Machine Learning

DAVE3625

Umair Mehmood Imam

DAVE3625- INTRO TO A.I. BY UMAIR M.I

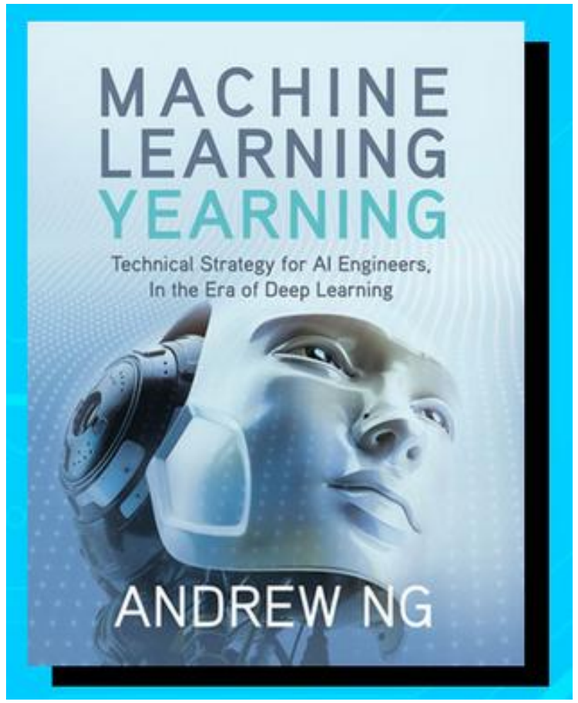
OSLO METROPOLITAN UNIVERSITY
STORBYUNIVERSITETET

OSLOMET

Germany - Registrations of new passenger cars

		Jun. 2021	Jun. 2020	y-o-y	Jan.-Jun. 2021	Jan.-Jun. 2020	y-o-y
VW	Total	57,159	39,126	46.1%	276,486	223,227	23.9%
	Share	20.8%	17.8%	3.1 pt	19.9%	18.4%	1.4 pt
BMW	Total	21,847	18,075	20.9%	118,388	98,839	19.8%
	Share	8.0%	8.2%	-0.2 pt	8.5%	8.2%	0.3 pt
Audi	Total	20,124	15,597	29.0%	106,397	98,623	7.9%
	Share	7.3%	7.1%	0.3 pt	7.6%	8.1%	-0.5 pt
Mercedes	Total	17,723	21,970	-19.3%	116,646	117,355	-0.6%
	Share	6.5%	10.0%	-3.5 pt	8.4%	9.7%	-1.3 pt
Opel	Total	16,535	9,764	69.3%	84,719	60,820	39.3%
	Share	6.0%	4.4%	1.6 pt	6.1%	5.0%	1.1 pt
Skoda	Total	15,288	13,595	12.5%	84,541	71,868	17.6%
	Share	5.6%	6.2%	-0.6 pt	6.1%	5.9%	0.1 pt
Seat	Total	14,626	8,993	62.6%	65,707	50,007	31.4%
	Share	5.3%	4.1%	1.3 pt	4.7%	4.1%	0.6 pt
Ford	Total	11,764	16,145	-27.1%	69,781	84,007	-16.9%
	Share	4.3%	7.3%	-3.0 pt	5.0%	6.9%	-1.9 pt
Hyundai	Total	11,641	6,414	81.5%	49,205	38,148	29.0%
	Share	4.2%	2.9%	1.3 pt	3.5%	3.2%	0.4 pt
Renault	Total	10,407	10,642	-2.2%	52,661	47,879	10.0%
	Share	3.8%	4.8%	-1.0 pt	3.8%	4.0%	-0.2 pt
Fiat	Total	9,364	9,141	2.4%	44,514	39,095	13.9%
	Share	3.4%	4.1%	-0.7 pt	3.2%	3.2%	0.0 pt
Toyota	Total	7,507	4,973	51.0%	34,623	29,729	16.5%
	Share	2.7%	2.3%	0.5 pt	2.5%	2.5%	0.0 pt
Kia	Total	6,583	4,110	60.2%	30,632	25,899	18.3%
	Share	2.4%	1.9%	0.5 pt	2.2%	2.1%	0.1 pt
Peugeot	Total	5,835	4,815	21.2%	29,360	23,718	23.8%
	Share	2.1%	2.2%	-0.1 pt	2.1%	2.0%	0.2 pt
Citroen	Total	5,350	4,008	33.5%	25,671	22,780	12.7%
	Share	2.0%	1.8%	0.1 pt	1.8%	1.9%	0.0 pt
Mazda	Total	4,748	2,986	59.0%	19,323	17,179	12.5%
	Share	1.7%	1.4%	0.4 pt	1.4%	1.4%	0.0 pt
Tesla	Total	4,466	841	431.0%	13,768	5,103	169.8%
	Share	1.6%	0.4%	1.2 pt	1.0%	0.4%	0.6 pt
Mini	Total	4,121	3,239	27.2%	22,761	16,944	34.3%
	Share	1.5%	1.5%	0.0 pt	1.6%	1.4%	0.2 pt

Resources for Machine learning



Course slide from
DIKU 004 - Supervised Machine Learning.pdf

Some core and important concepts

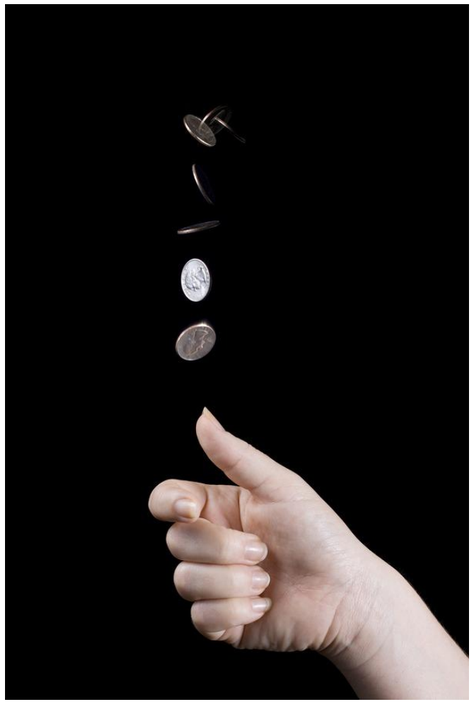
Probability

How likely something is to happen

Example 1: Coin Flipping

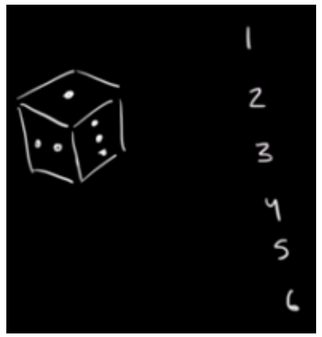
P(H) = (number of ways it can happen) / (Total number of outcomes)

P(H) = ½ = 0.5 or 50%



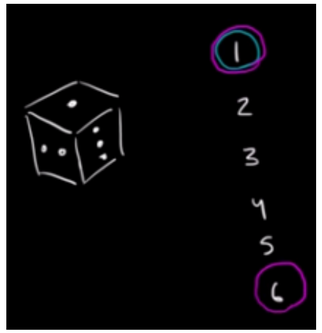
Ref image : https://www.bellevuerarecoins.com/history-coin-flip/

Another example



What's the probability of rolling a one?

$$P(1) = \frac{1}{6}$$



What's the probability of rolling a one or a six?

$$P(1 \text{ or } 6) = \frac{2}{6} = \frac{1}{3}$$

Ref: Khan Academy

Uses of Probability

C11 =PROB(B4:B7,C4:C7,C9,C10)

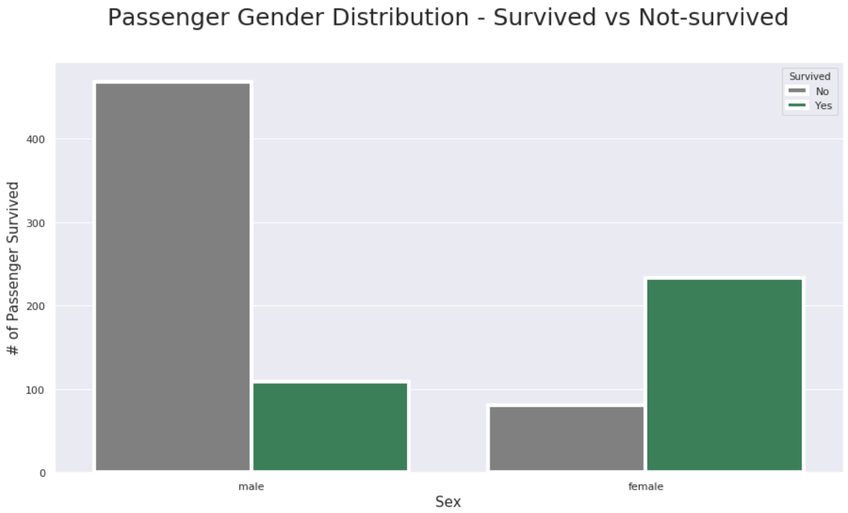
	A	B	C
1			
2			
3		Product Sales	Probability
4		20	0.23
5		50	0.50
6		80	0.18
7		110	0.09
8			
9		Lower limit	50
10		Upper limit	80
11		Probability	0.68
12			

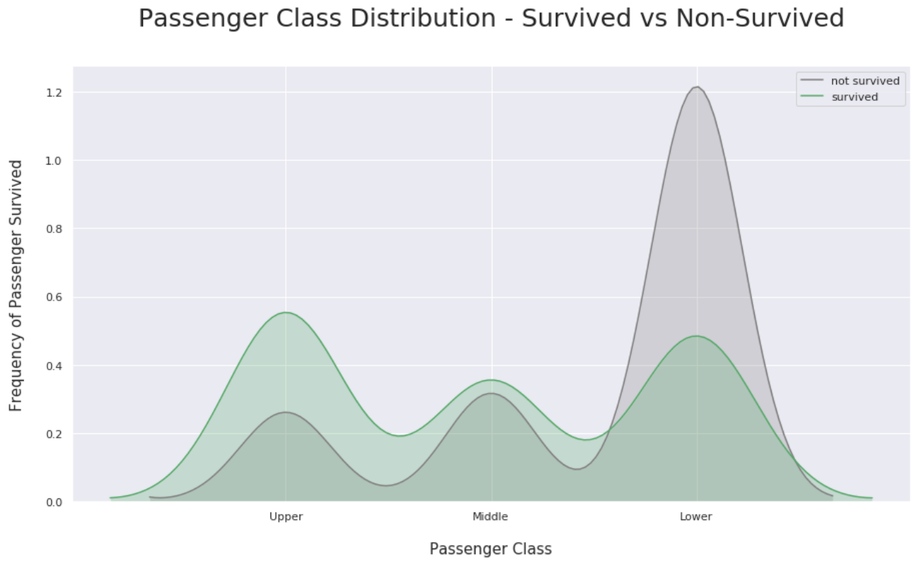
Statistics

Statistics is a traditional field, broadly defined as a *branch of mathematics dealing with data collection, organization, analysis, interpretation and presentation*

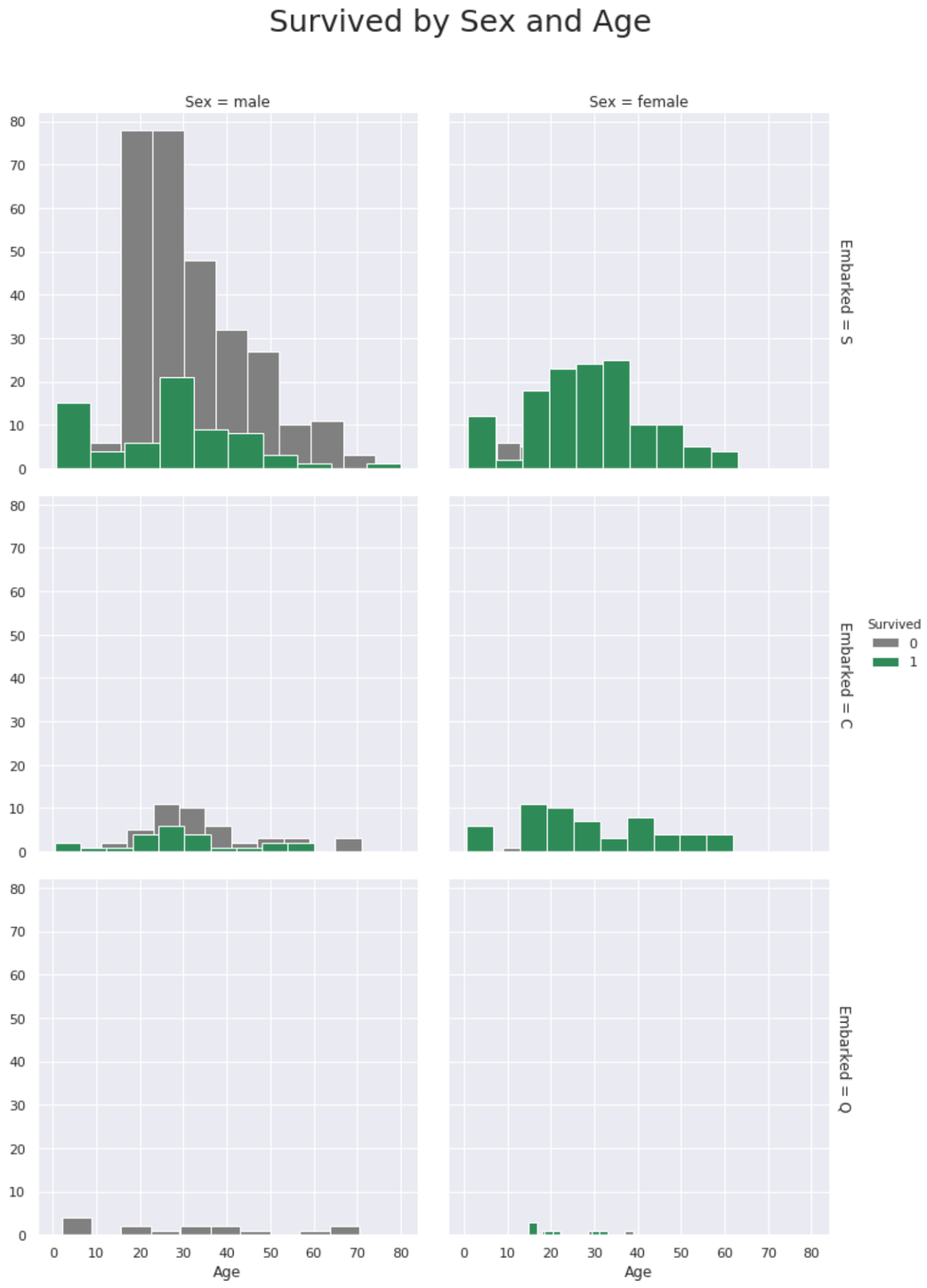
(Ref: Wikipedia)

Titanic data set



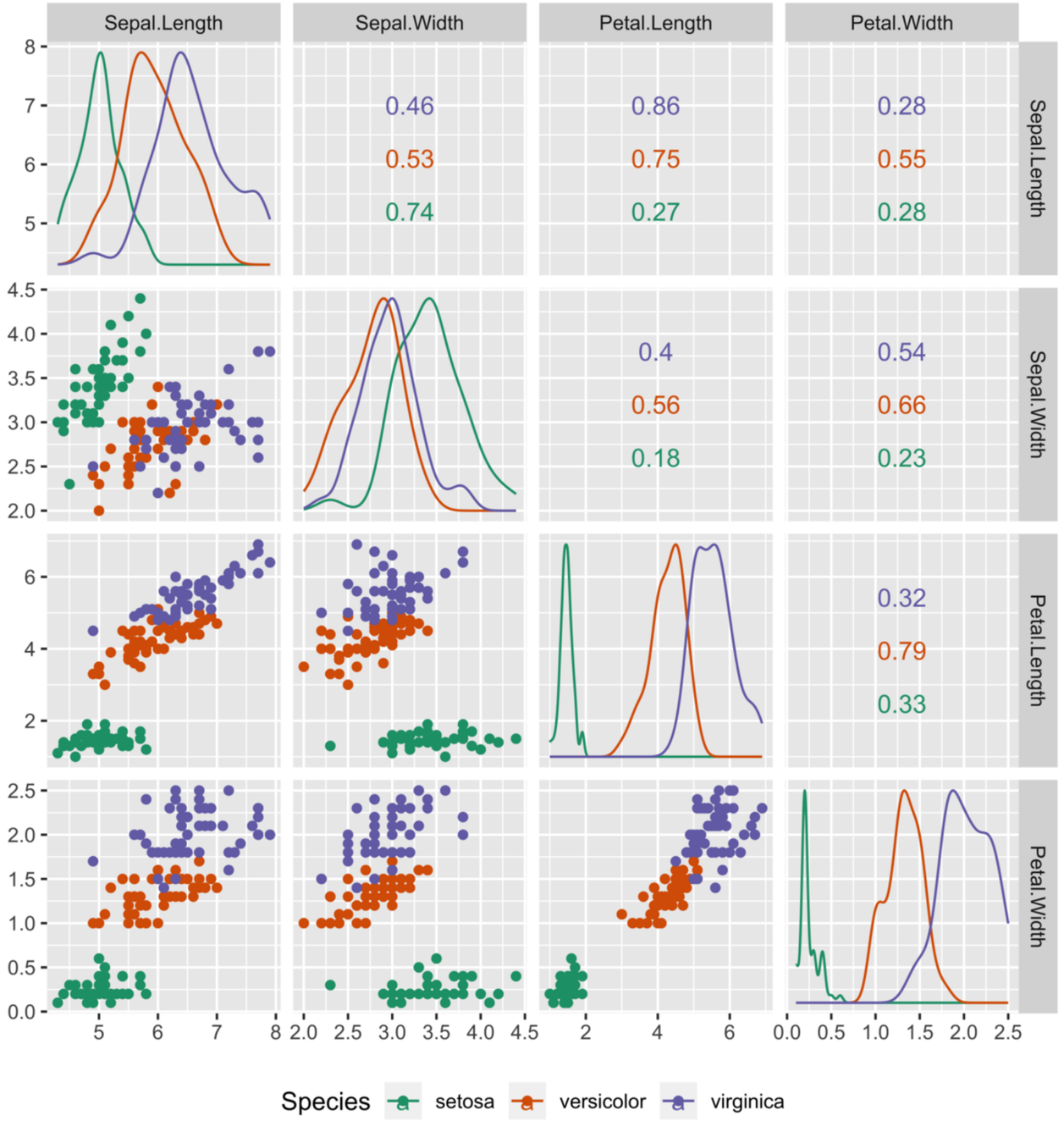


- Most passengers seem to be boarded on Southampton(S).
- More than 60% of the passengers died boarded on Southampton.
- More than 60% of the passengers lived boarded on Cherbourg(C).
- Pretty much every male that boarded on Queenstown(Q) did not survive.
- There were very few females boarded on Queenstown, however, most of them survived.



OSLOMET

Scatter plots



What is Machine Learning

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.

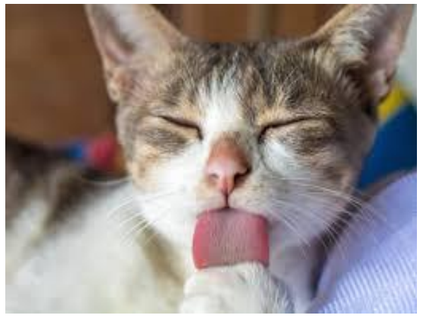
Uses of M.L

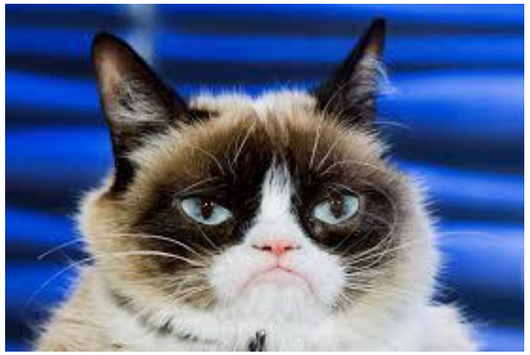
Example: Search for cats on google. How does this search work ?

- Google first gets a large quantity of examples of photos labeled “cat”
- Then the Machine learning algorithm looks for patterns of pixels and patterns of colors that will help it predict if the image is of “cat”.
- At first, Google’s computers make a random guess of what patterns are good in order to identify an image of a cat.
- If it makes a mistake, then a set of adjustments are made (by humans) in order for the algorithm to get it right.
- In the end the algorithm will learn such patterns and improve its output

Q. How would you build such a M.L algorithm ?

1. Collect data
2. Label the data



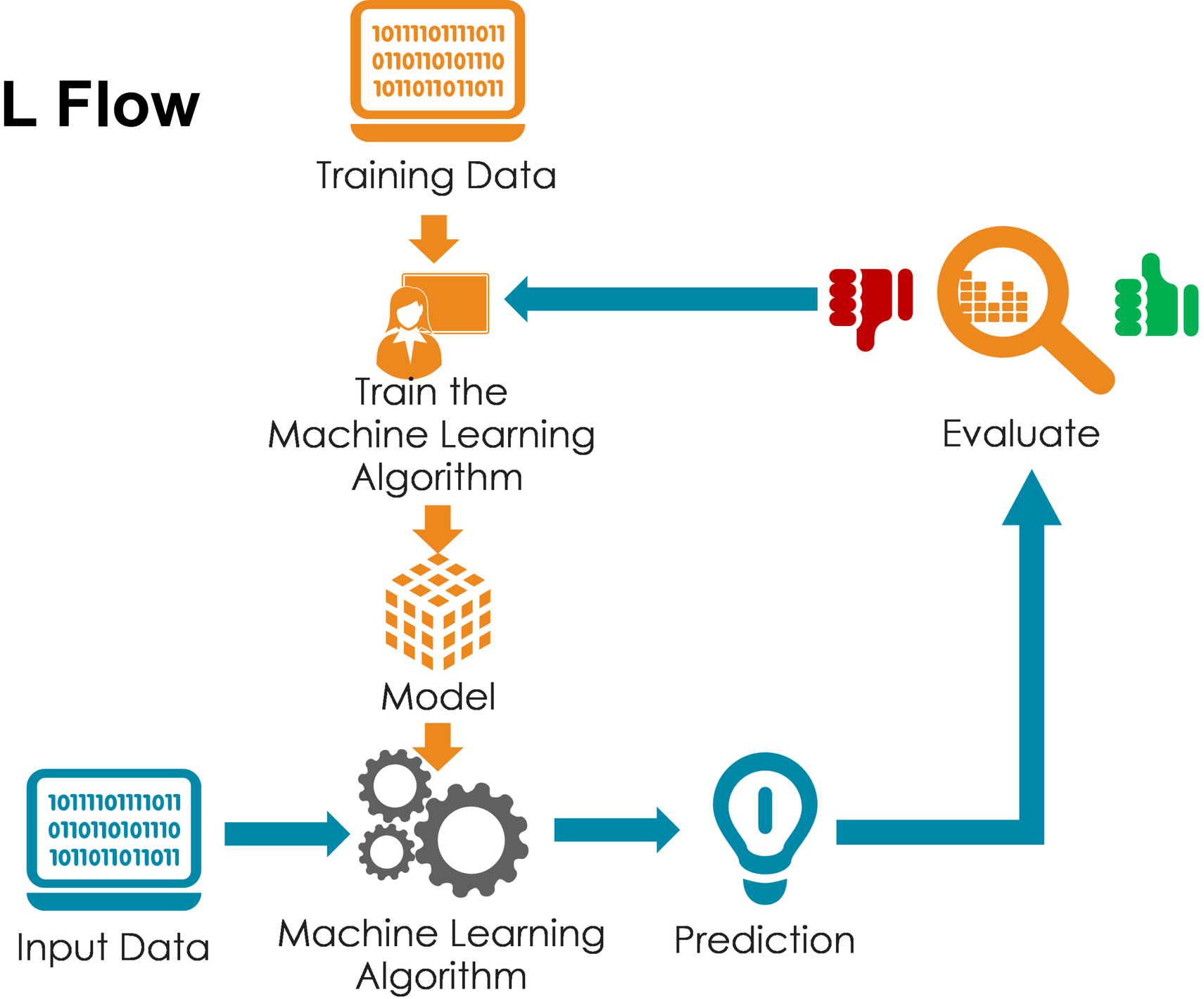


3. Train the machine learning model with data

The model will look for patterns in the images in order to identify cats and lions

4. Once the model is trained, we give it different input images to test and see if the model gives us the right answer.

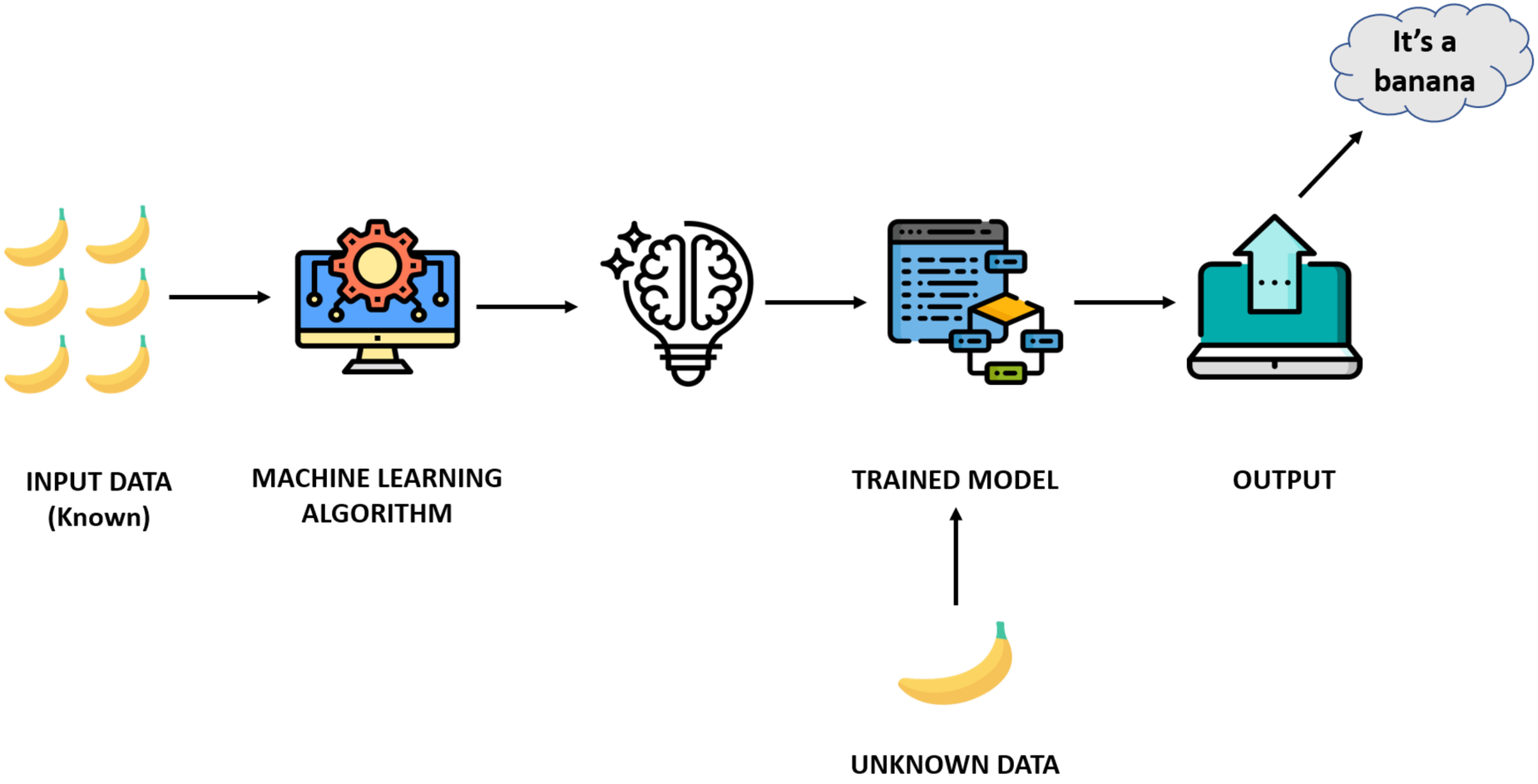
Typical M.L Flow

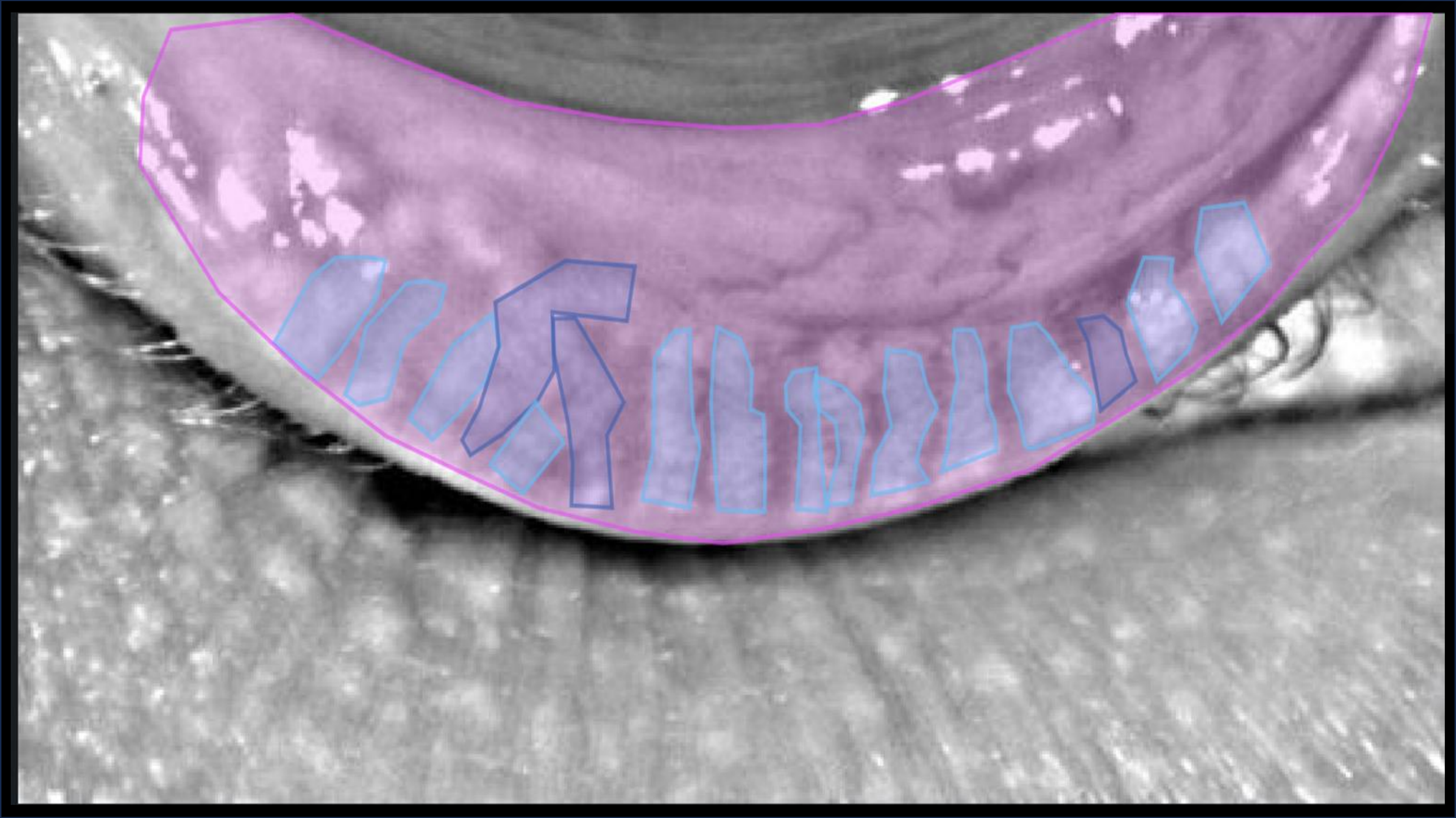


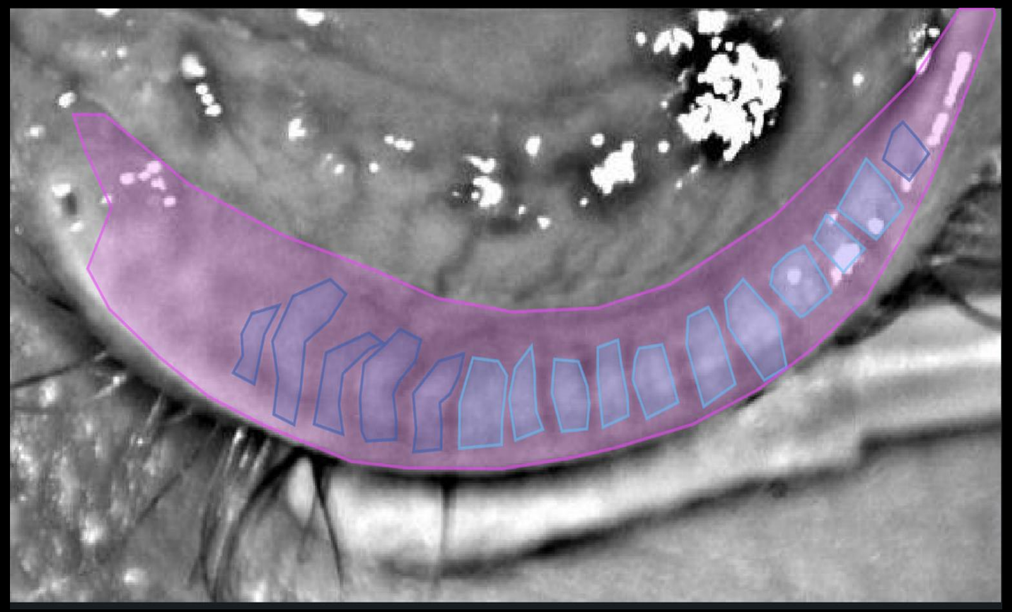
Types of Machine Learning algorithms

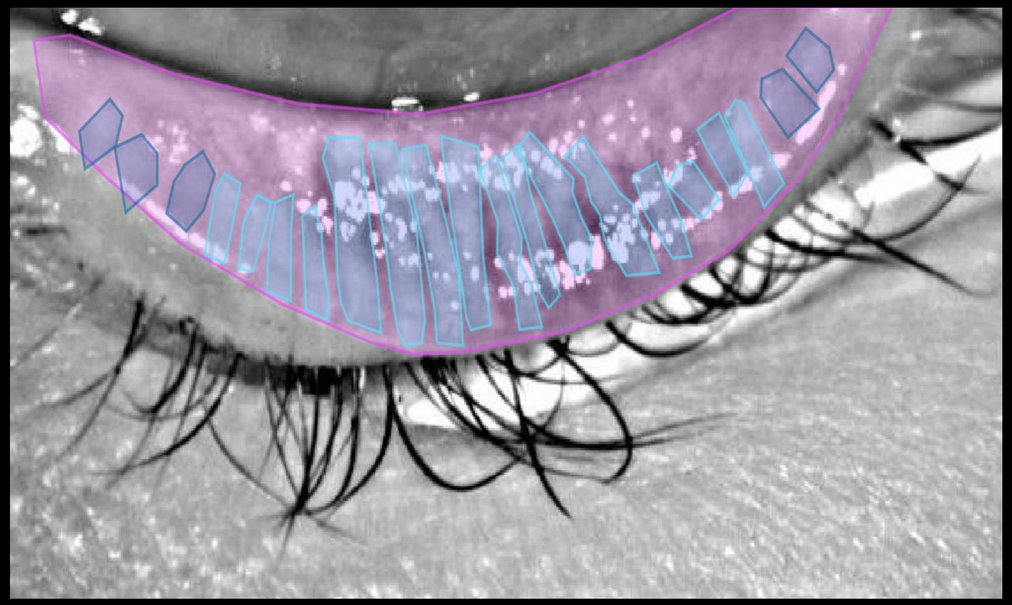
- Supervised learning
- Unsupervised learning
- Reinforcement learning
- Recommender systems

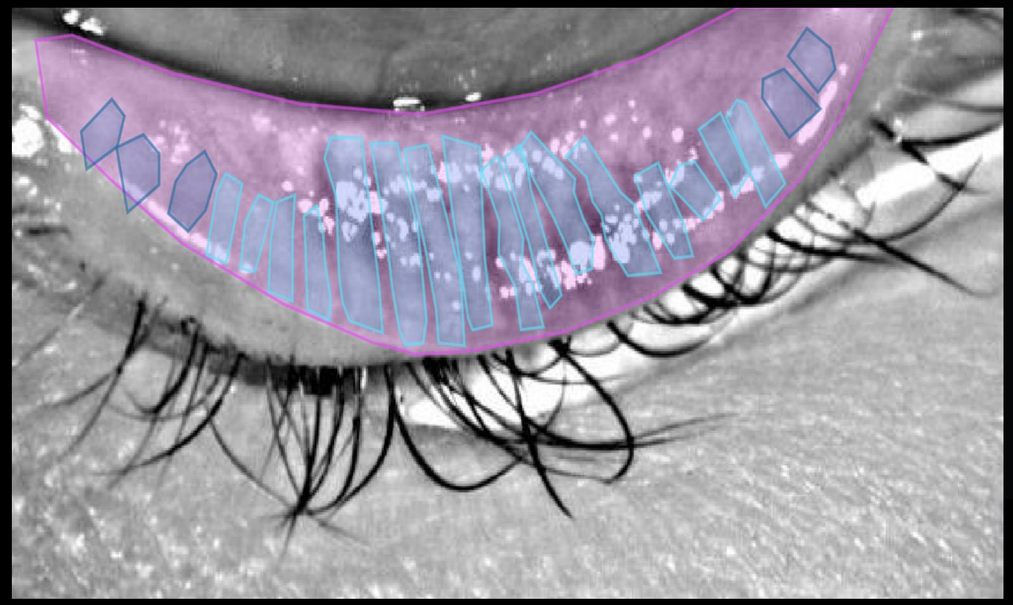
Supervised learning









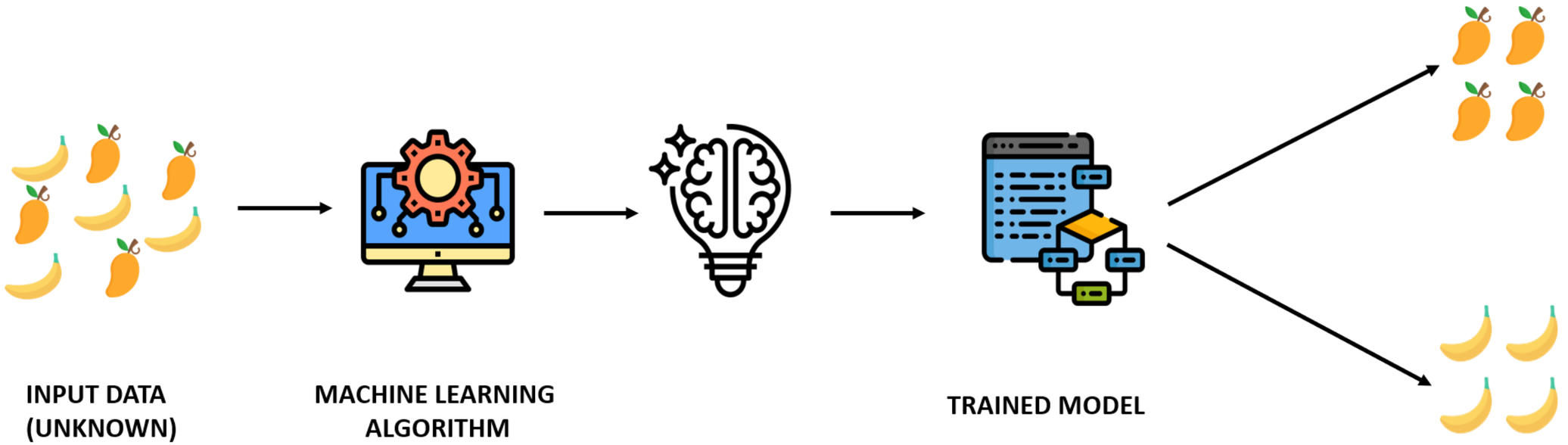


Health AI (labeling sections to identify dry eye disease)

simula

bineric data

Unsupervised learning



SKOLEELEVER MÅ STÅ I BUSSEN: ANITA AASHEIM ER REDD DET SKAL SKJE EN ULYKKE

– Vil utvilsomt kunne få st

Flere foreldre er redde for at det skal skje en ulykke med busser hvor skoleelever står i midtgangen.

NANNESTAD

Sporveien har ikke noe valg – må kjøre med slitte seter

Oslos gamle trikker skal fases ut i 2024, men det er ikke derfor Spor kjører med ødelagte og slitte seter.

Wasim Riaz

Føler seg ikke hørt

Nordre Aker Budstikke – Nyheter – 12. Jul 2021 07:01

By: Christian Fredrik Borg

– Ren lotto om du kommer med skolebussen

– Det kan ikke være slik at man ikke vet om man får plass på bussen eller ikke, sier Liv Gustavsen (Frp), som nå tar bussproblemene ved Hvam videregående skole opp i fylkestinget.

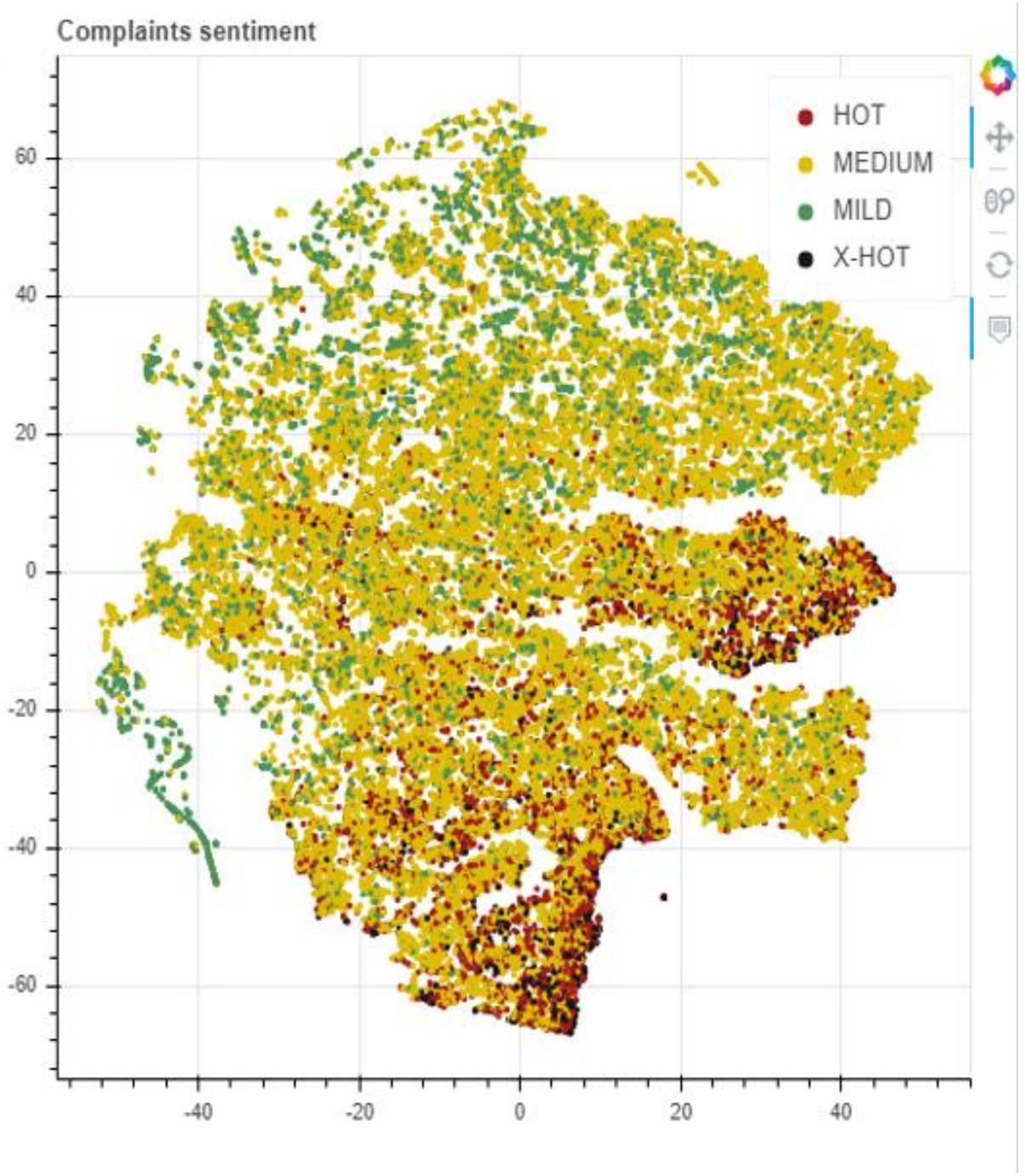
rike, og at denne derfor burde

Slutt for busstopp: – Håpløs avgjørelse

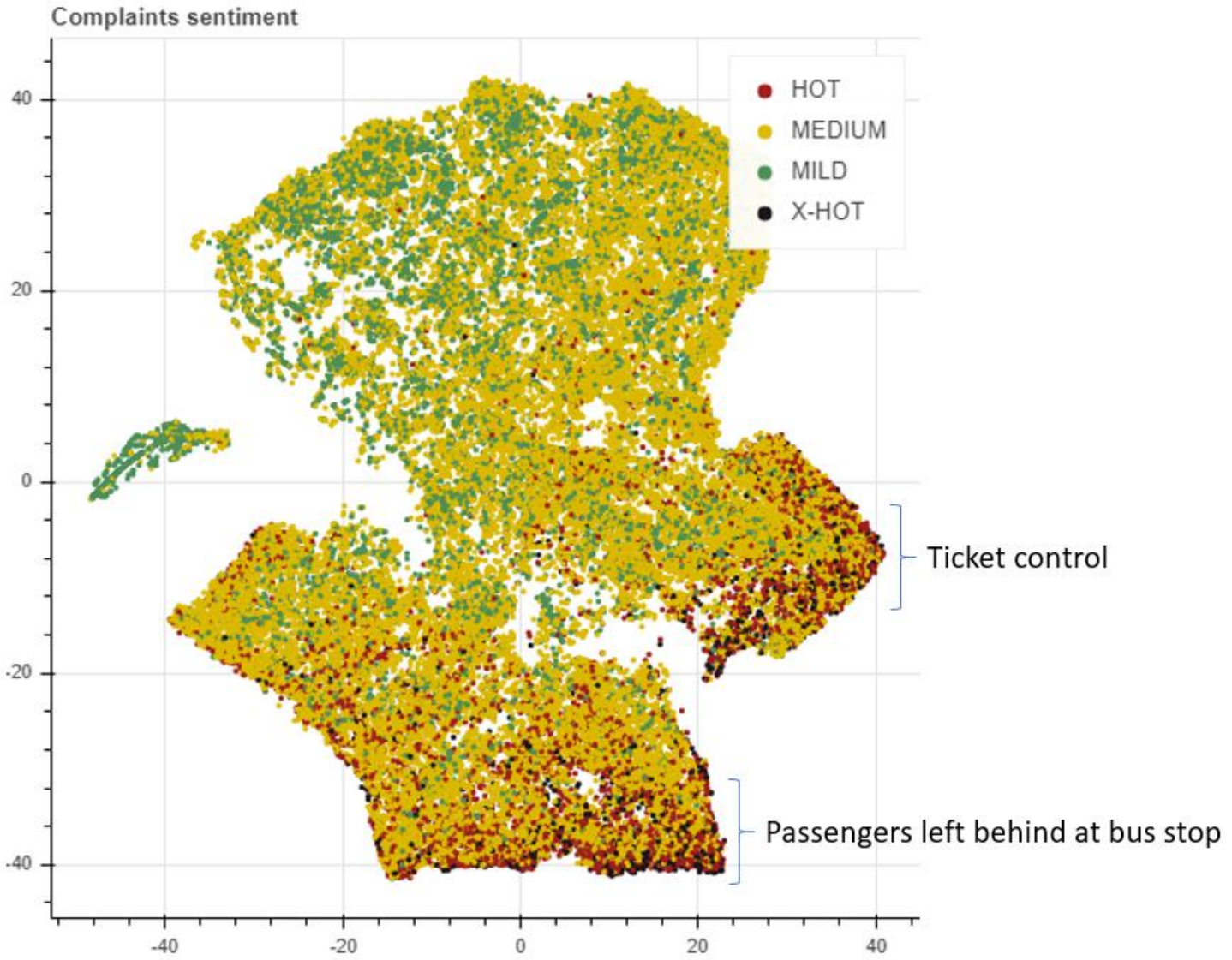
Bussen stoppet ikke - Anine kom for sent på skolen

Ruter#

Step 1: Analyzing how angry the users are (sentiment analysis)

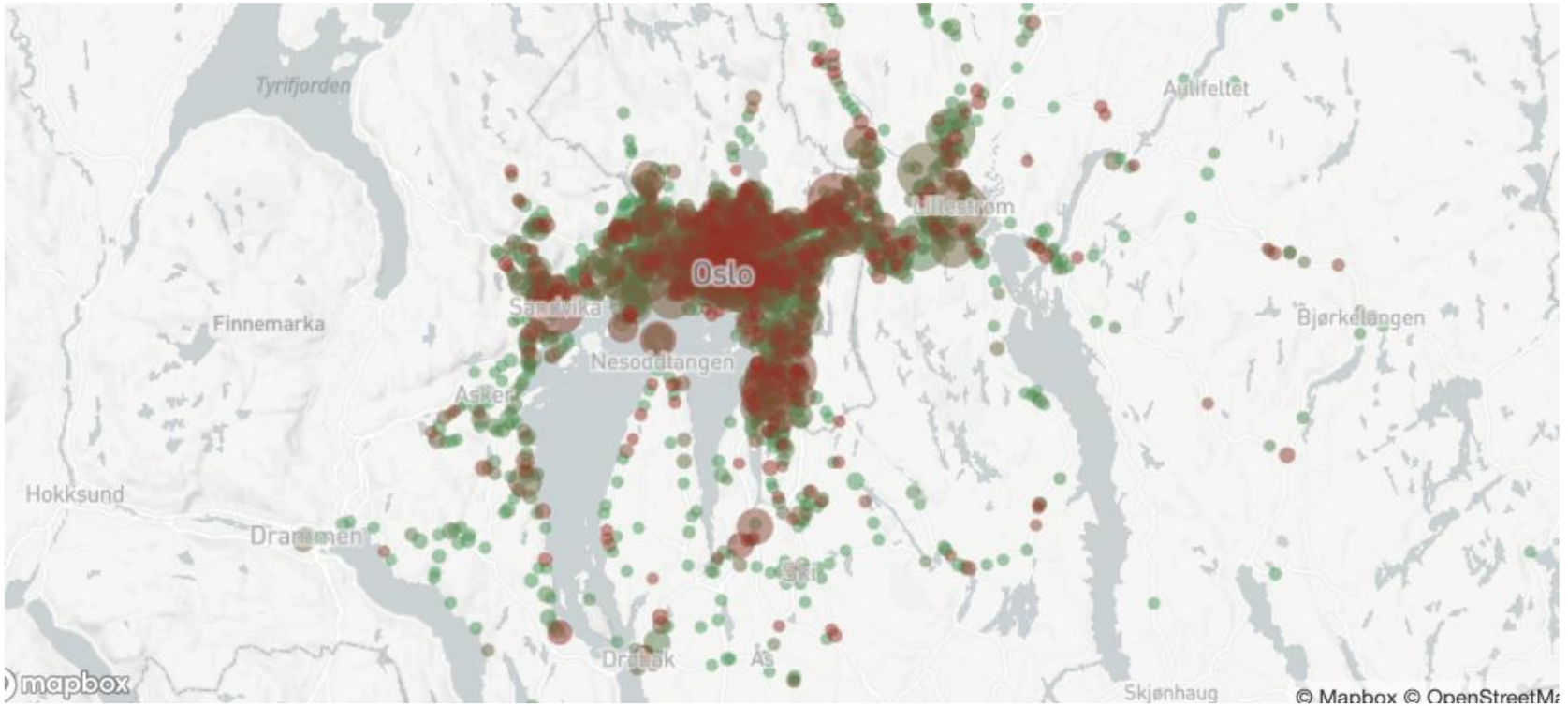

Complaints sentiment
HOT
MEDIUM
MILD
X-HOT
60
40
20
0
-20
-40
-60
-40
-20
0
20
40

Ruter#



Sentiment Analysis
(Nor-BERT)

Ruter#


Tyrifjorden
Oslo
Finnemarka
Nesoddtangen
Asker
Hokksund
Drammen
Bjørkelangen
Skjønhaug
mapbox

Ruter#

Reinforcement learning

- Reinforcement Learning is a type of machine learning technique that enables an agent to learn in an interactive environment by trial and error using feedback from its own actions and experiences.

- Whenever the model predicts or produces a result, it is penalized if the prediction is wrong or rewarded if the prediction is correct.

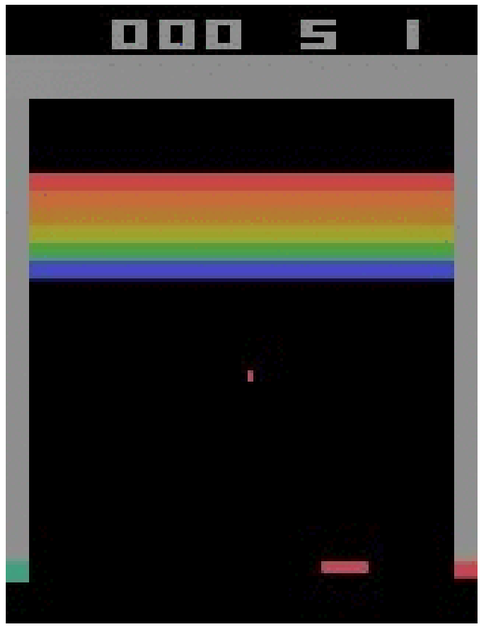
Challenges with Reinforcement learning

- The main challenge in reinforcement learning lays in preparing the simulation environment, which is highly dependant on the task to be performed.
- Scaling and tweaking the neural network controlling the agent is another challenge. There is no way to communicate with the network other than through the system of rewards and penalties.

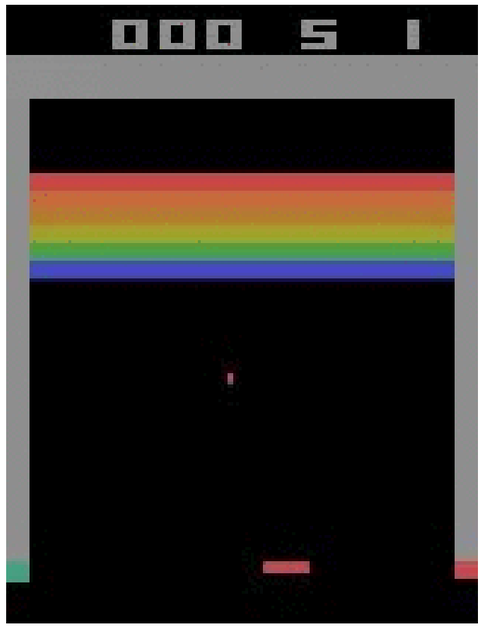
There are agents that will optimize the prize without performing the task it was designed for. An interesting example can be found in the OpenAI video (below), where the agent learned to gain rewards, but not to complete the race.

Breakout game - Atari

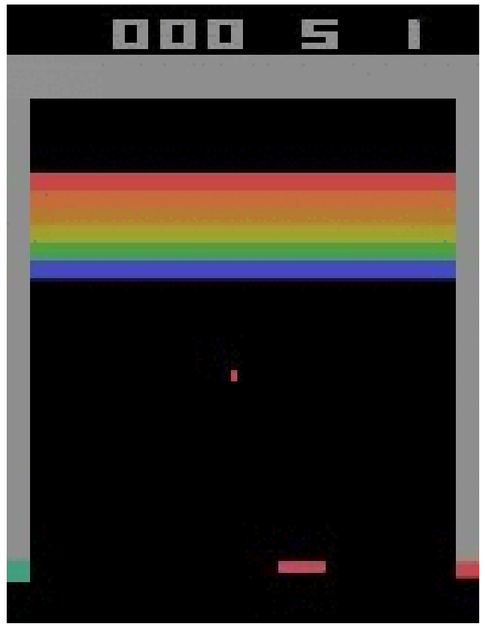
Initial performance



After 15 min of training

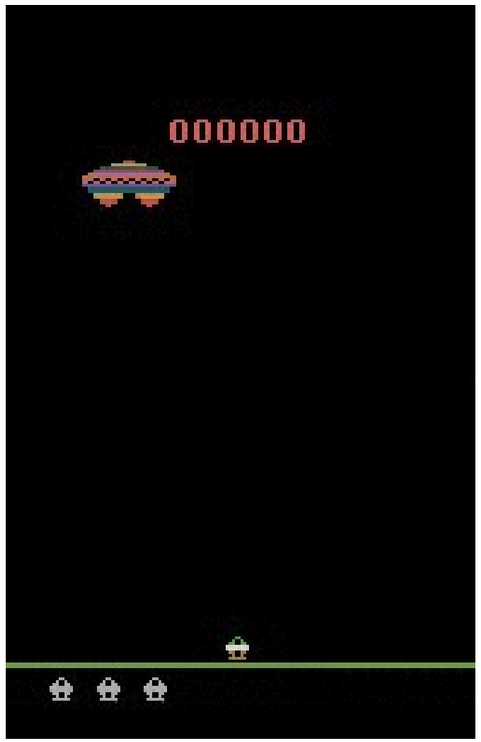


After 30 min of training

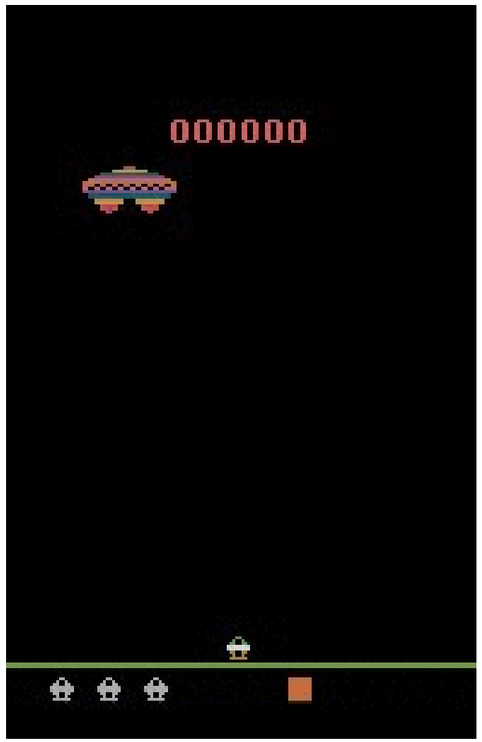


Assault game - Atari

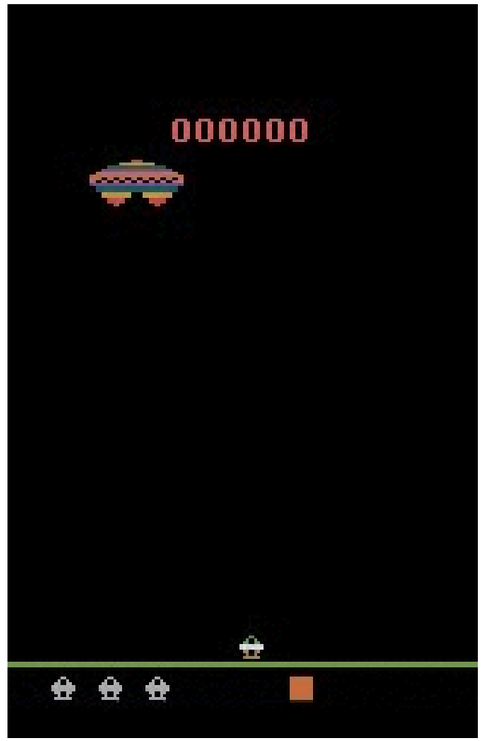
Initial performance

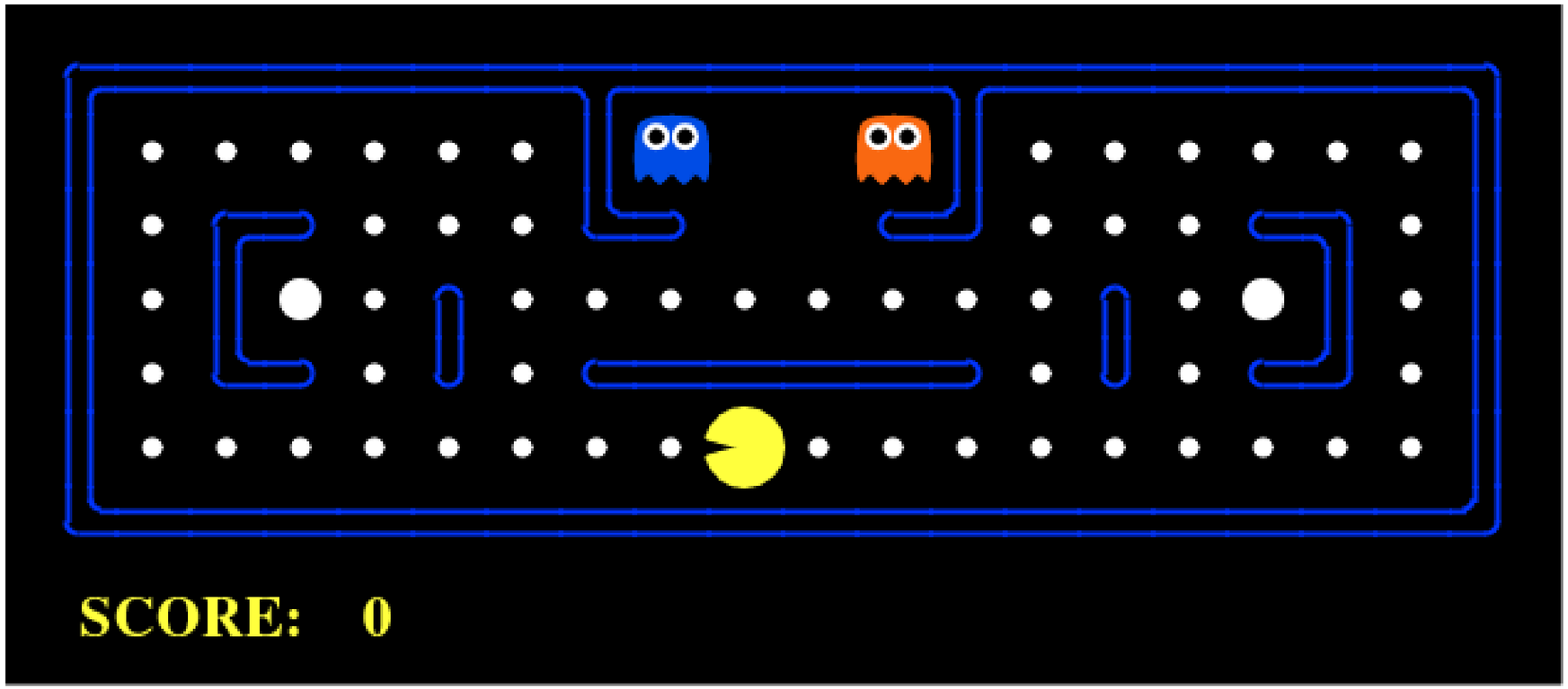


After 15 min of training



After 30 min of training




SCORE: 0

Wanna try some A.I reinforcement learning in games ?

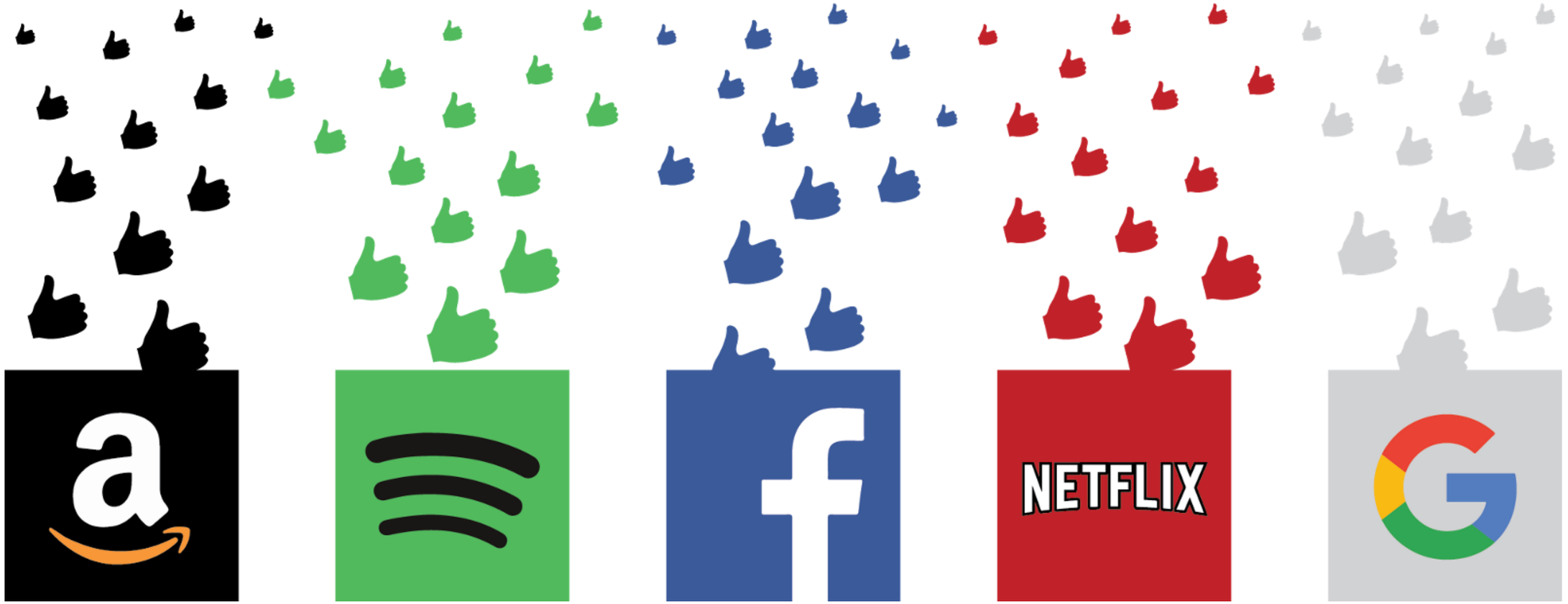
https://gym.openai.com/

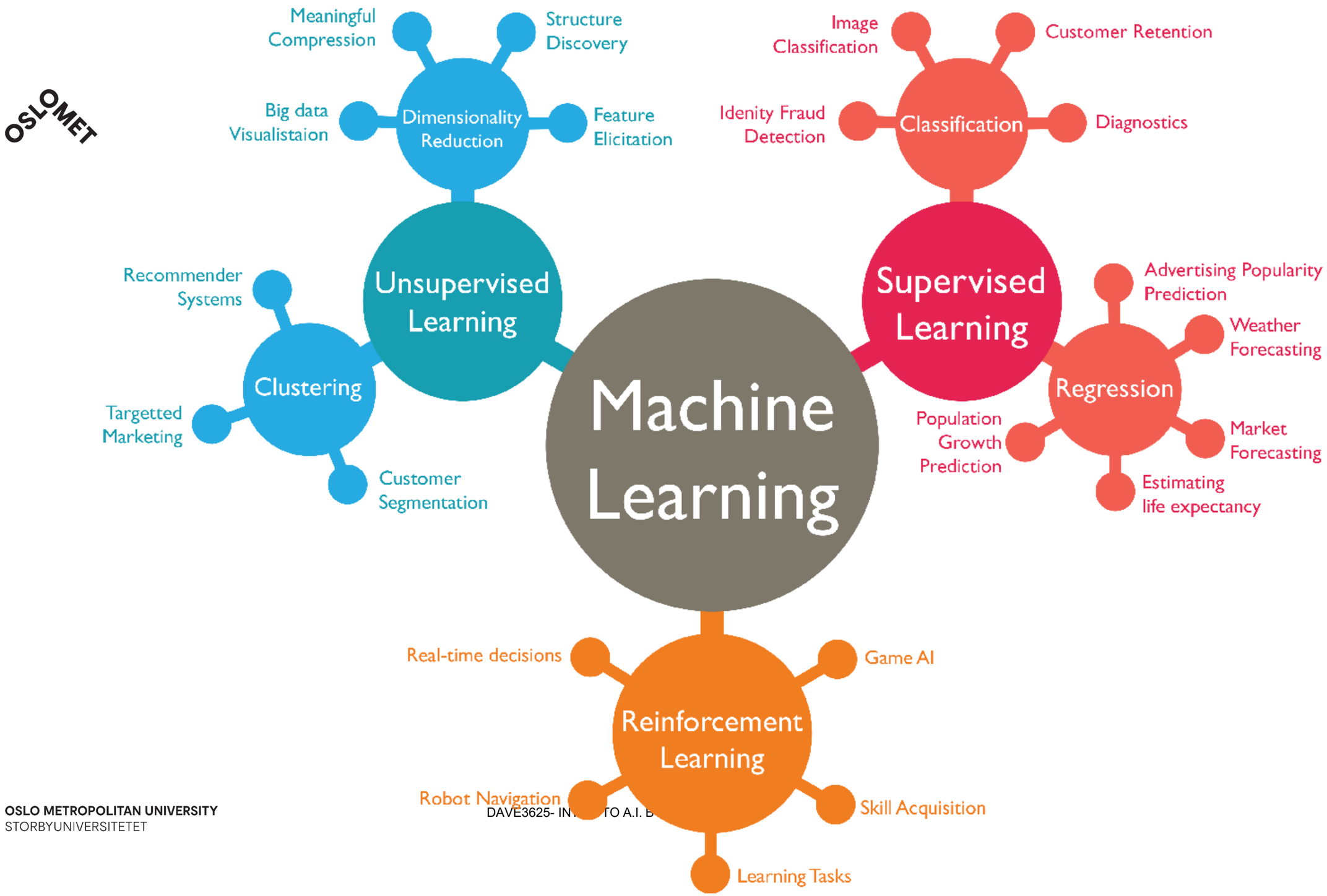
Reccomender systems

- *Recommender systems* are an important class of machine learning algorithms that offer "relevant" suggestions to users.

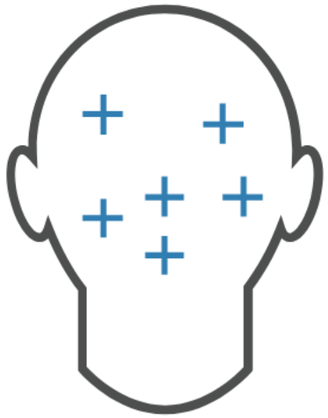
- These systems predict the most likely product that the users are most likely to purchase and are of interest to them

OSLOMET




Machine Learning
Unsupervised Learning
Supervised Learning
Reinforcement Learning
Dimensionality Reduction
Clustering
Classification
Regression
Meaningful Compression
Structure Discovery
Big data Visualistaion
Feature Elicitation
Recommender Systems
Targetted Marketing
Customer Segmentation
Image Classification
Customer Retention
Idenity Fraud Detection
Diagnostics
Advertising Popularity Prediction
Weather Forecasting
Population Growth Prediction
Market Forecasting
Estimating life expectancy
Real-time decisions
Game AI
Robot Navigation
Skill Acquisition
Learning Tasks

When should we use machine learning ?



Hand-written rules and equations are too complex—as in face recognition and speech recognition.

The rules of a task are constantly changing—as in fraud detection from transaction records.

The nature of the data keeps changing, and the program needs to adapt—as in automated trading, energy demand forecasting, and predicting shopping trends.