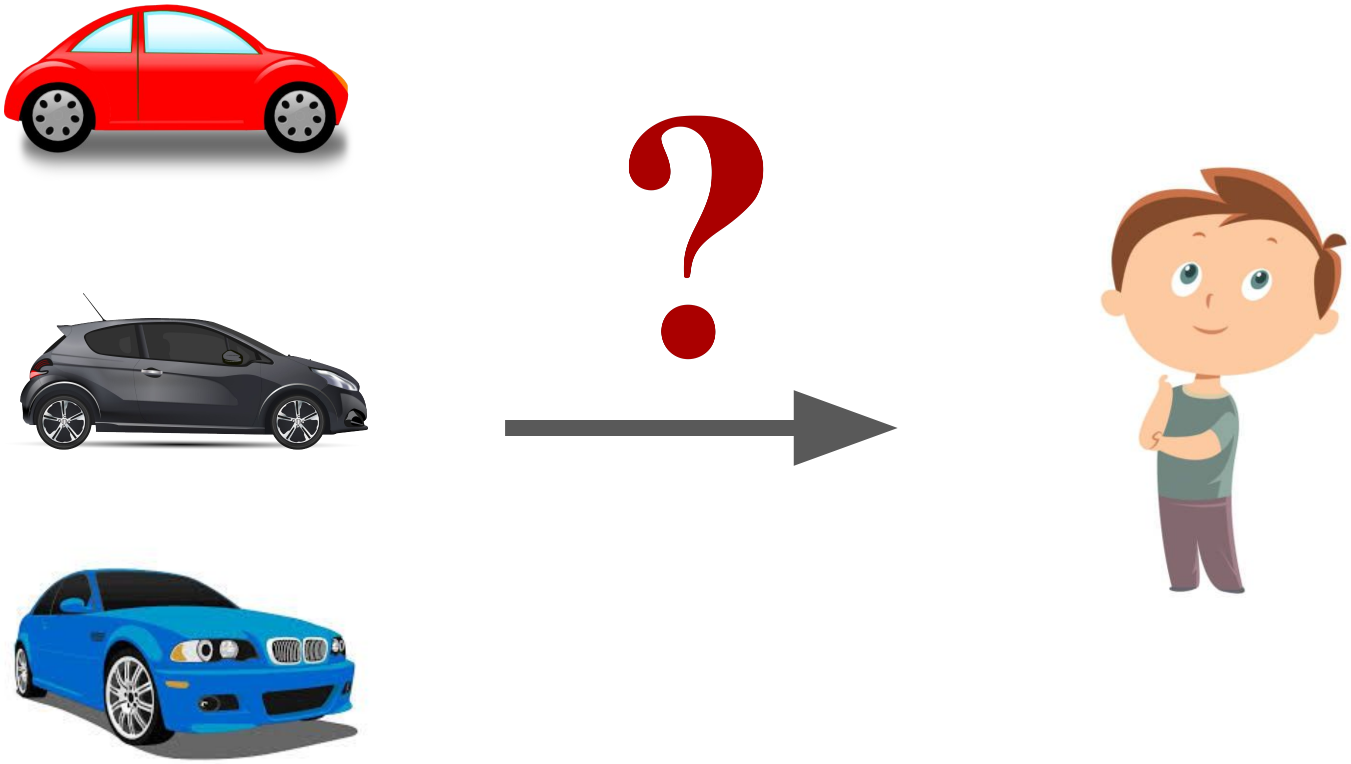
Voice assistant

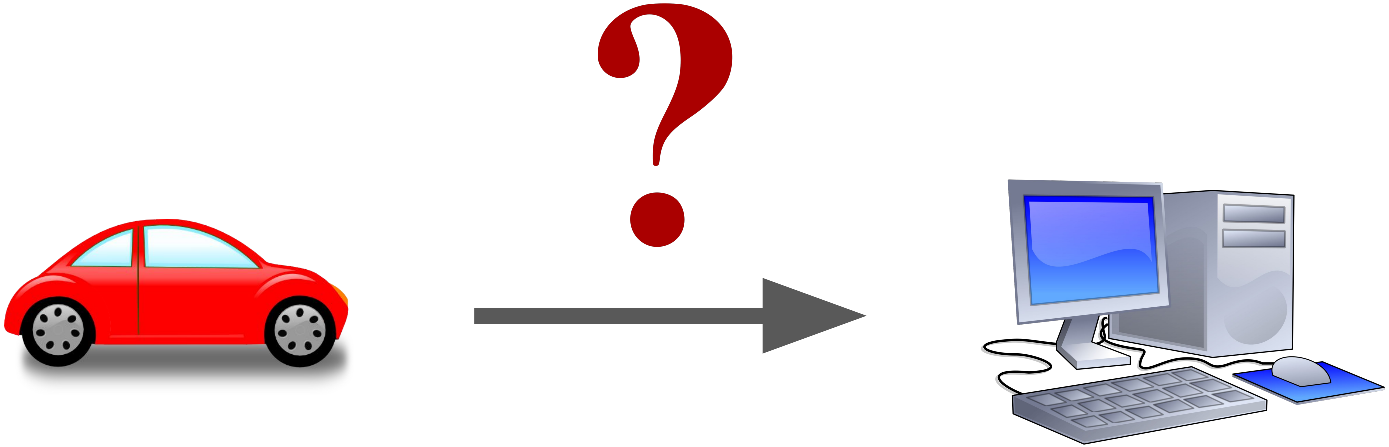
Vacuum robot

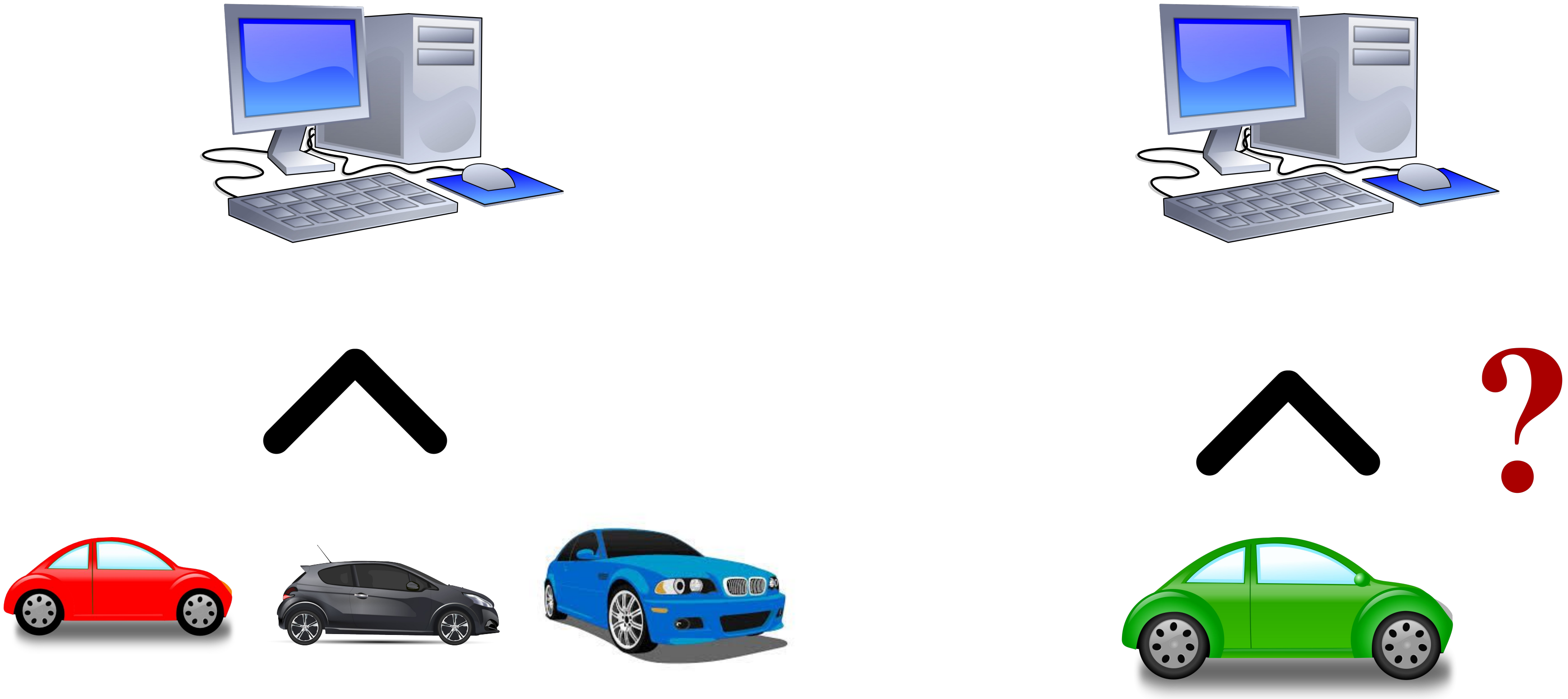
Movies suggestion

What is Machine Learning?

Machine learning is a study of computer algorithms that improve automatically through experience.







Training Data

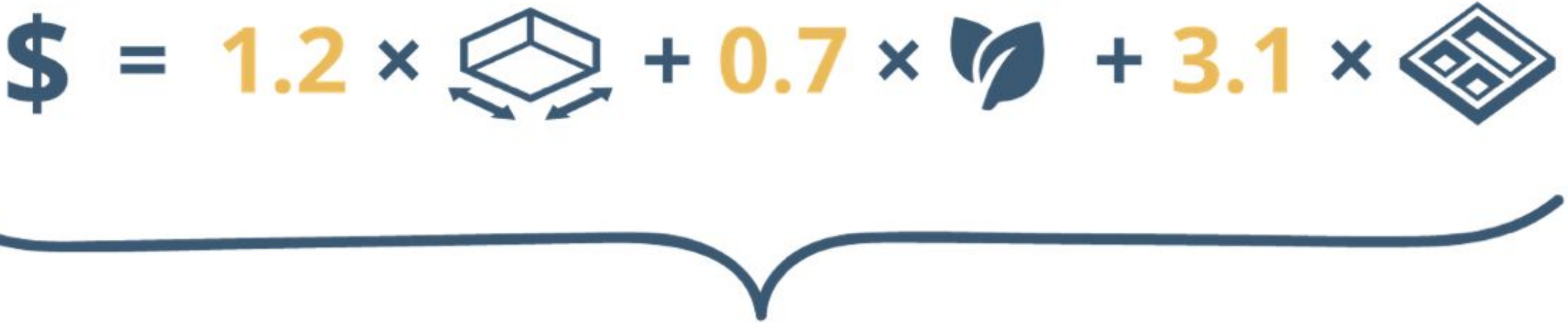
Test Data

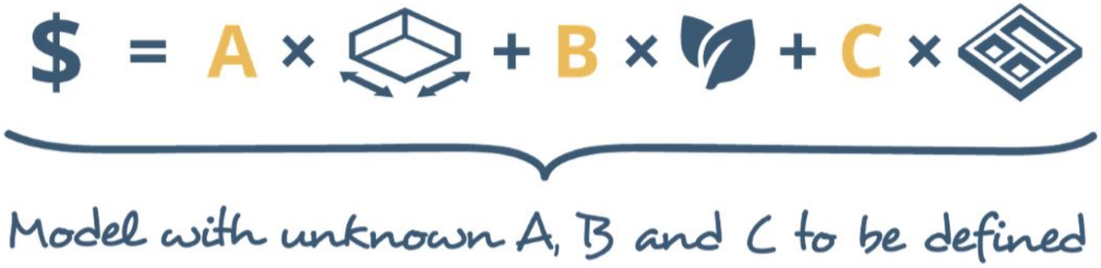
$

Size of house

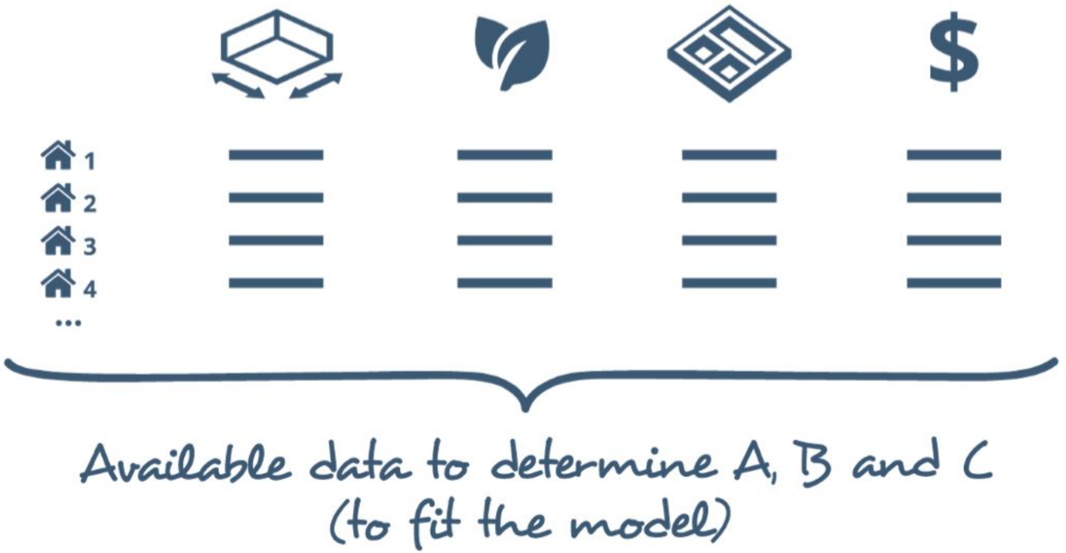
Size of garden

Number of rooms

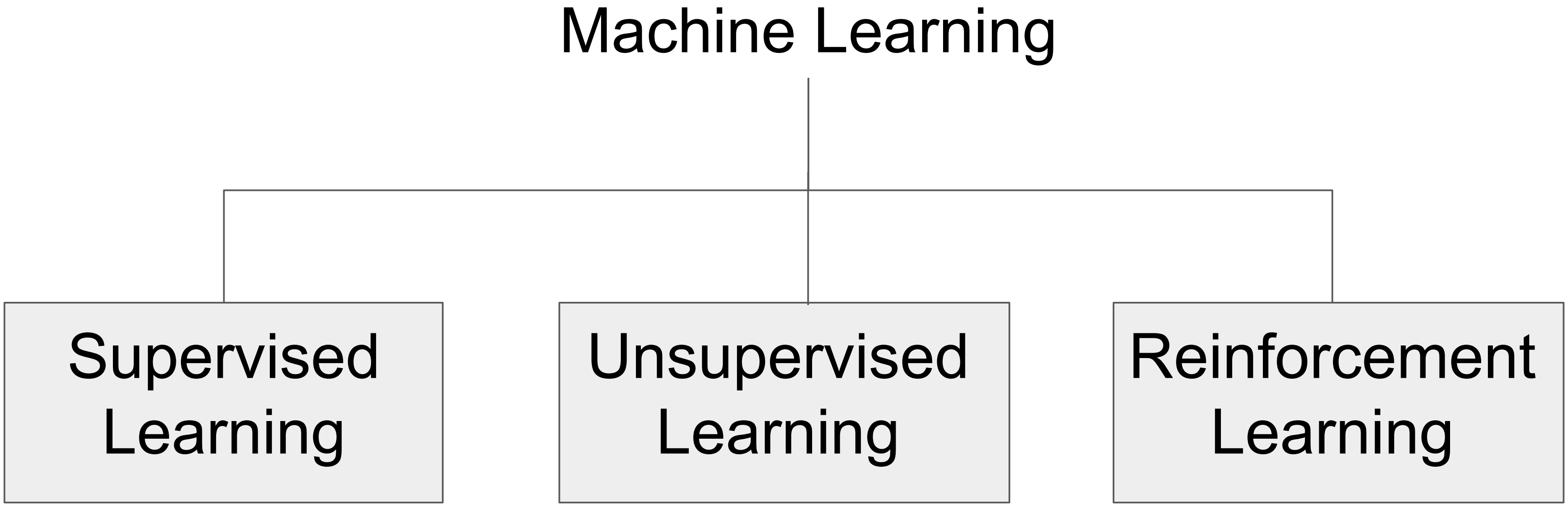



$ = A × + B × + C ×
Model with unknown A, B and C to be defined

< Model


$
1
2
3
4
...
Available data to determine A, B and C
(to fit the model)

< Data



Machine Learning

Supervised Learning

Unsupervised Learning

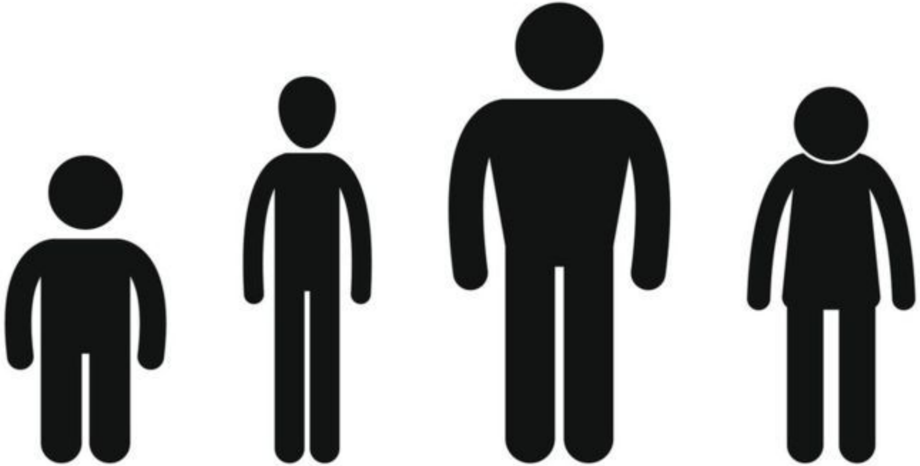
Reinforcement Learning

Supervised Learning

A supervised model is trained on a labeled dataset of (feature, label) pairs.

Regression Model - numerical label

Problem: Predict weight (number) given height and age

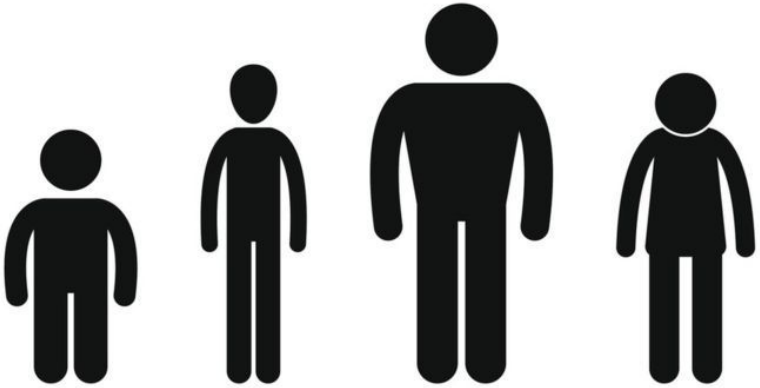


Features:

Height, Age

Label:

Weight



Height:	1.50	1.70	2.10	1.55	1.62
Age:	10	24	40	20	30
Weight:	40	58	80	45	**?**

Training data

Test data

Classification Model - categorical label

Problem: Predict if the object is an apple or not (True/False) given color and shape.

Features:

Color, Shape

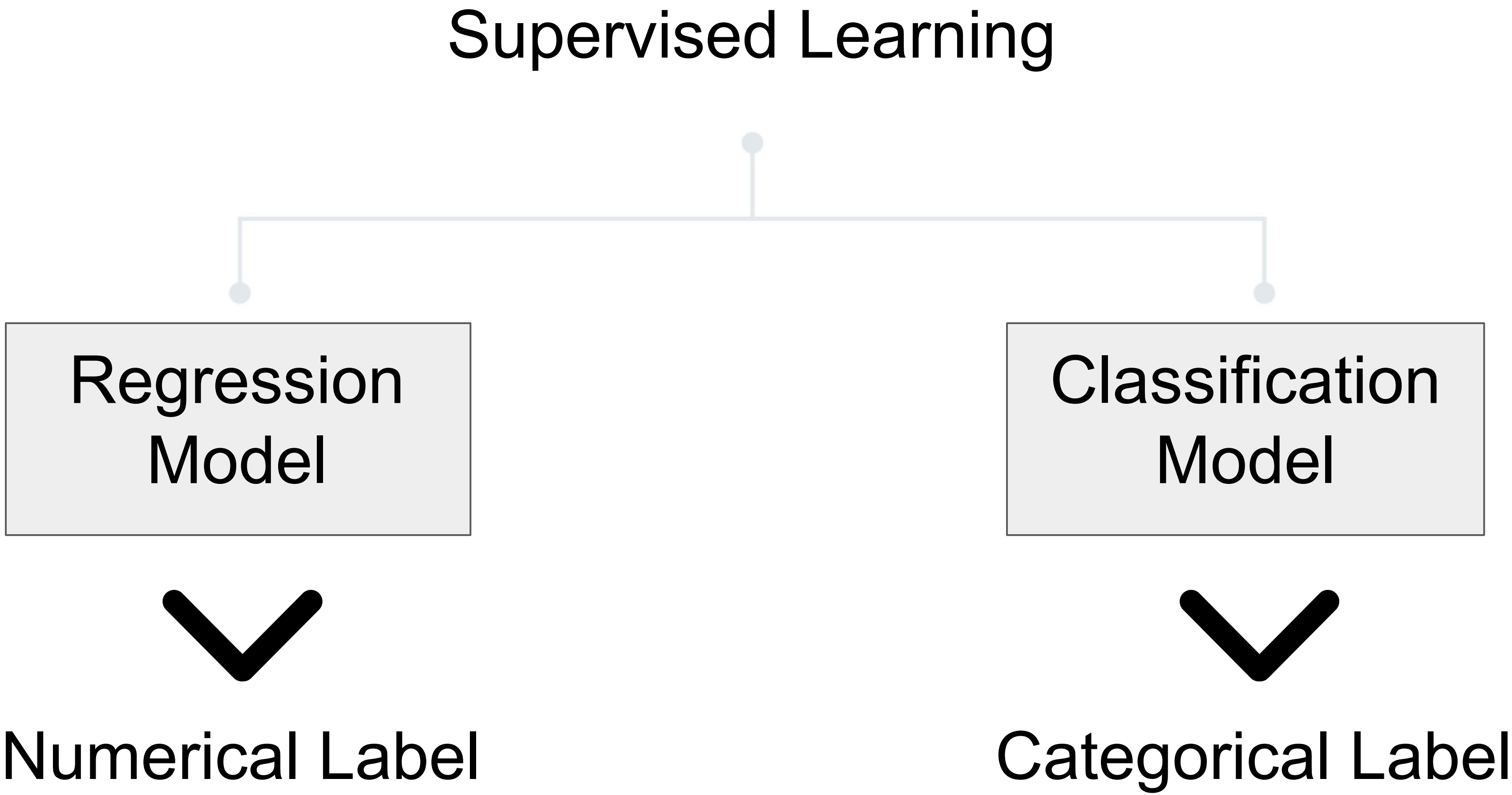
Label:

True/False

Color:	red	yellow	green	yellow	Green
Shape:	round	round	oval	oval	round
Apple:	True	True	False	False	**?**

Training data

Test data



Supervised Learning

Regression Model

Numerical Label

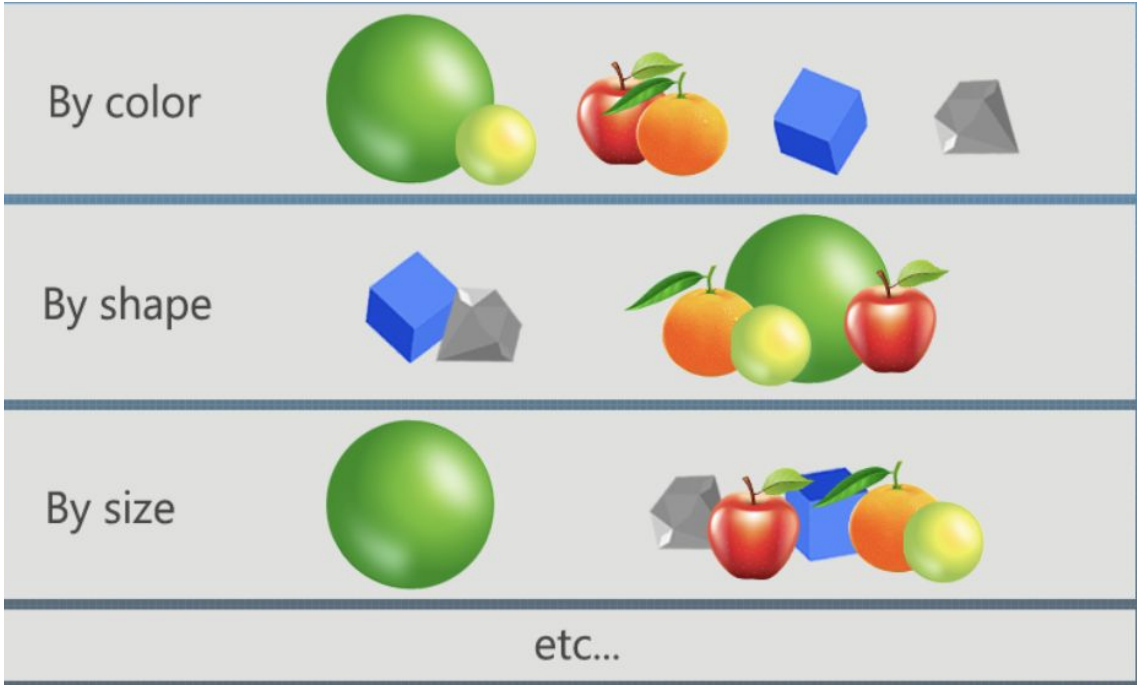
Classification Model

Categorical Label

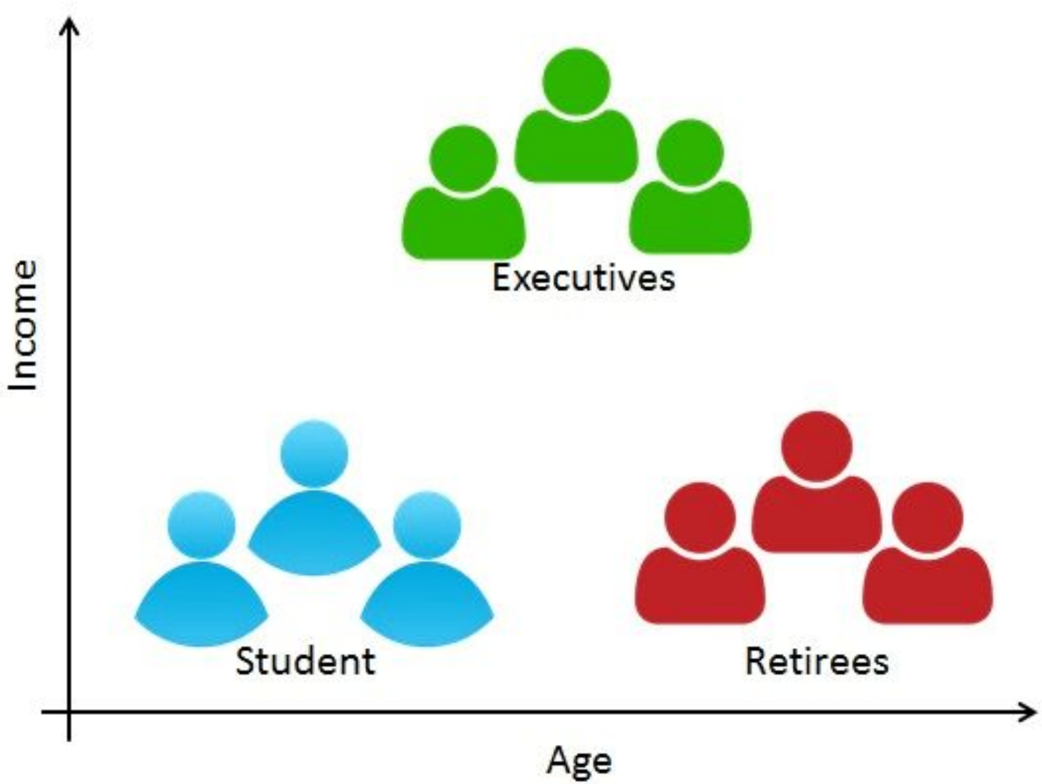
Unsupervised Learning

An unsupervised model is trained on a unlabeled dataset that contains only features but with NO labels

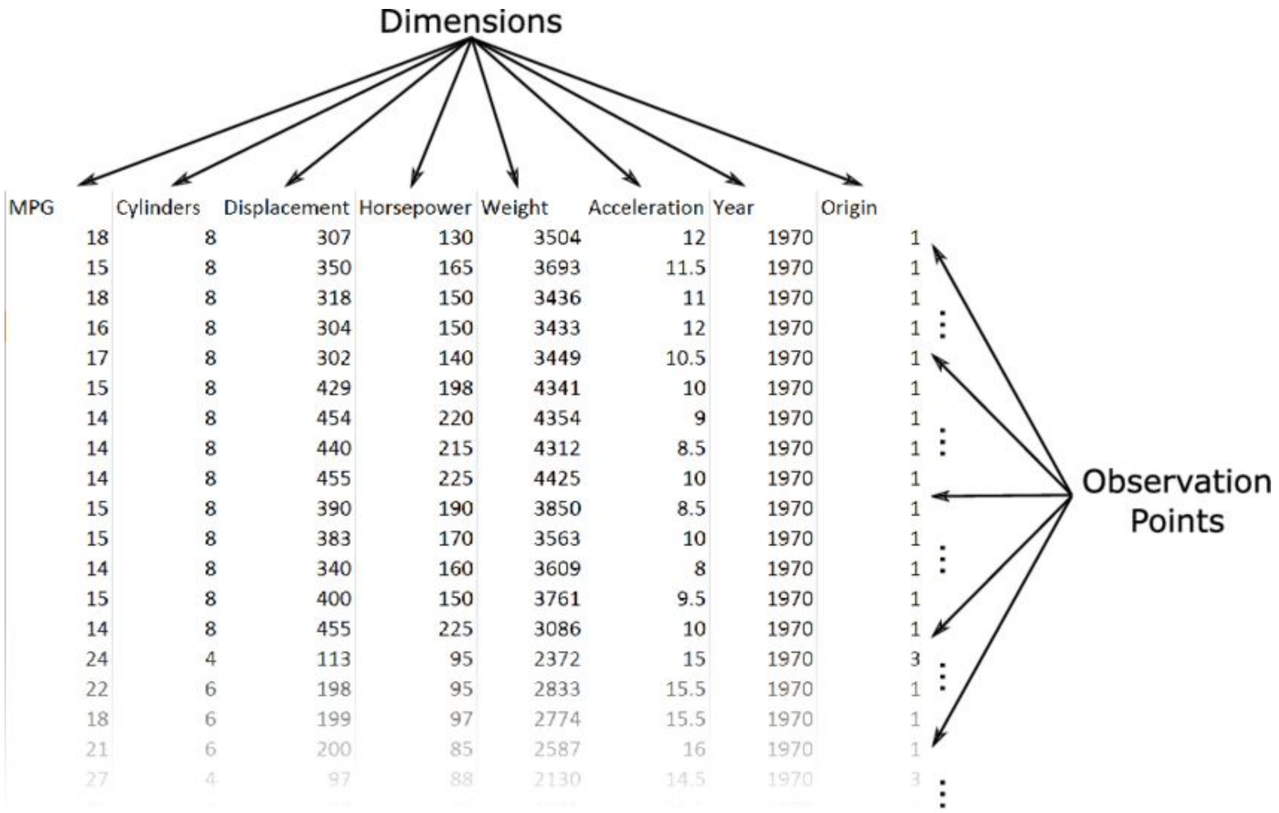
Clustering model - group similar instances together



Clustering model - customer segmentation given income and age

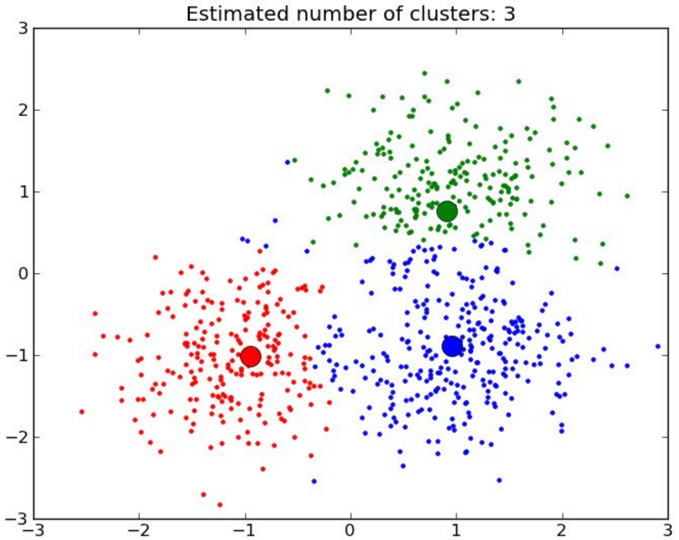


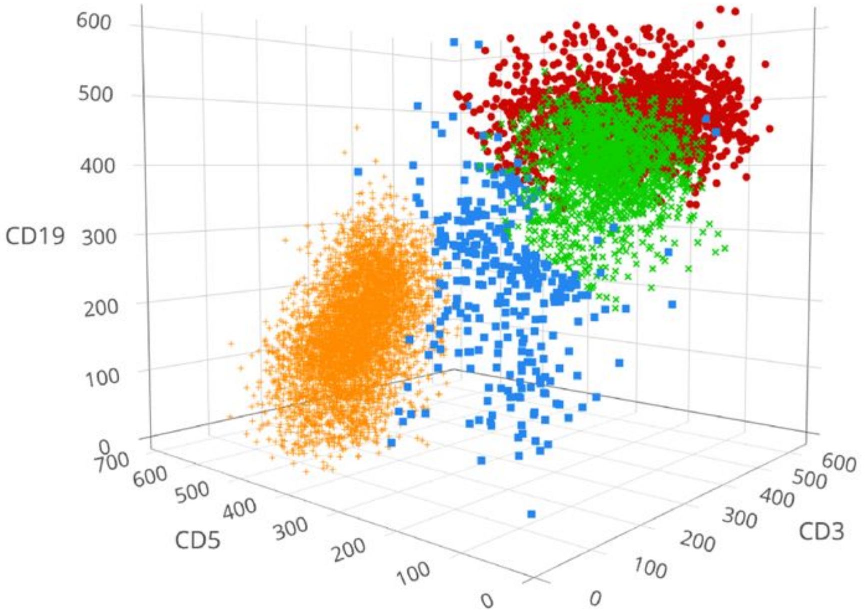
Dimension reduction model - express data with 2-3 dimensions



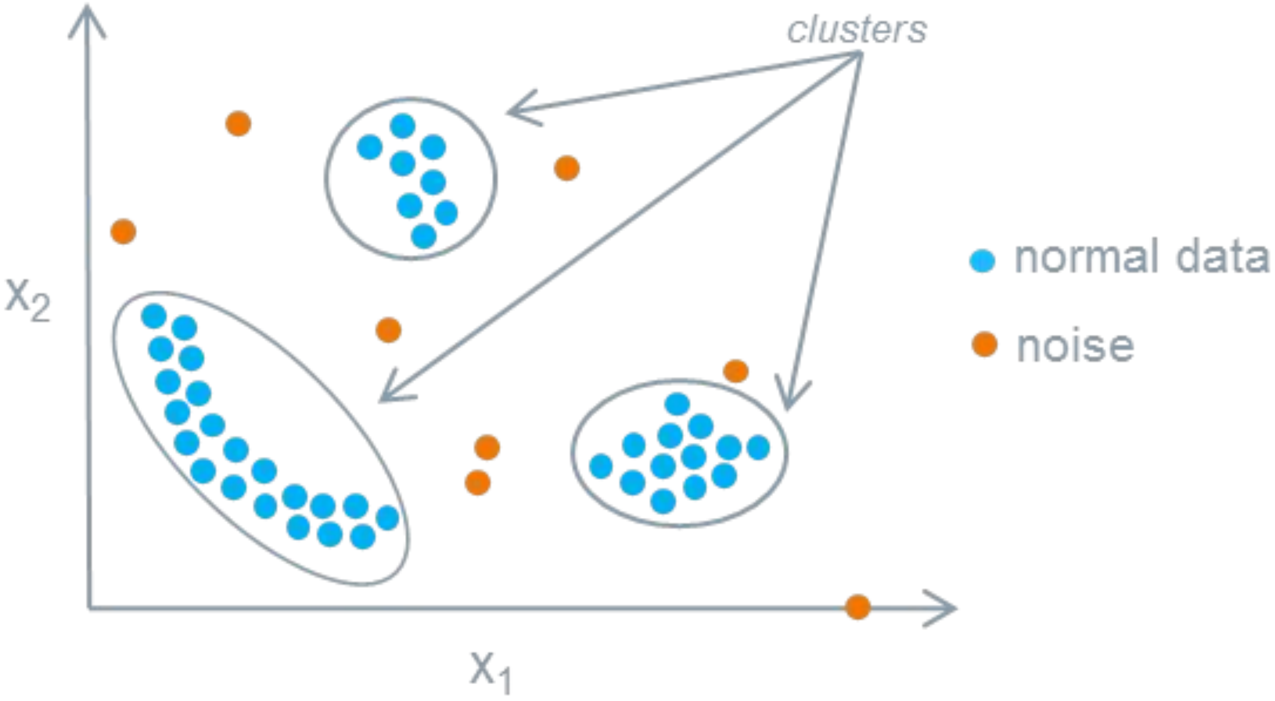
MPG	Cylinders	Displacement	Horsepower	Weight	Acceleration	Year	Origin
18	8	307	130	3504	12	1970	1
15	8	350	165	3693	11.5	1970	1
18	8	318	150	3436	11	1970	1
16	8	304	150	3433	12	1970	1
17	8	302	140	3449	10.5	1970	1
15	8	429	198	4341	10	1970	1
14	8	454	220	4354	9	1970	1
14	8	440	215	4312	8.5	1970	1
14	8	455	225	4425	10	1970	1
15	8	390	190	3850	8.5	1970	1
15	8	383	170	3563	10	1970	1
14	8	340	160	3609	8	1970	1
15	8	400	150	3761	9.5	1970	1
14	8	455	225	3086	10	1970	1
24	4	113	95	2372	15	1970	3
22	6	198	95	2833	15.5	1970	1
18	6	199	97	2774	15.5	1970	1
21	6	200	85	2587	16	1970	1
27	4	97	88	2130	14.5	1970	3

Dimension reduction model - express data with 2-3 dimensions

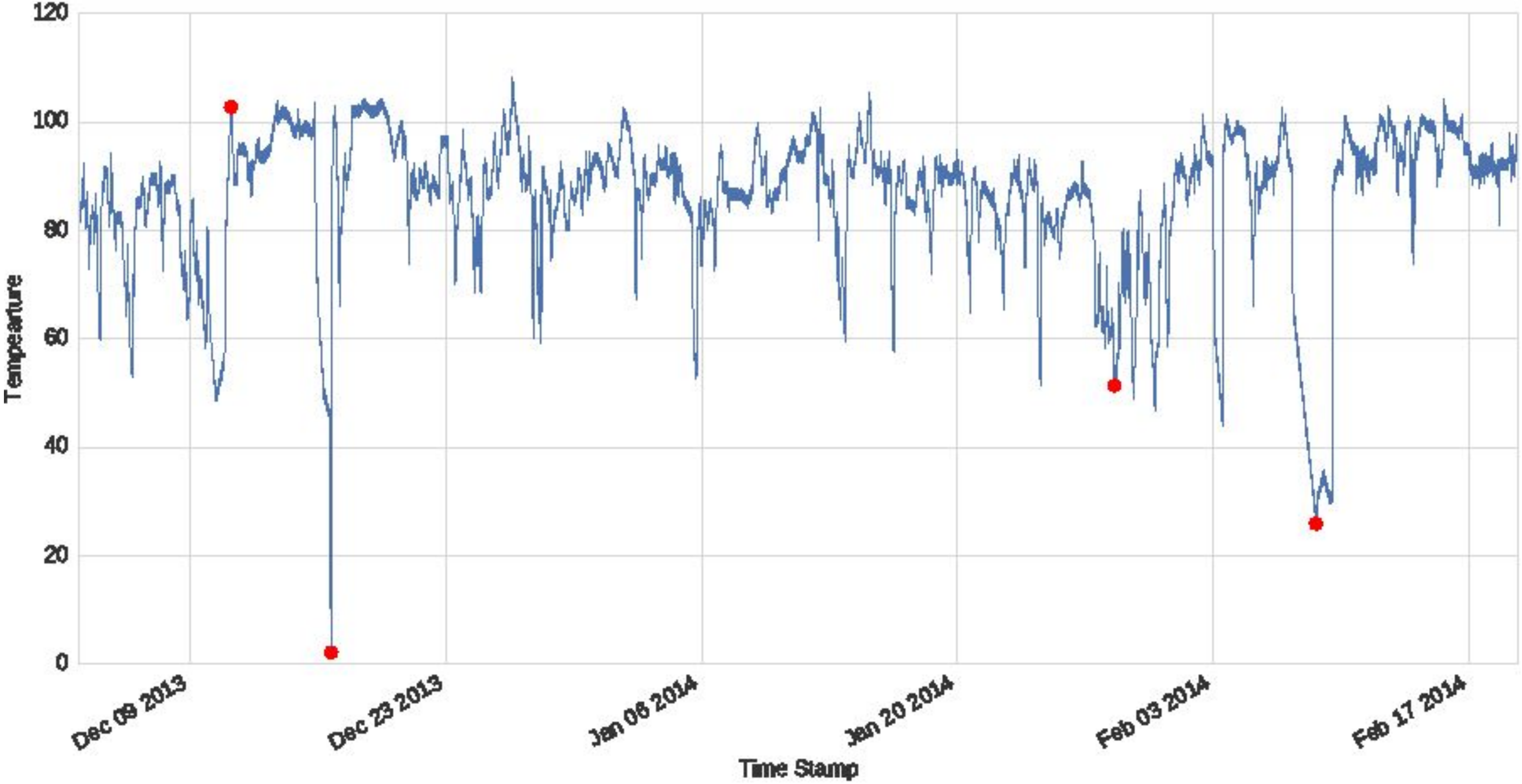




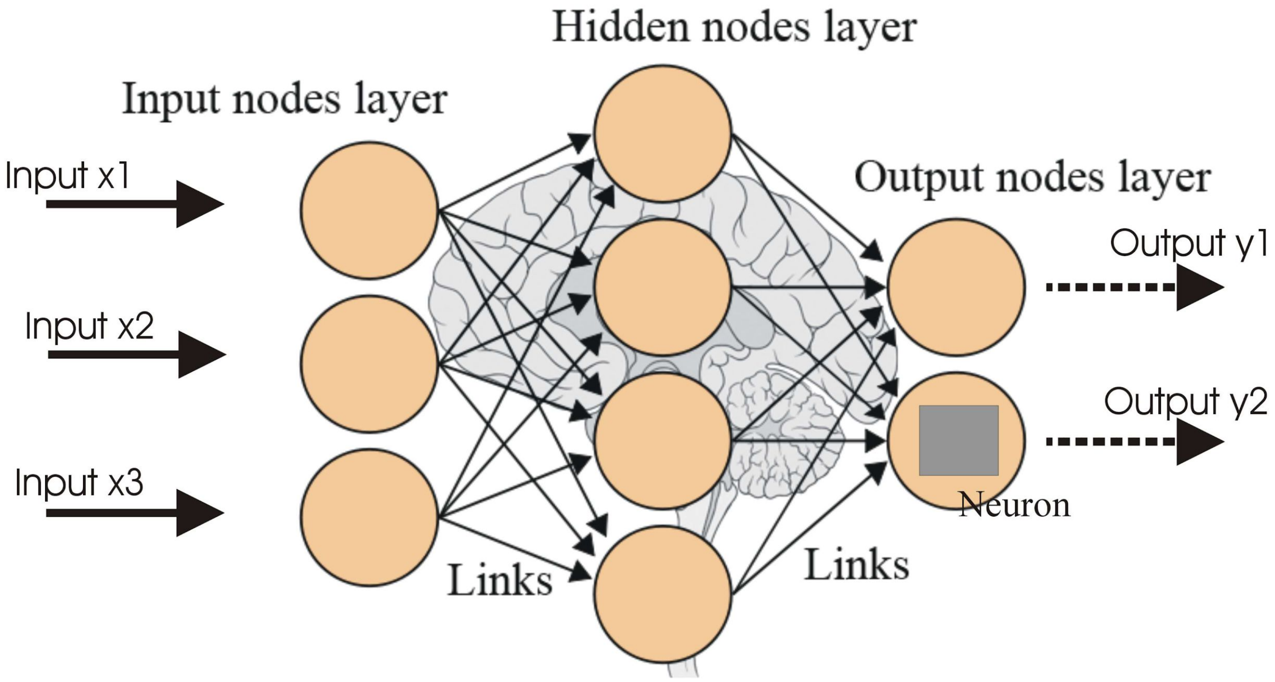
Anomaly detection - finding outliers



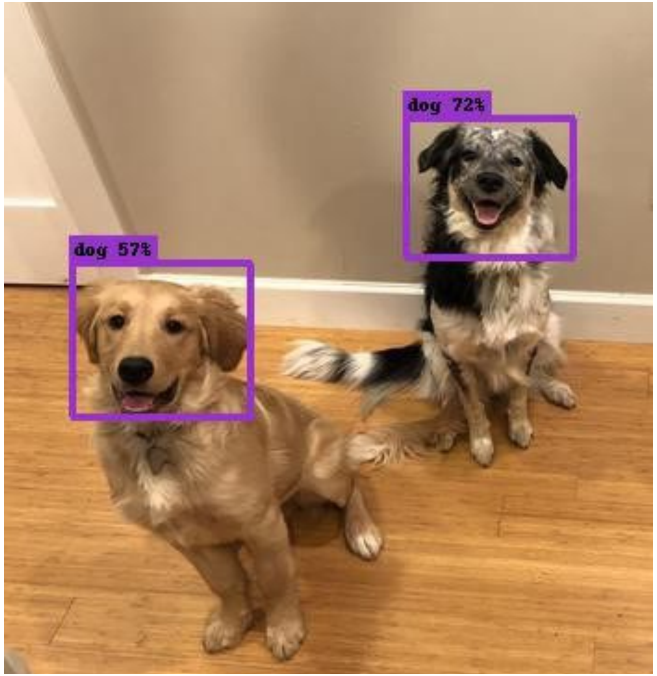
Anomaly detection - finding outliers

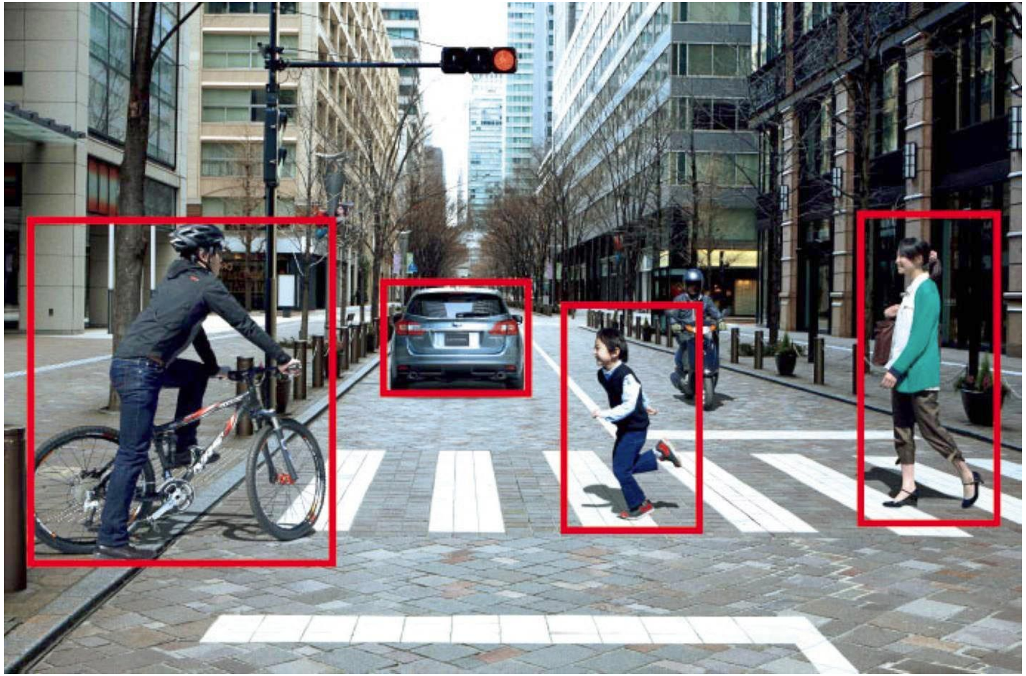


Neural Networks



Convolutional Neural Networks

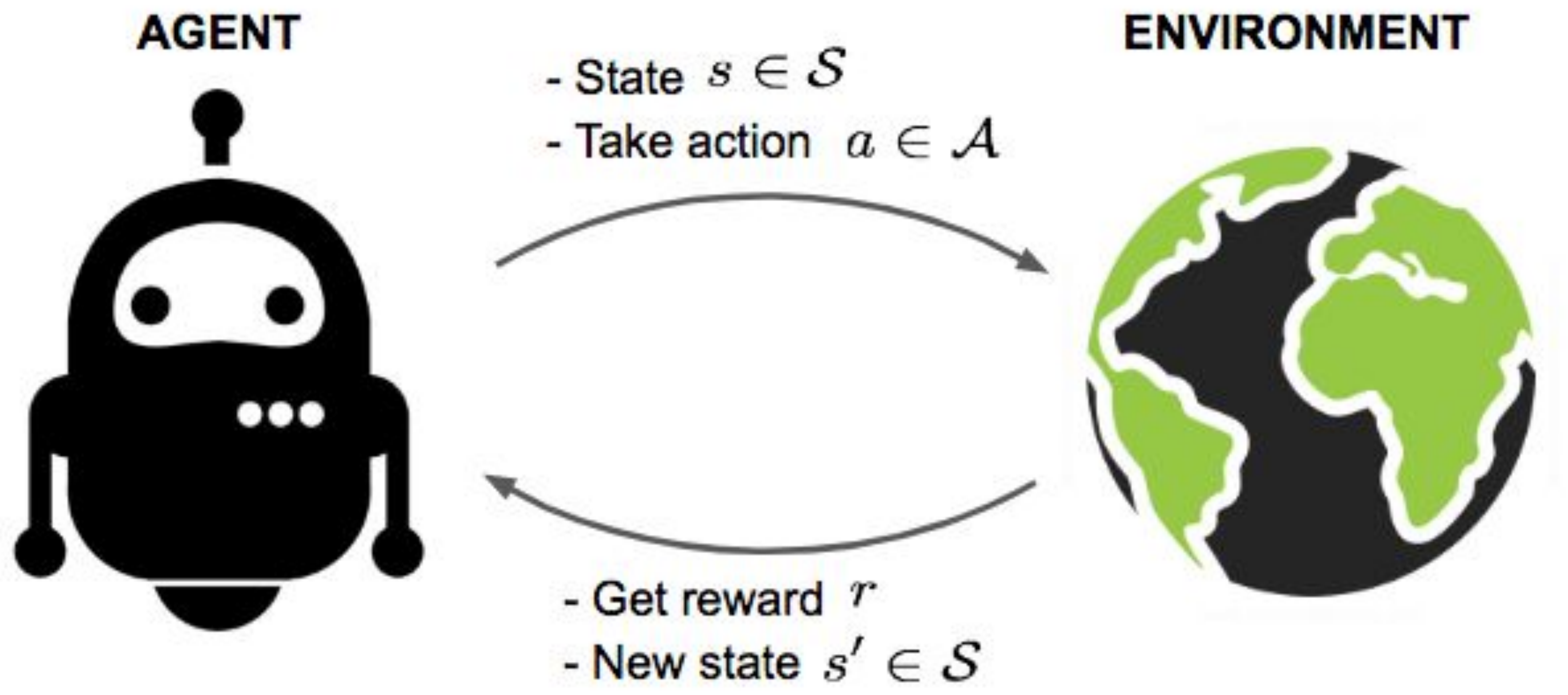




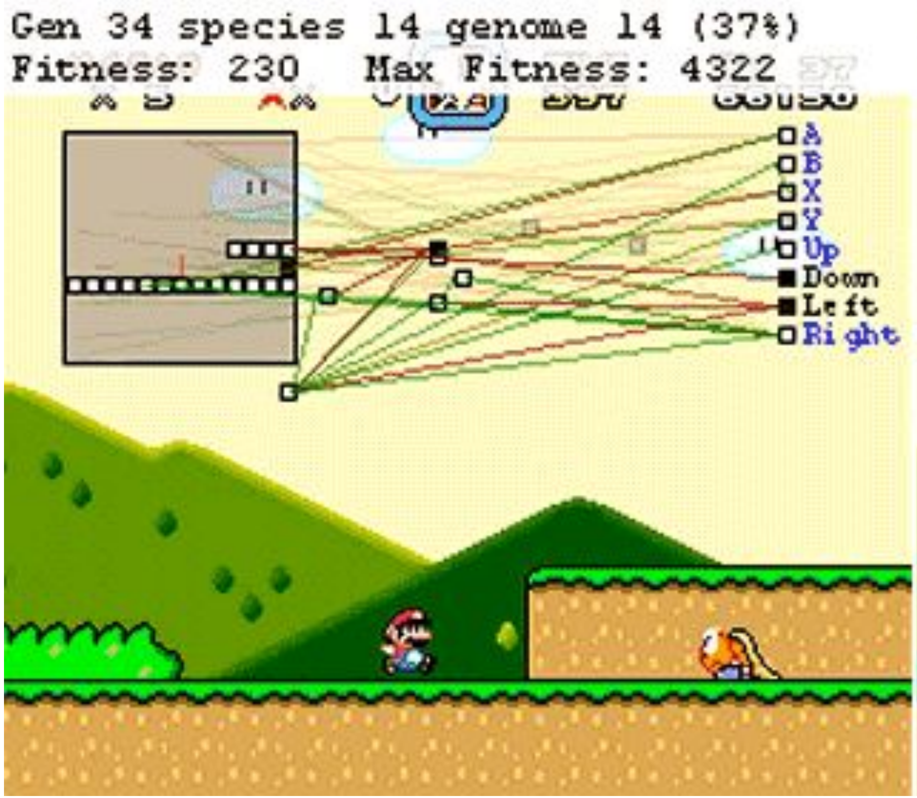
Reinforcement Learning

Train a machine learning model to generate a sequence of decisions

(Model)

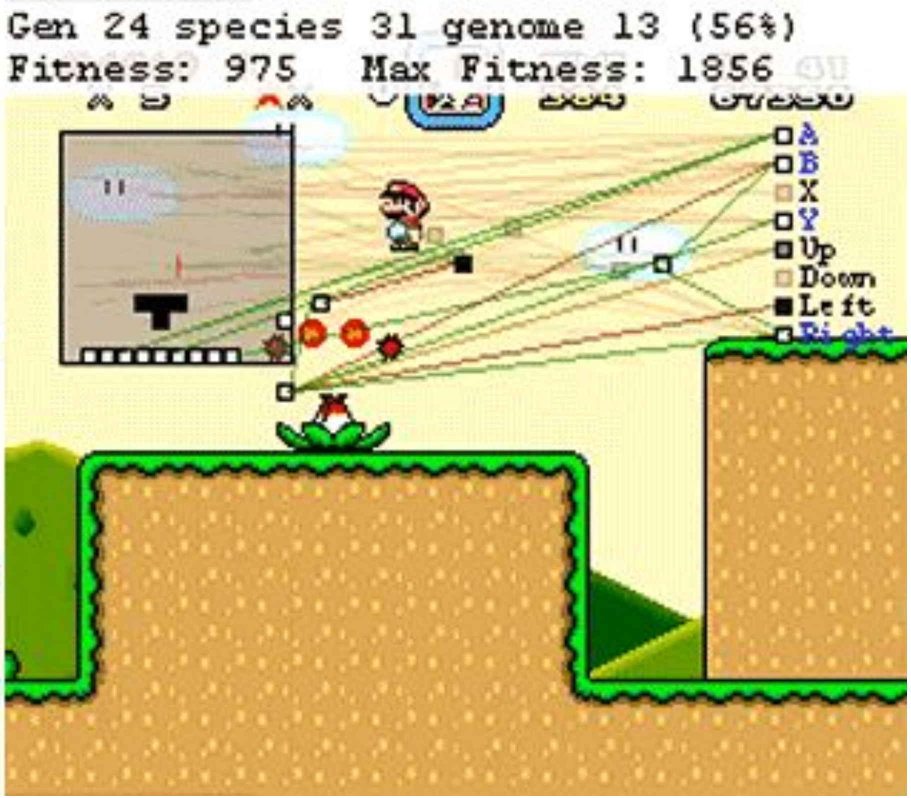

AGENT
ENVIRONMENT
- State $s \in \mathcal{S}$
- Take action $a \in \mathcal{A}$
- Get reward r
- New state $s' \in \mathcal{S}$

Mario game



Gen 34 species 14 genome 14 (37%)

Fitness: 230 Max Fitness: 4322

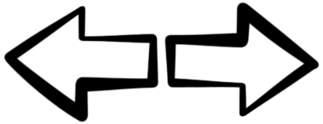


Gen 24 species 31 genome 13 (56%)

Fitness: 975 Max Fitness: 1856

Autonomous car navigation

Environment: street model

Set of actions: 

Scoring: penalty/reward

https://www.youtube.com/watch?v=3ROVzjkkCIA

Thank you!