KOÇ UNIVERSITY

Machine Learning & Computer Vision

Buket Yüksel

What is Computer Vision?

- Computer vision is the process of understanding digital images and videos using computers.
- It seeks to automate tasks that human vision can achieve. This involves methods of acquiring, processing, analysing, and understanding digital images, and extraction of data from the real world to produce information.
- It also has sub-domains such as object recognition, video tracking, and motion estimation, thus having applications in medicine, navigation, and object modelling.

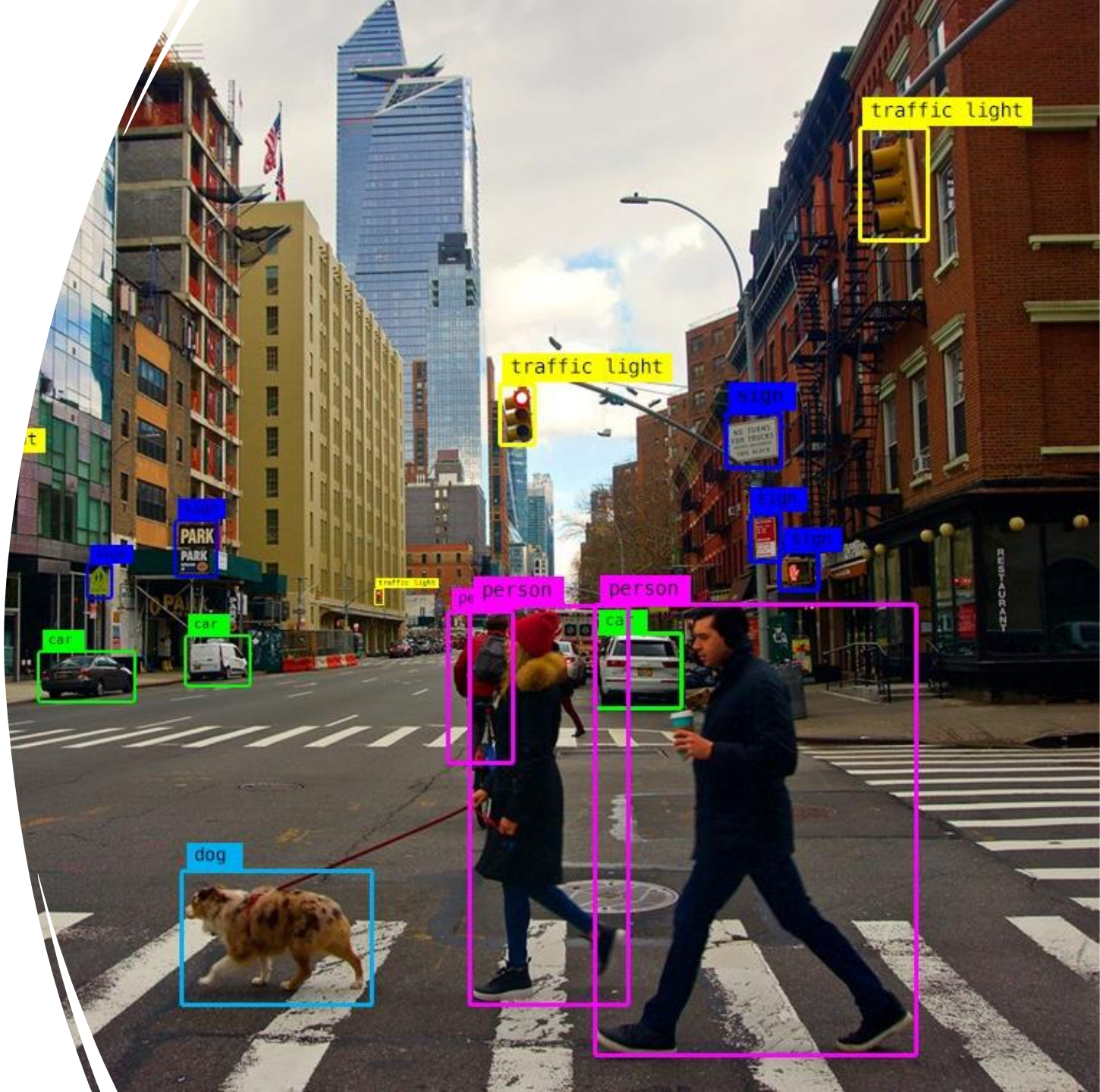
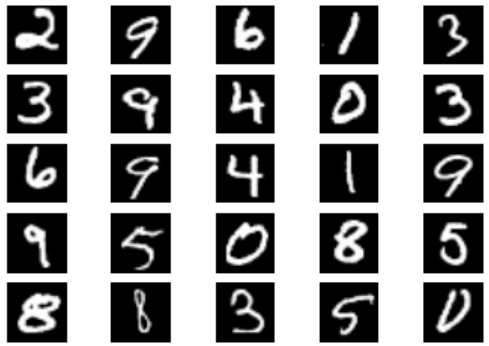


Image classification

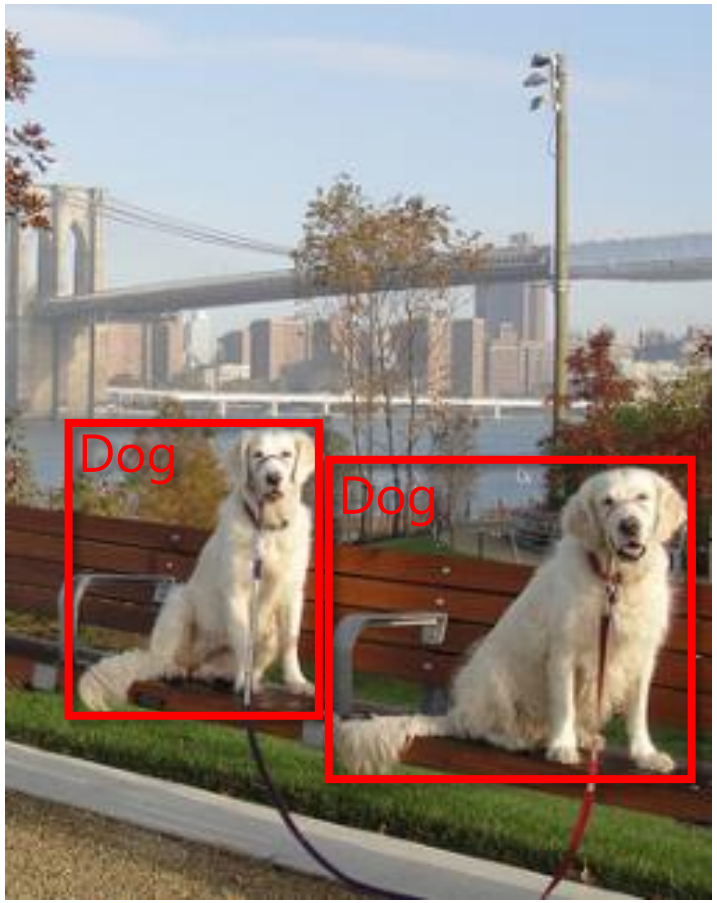
- K classes
- Task: assign correct class label to the whole image



Digit classification

Object recognition

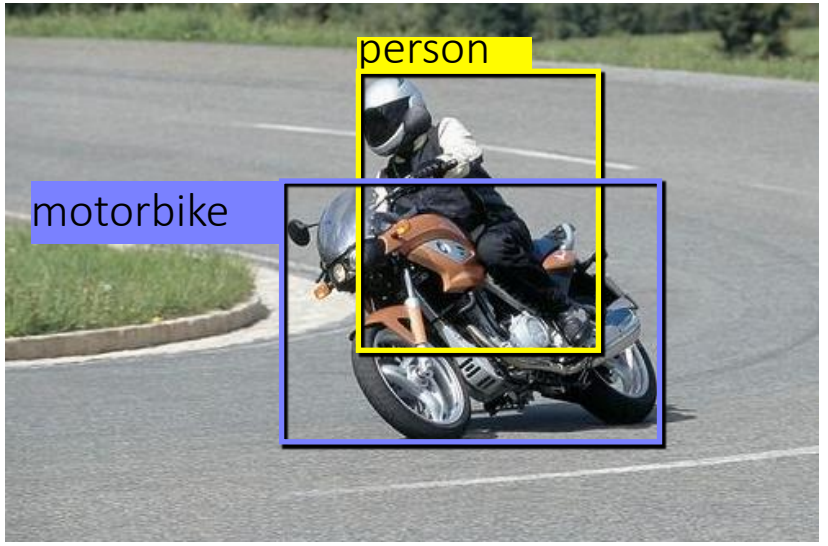
Classification vs. Detection



Problem formulation

{ airplane, bird, motorbike, person, sofa }

Input



Desired output

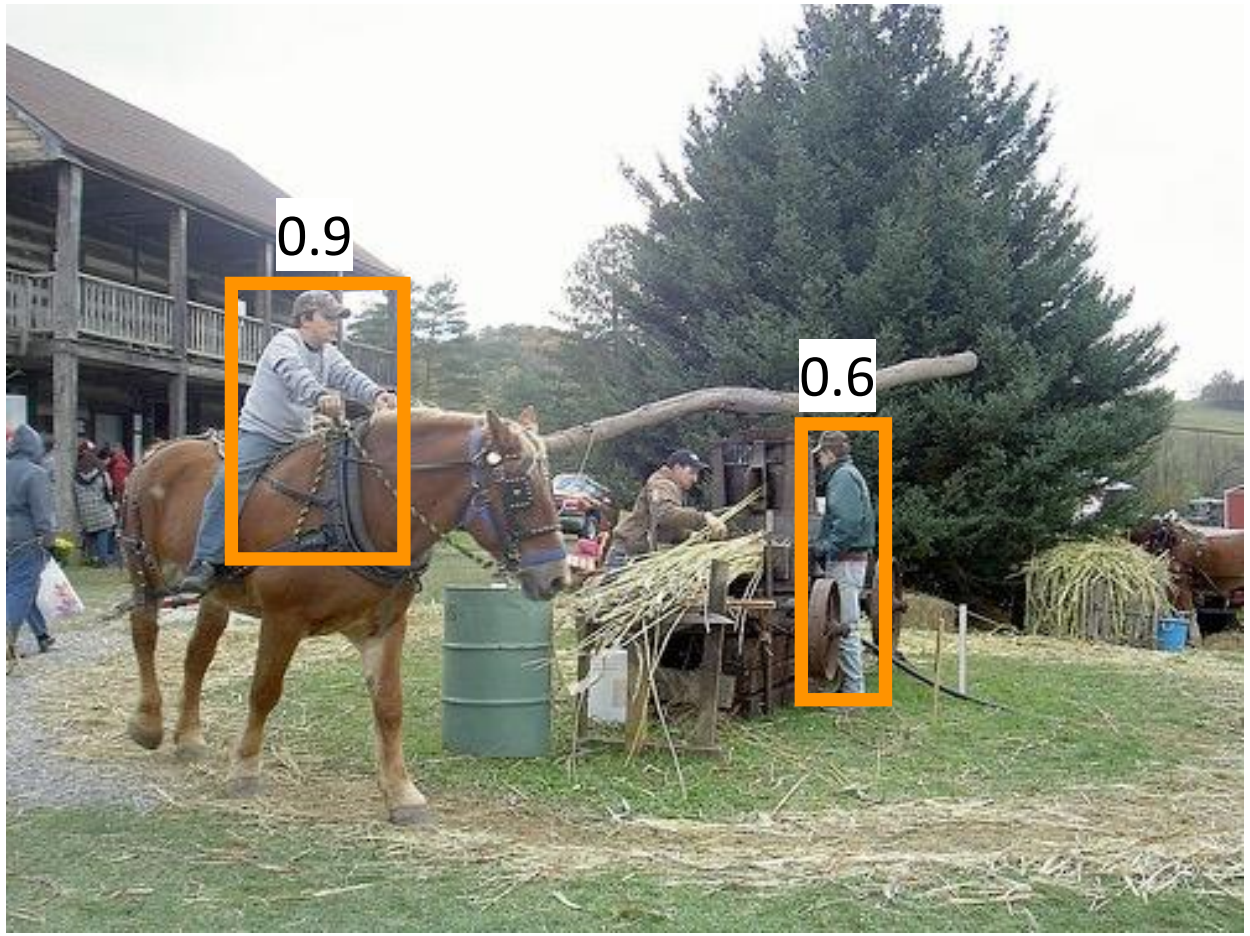
Evaluating a detector

Test image (previously unseen)

First detection ...

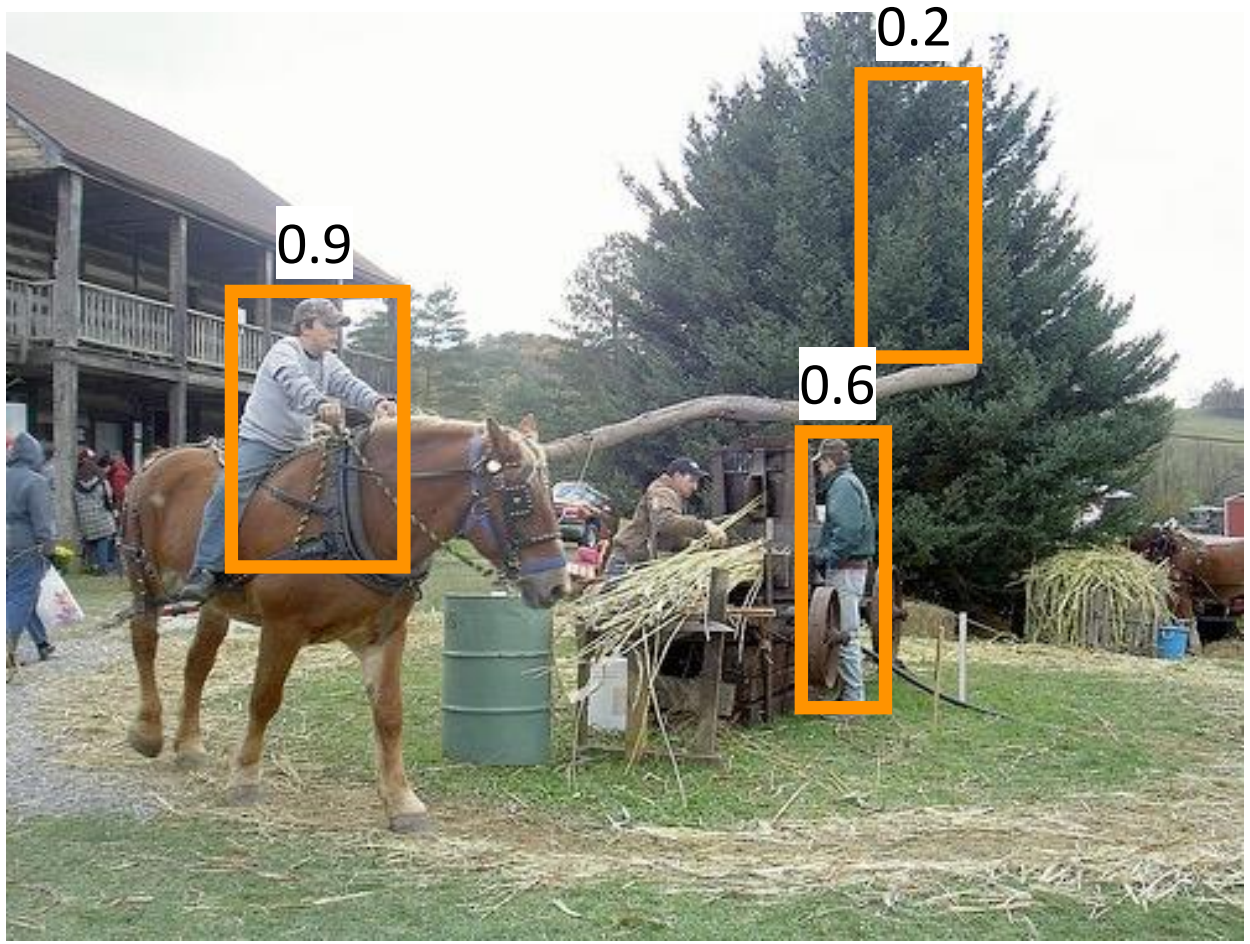
‘person’ detector predictions

Second detection ...



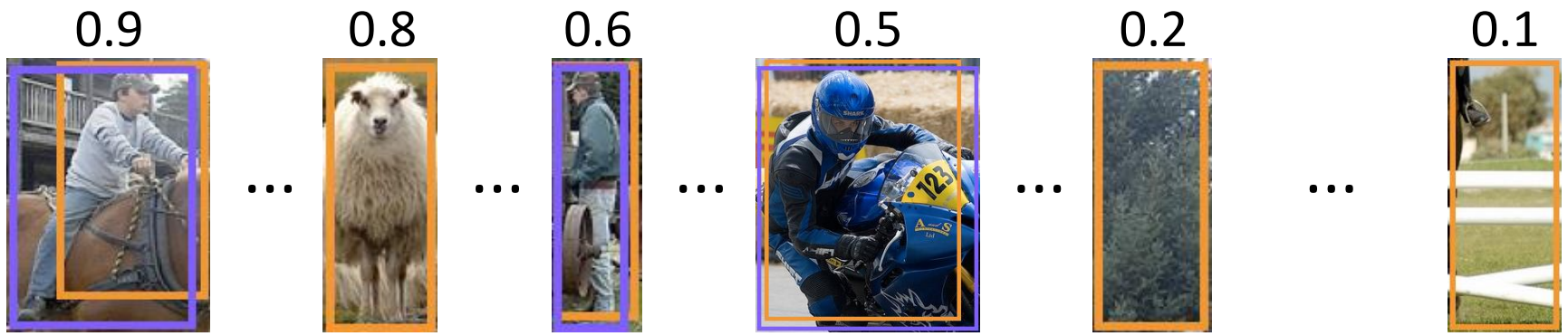
‘person’ detector predictions

Third detection …



‘person’ detector predictions

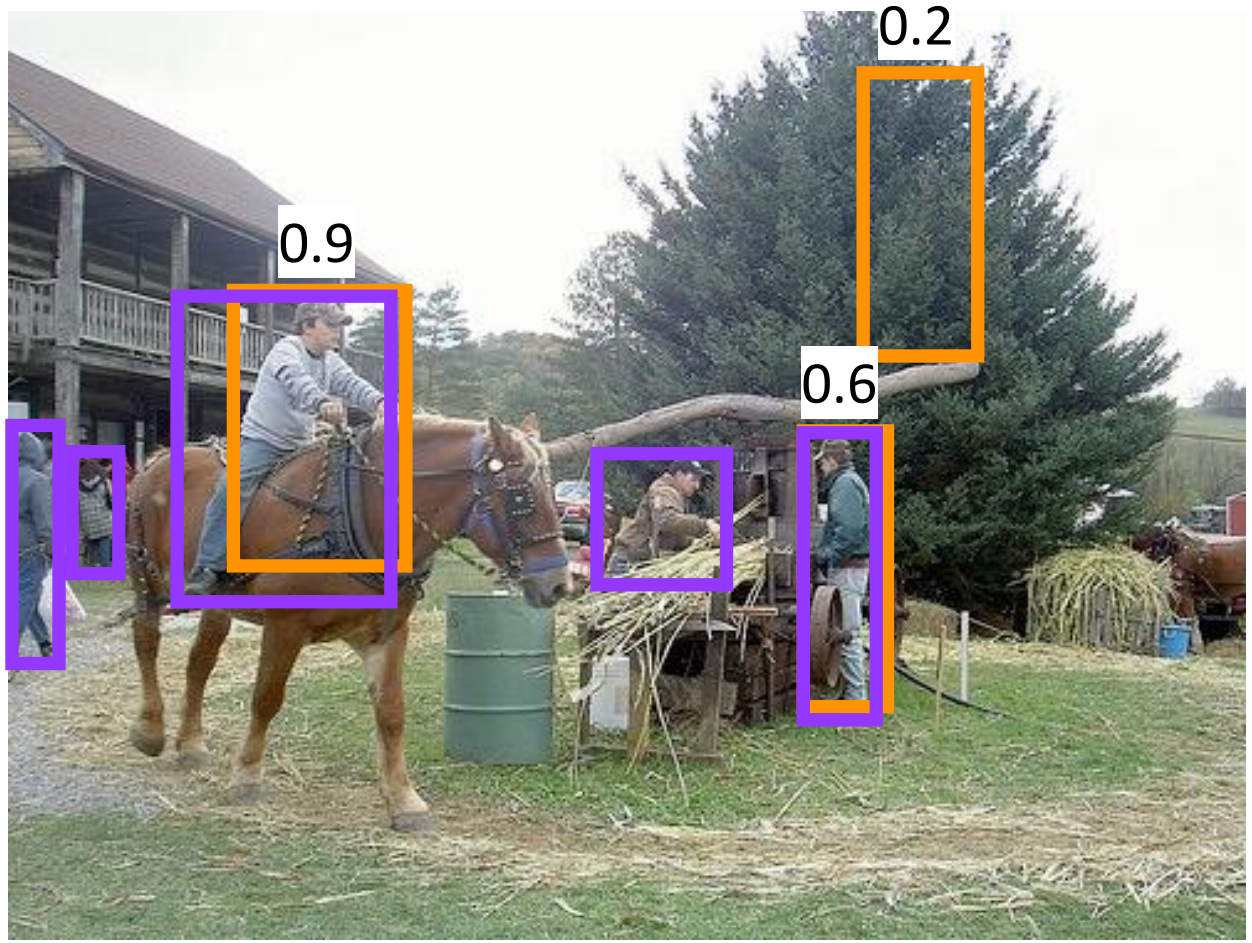
Sort by confidence



true positive
(high overlap)

false positive
(no overlap, low overlap, or duplicate)

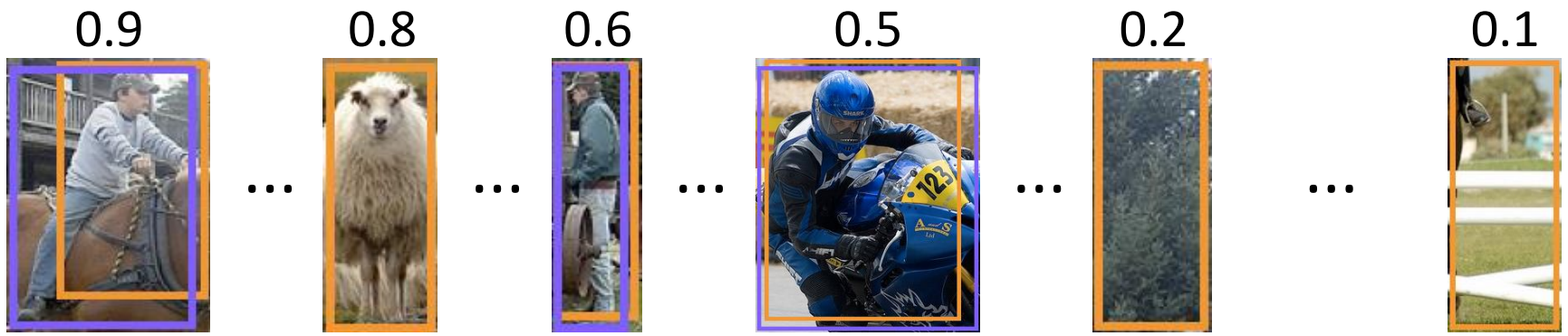
Compare to ground truth



‘person’ detector predictions

ground truth ‘person’ boxes

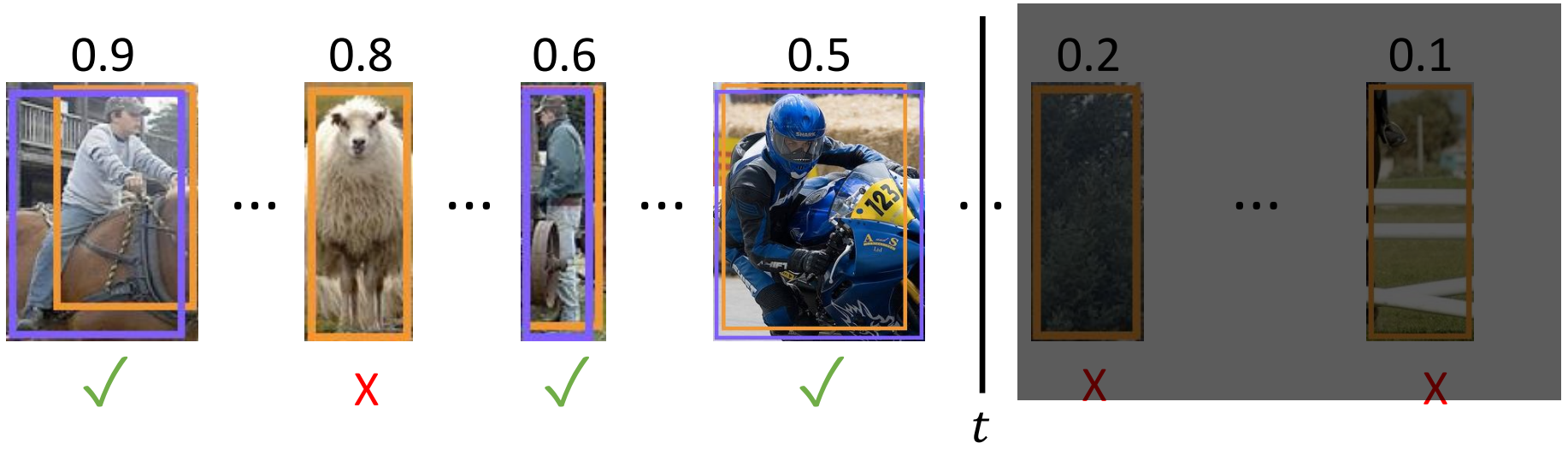
Sort by confidence



true positive
(high overlap)

false positive
(no overlap, low overlap, or duplicate)

Evaluation metric

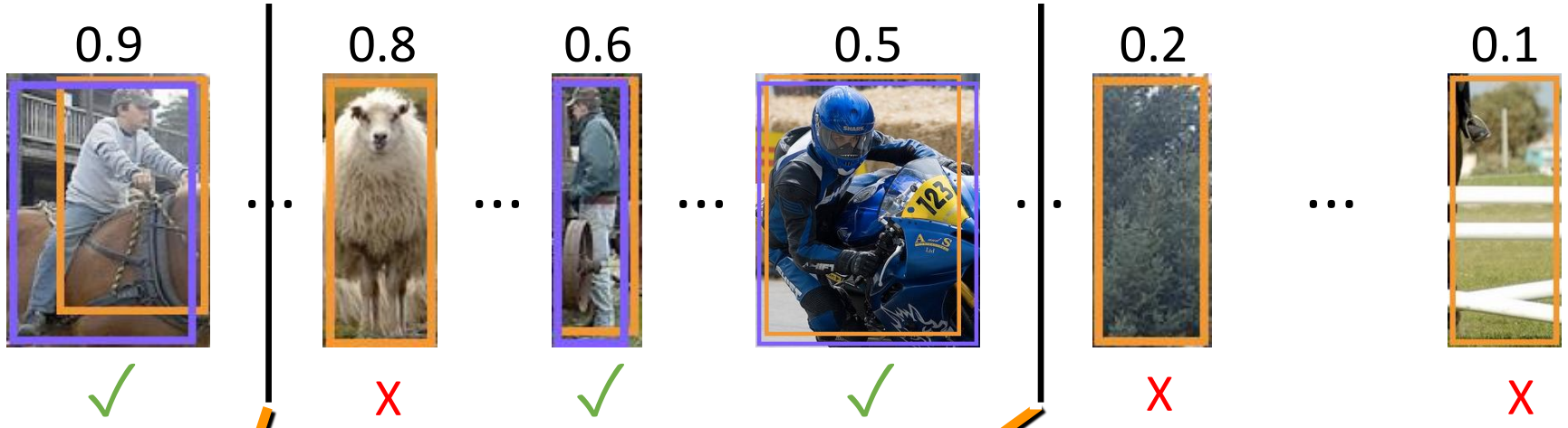


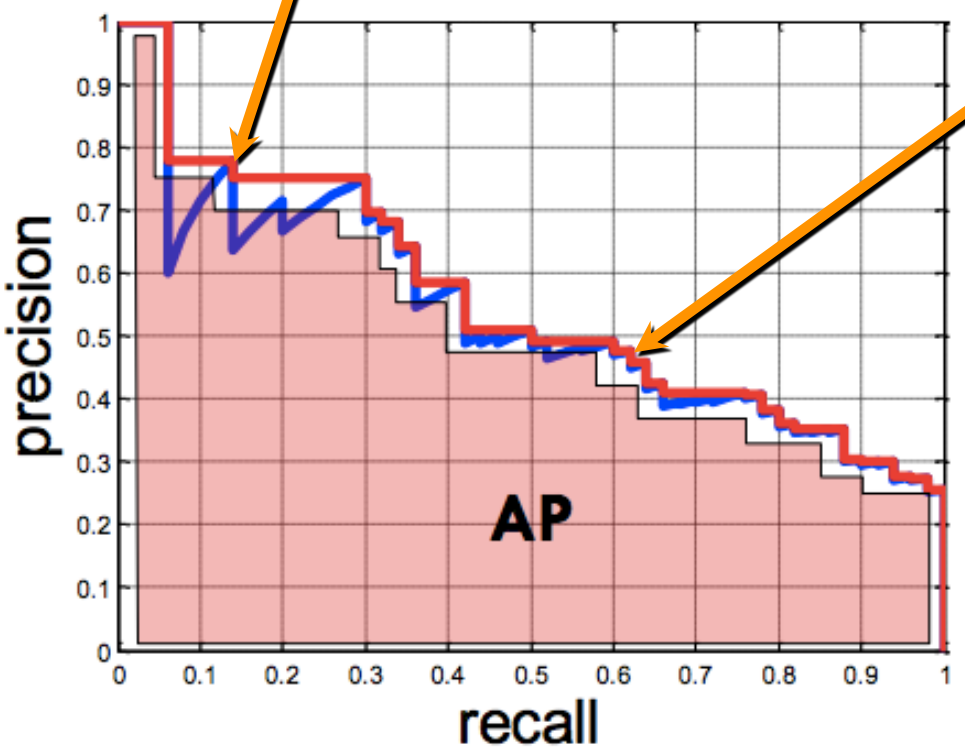
$$precision@t = \frac{\#true\ positives@t}{\#true\ positives@t + \#false\ positives@t}$$

$$\frac{\checkmark}{\checkmark + X}$$

$$recall@t = \frac{\#true\ positives@t}{\#ground\ truth\ objects}$$

Evaluation metric





Average Precision (AP)

0% is worst

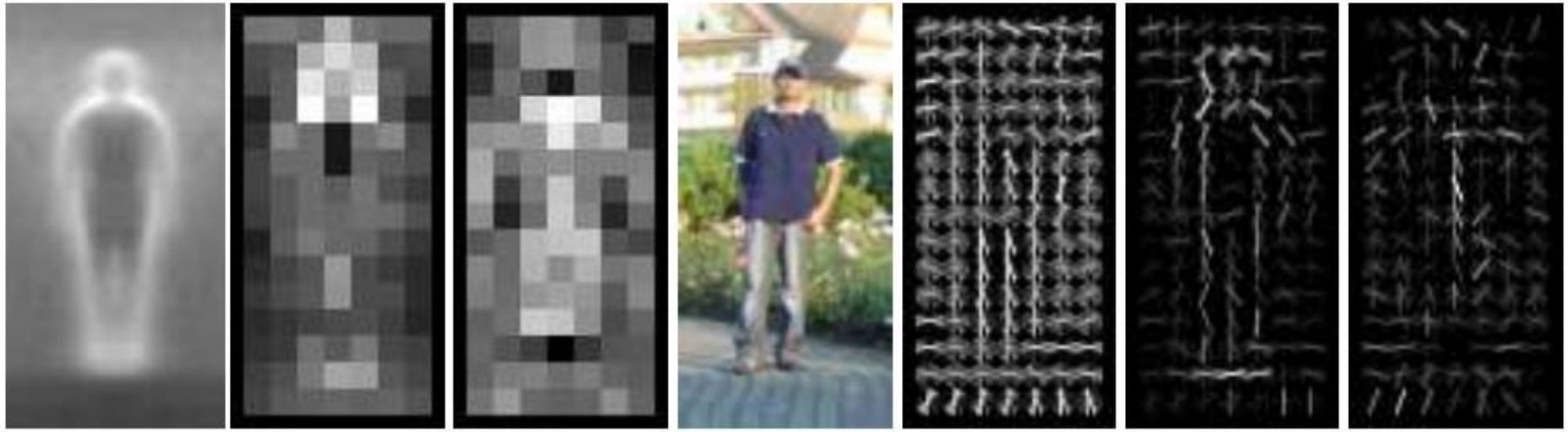
100% is best

mean AP over classes (mAP)

Pedestrians

Histograms of Oriented Gradients for Human Detection,
Dalal and Triggs, CVPR 2005

AP ~77%
More sophisticated methods: AP ~90%



(a) average gradient image over training examples
(b) each “pixel” shows max positive SVM weight in the block centered on that pixel
(c) same as (b) for negative SVM weights
(d) test image
(e) its R-HOG descriptor
(f) R-HOG descriptor weighted by positive SVM weights
(g) R-HOG descriptor weighted by negative SVM weights

Why did it work?

Average gradient image

Quiz time

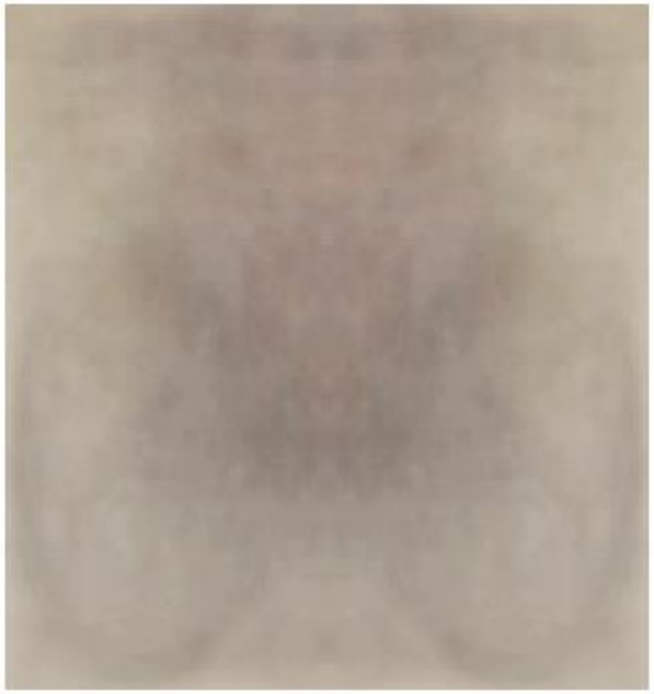
Warm up

This is an average image of which object class?

Warm up

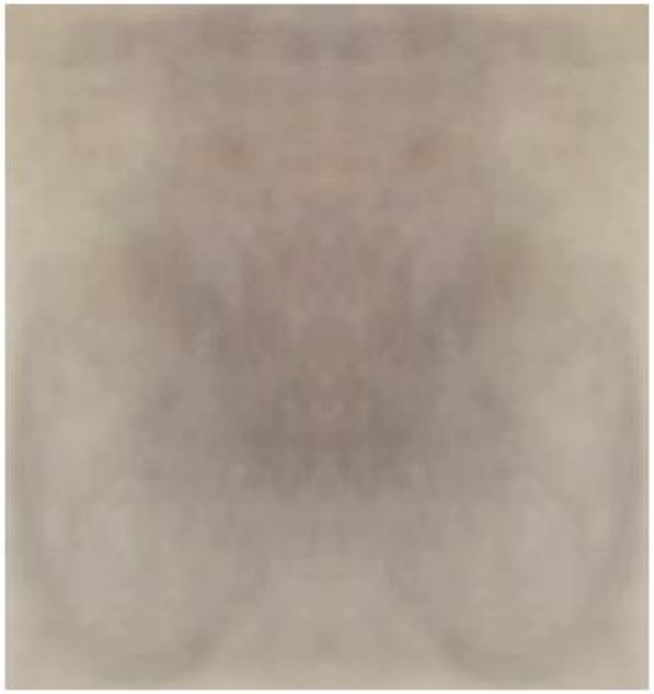
pedestrian

A little harder



?

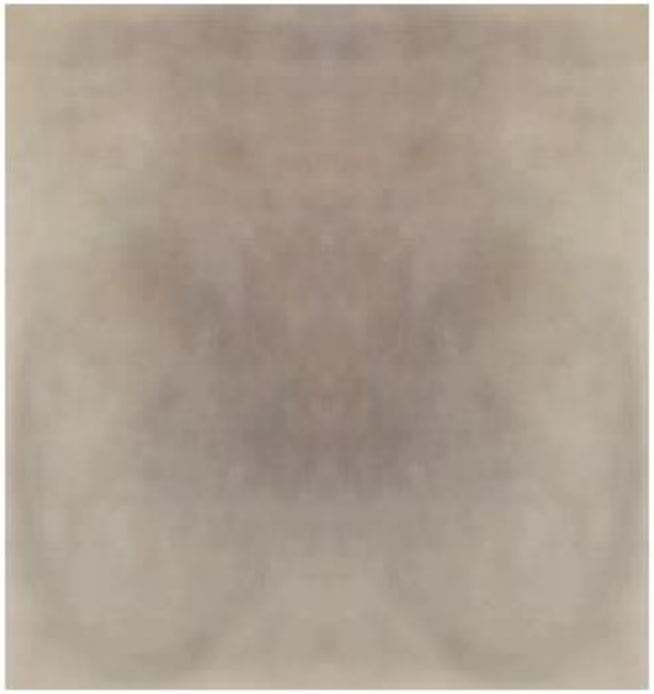
A little harder



?

Hint: airplane, bicycle, bus, car, cat, chair, cow, dog, dining table

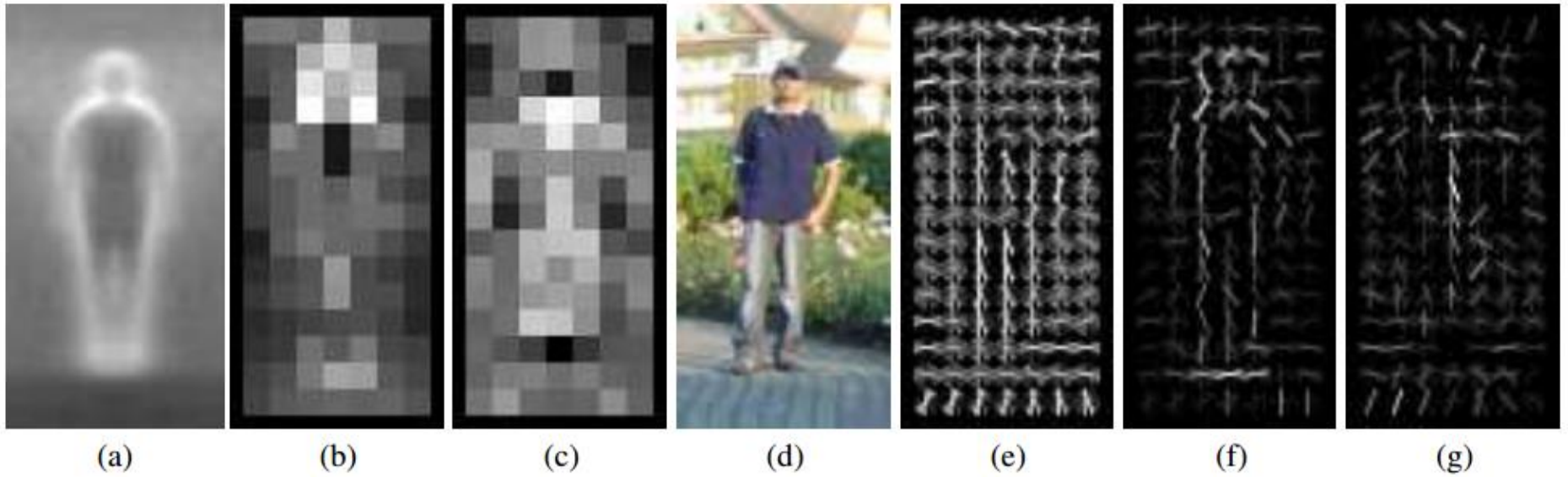
A little harder



bicycle (PASCAL)

The key to SVMs

- *It's all about the features*



Histograms of Oriented Gradients for Human Detection,
Dalal and Triggs, CVPR 2005

Core idea of “deep learning”

- Input: the “*raw*” signal (image, waveform, …)

- Features: hierarchy of features is *learned* from the raw input

Classical Machine Learning

Supervised Learning

(Pre Categorized Data)

Predications & Predictive Models

Classification

(Divide the socks by Color)

Eg. Identity Fraud Detection

Regression

(Divide the Ties by Length)

Eg. Market Forecasting

Unsupervised Learning

(Unlabelled Data)

Pattern/ Structure Recognition

Clustering

(Divide by Similarity)

Eg. Targeted Marketing

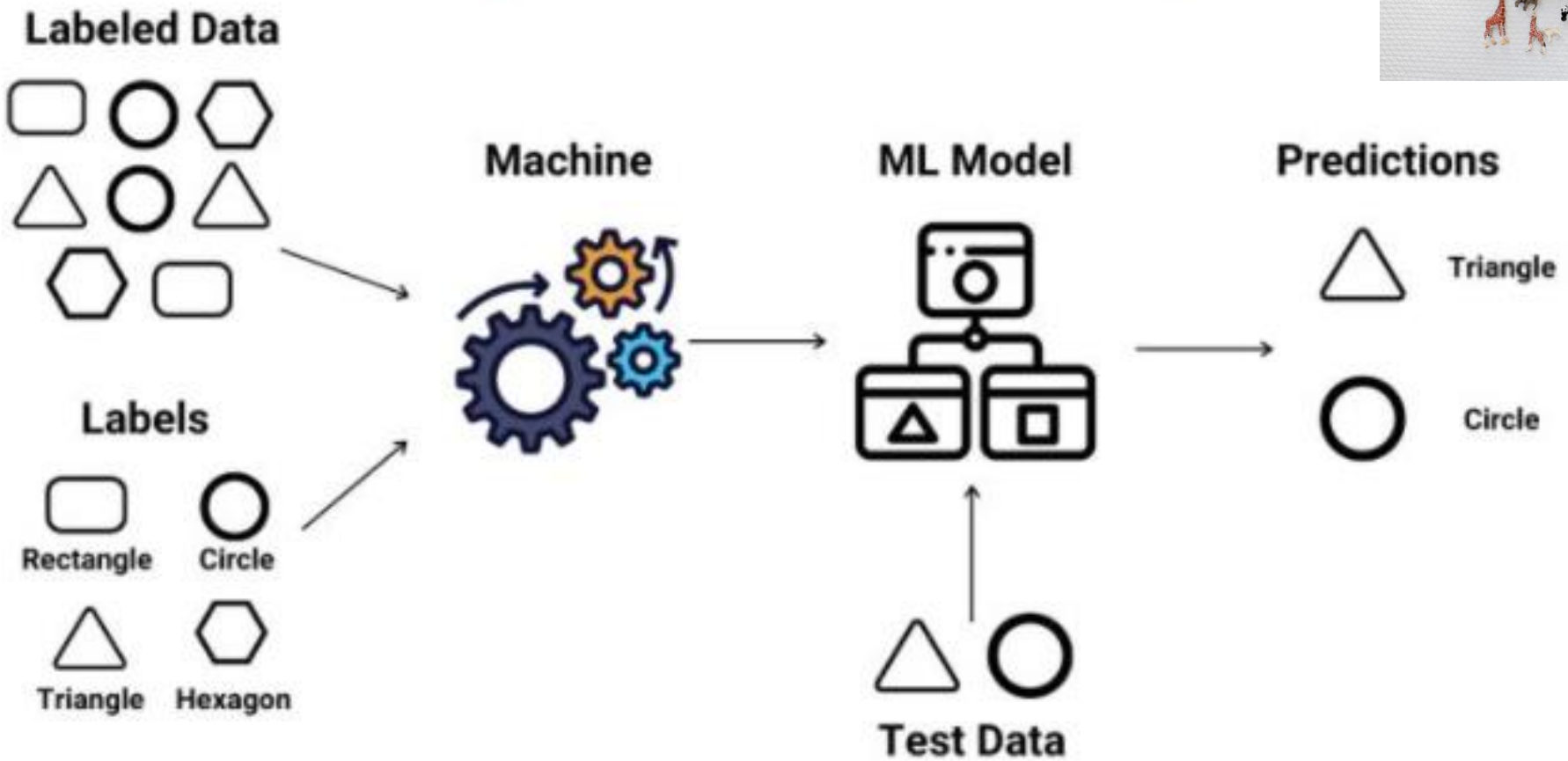
Association

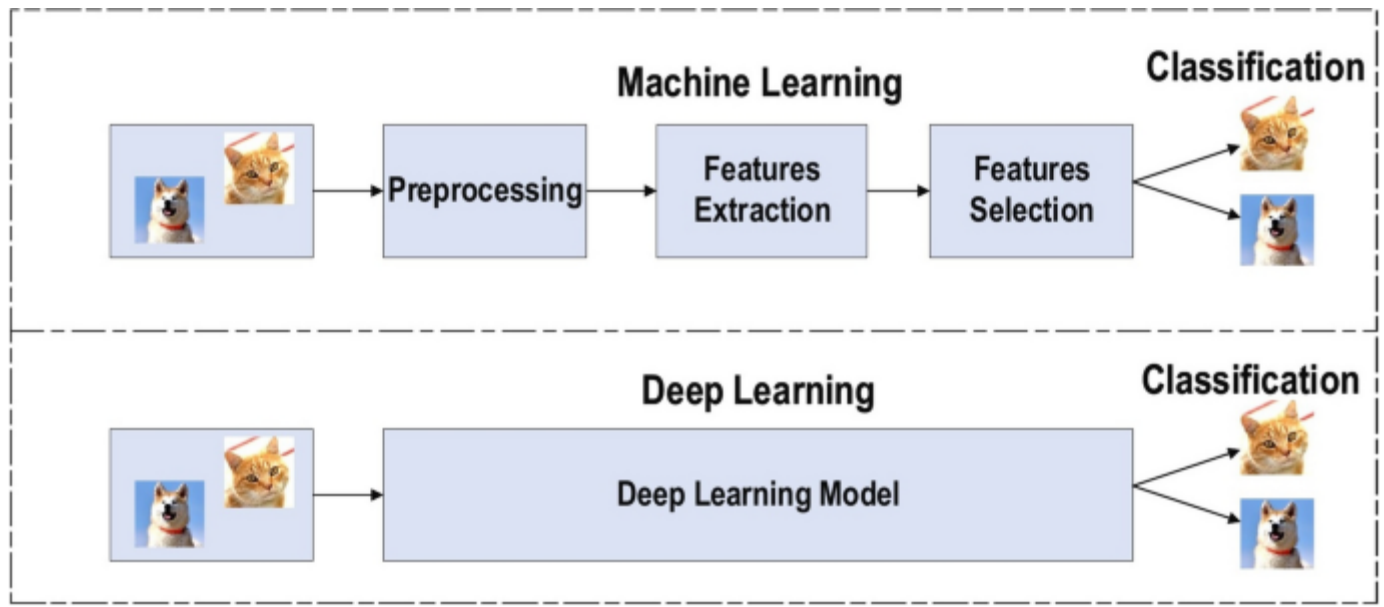
(Identify Sequences)

Eg. Customer Recommendation

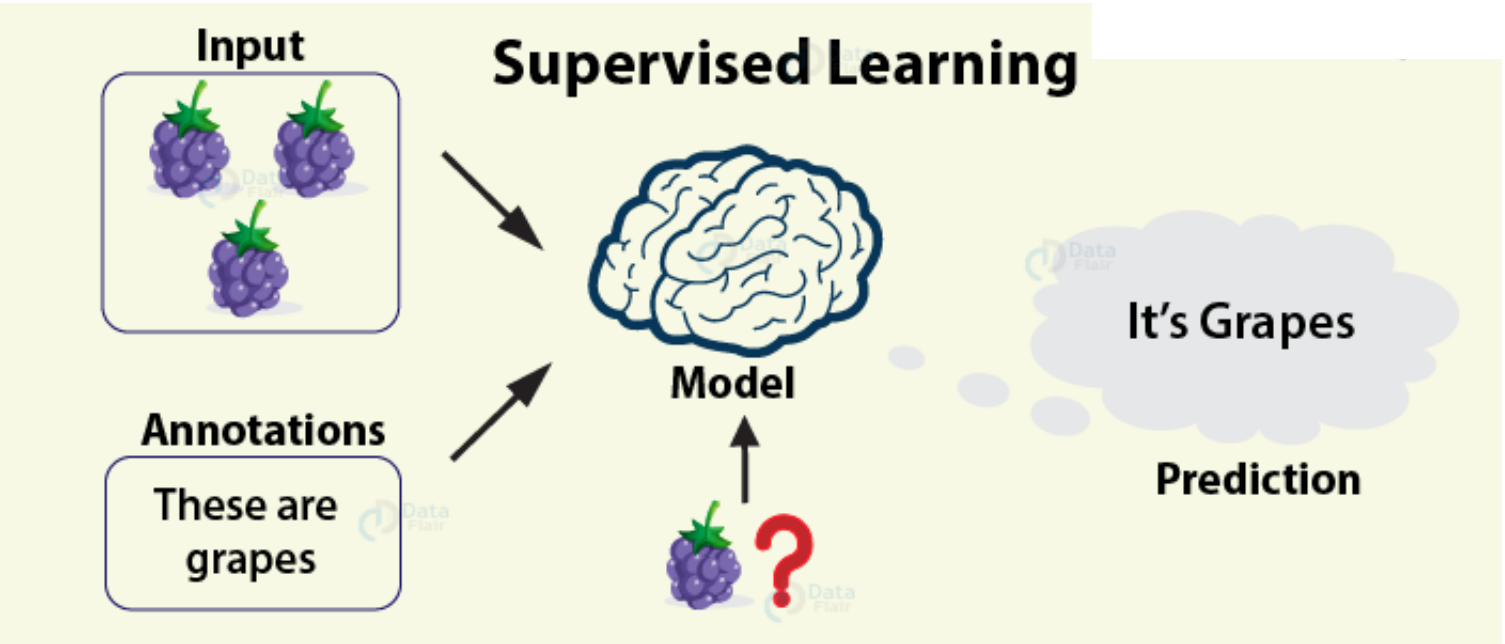
Why do we call it Supervised Learning?

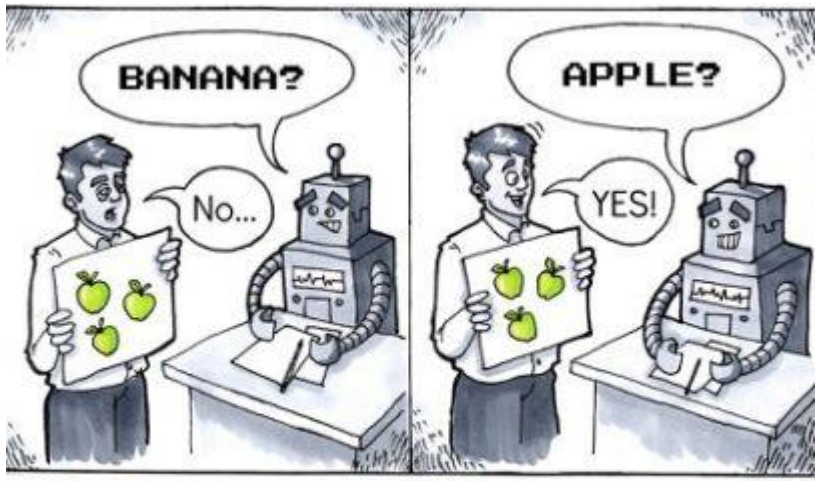
Supervised Learning




Machine Learning
Classification
Preprocessing
Features Extraction
Features Selection
Deep Learning
Classification
Deep Learning Model

Supevised VS Unsupevised

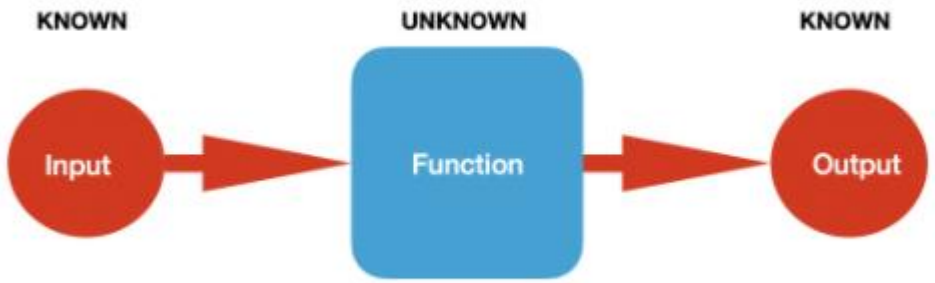




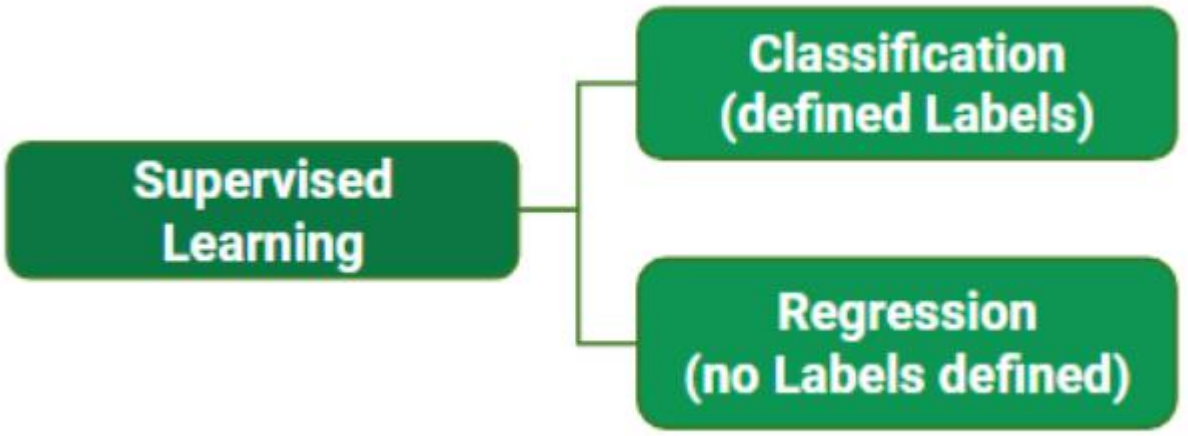
Supervised Learning

Supervised Learning

- Algorithm that learn from training dataset can be thought of as a guide supervising the learning process.
- We already know the correct answers! The algorithm iteratively makes predictions on the training data and corrected by the guide.
- Learning stops when algorithm achieves an acceptable level of performance.



Types of Supervised Learning



Discrete Values

Continuous Values!

User ID	Gender	Age	Salary	Purchased
15624510	Male	19	19000	0
15810944	Male	35	20000	1
15668575	Female	26	43000	0
15603246	Female	27	57000	0
15804002	Male	19	76000	1
15728773	Male	27	58000	1
15598044	Female	27	84000	0
15694829	Female	32	150000	1
15600575	Male	25	33000	1
15727311	Female	35	65000	0
15570769	Female	26	80000	1
15606274	Female	26	52000	0
15746139	Male	20	86000	1
15704987	Male	32	18000	0
15628972	Male	18	82000	0
15697686	Male	29	80000	0
15733883	Male	47	25000	1

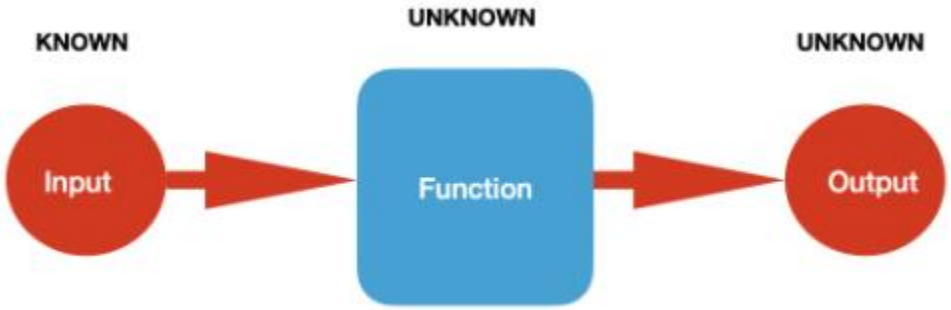
Figure A: CLASSIFICATION

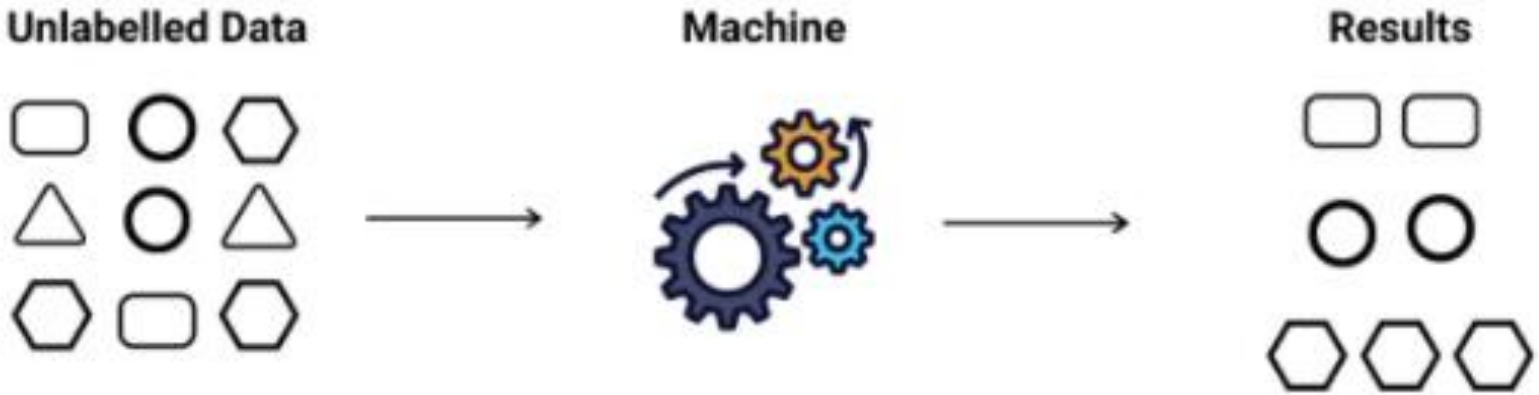
Temperature	Pressure	Relative Humidity	Wind Direction	Wind Speed
10.69261758	986.882019	54.19337313	195.7150879	3.278597116
13.59184184	987.8729248	48.0648859	189.2951202	2.909167767
17.70494885	988.1119385	39.11965597	192.9273834	2.973036289
20.95430404	987.8500366	30.66273218	202.0752869	2.965289593
22.9278274	987.2833862	26.06723423	210.6589203	2.798230886
24.04233986	986.2907104	23.46918024	221.1188507	2.627005816
24.41475295	985.2338867	22.25082295	233.7911987	2.448749781
23.93361956	984.8914795	22.35178837	244.3504333	2.454271793
22.68800023	984.8461304	23.7538641	253.0864716	2.418341875
20.56425726	984.8380737	27.07867944	264.5071106	2.318677425
17.76400389	985.4262085	33.54900114	280.7827454	2.343950987
11.25680746	988.9386597	53.74139903	68.15406036	1.650191426
14.37810685	989.6819458	40.70884681	72.62069702	1.553469896
18.45114201	990.2960205	30.85038484	71.70604706	1.005017161
22.54895853	989.9562988	22.81738811	44.66042709	0.264133632
24.23155922	988.796875	19.74790765	318.3214111	0.329656571

Figure B: REGRESSION

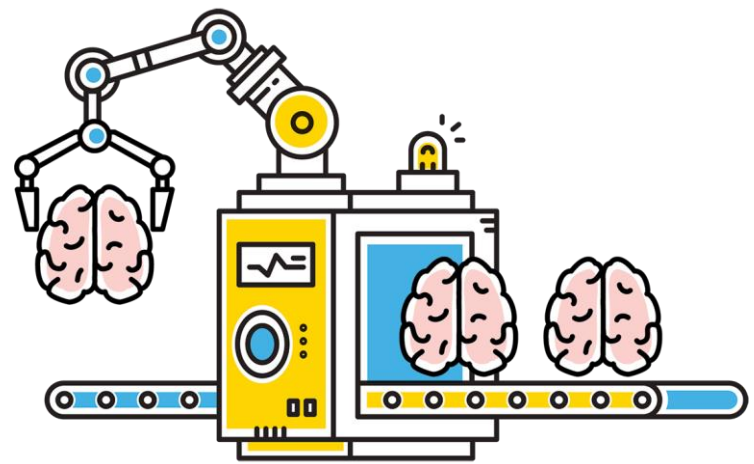
Unsupervised Learning

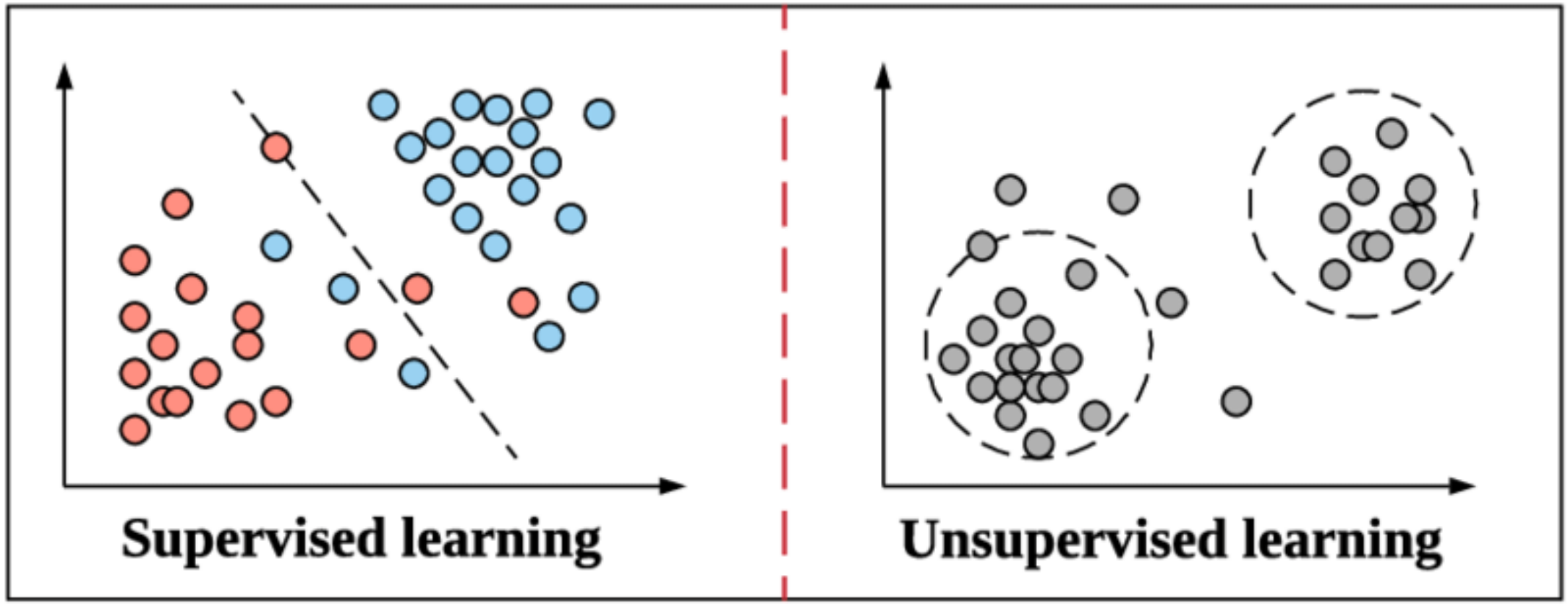
- We only have the input data to feed the model but no corresponding data
- We know the value of input but the output and the mapping function
- In such scenarios, machine learning algorithms find the function that finds similarity among different input data instances based on the similarity index, which is the output of unsupervised learning



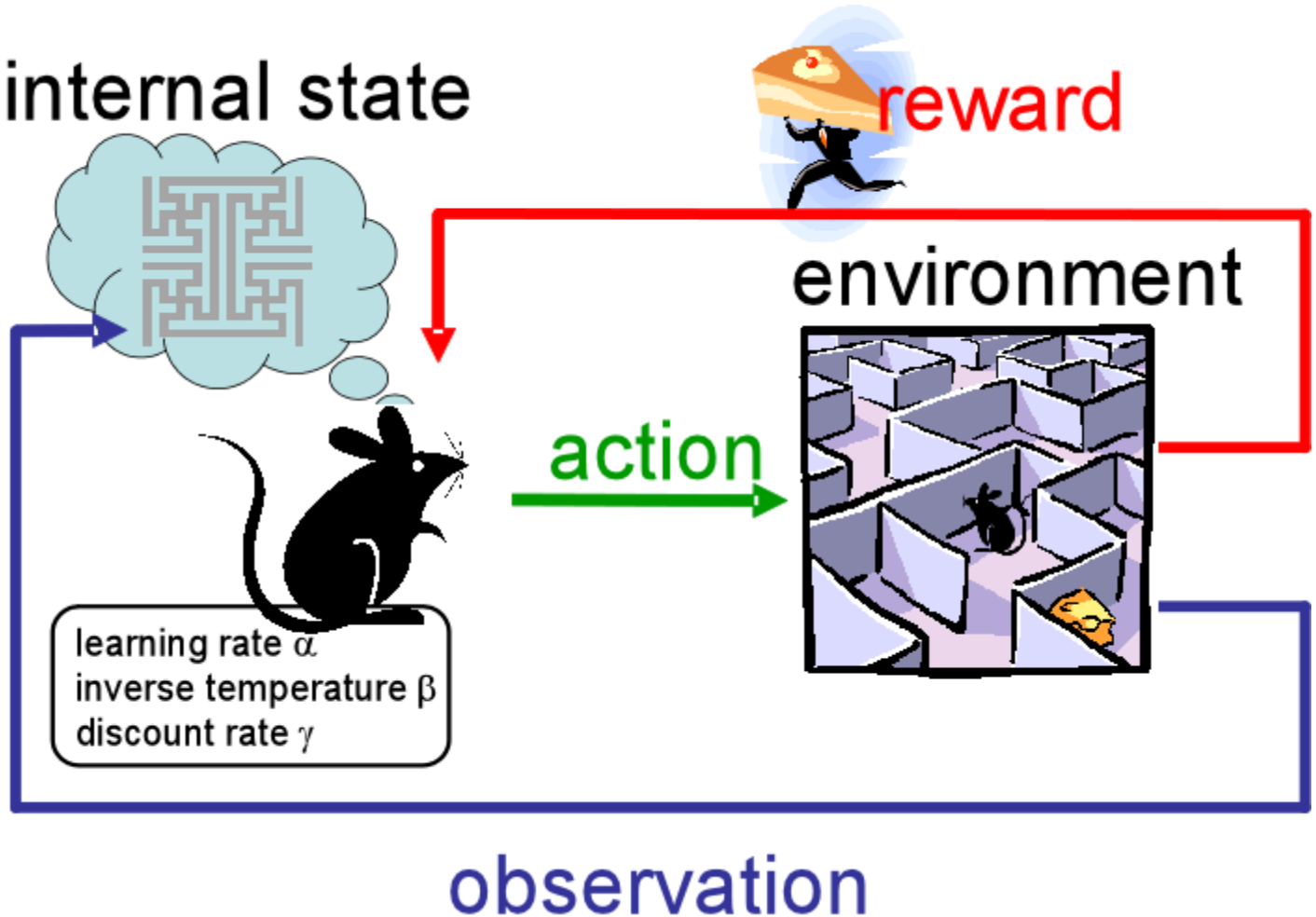

Unlabelled Data
Machine
Results


CUSTOMERS WHO BOUGHT THIS ITEM:
ALSO BOUGHT:
Piccolo




Supervised learning
Unsupervised learning

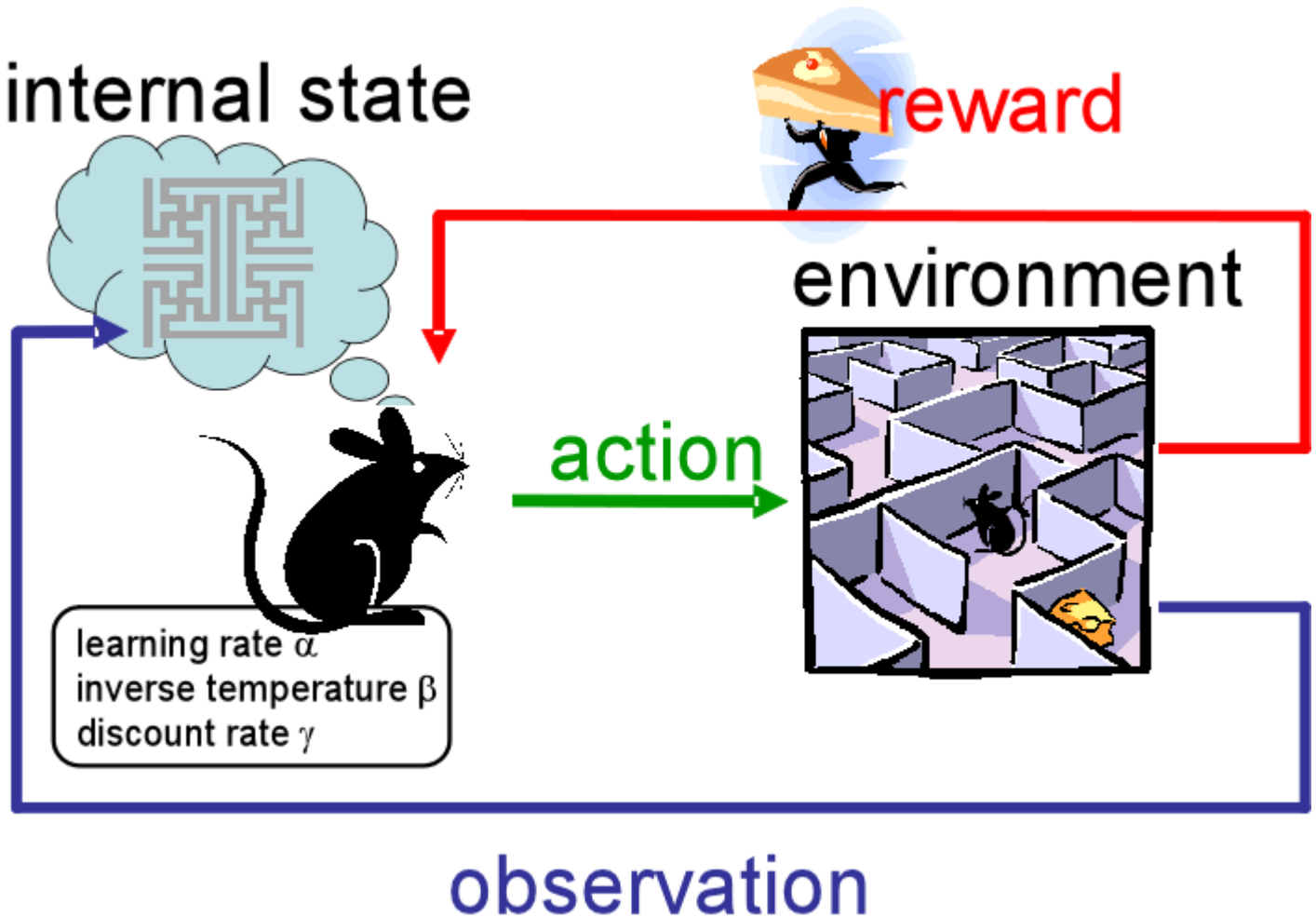
Reinforcement Learning

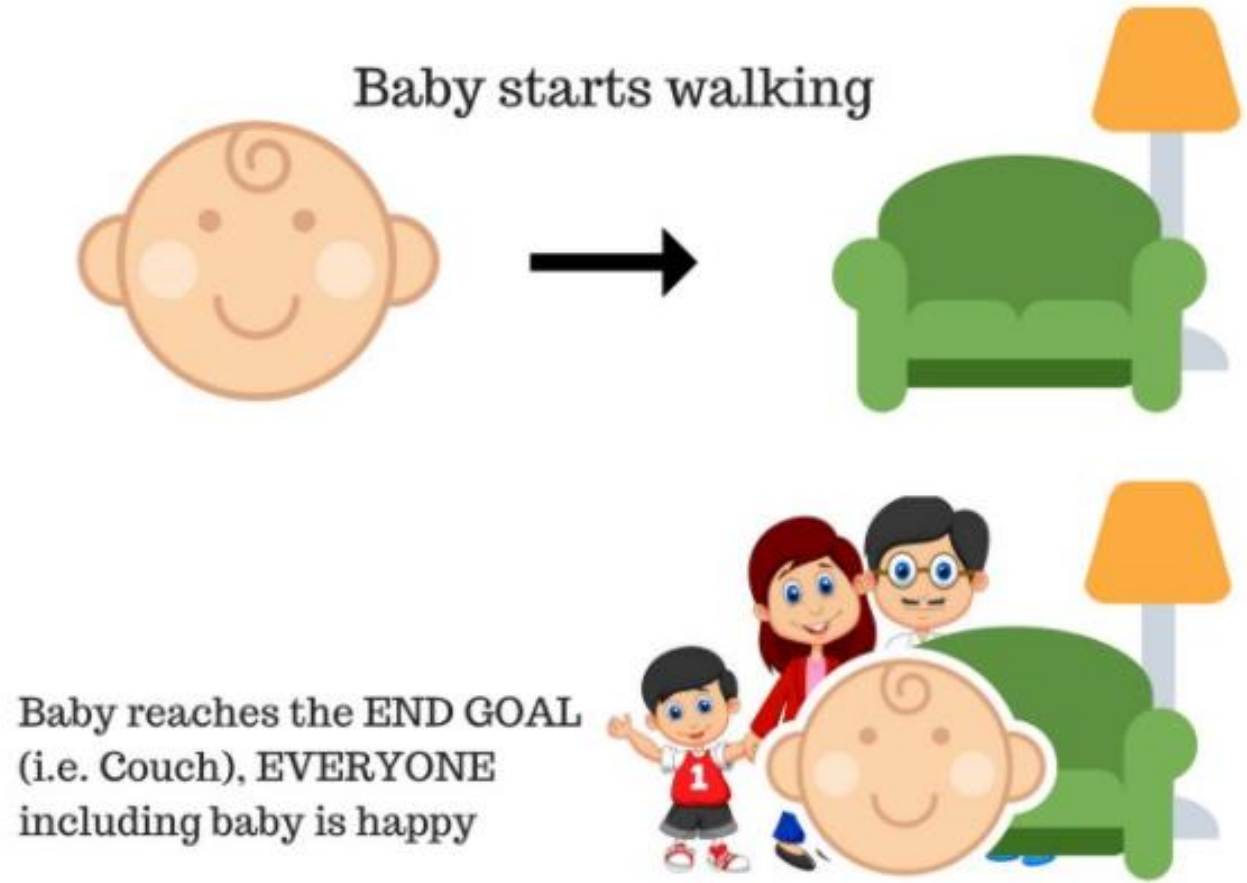


Reinforcement Learning

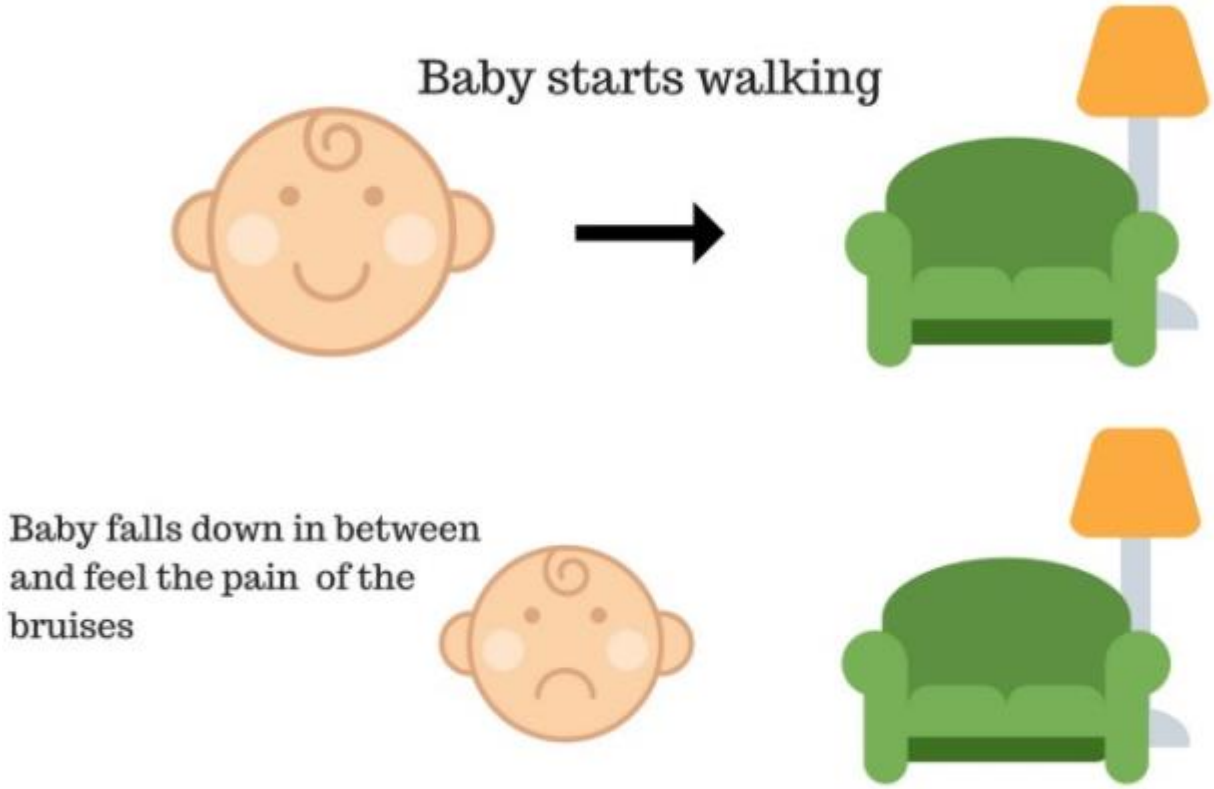
- Reinforcement Learning is an aspect of Machine learning where an agent learns to behave in an environment, by performing certain actions and observing the rewards/results which it get from those actions.

Reinforcement Learning





So, the baby is happy and receives appreciation from her parents. It's positive — the baby feels good *(Positive Reward +n)*.

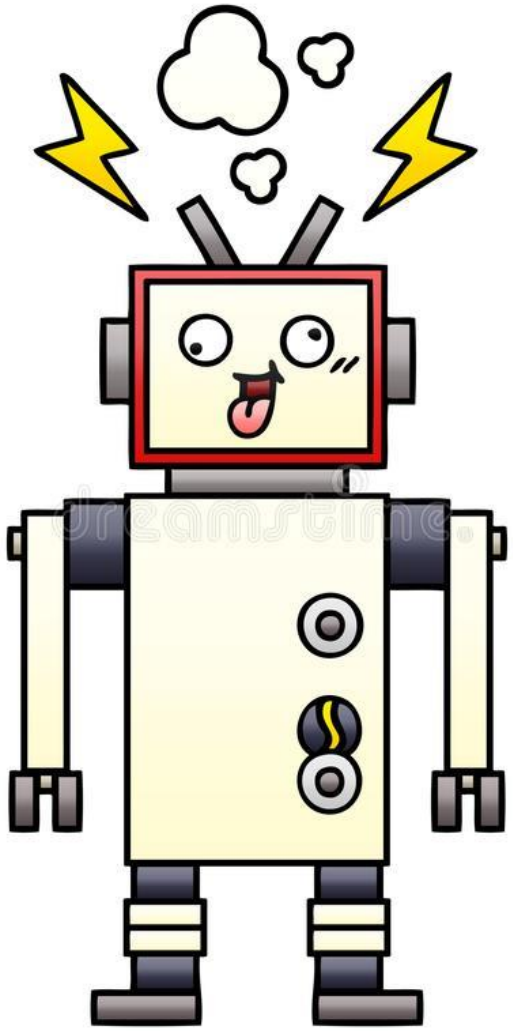


Ouch! The baby gets hurt and is in pain. It's negative — the baby cries *(Negative Reward -n).*

That's how we humans learn — by trail and error. Reinforcement learning is conceptually the same, but is a computational approach to learn by actions.

Difference between SL and RL?

- Supervised learning is **when a model learns from a labeled dataset with guidance**. ... Whereas reinforcement learning is when a machine or an agent interacts with its environment, performs actions, and learns by a trial-and-error method.



Things can get confusing and crazy! But machine learning is still fun!