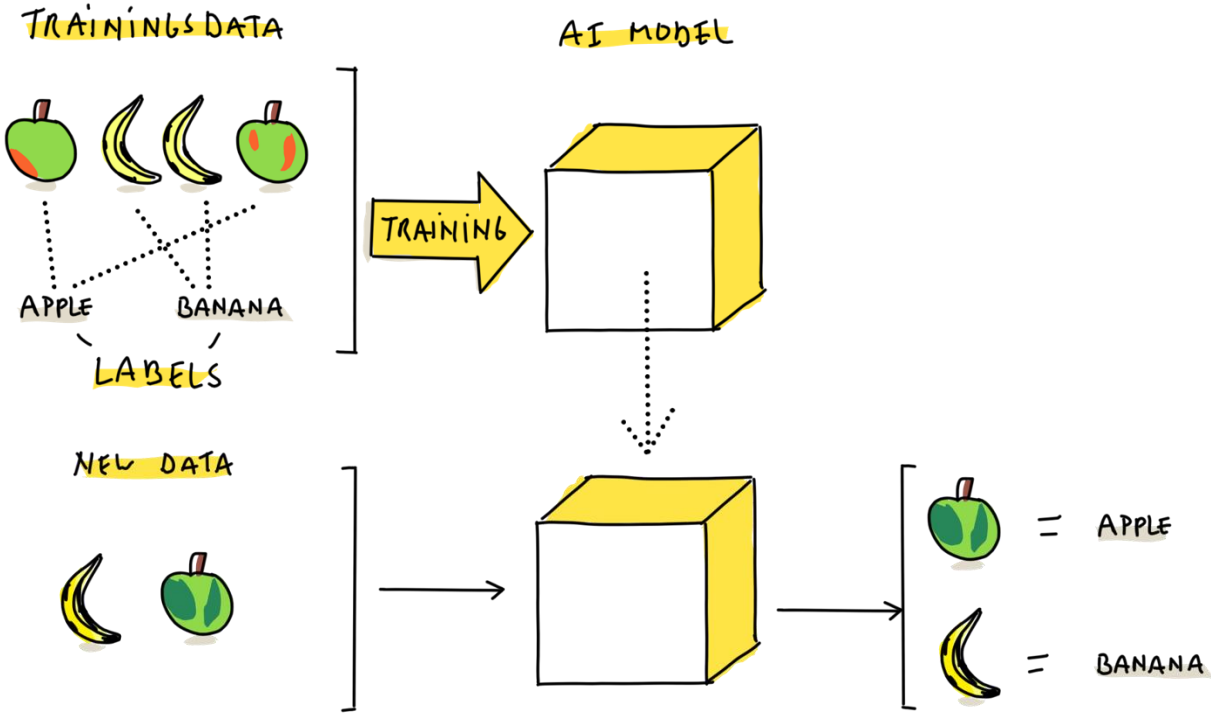
DEEL 2

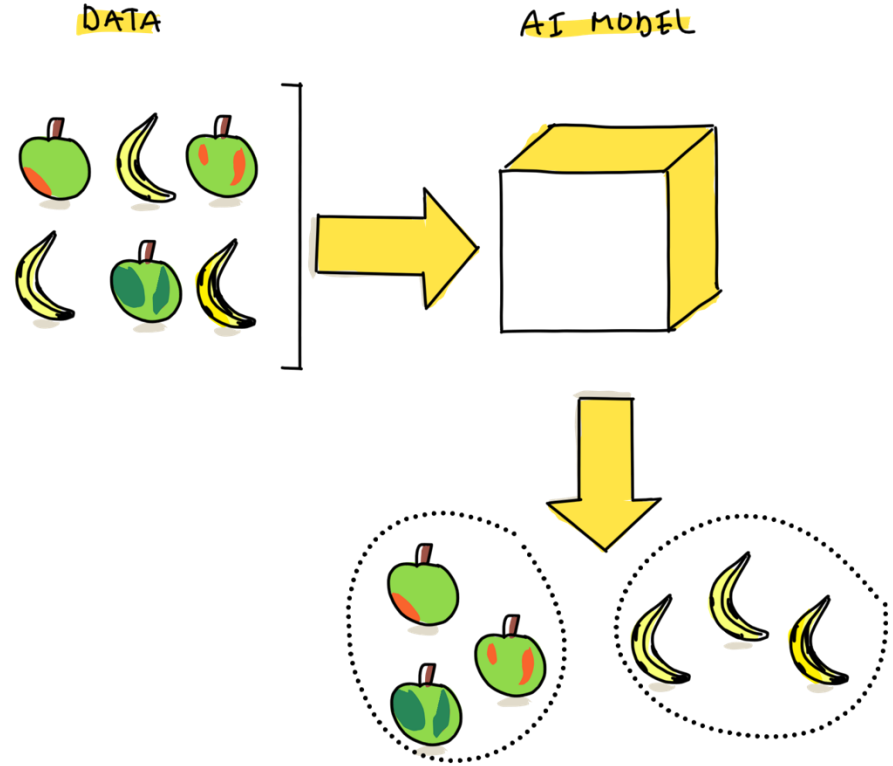
MACHINE LEARNING

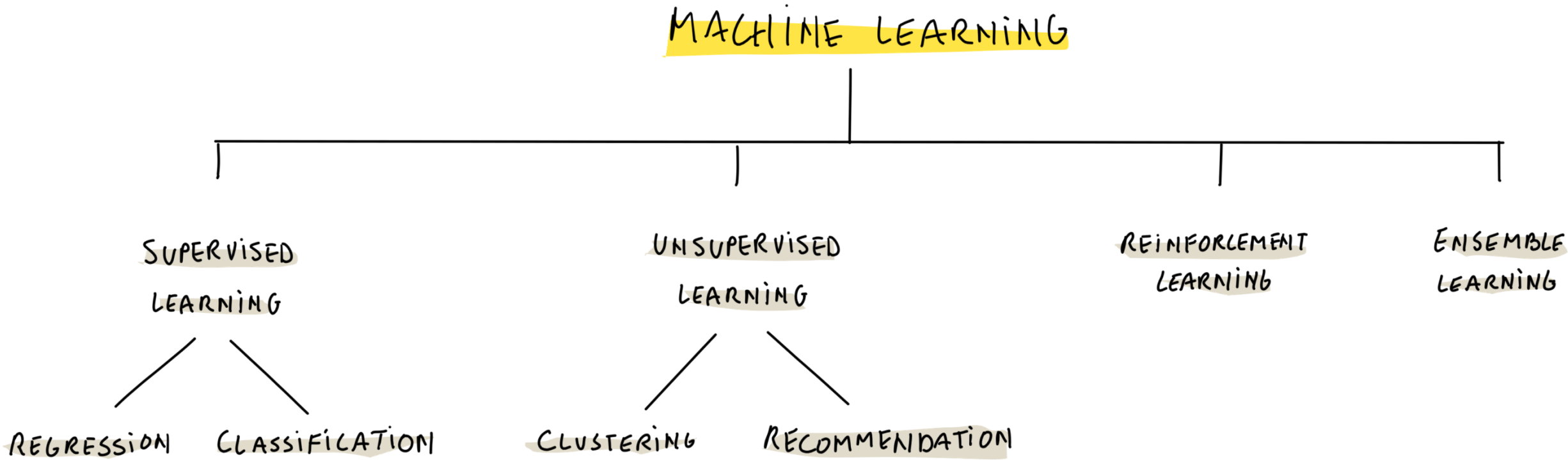
Machine learning is a **subset of AI**, which uses algorithms that **learn from data** to make predictions. These predictions can be generated through **supervised learning**, where algorithms learn patterns from existing data, or **unsupervised learning**, where they discover general patterns in data.

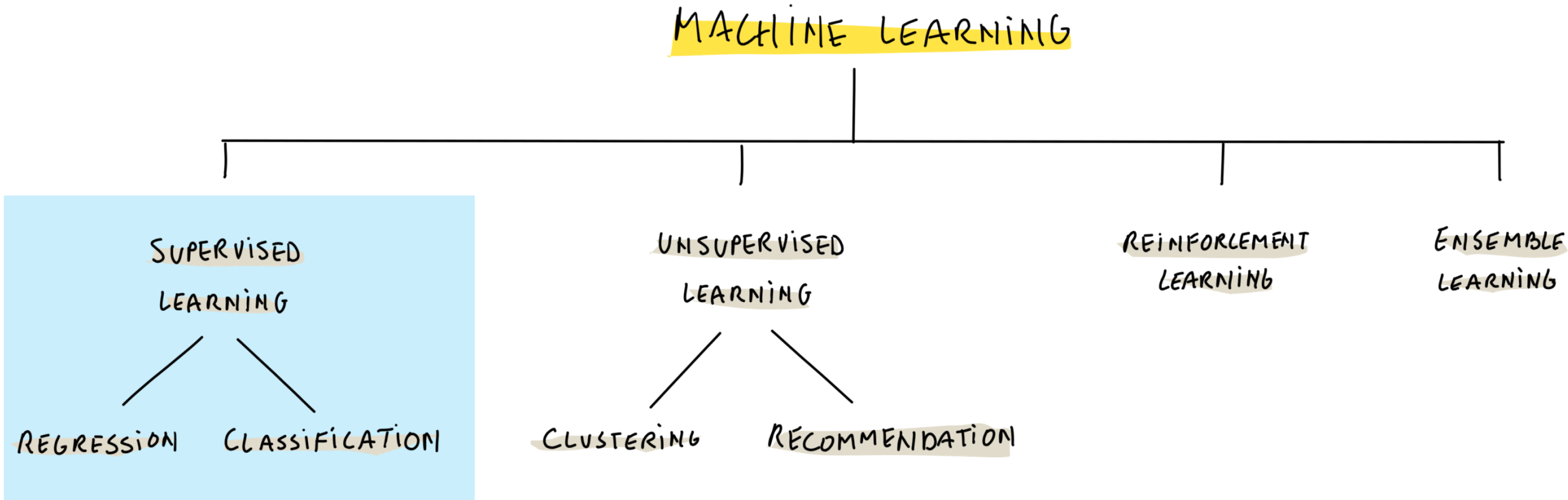
SUPERVISED LEARNING



UNSUPERVISED LEARNING

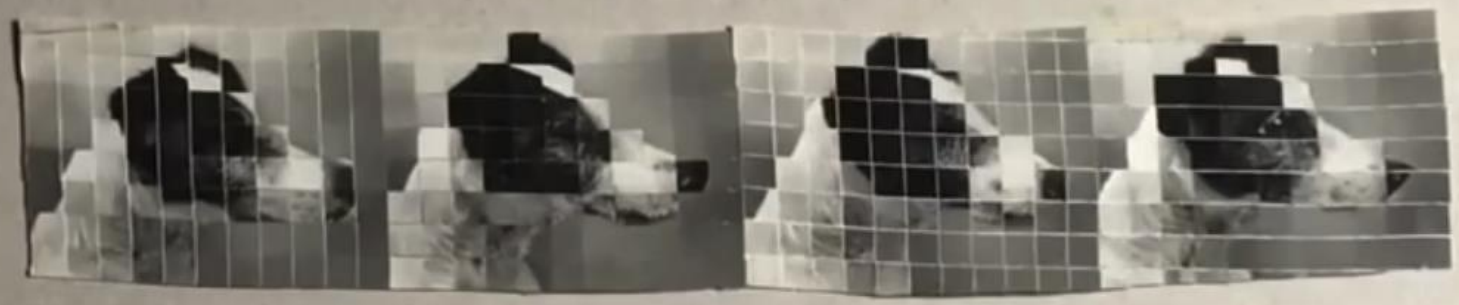



MACHINE LEARNING
SUPERVISED LEARNING
UNSUPERVISED LEARNING
REINFORCEMENT LEARNING
ENSEMBLE LEARNING
REGRESSION
CLASSIFICATION
CLUSTERING
RECOMMENDATION

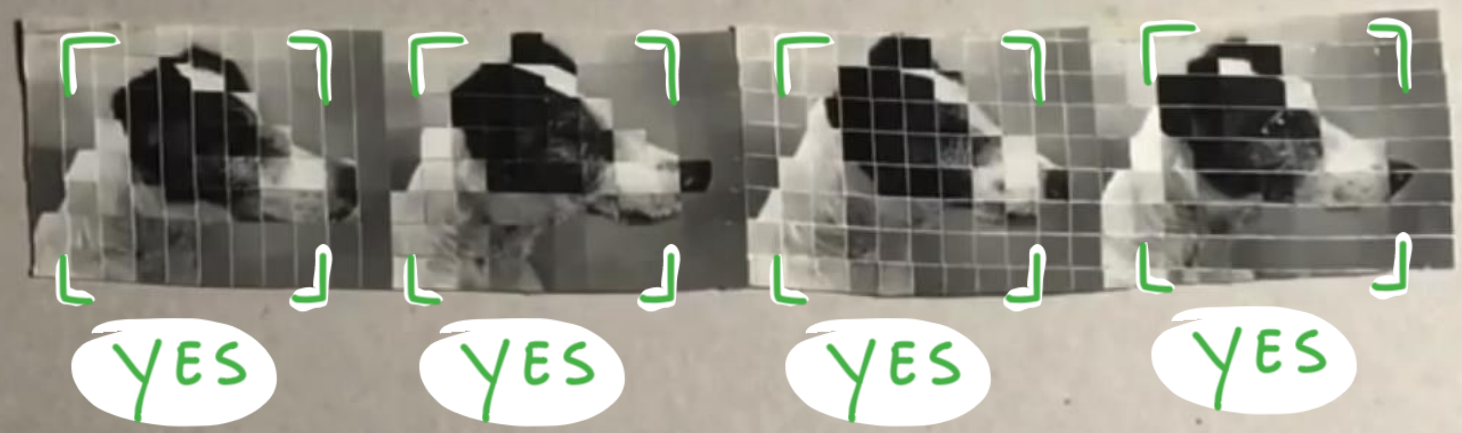

MACHINE LEARNING
SUPERVISED LEARNING
REGRESSION
CLASSIFICATION
UNSUPERVISED LEARNING
CLUSTERING
RECOMMENDATION
REINFORCEMENT LEARNING
ENSEMBLE LEARNING

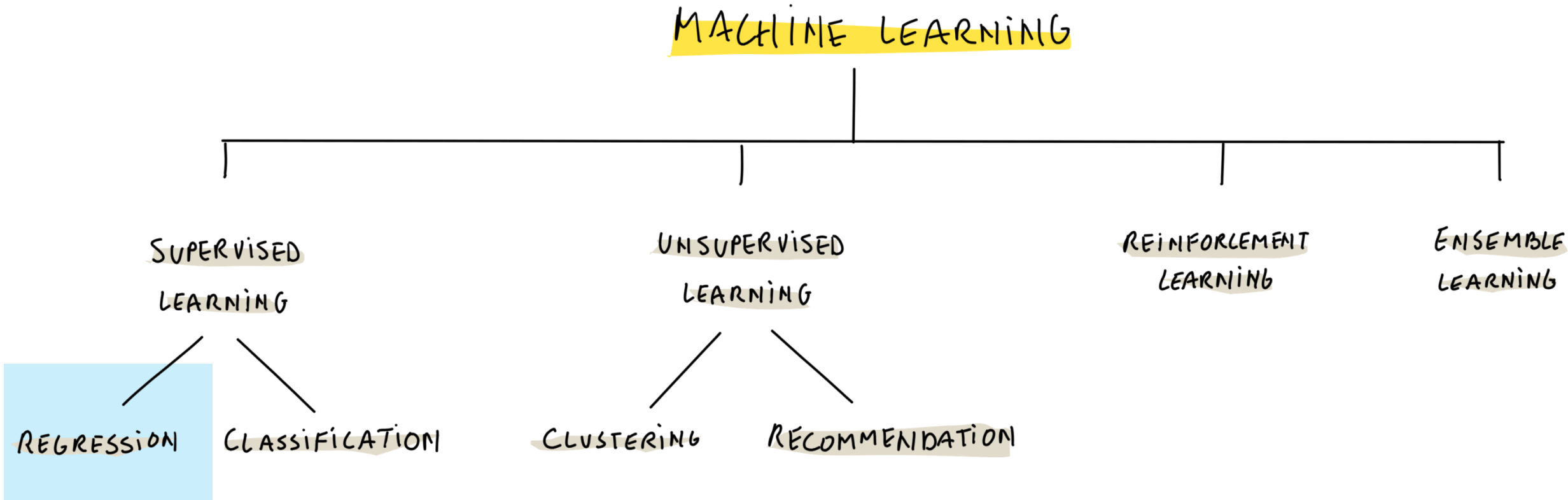
https://www.kensukekoike.com

Is this a Dog?



Is this a Dog?

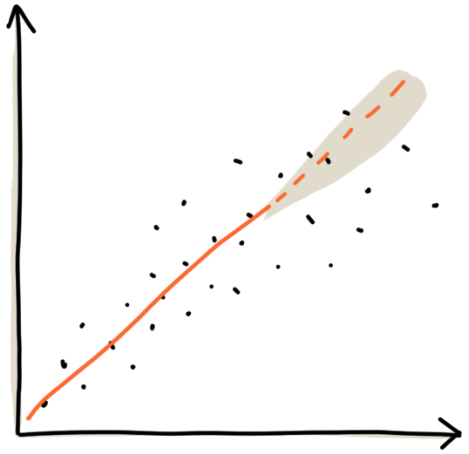



MACHINE LEARNING
SUPERVISED LEARNING
UNSUPERVISED LEARNING
REINFORCEMENT LEARNING
ENSEMBLE LEARNING
REGRESSION
CLASSIFICATION
CLUSTERING
RECOMMENDATION

LINEAR REGRESSION

GOAL

PREDICTING
CONTINUOUS
VARIABLES



USED IN FORECASTING

Predicting future sales by looking at what happened before.

Predicting the price of a house based on house criteria

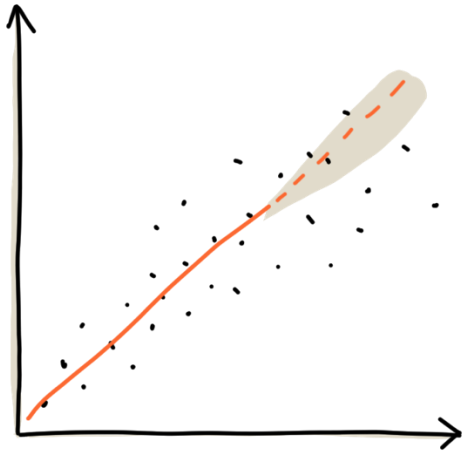
Predicting how many people may buy something because of ads

Predicting the needed stock in a supply chain process

LINEAR REGRESSION

GOAL

PREDICTING
CONTINUOUS
VARIABLES



USED IN FORECASTING

Predicting future sales by looking at what happened before.

Predicting the price of a house based on house criteria

Predicting how many people may buy something because of ads

Predicting the needed stock in a supply chain process

“Predicting the price of a house based on house criteria”

	house age	distance to the nearest MRT station	number of convenience stores	house price of unit area
0	32.0	84.87882	10	37.9
1	19.5	306.59470	9	42.2
2	13.3	561.98450	5	47.3
3	13.3	561.98450	5	54.8
4	5.0	390.56840	5	43.1
5	7.1	2175.03000	3	32.1
6	34.5	623.47310	7	40.3
7	20.3	287.60250	6	46.7
8	31.7	5512.03800	1	18.8
9	17.9	1783.18000	3	22.1

INPUT

OUTPUT

	house age	distance to the nearest MRT station	number of convenience stores	house price of unit area
0	32.0	84.87882	10	37.9
1	19.5	306.59470	9	42.2
2	13.3	561.98450	5	47.3
3	13.3	561.98450	5	54.8
4	5.0	390.56840	5	43.1
5	7.1	2175.03000	3	32.1
6	34.5	623.47310	7	40.3
7	20.3	287.60250	6	46.7
8	31.7	5512.03800	1	18.8
9	17.9	1783.18000	3	22.1

INPUT

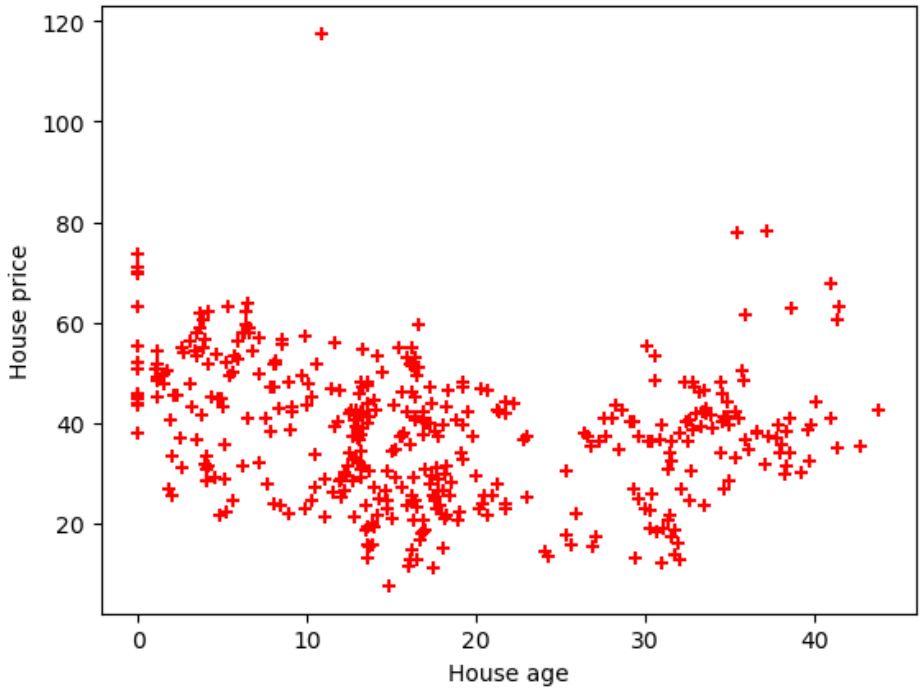
OUTPUT

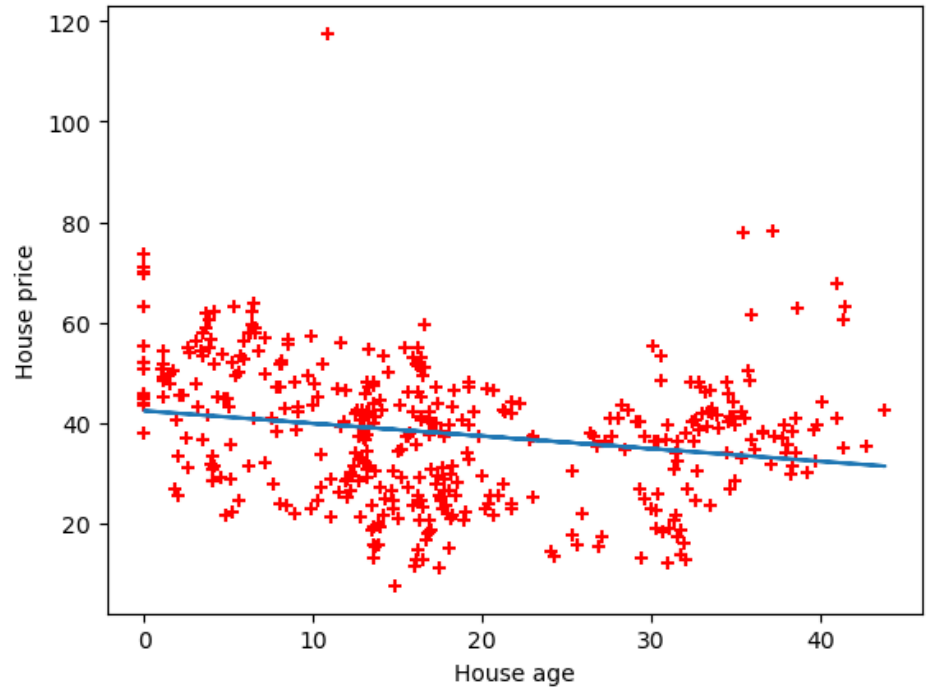
	house age	distance to the nearest MRT station	number of convenience stores	house price of unit area
0	32.0	84.87882	10	37.9
1	19.5	306.59470	9	42.2
2	13.3	561.98450	5	47.3
3	13.3	561.98450	5	54.8
4	5.0	390.56840	5	43.1
5	7.1	2175.03000	3	32.1
6	34.5	623.47310	7	40.3
7	20.3	287.60250	6	46.7
8	31.7	5512.03800	1	18.8
9	17.9	1783.18000	3	22.1

DATA

TRAINING

AI MODEL





WHY IS THIS AI ???

$$F(\underbrace{\text{HOUSE AGE}}) = \underbrace{\text{HOUSE PRICE}}$$

$$\underbrace{m}_{\text{SLOPE}} \cdot x + \underbrace{b}_{\text{INTERCEPT}} = y$$

SLOPE: -0.2514

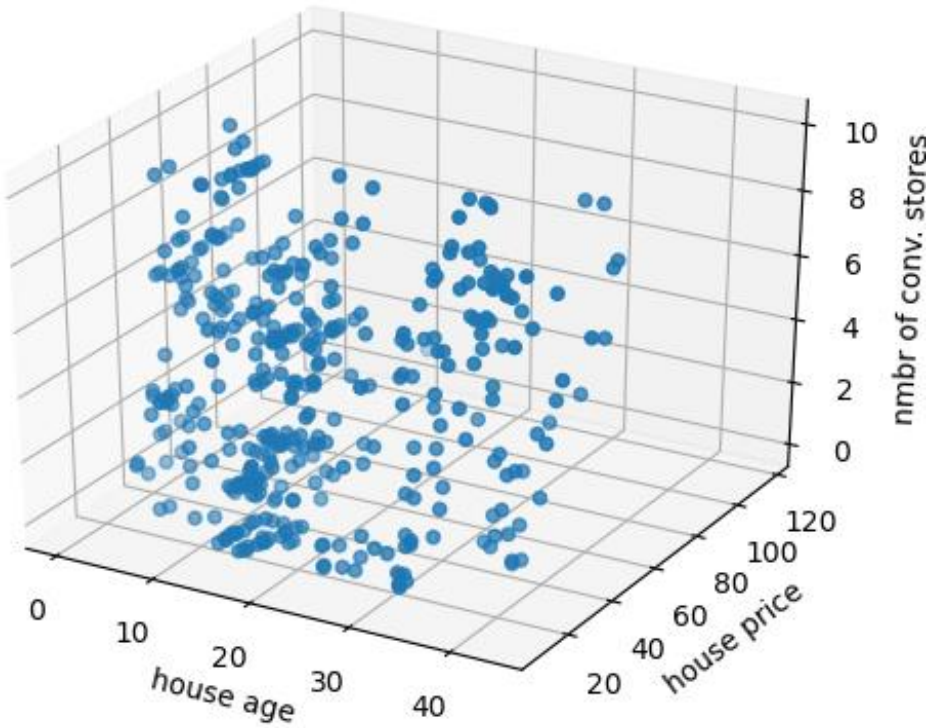
INTERCEPT: 42.4346

INPUT

OUTPUT

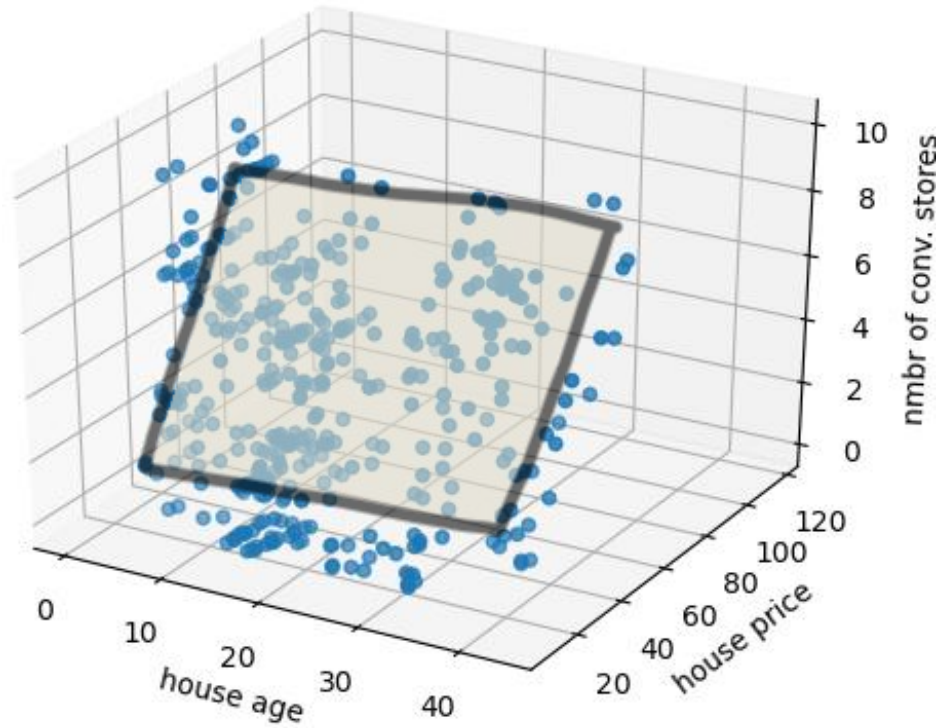
	house age	distance to the nearest MRT station	number of convenience stores	house price of unit area
0	32.0	84.87882	10	37.9
1	19.5	306.59470	9	42.2
2	13.3	561.98450	5	47.3
3	13.3	561.98450	5	54.8
4	5.0	390.56840	5	43.1
5	7.1	2175.03000	3	32.1
6	34.5	623.47310	7	40.3
7	20.3	287.60250	6	46.7
8	31.7	5512.03800	1	18.8
9	17.9	1783.18000	3	22.1

DATA



TRAINING

AI MODEL



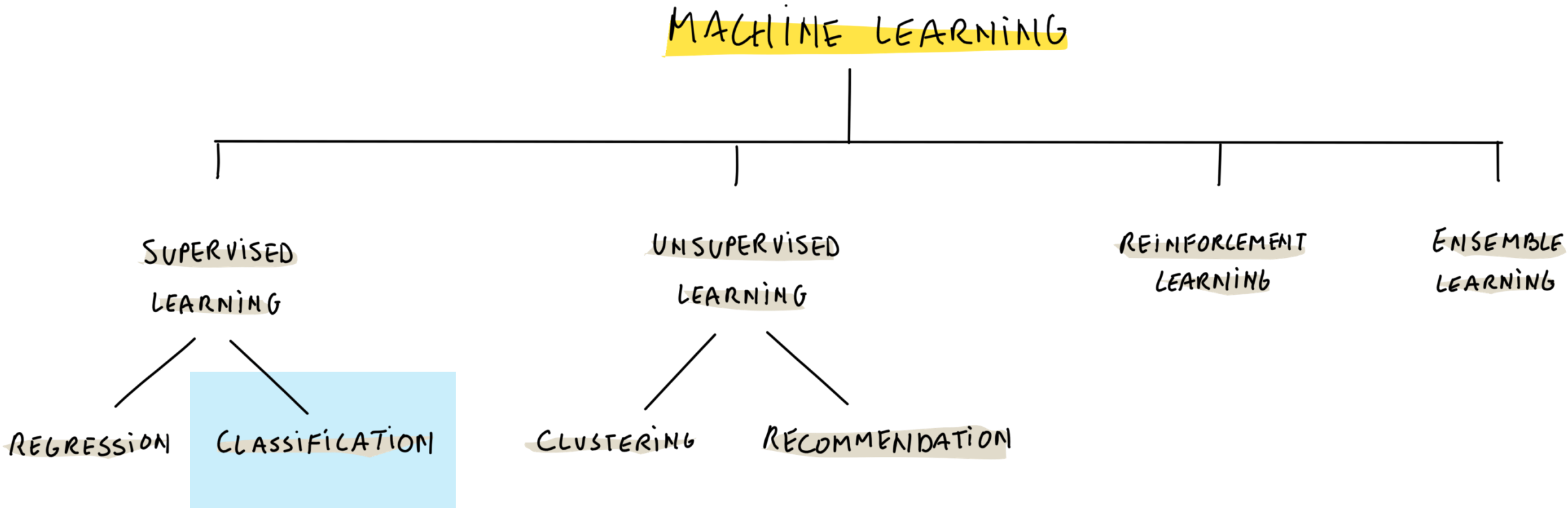
F(HOUSE AGE, NMBR OF CONV. STORES) = HOUSE PRICE

INPUT

OUTPUT

	house age	distance to the nearest MRT station	number of convenience stores	house price of unit area
0	32.0	84.87882	10	37.9
1	19.5	306.59470	9	42.2
2	13.3	561.98450	5	47.3
3	13.3	561.98450	5	54.8
4	5.0	390.56840	5	43.1
5	7.1	2175.03000	3	32.1
6	34.5	623.47310	7	40.3
7	20.3	287.60250	6	46.7
8	31.7	5512.03800	1	18.8
9	17.9	1783.18000	3	22.1

F(HOUSE AGE, NMBR OF CONV. STORES, DISTANCE MRT STATION) = HOUSE PRICE


MACHINE LEARNING
SUPERVISED LEARNING
UNSUPERVISED LEARNING
REINFORCEMENT LEARNING
ENSEMBLE LEARNING
REGRESSION
CLASSIFICATION
CLUSTERING
RECOMMENDATION

CLASSIFICATION

GOAL

PREDICTING THE CORRECT LABEL

USE CASES

- **Customer Segmentation**
- **Fraud Detection**
- **Spam Filtering**
- **Medical Diagnosis**

CLASSIFICATION

GOAL

PREDICTING THE CORRECT LABEL

USE CASES

- **Customer Segmentation**
- **Fraud Detection**
- **Spam Filtering**
- **Medical Diagnosis**

“Who will Buy a Life Insurance?”

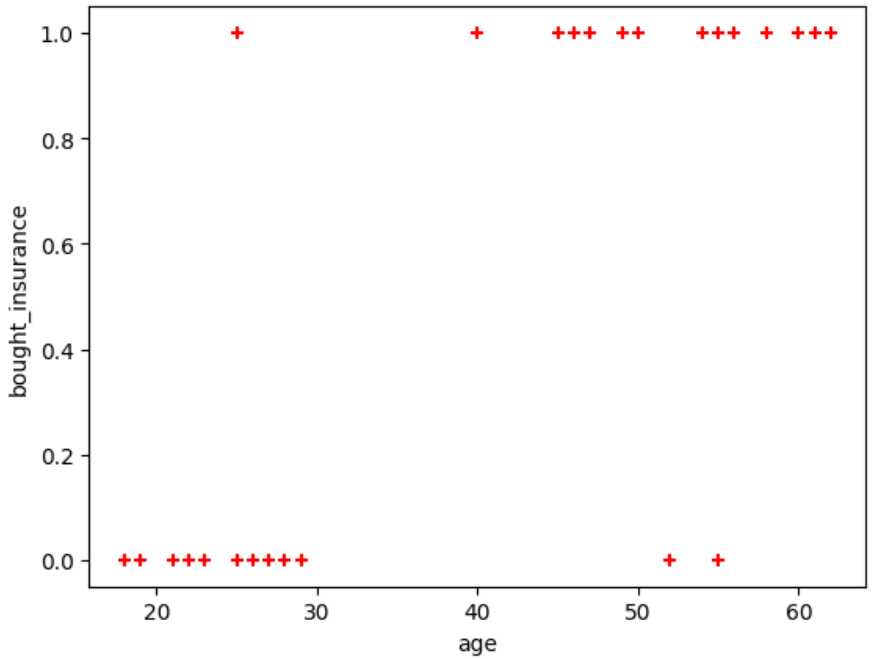
	age	bought_insurance
0	22	0
1	25	0
2	47	1
3	52	0
4	46	1
5	56	1
6	55	0
7	60	1
8	62	1
9	61	1

INPUT

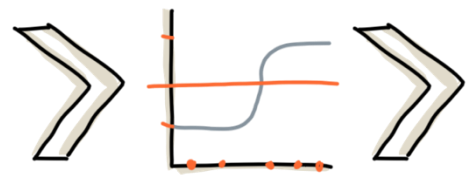
OUTPUT

	age	bought_insurance
0	22	0
1	25	0
2	47	1
3	52	0
4	46	1
5	56	1
6	55	0
7	60	1
8	62	1
9	61	1

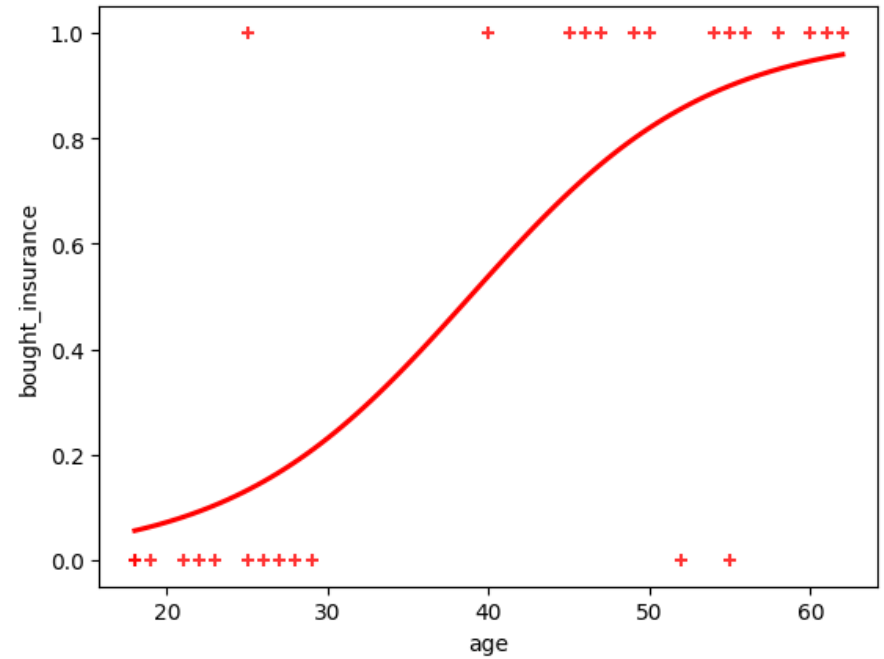
DATA



TRAINING



AI MODEL



F(AGE) = BOUGHT INSURANCE

↳ LOGISTIC REGRESSION

=> SIGMOID FUNCTION

CLASSIFICATION

GOAL

PREDICTING THE CORRECT LABEL

USE CASES

- **Customer Segmentation**
- **Fraud Detection**
- **Spam Filtering**
- **Medical Diagnosis**

“Who Will Develop Diabetes?”

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1
5	5	116	74	0	0	25.6	0.201	30	0
6	3	78	50	32	88	31.0	0.248	26	1
7	10	115	0	0	0	35.3	0.134	29	0
8	2	197	70	45	543	30.5	0.158	53	1
9	8	125	96	0	0	0.0	0.232	54	1

INPUT

OUTPUT

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1
5	5	116	74	0	0	25.6	0.201	30	0
6	3	78	50	32	88	31.0	0.248	26	1
7	10	115	0	0	0	35.3	0.134	29	0
8	2	197	70	45	543	30.5	0.158	53	1
9	8	125	96	0	0	0.0	0.232	54	1

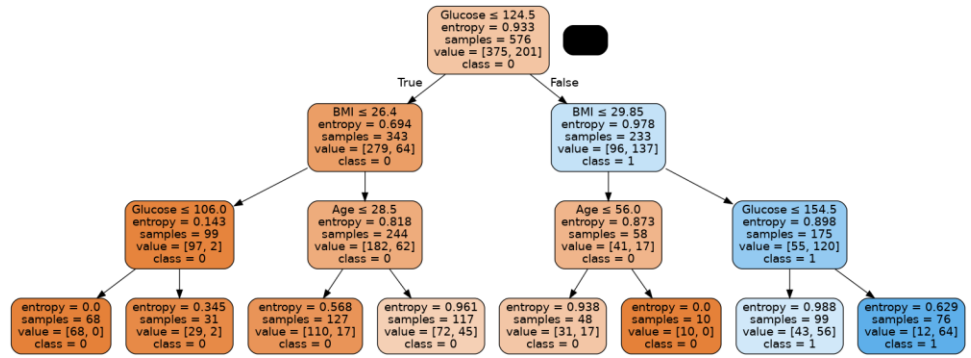
F(...) = OUTCOME

DATA

TRAINING

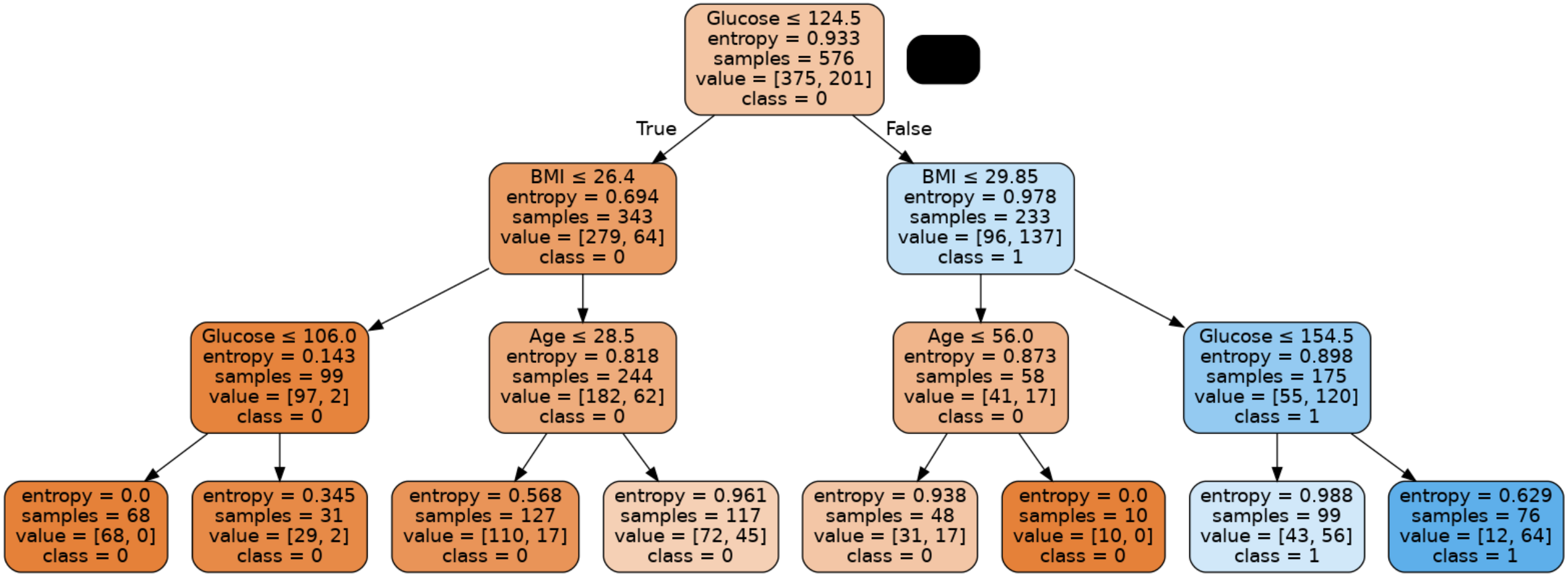
AI MODEL

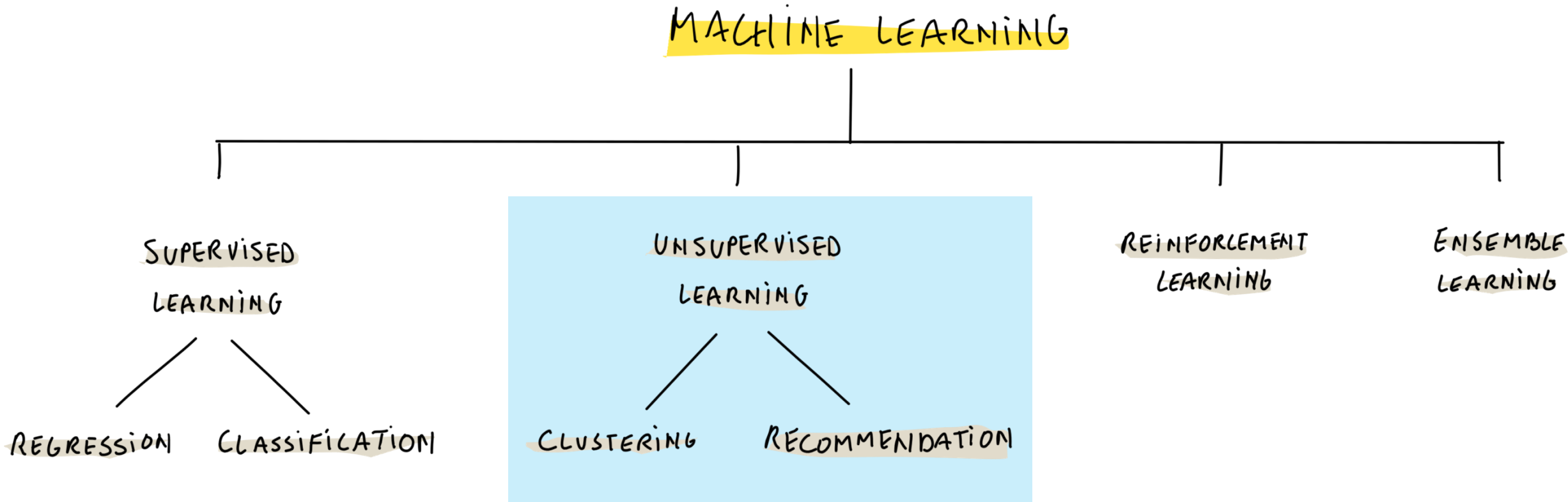
	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1
5	5	116	74	0	0	25.6	0.201	30	0
6	3	78	50	32	88	31.0	0.248	26	1
7	10	115	0	0	0	35.3	0.134	29	0
8	2	197	70	45	543	30.5	0.158	53	1
9	8	125	96	0	0	0.0	0.232	54	1



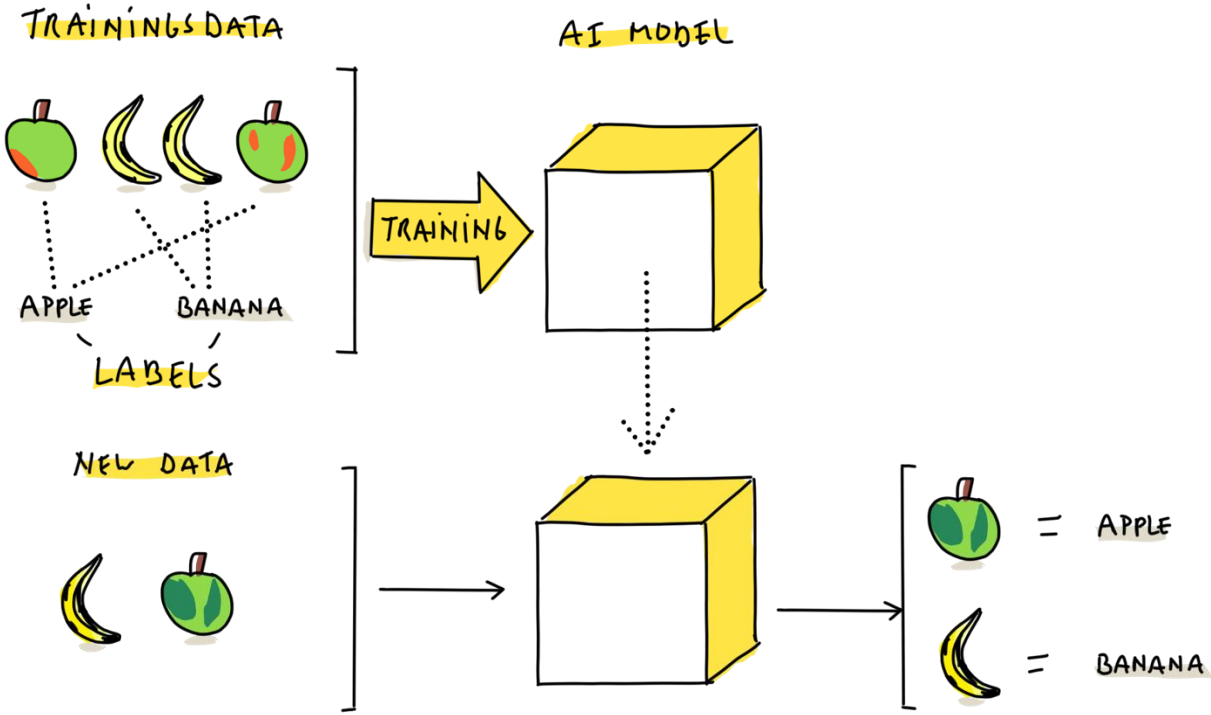
F(...) = OUTCOME

└ DECISION TREE

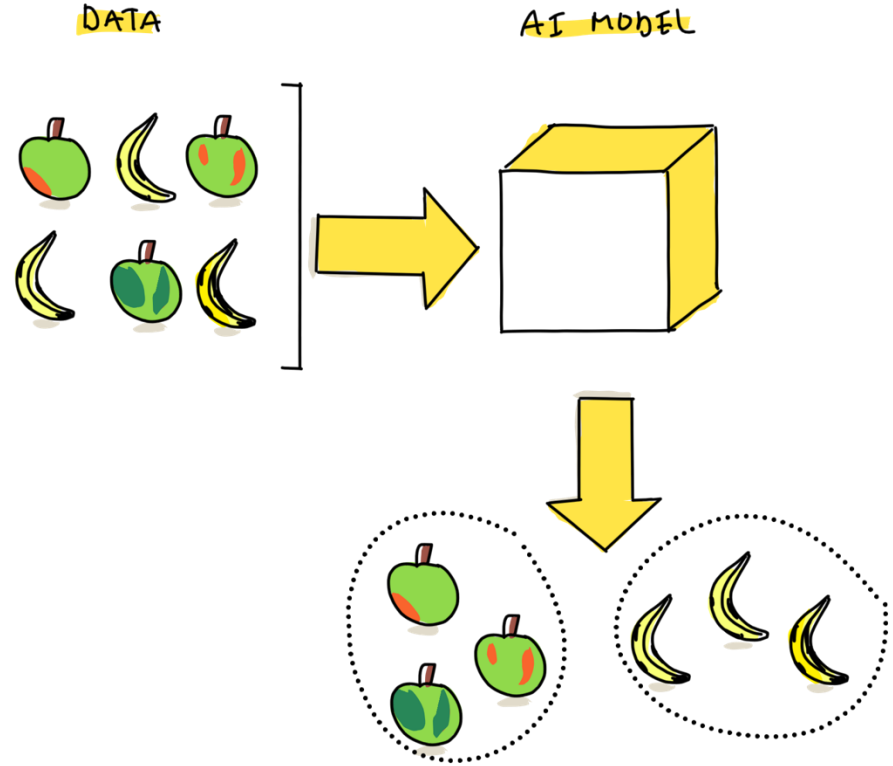

Glucose ≤ 124.5
entropy = 0.933
samples = 576
value = [375, 201]
class = 0
True
False
BMI ≤ 26.4
entropy = 0.694
samples = 343
value = [279, 64]
class = 0
BMI ≤ 29.85
entropy = 0.978
samples = 233
value = [96, 137]
class = 1
Glucose ≤ 106.0
entropy = 0.143
samples = 99
value = [97, 2]
class = 0
Age ≤ 28.5
entropy = 0.818
samples = 244
value = [182, 62]
class = 0
Age ≤ 56.0
entropy = 0.873
samples = 58
value = [41, 17]
class = 0
Glucose ≤ 154.5
entropy = 0.898
samples = 175
value = [55, 120]
class = 1
entropy = 0.0
samples = 68
value = [68, 0]
class = 0
entropy = 0.345
samples = 31
value = [29, 2]
class = 0
entropy = 0.568
samples = 127
value = [110, 17]
class = 0
entropy = 0.961
samples = 117
value = [72, 45]
class = 0
entropy = 0.938
samples = 48
value = [31, 17]
class = 0
entropy = 0.0
samples = 10
value = [10, 0]
class = 0
entropy = 0.988
samples = 99
value = [43, 56]
class = 1
entropy = 0.629
samples = 76
value = [12, 64]
class = 1


MACHINE LEARNING
SUPERVISED LEARNING
REGRESSION
CLASSIFICATION
UNSUPERVISED LEARNING
CLUSTERING
RECOMMENDATION
REINFORCEMENT LEARNING
ENSEMBLE LEARNING

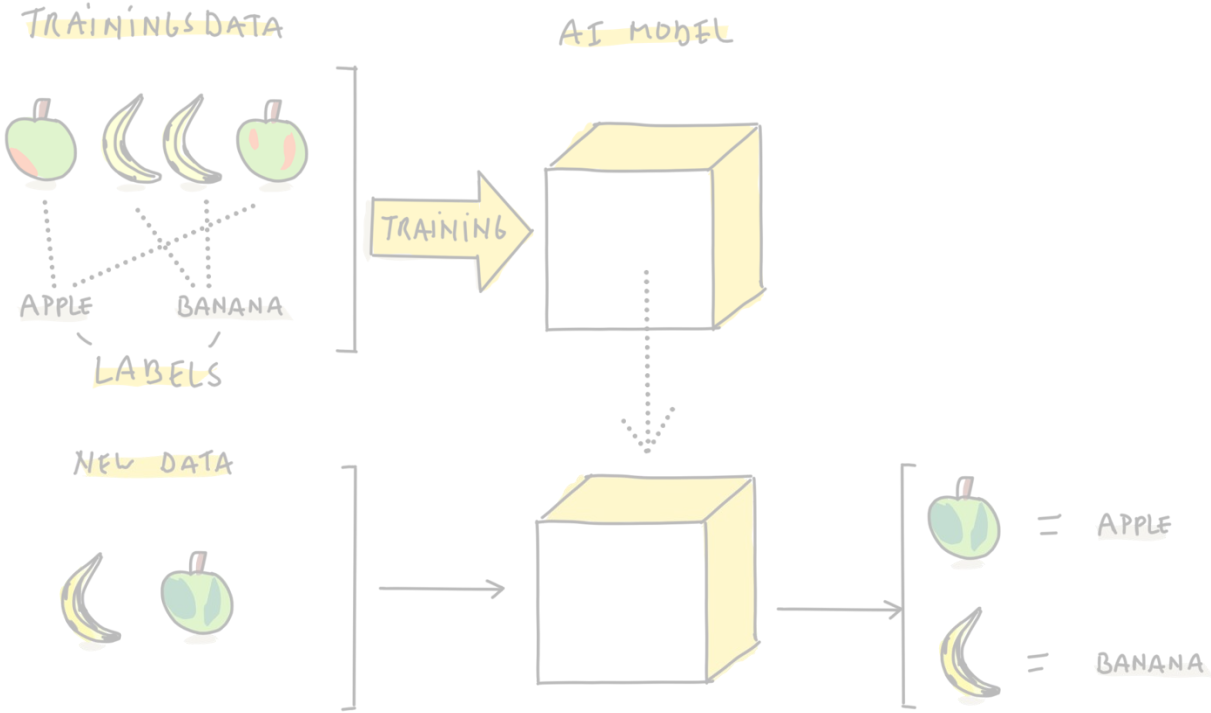
SUPERVISED LEARNING



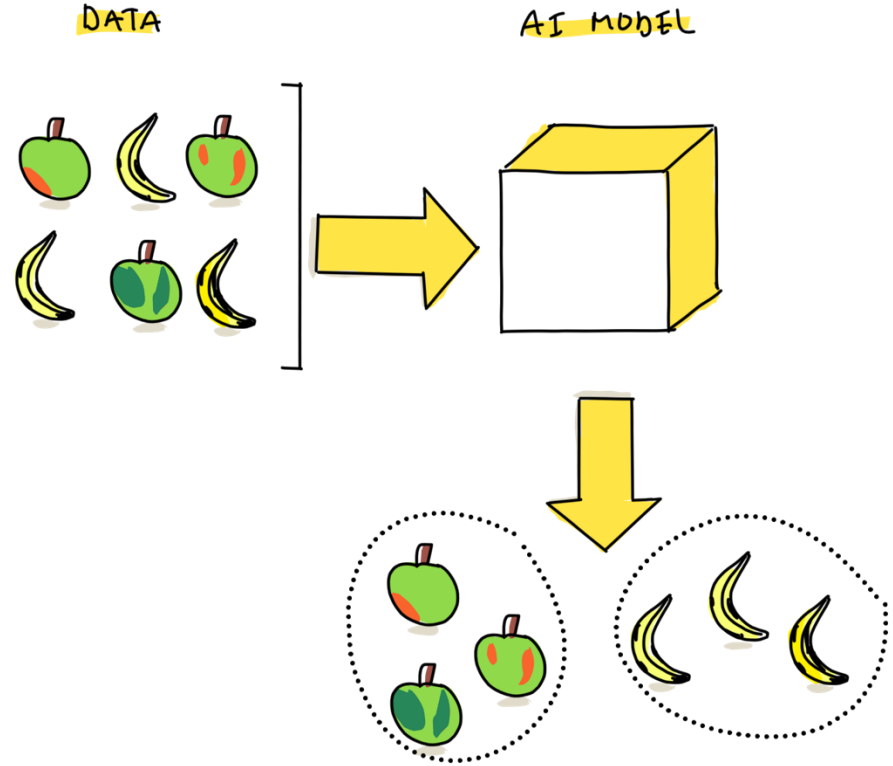
UNSUPERVISED LEARNING

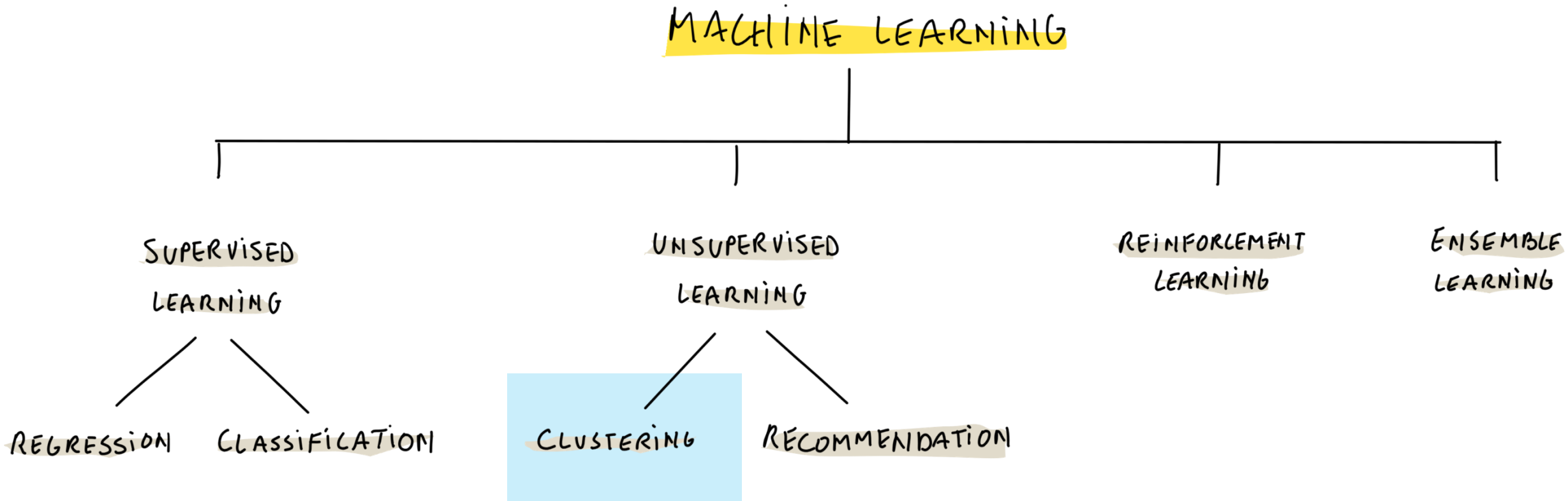


SUPERVISED LEARNING



UNSUPERVISED LEARNING




MACHINE LEARNING
SUPERVISED LEARNING
UNSUPERVISED LEARNING
REINFORCEMENT LEARNING
ENSEMBLE LEARNING
REGRESSION
CLASSIFICATION
CLUSTERING
RECOMMENDATION

CLUSTERING

GOAL

DIVIDING YOUR DATASET IN LOGICAL GROUPS

USE CASES

Customer Segmentation

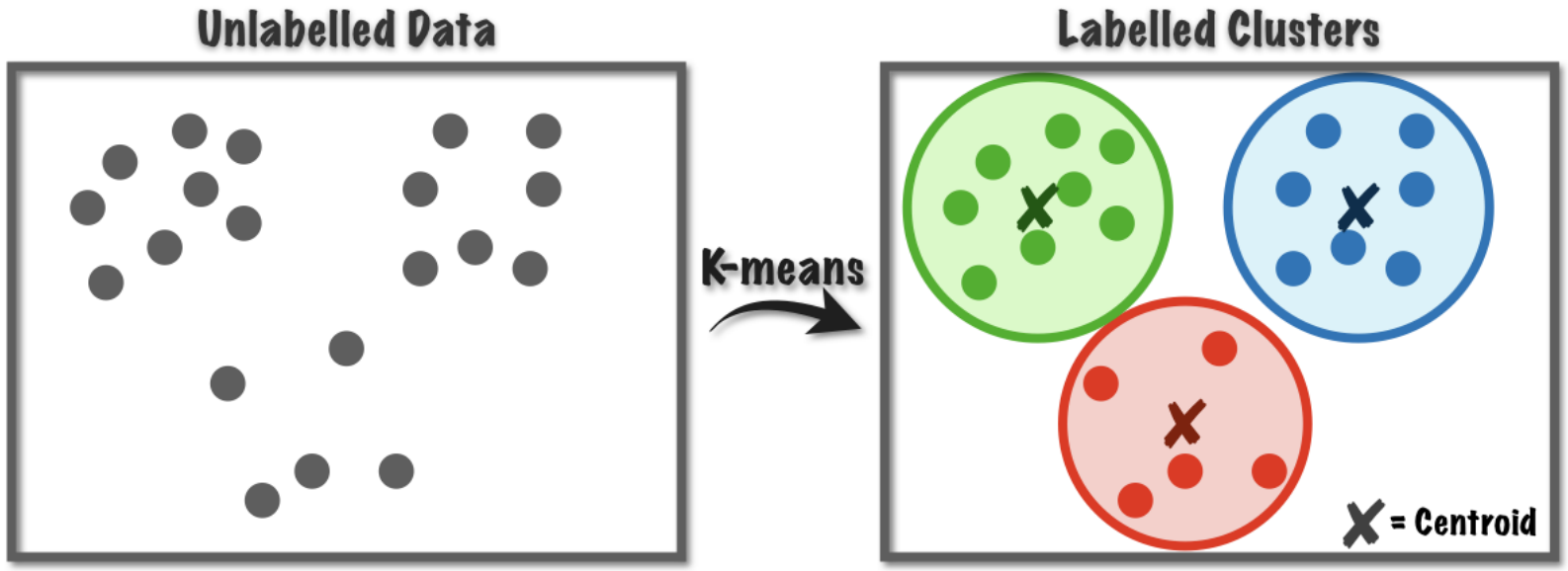
Clustering patients with similar symptoms and medical histories

Clustering transaction to flag transactions that don’t fit into any existing cluster as potential fraud.

K-means clustering

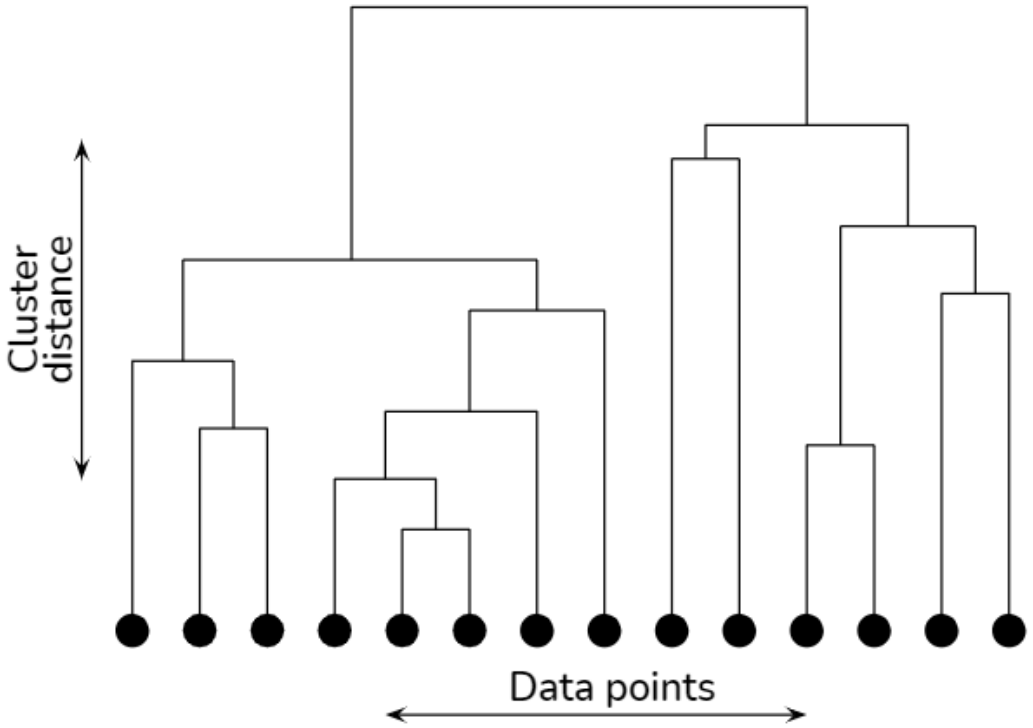
Input: K = the expected number of clusters

Output: exactly K clusters



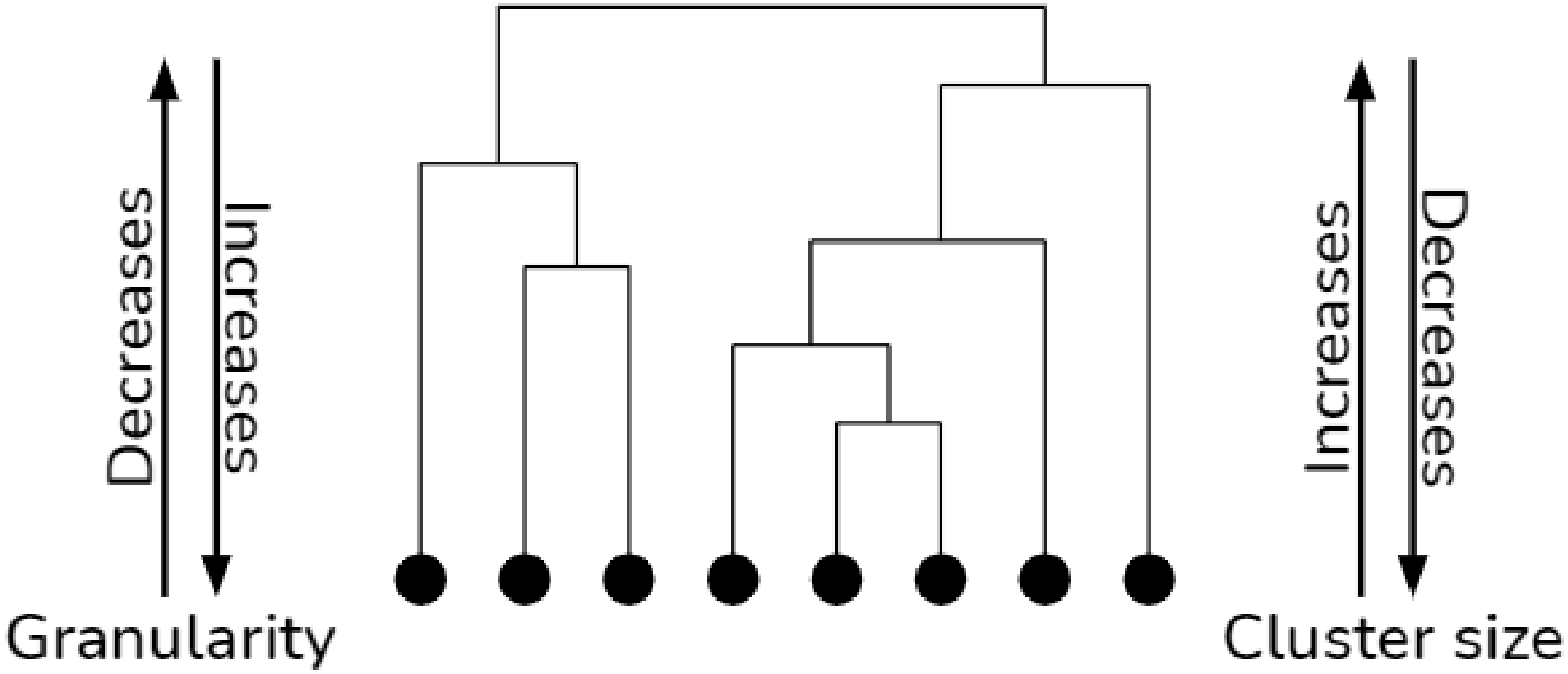
Hierarchical clustering

A dendrogram is a tree-like structure that explains the relationship between all the data points in the system.



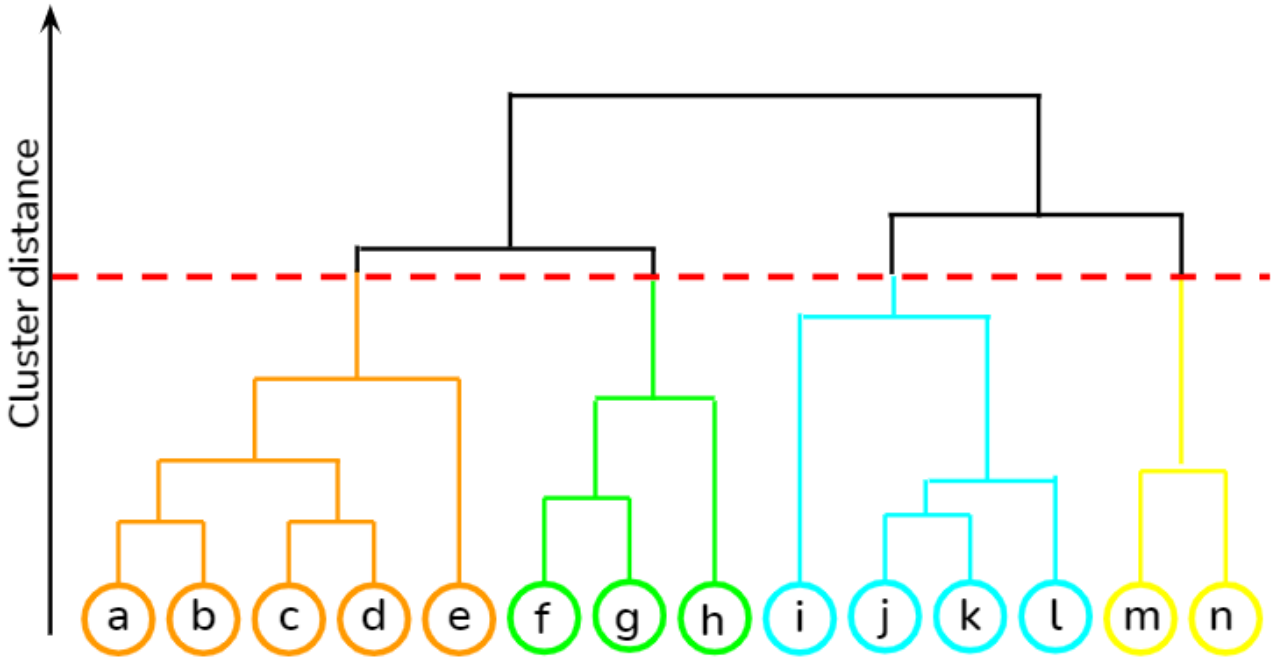
[Hierarchical Clustering Explained]

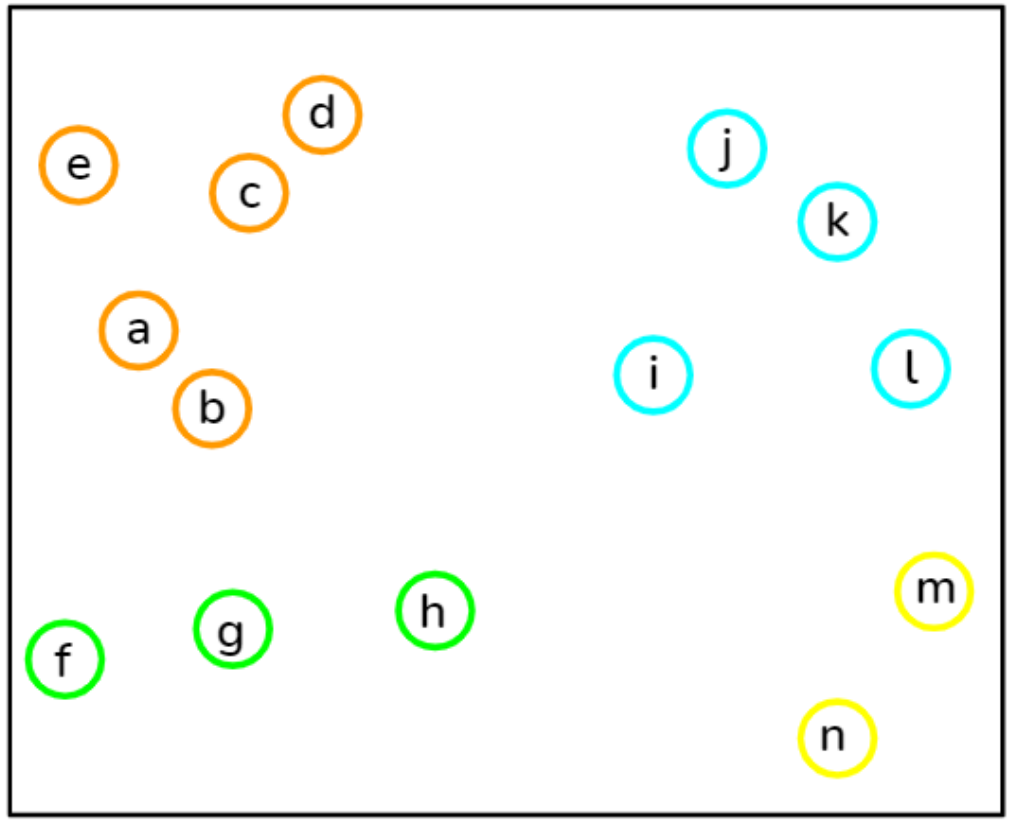
Hierarchical clustering



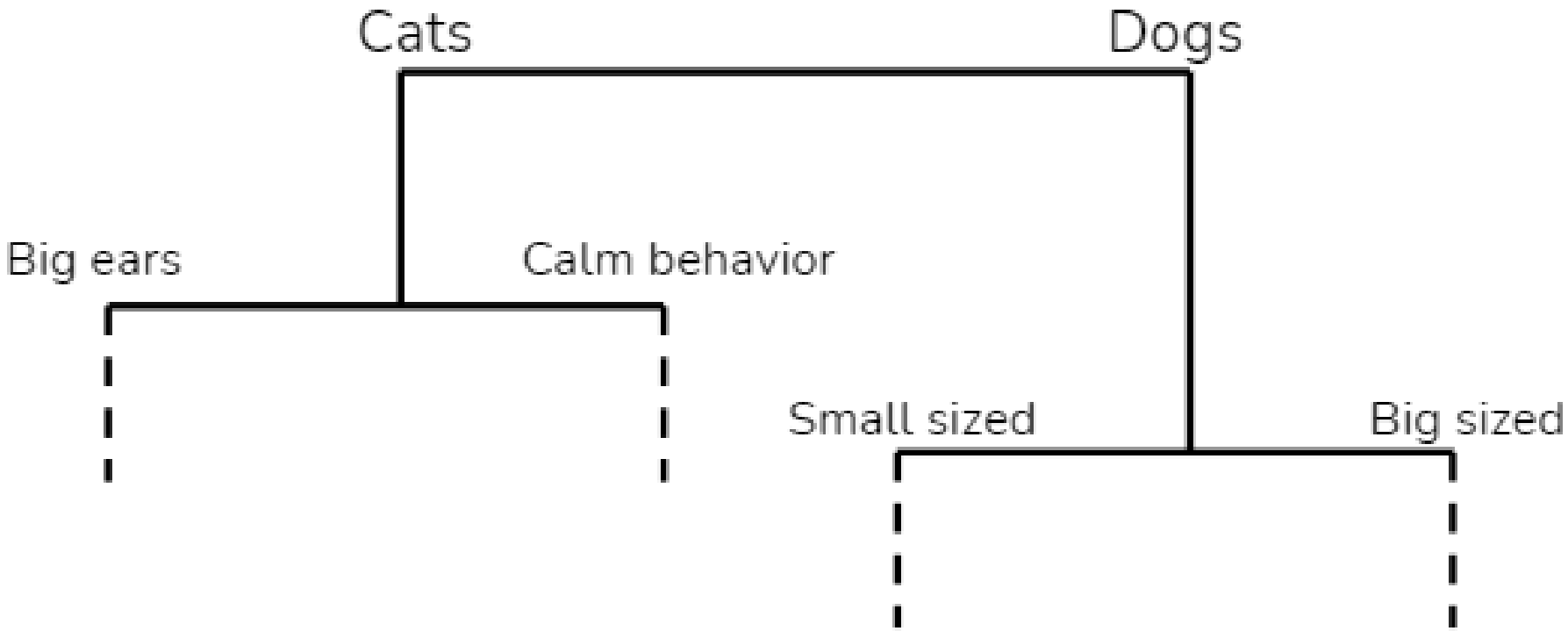
[Hierarchical Clustering Explained]

Hierarchical clustering

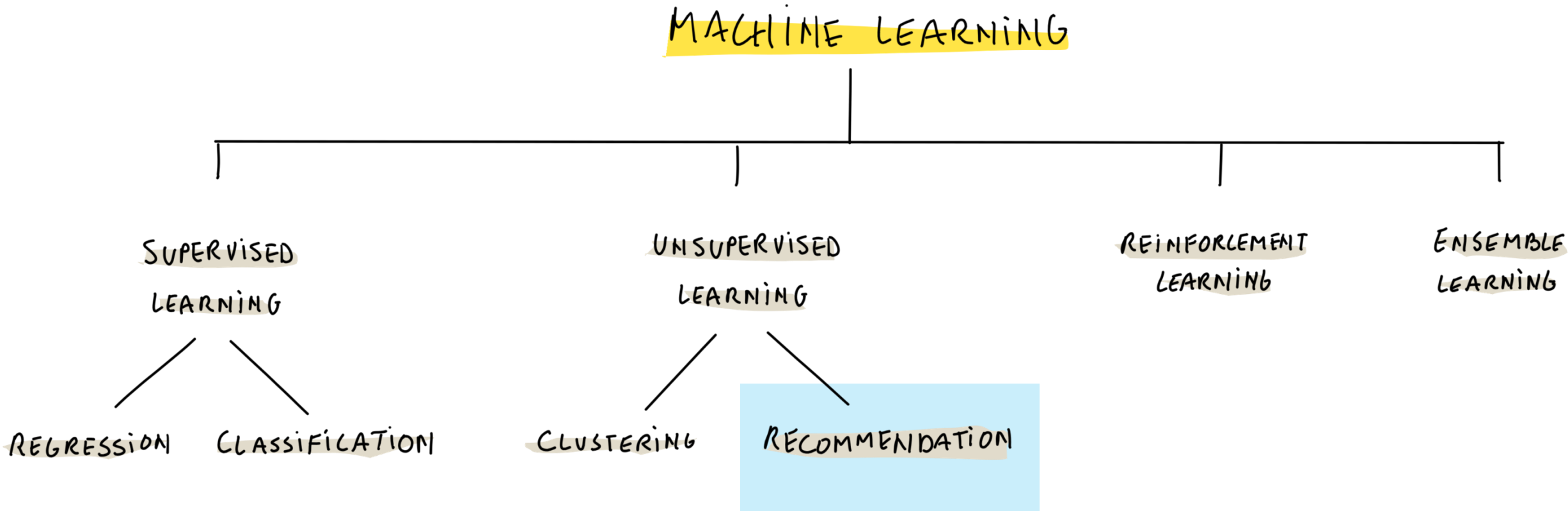




Hierarchical clustering



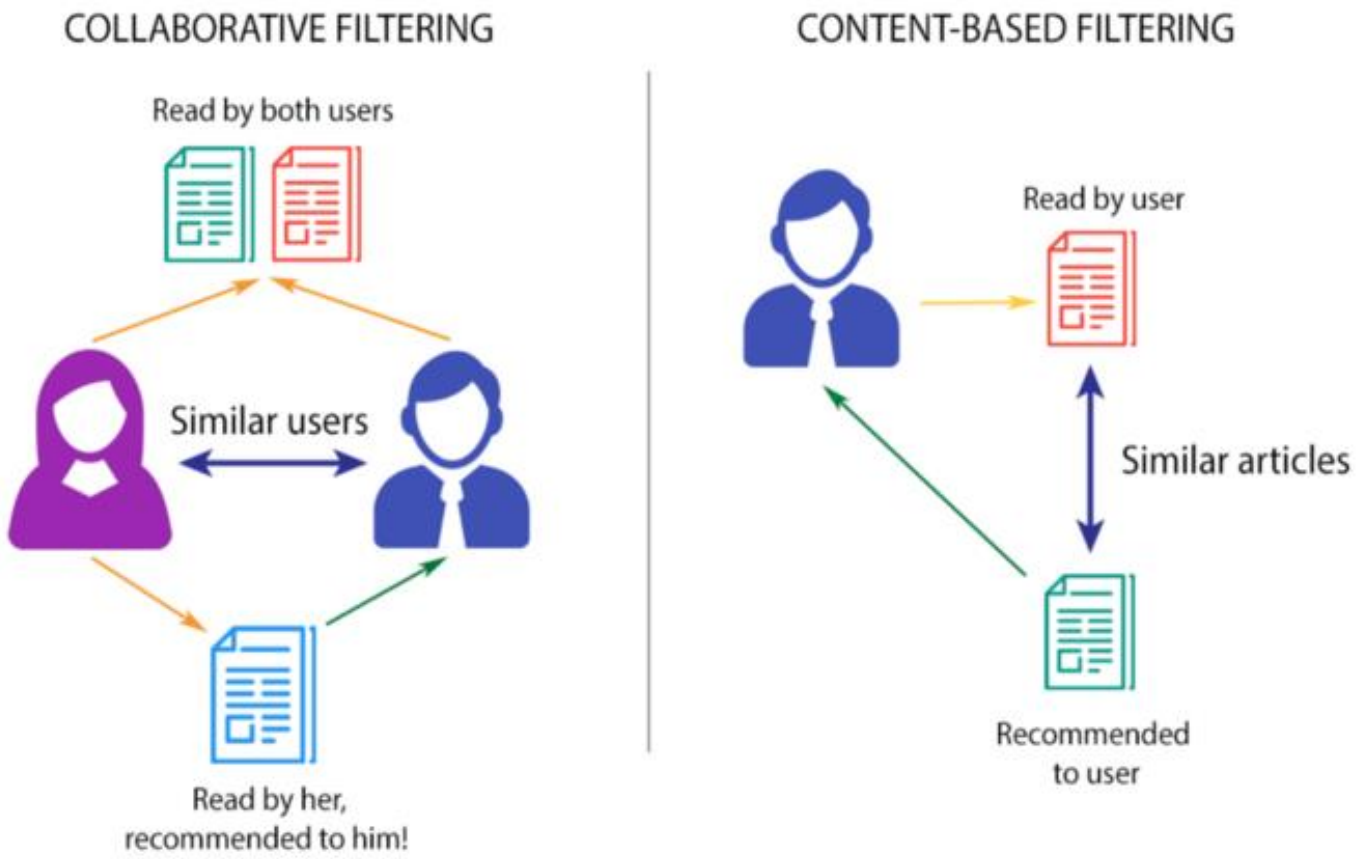
[Hierarchical Clustering Explained]


MACHINE LEARNING
SUPERVISED LEARNING
UNSUPERVISED LEARNING
REINFORCEMENT LEARNING
ENSEMBLE LEARNING
REGRESSION
CLASSIFICATION
CLUSTERING
RECOMMENDATION

Recommendation Systems

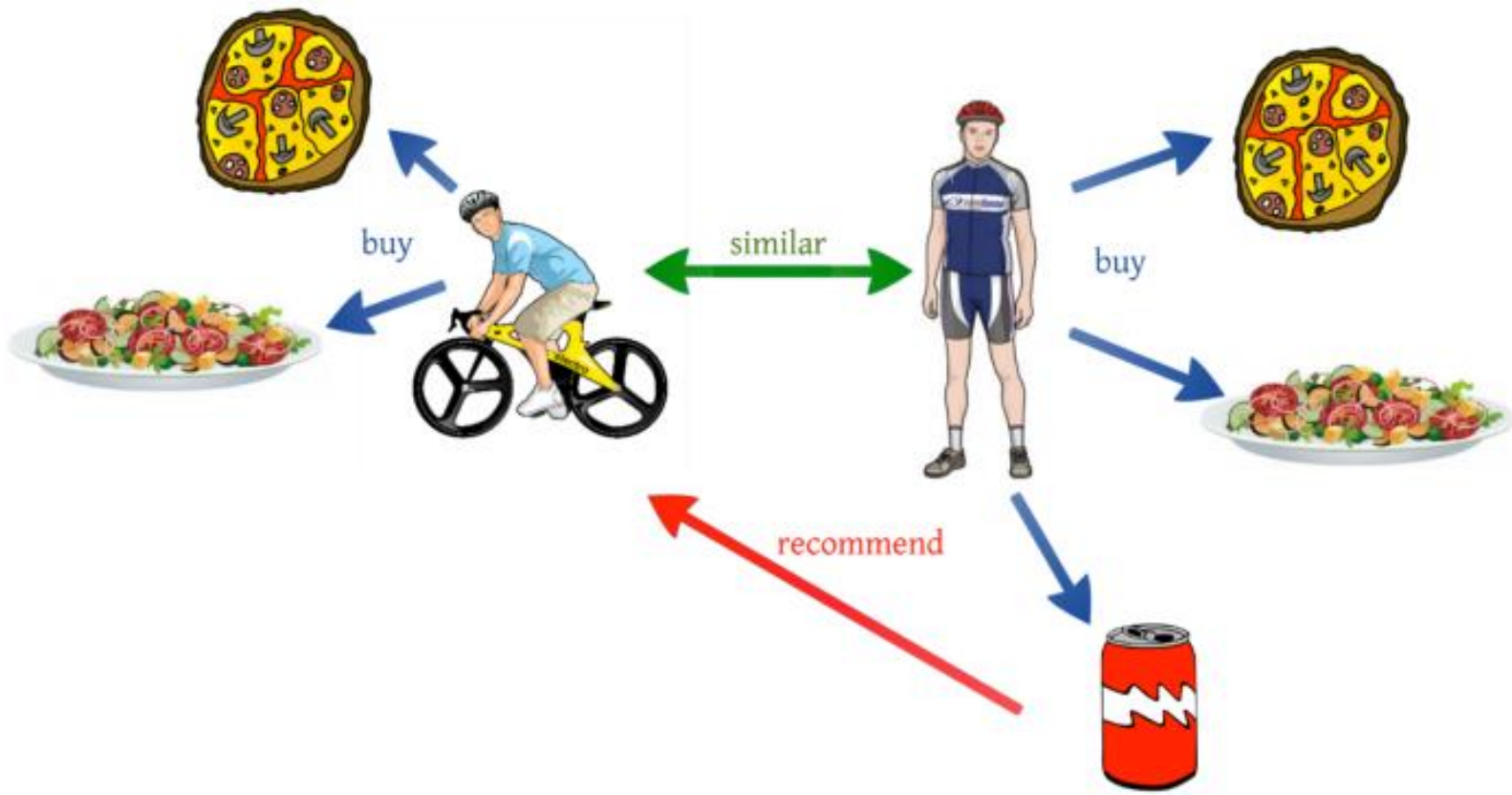
Recommendation systems typically help customers to buy the most suitable product for them.

Based on “Associations”



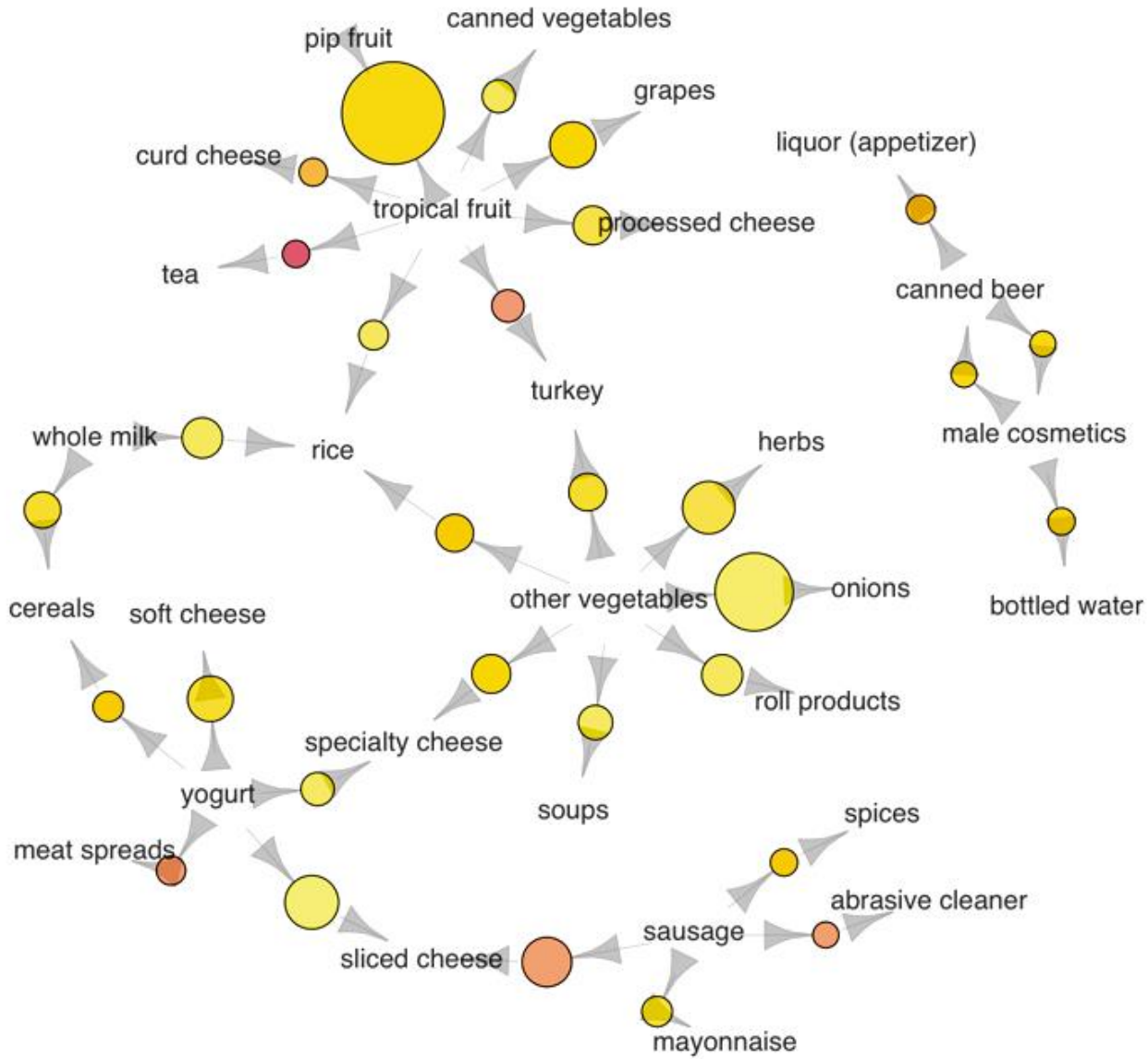
[Brief on Recommender Systems]

Recommendation Systems



[Brief on Recommender Systems]

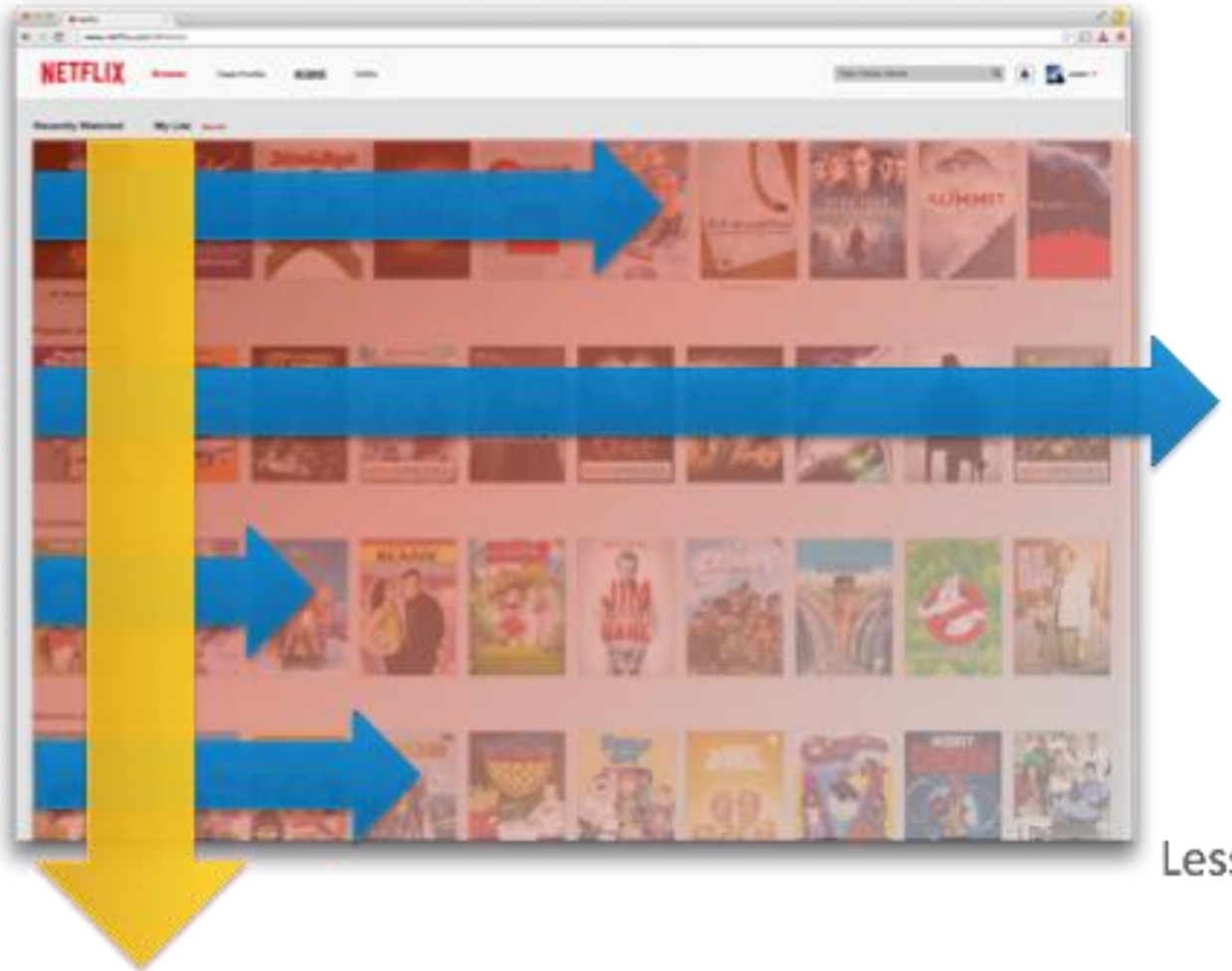
Recommendation Systems



[Association Rules]

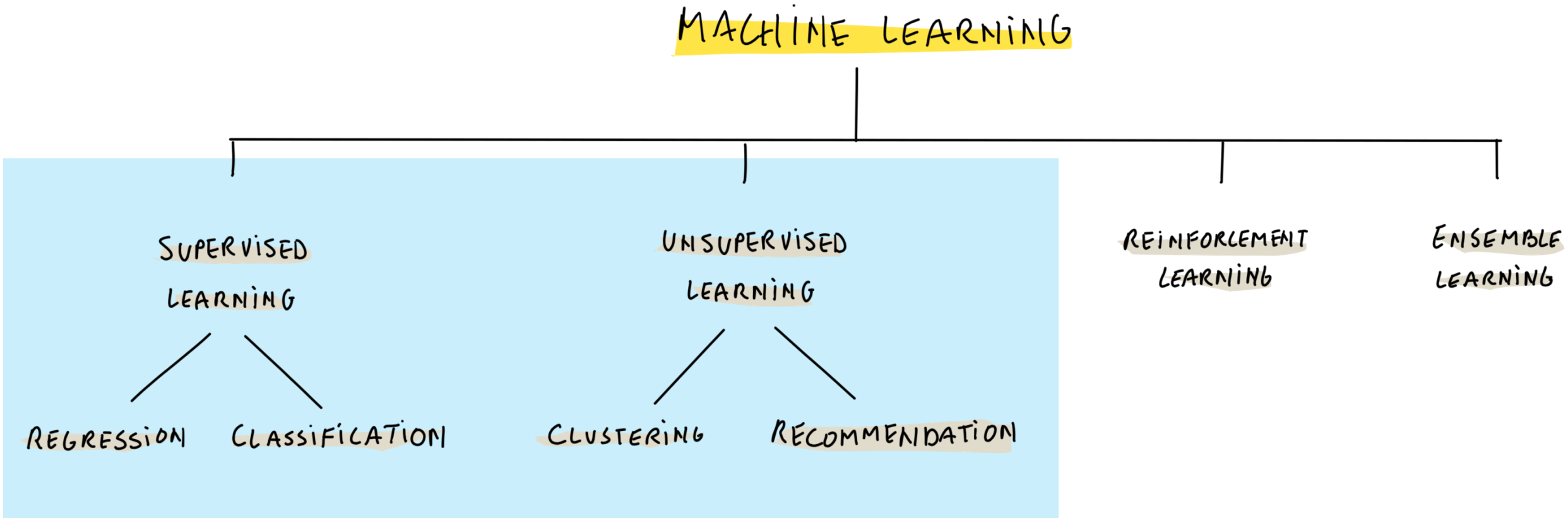
Recommender Systems: Netflix

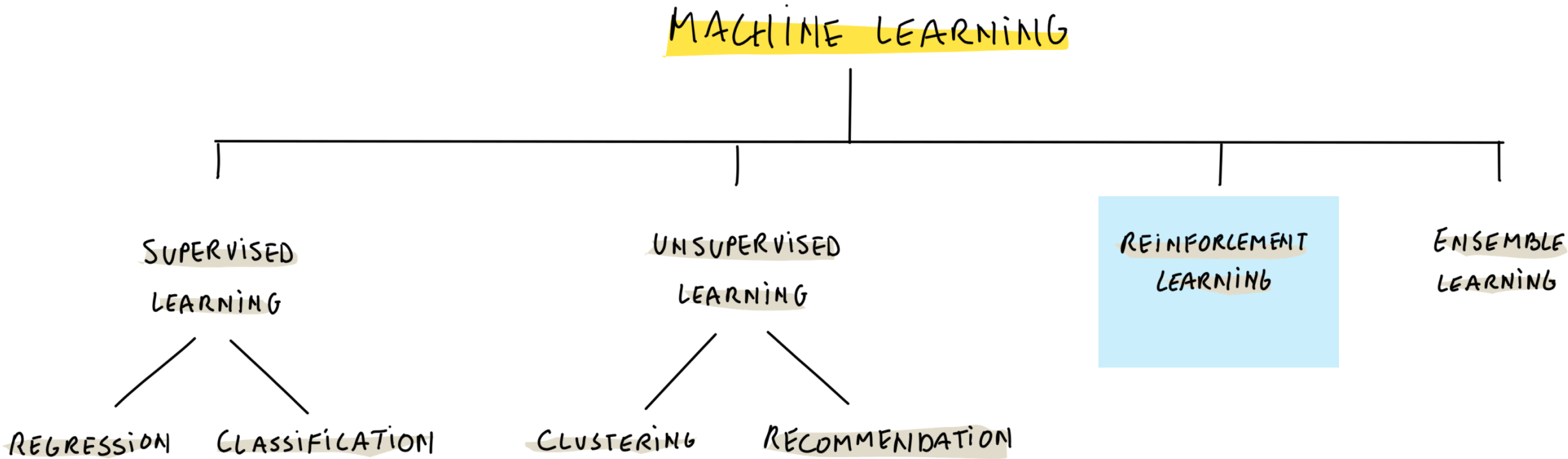
More likely to see



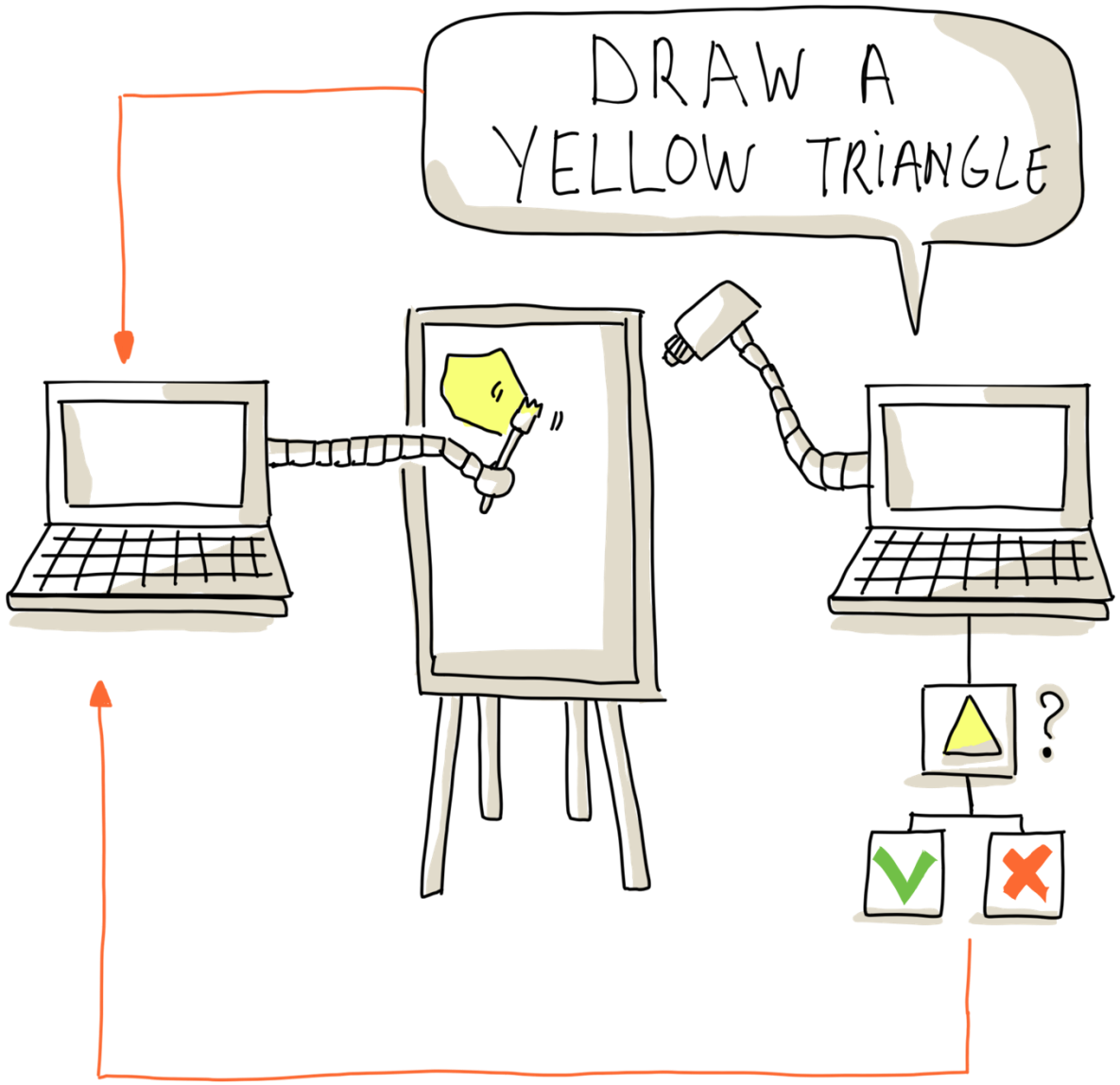
Less likely

[Brief on Recommender Systems]

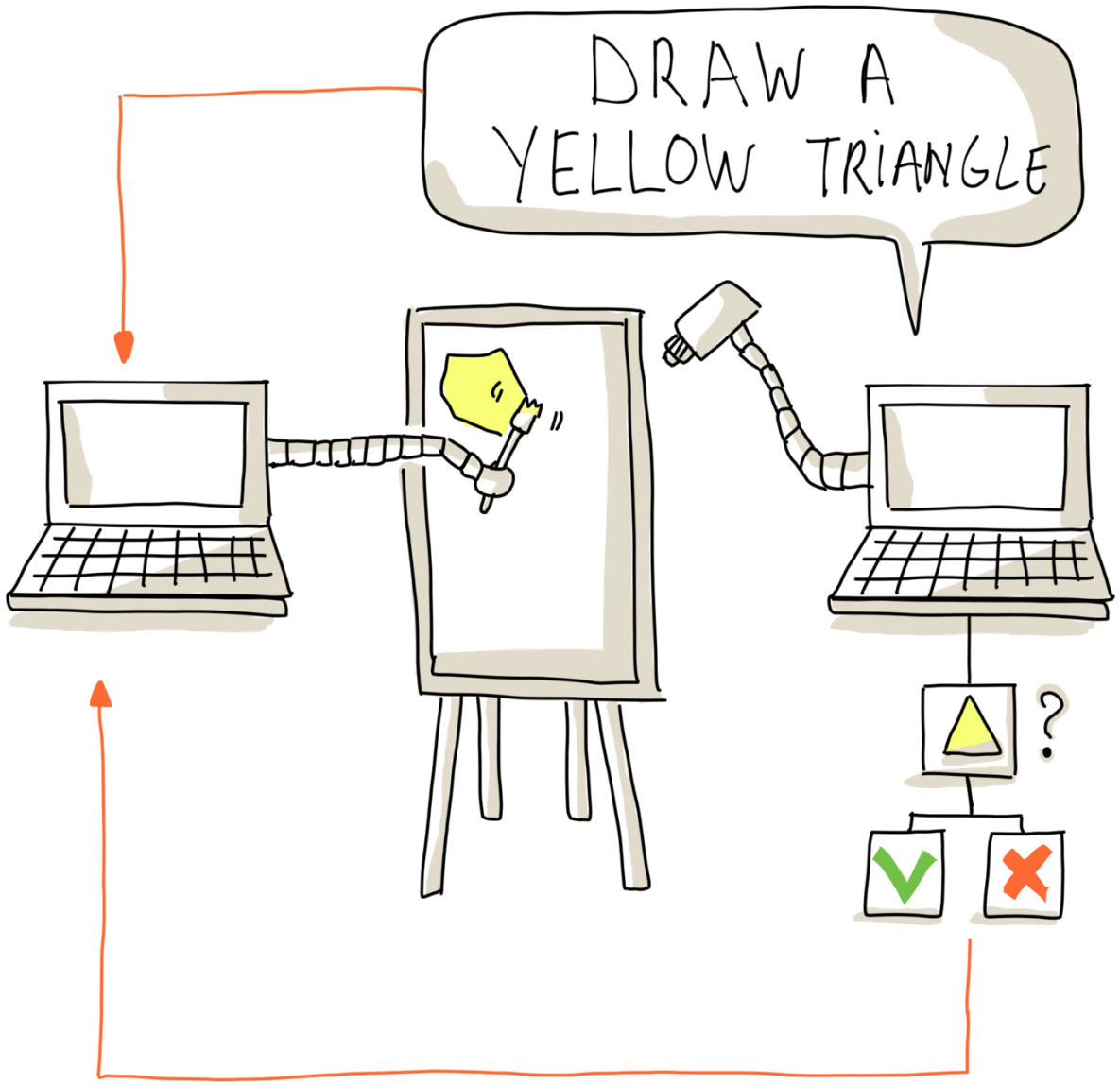

MACHINE LEARNING
SUPERVISED LEARNING
UNSUPERVISED LEARNING
REINFORCEMENT LEARNING
ENSEMBLE LEARNING
REGRESSION
CLASSIFICATION
CLUSTERING
RECOMMENDATION


MACHINE LEARNING
SUPERVISED LEARNING
UNSUPERVISED LEARNING
REINFORCEMENT LEARNING
ENSEMBLE LEARNING
REGRESSION
CLASSIFICATION
CLUSTERING
RECOMMENDATION

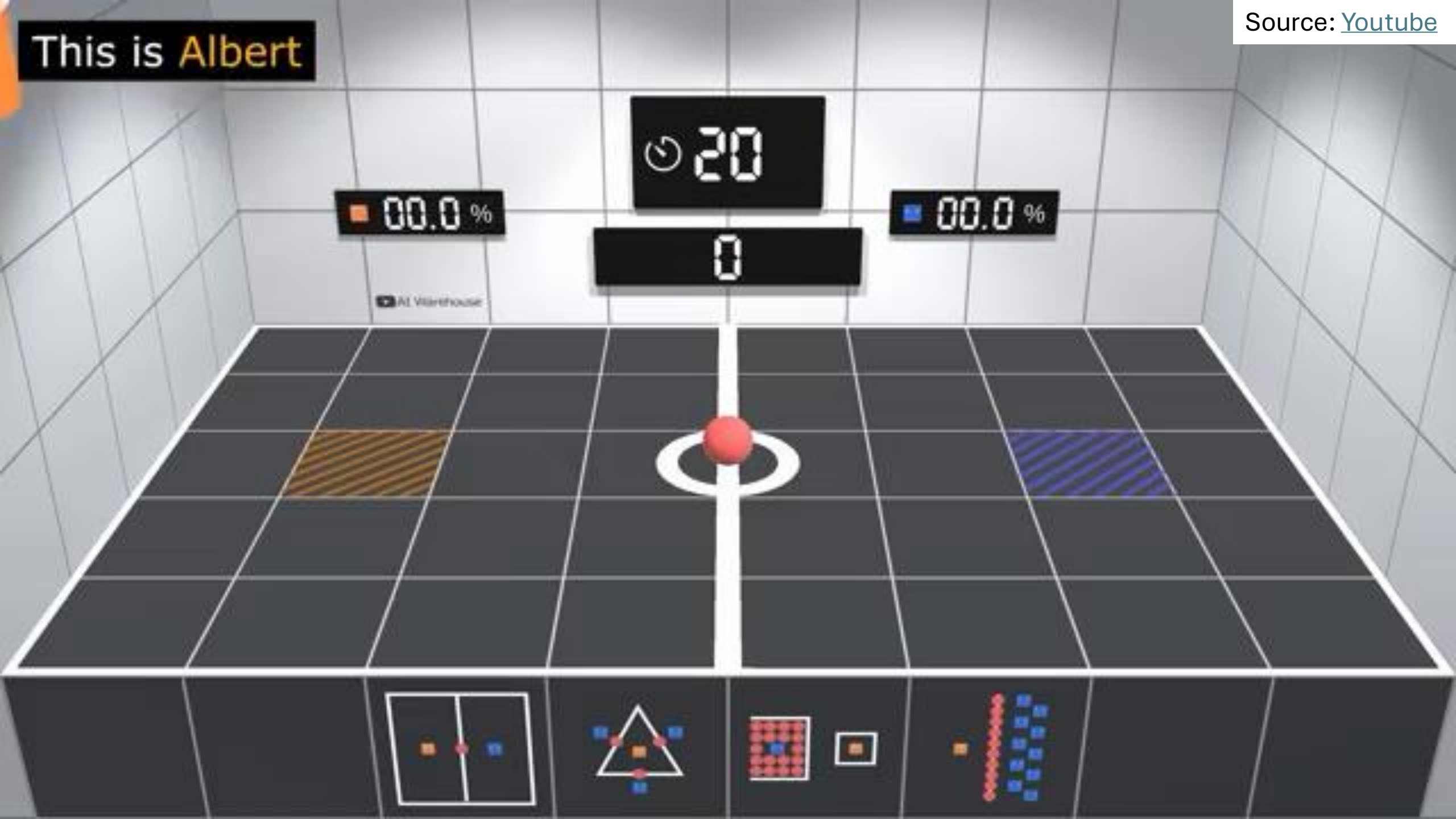
Reinforcement Learning



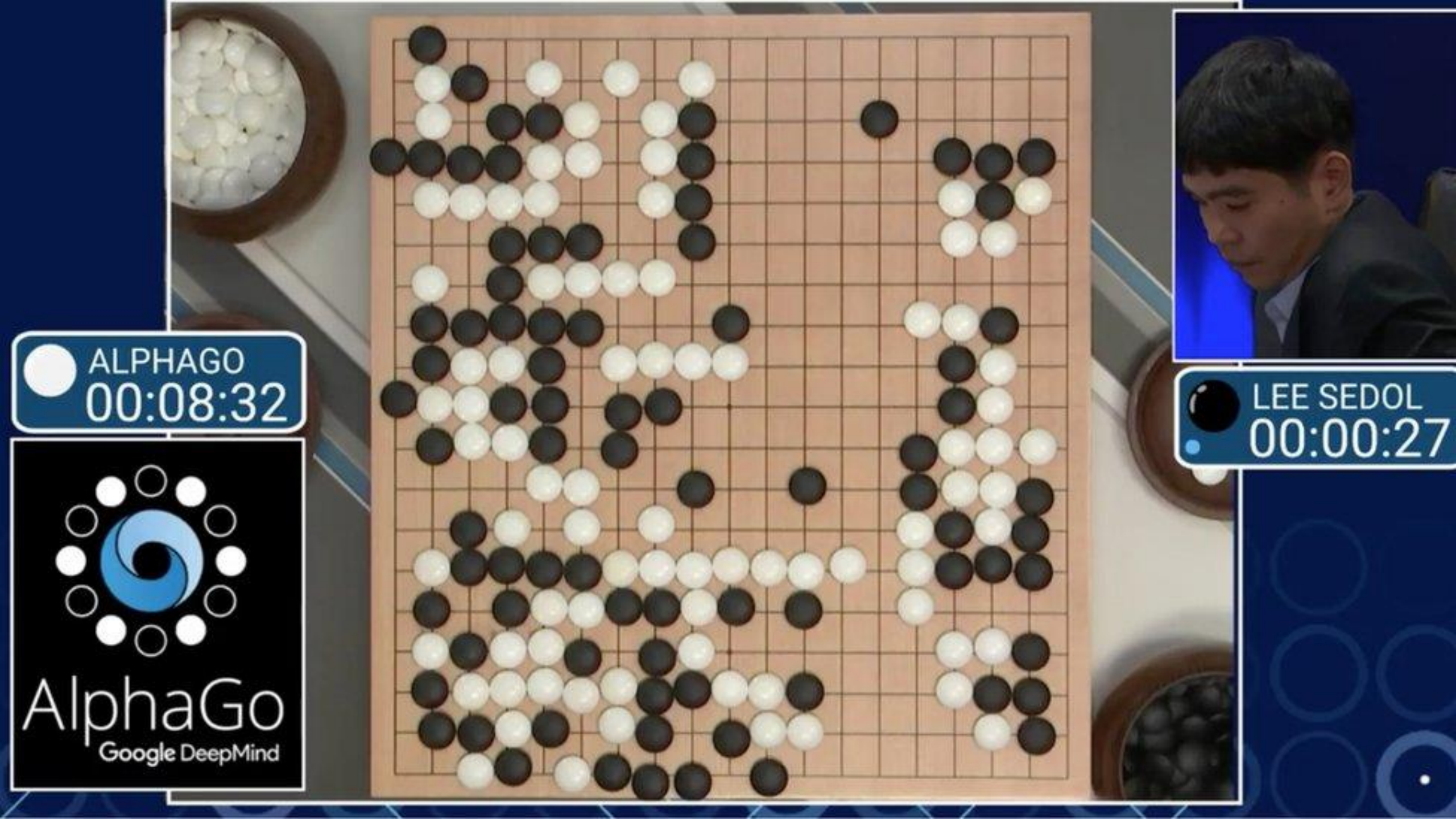
Reinforcement Learning

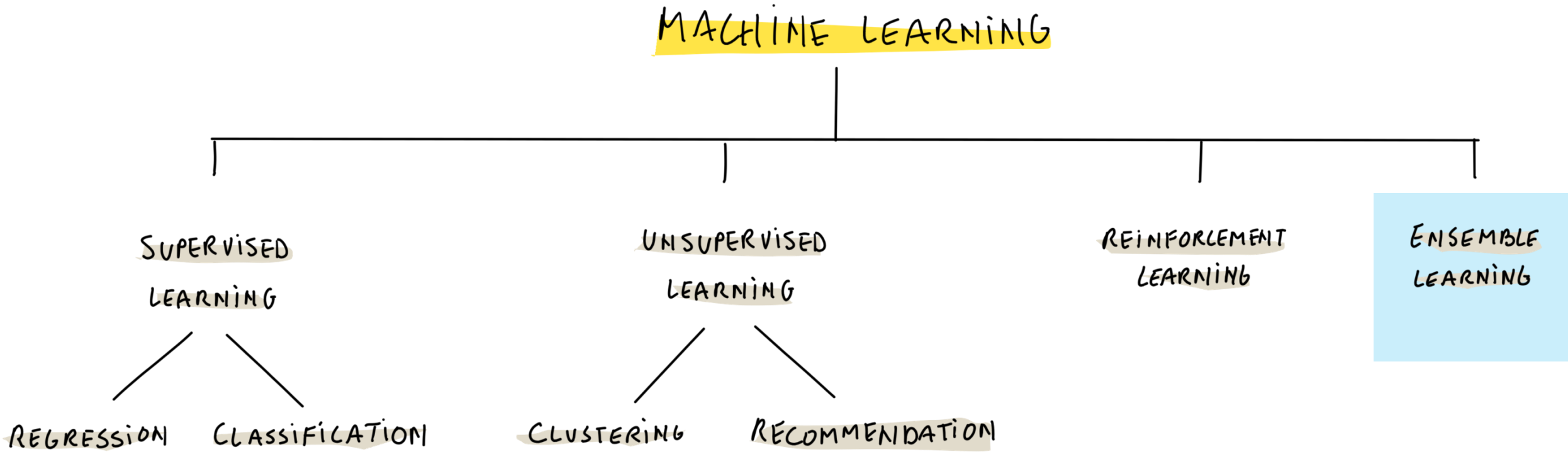


A machine learning training method based on **rewarding desired behaviors and punishing undesired ones**. In general, a reinforcement learning agent -- the entity being trained -- is able to **perceive and interpret its environment, take actions and learn through trial and error**.

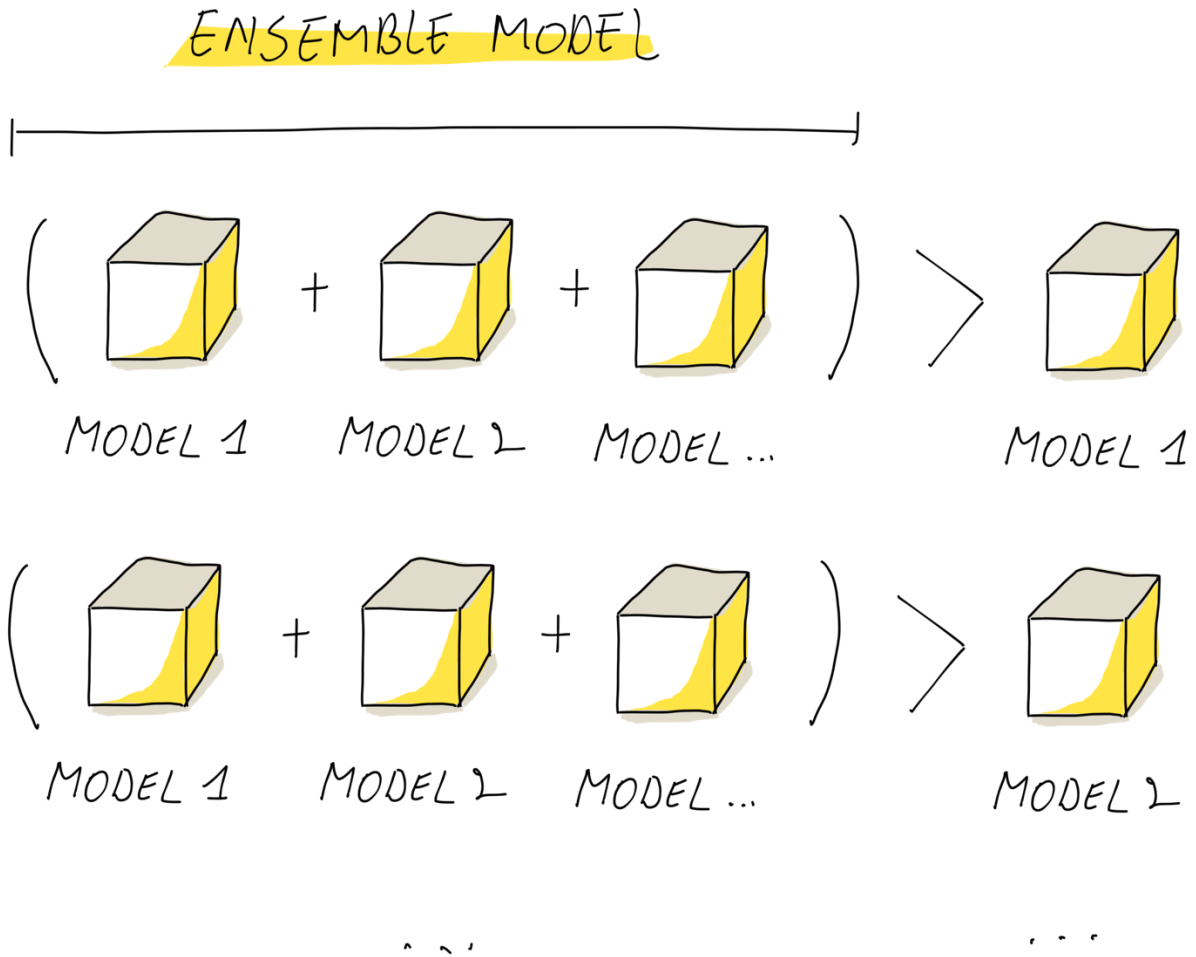


Source: Youtube

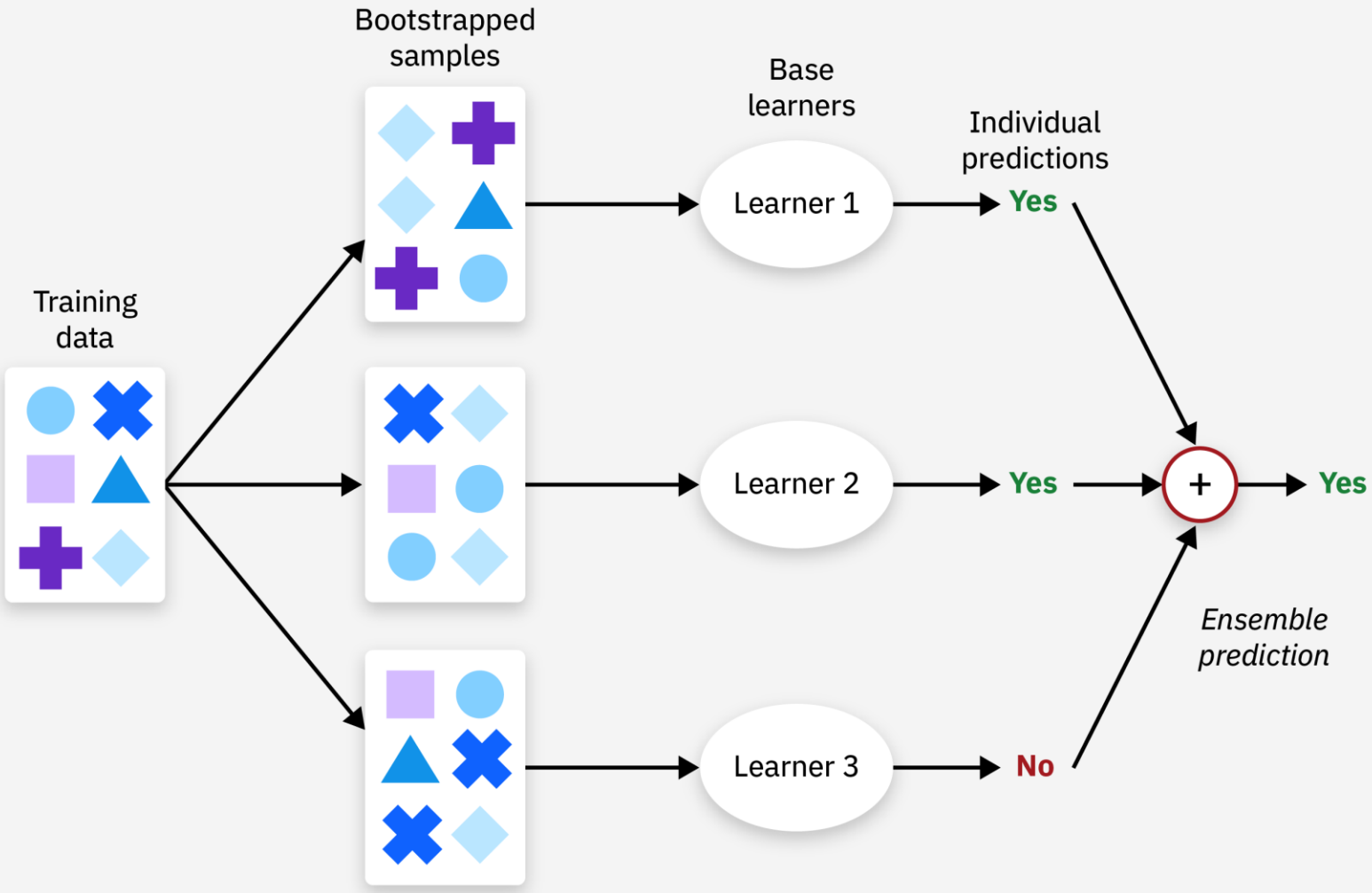

ALPHAGO
00:08:32
AlphaGo
Google DeepMind
LEE SEDOL
00:00:27


MACHINE LEARNING
SUPERVISED LEARNING
UNSUPERVISED LEARNING
REINFORCEMENT LEARNING
ENSEMBLE LEARNING
REGRESSION
CLASSIFICATION
CLUSTERING
RECOMMENDATION

Ensemble Models

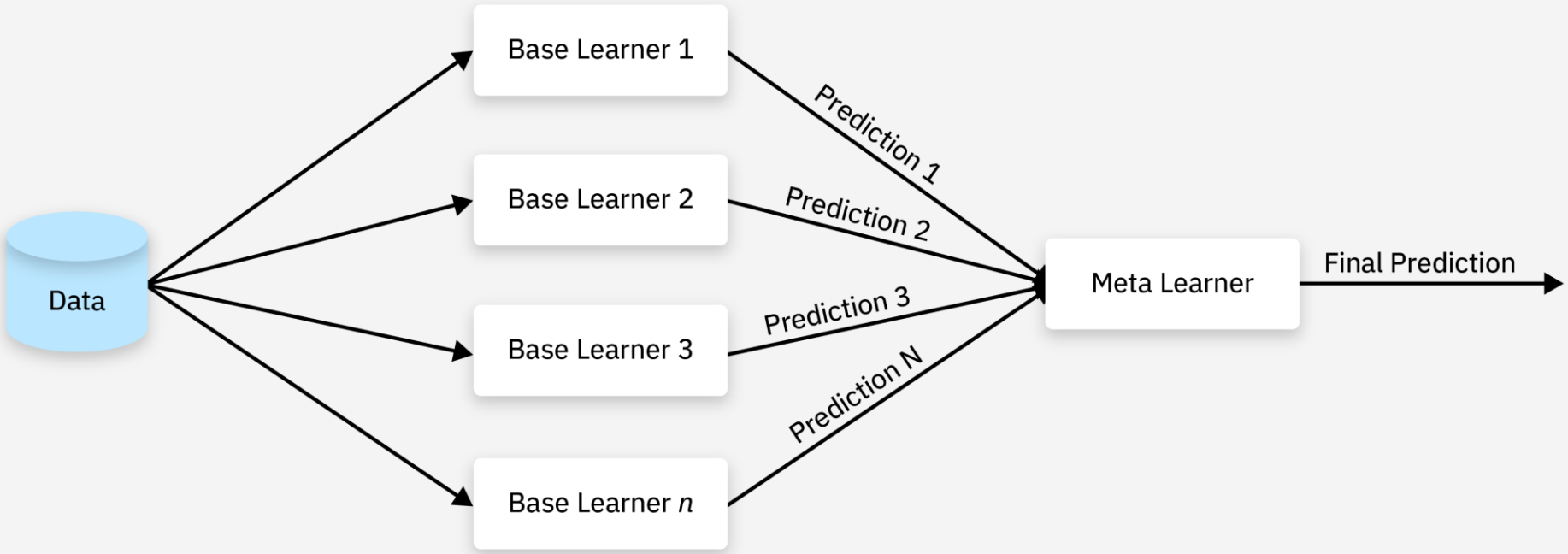


Technique 1: **Bagging**



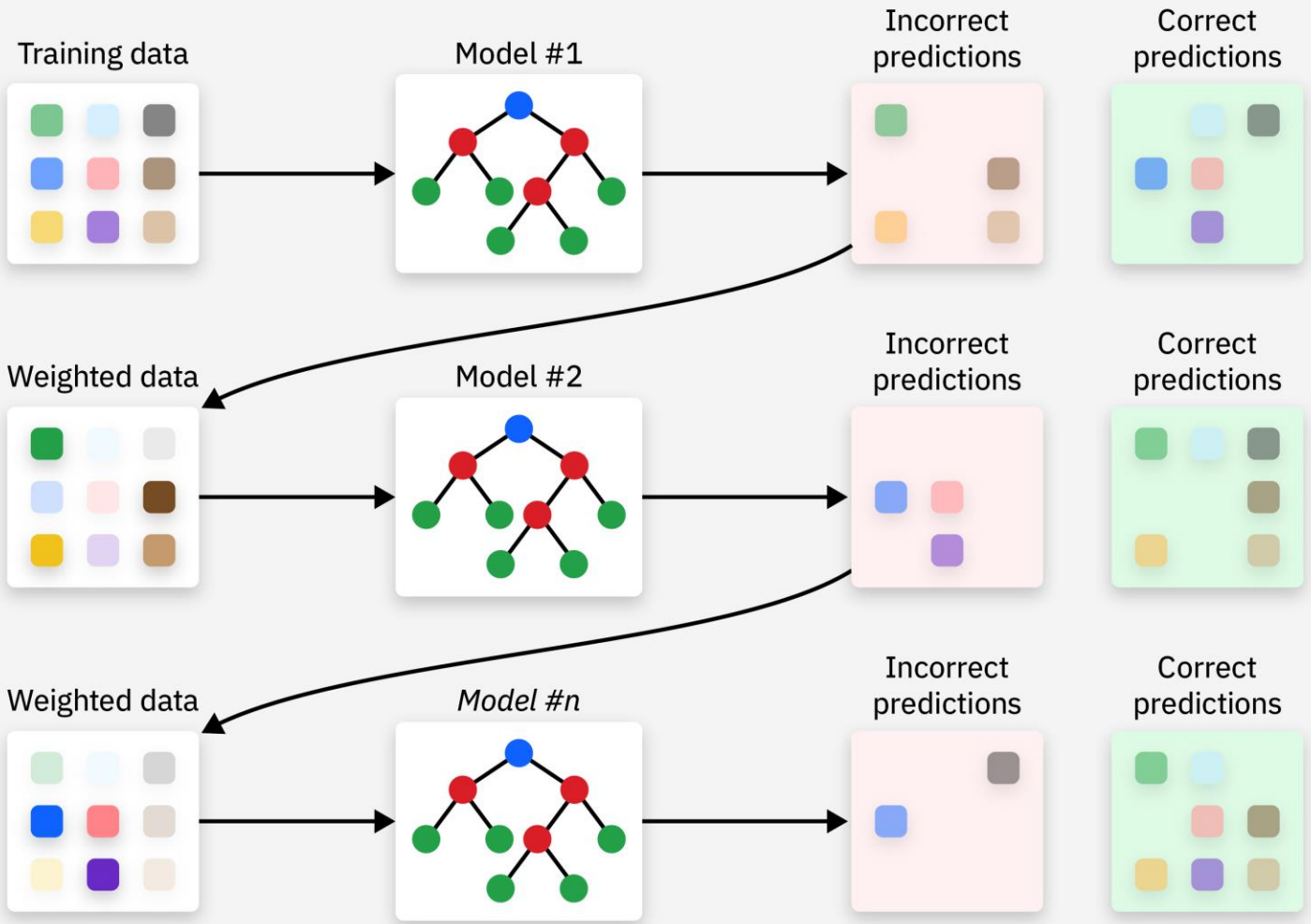
[Ensemble Learning, IBM]

Technique 2: **Stacking**



[Ensemble Learning, IBM]

Technique 3: **Boosting**

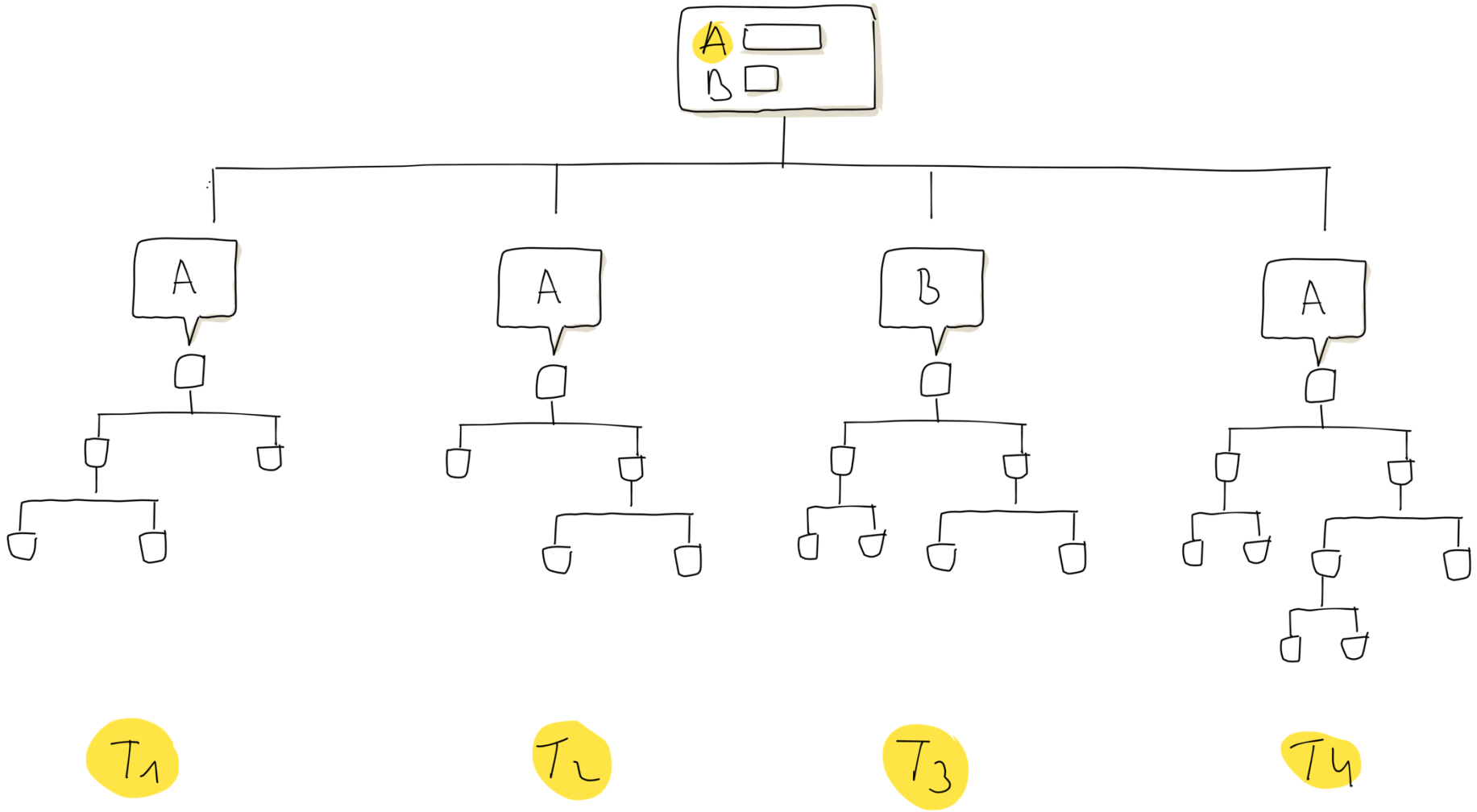


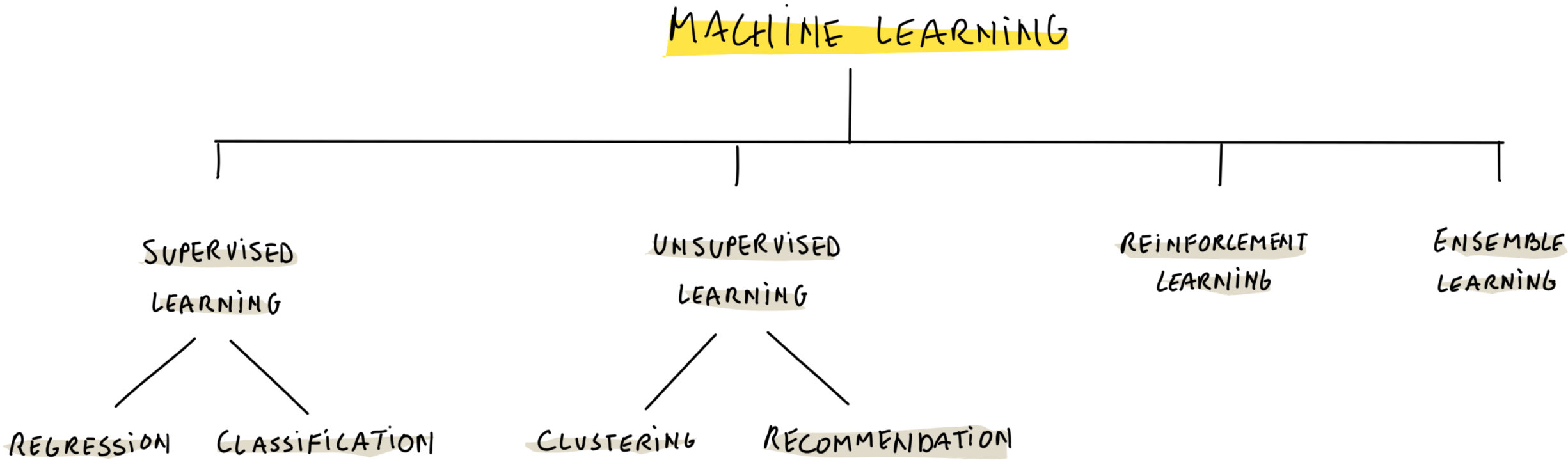
[Ensemble Learning, IBM]

Popular Ensemble Learning Algorithms

- eXtreme Gradient Boosting (XGBoost)
- ADA Boost
- Random Forests

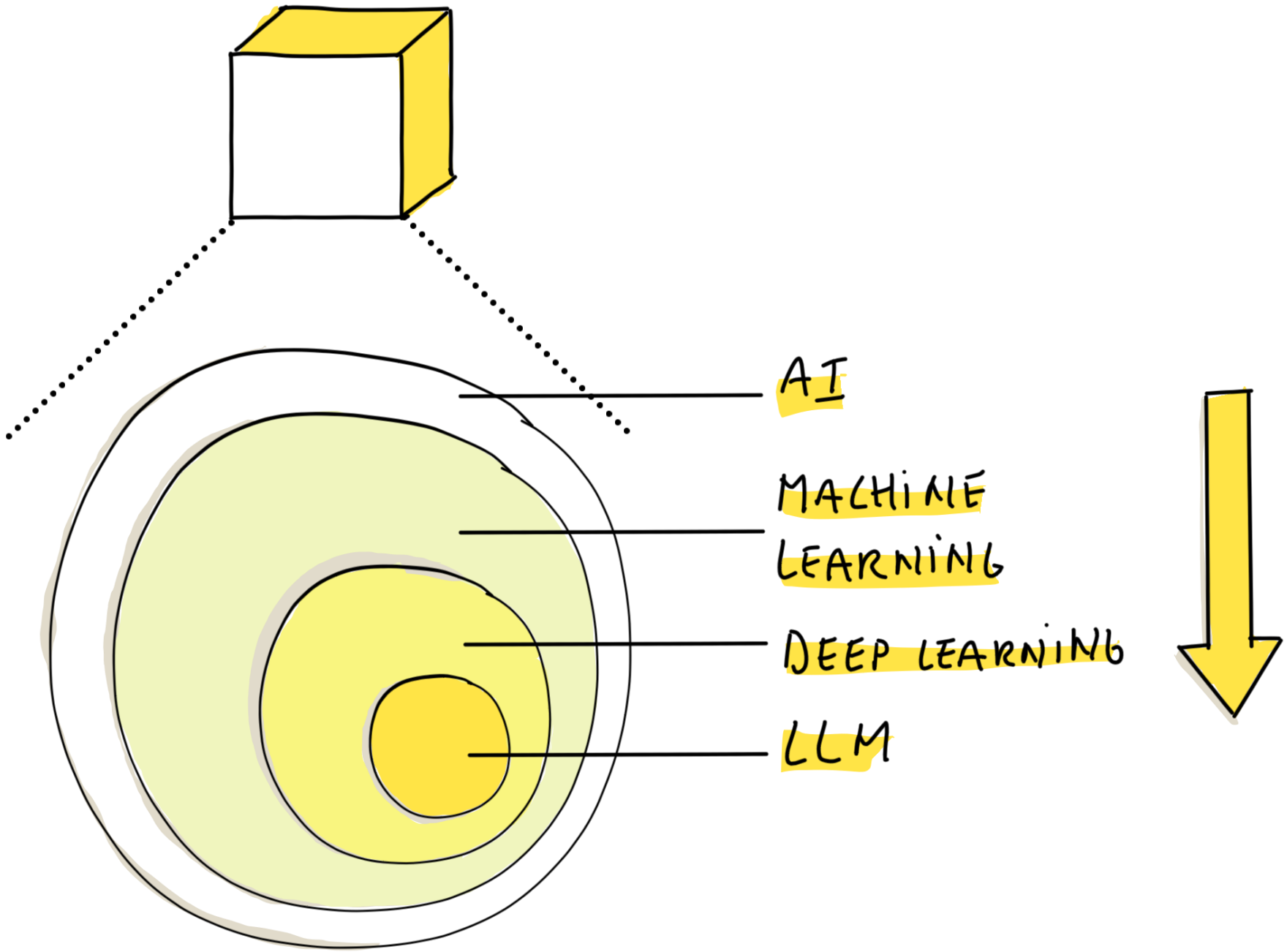
Example Model: Random Forests

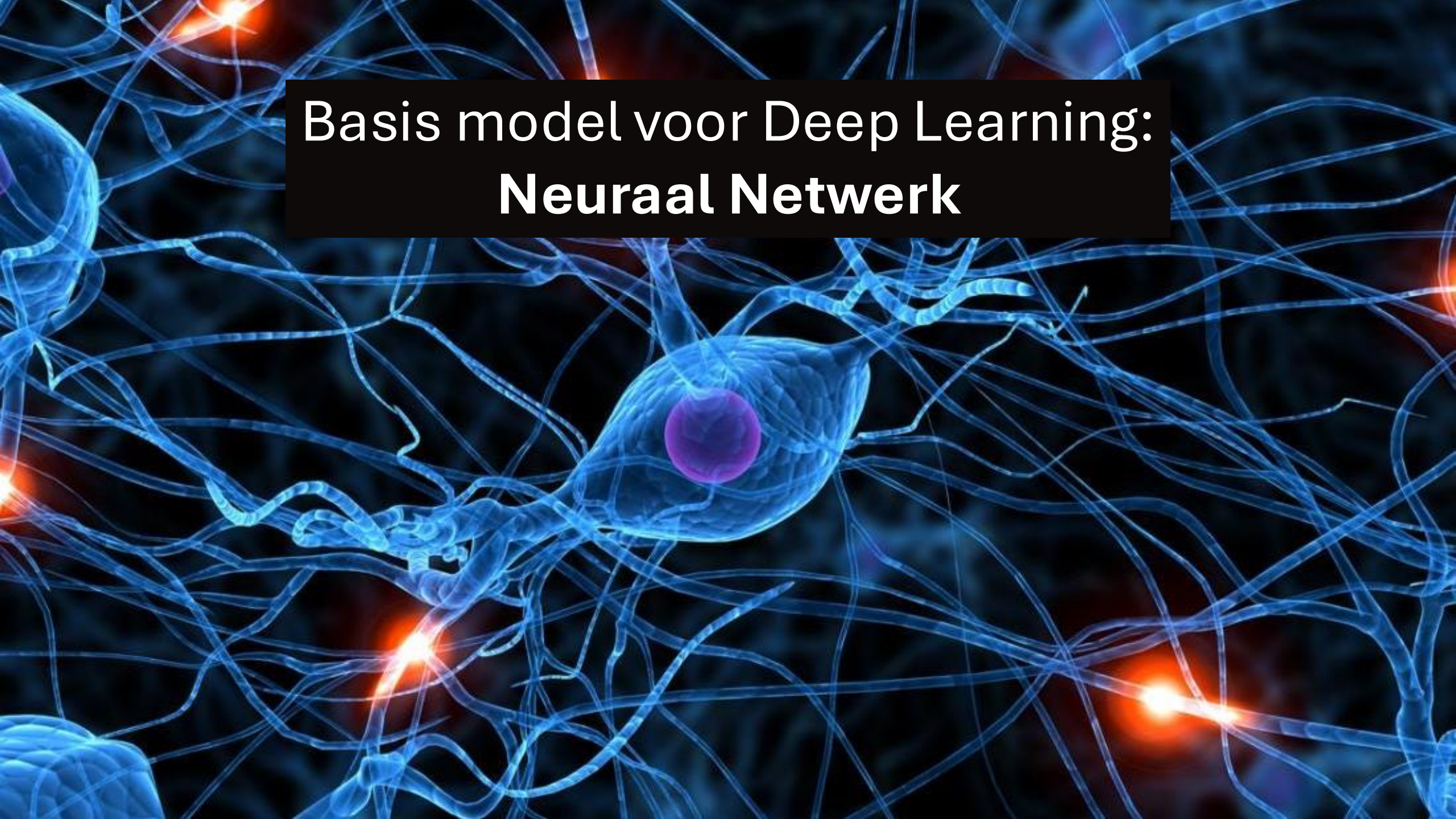


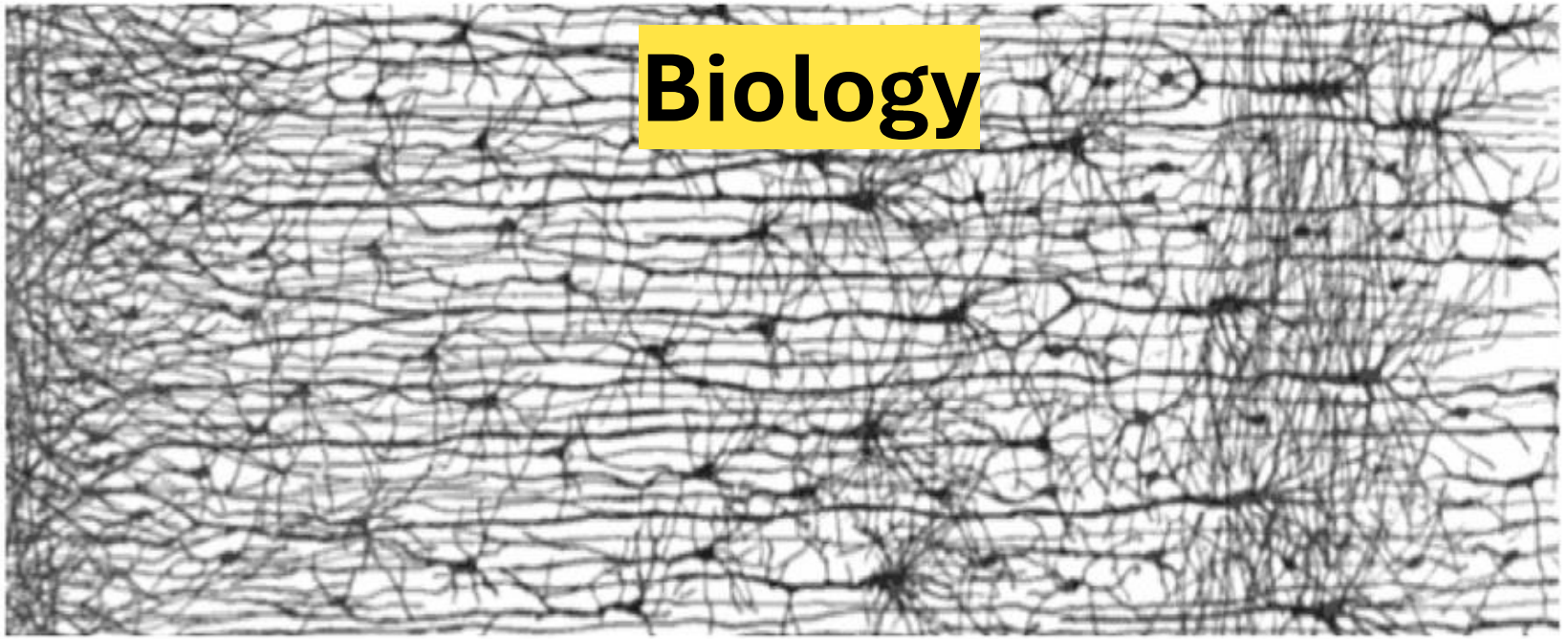

MACHINE LEARNING
SUPERVISED LEARNING
UNSUPERVISED LEARNING
REINFORCEMENT LEARNING
ENSEMBLE LEARNING
REGRESSION
CLASSIFICATION
CLUSTERING
RECOMMENDATION

DEEL 3

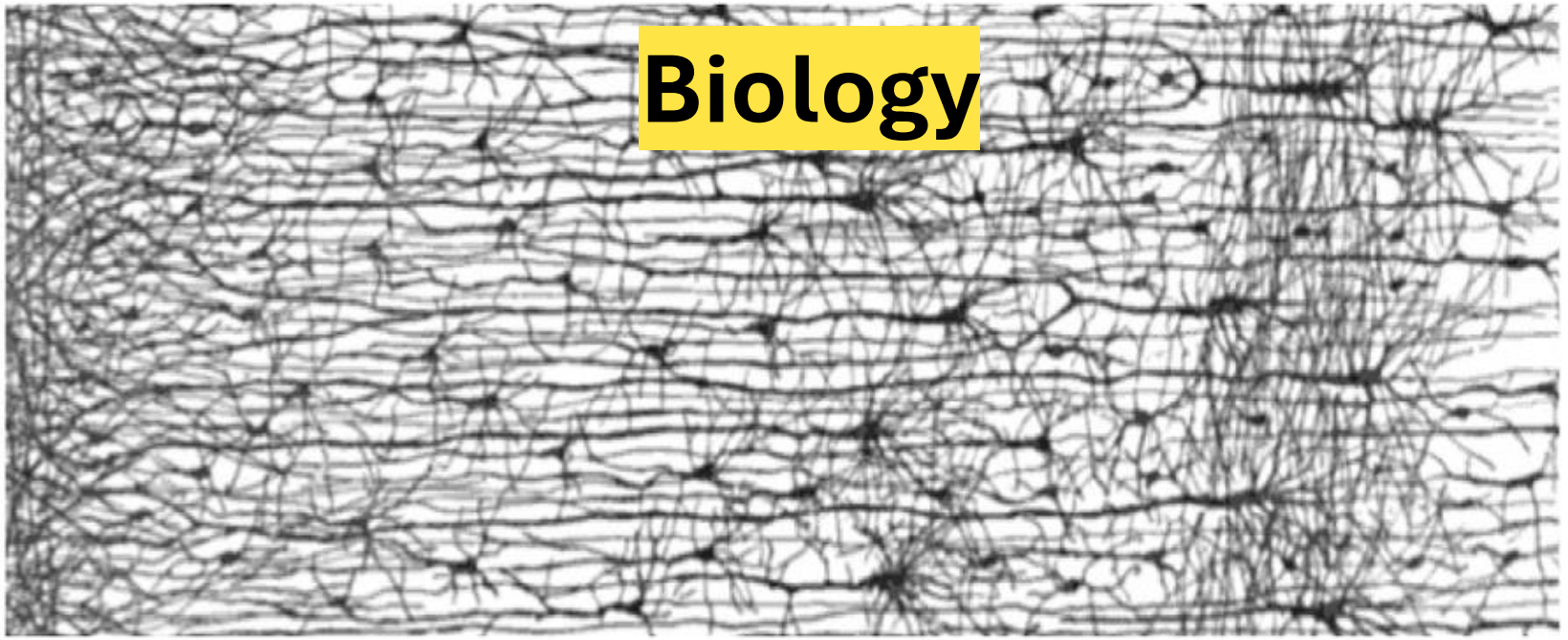
DEEP LEARNING

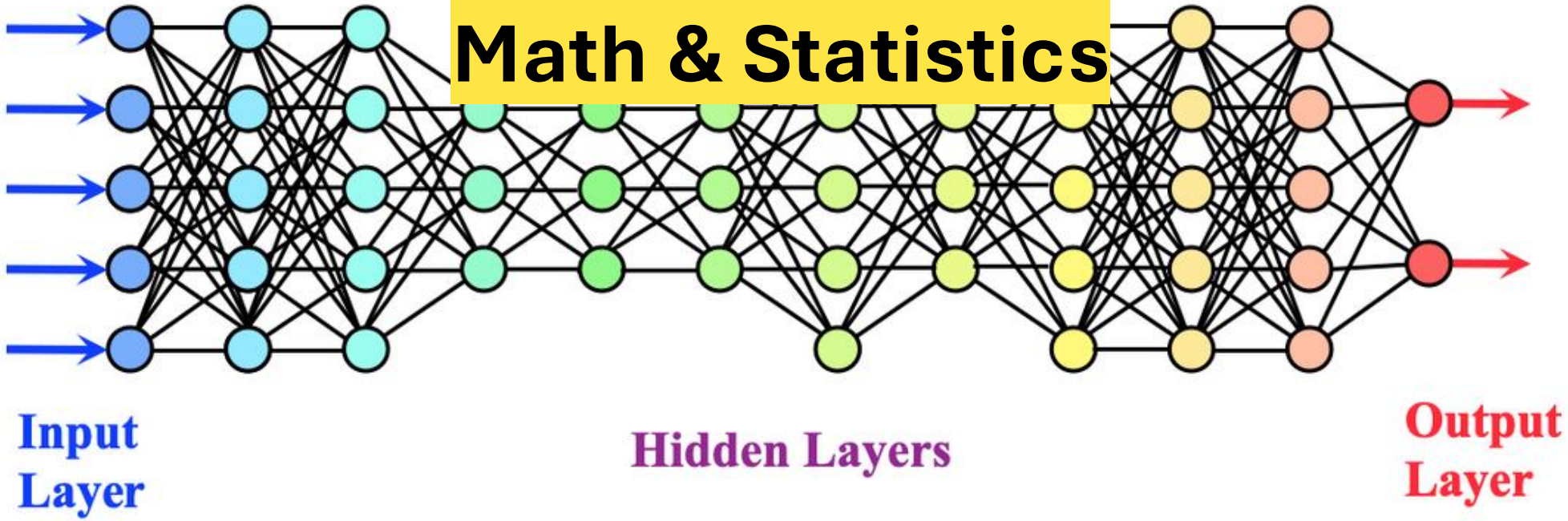

AI
MACHINE LEARNING
DEEP LEARNING
LLM


Basis model voor Deep Learning:
Neuraal Netwerk



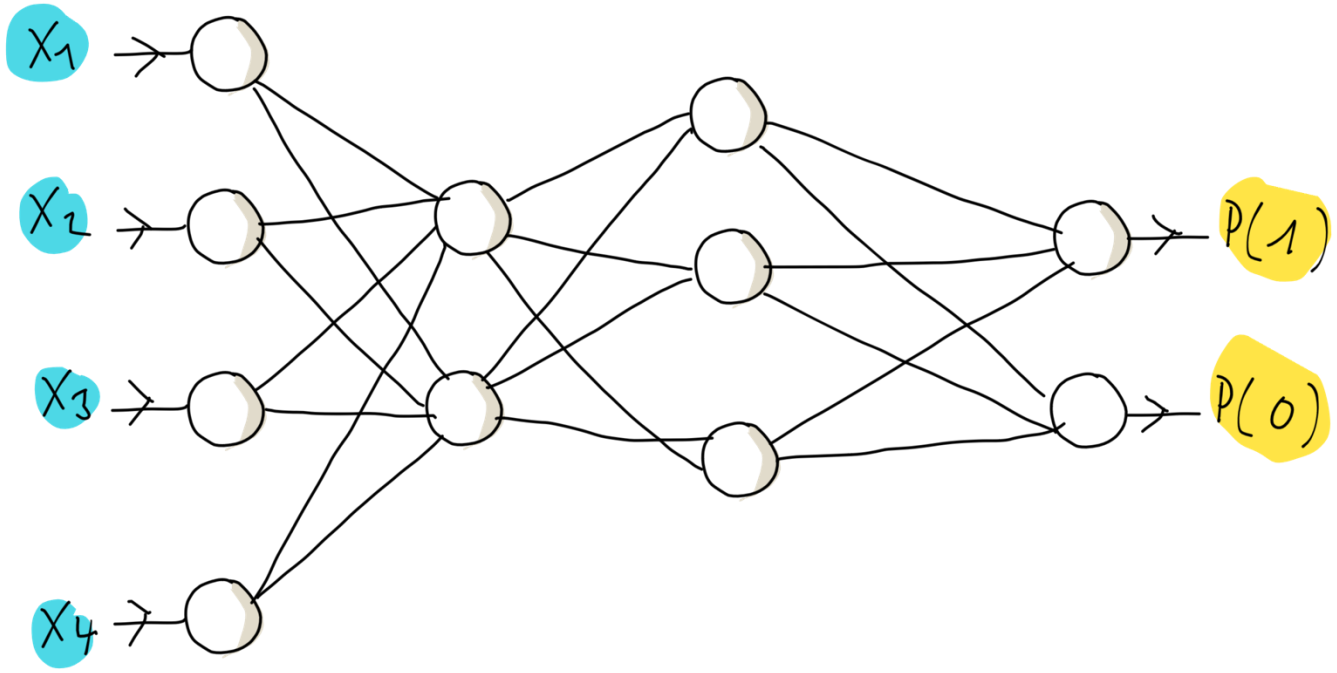
Biology


Biology


Math & Statistics
Input Layer
Hidden Layers
Output Layer

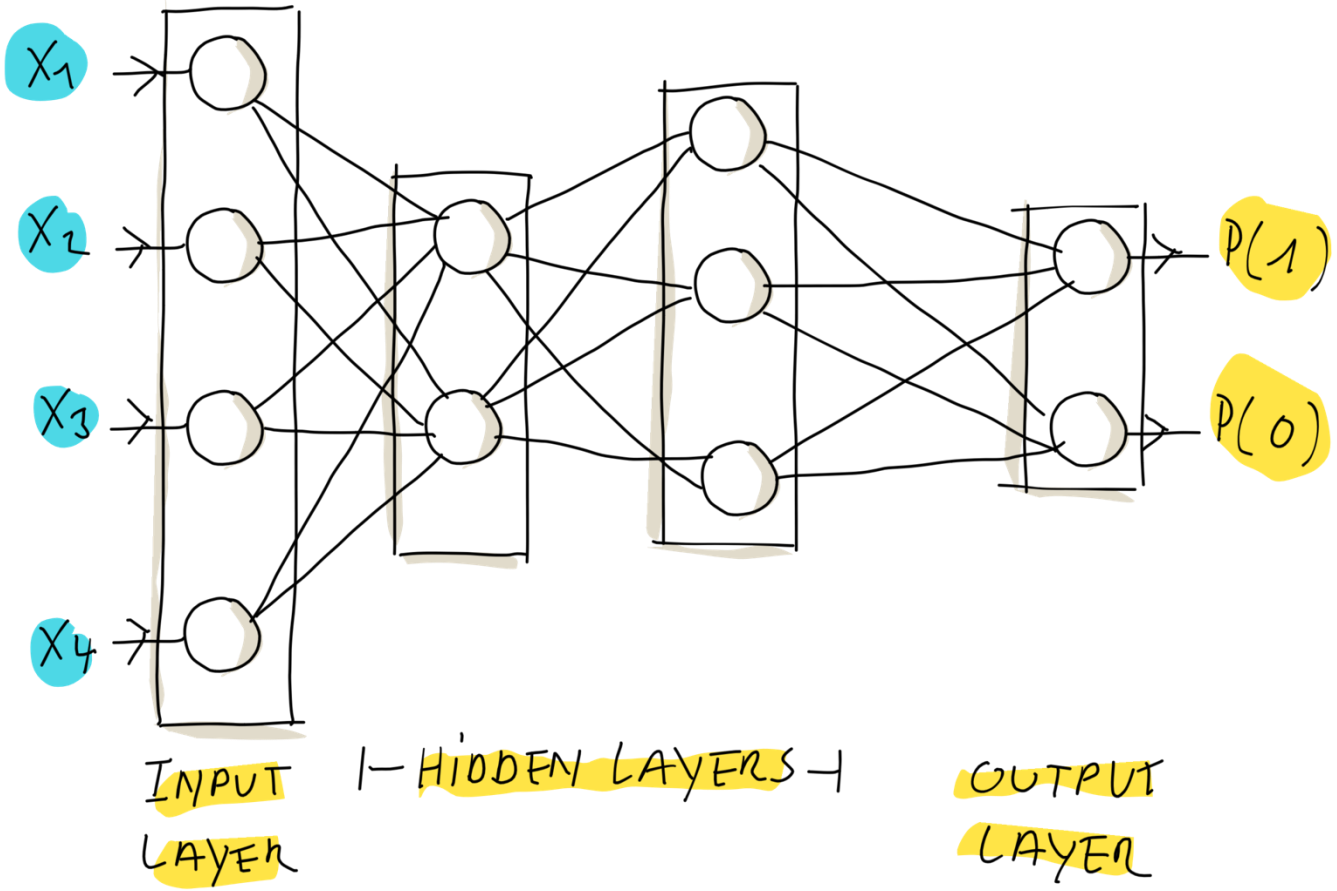
ANN: Artificial Neural Network

X_1	X_2	X_3	X_4	y
0.1	5.2	-1	T	1
0.2	4.9	-3	T	0
0.4	8	-4	F	0
0.3	3	-5	F	1
0.9	2	-2	T	0

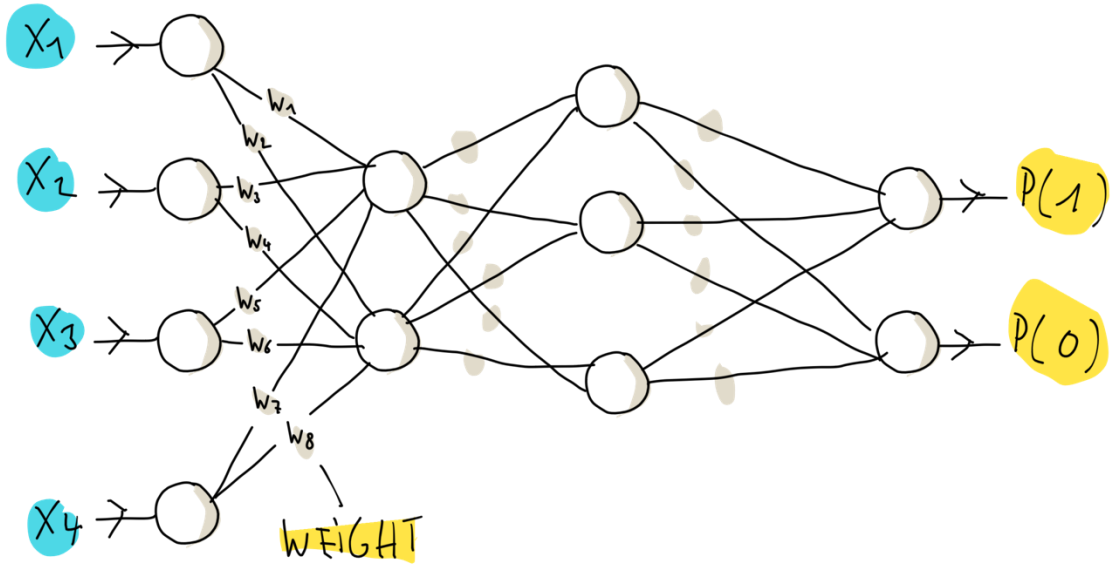


ANN: Artificial Neural Network

X_1	X_2	X_3	X_4	y
0.1	5.2	-1	T	1
0.2	4.9	-3	T	0
0.4	8	-4	F	0
0.3	3	-5	F	1
0.9	2	-2	T	0



ANN Training



1 ——— FORWARD ———→ EVALUATE (2)

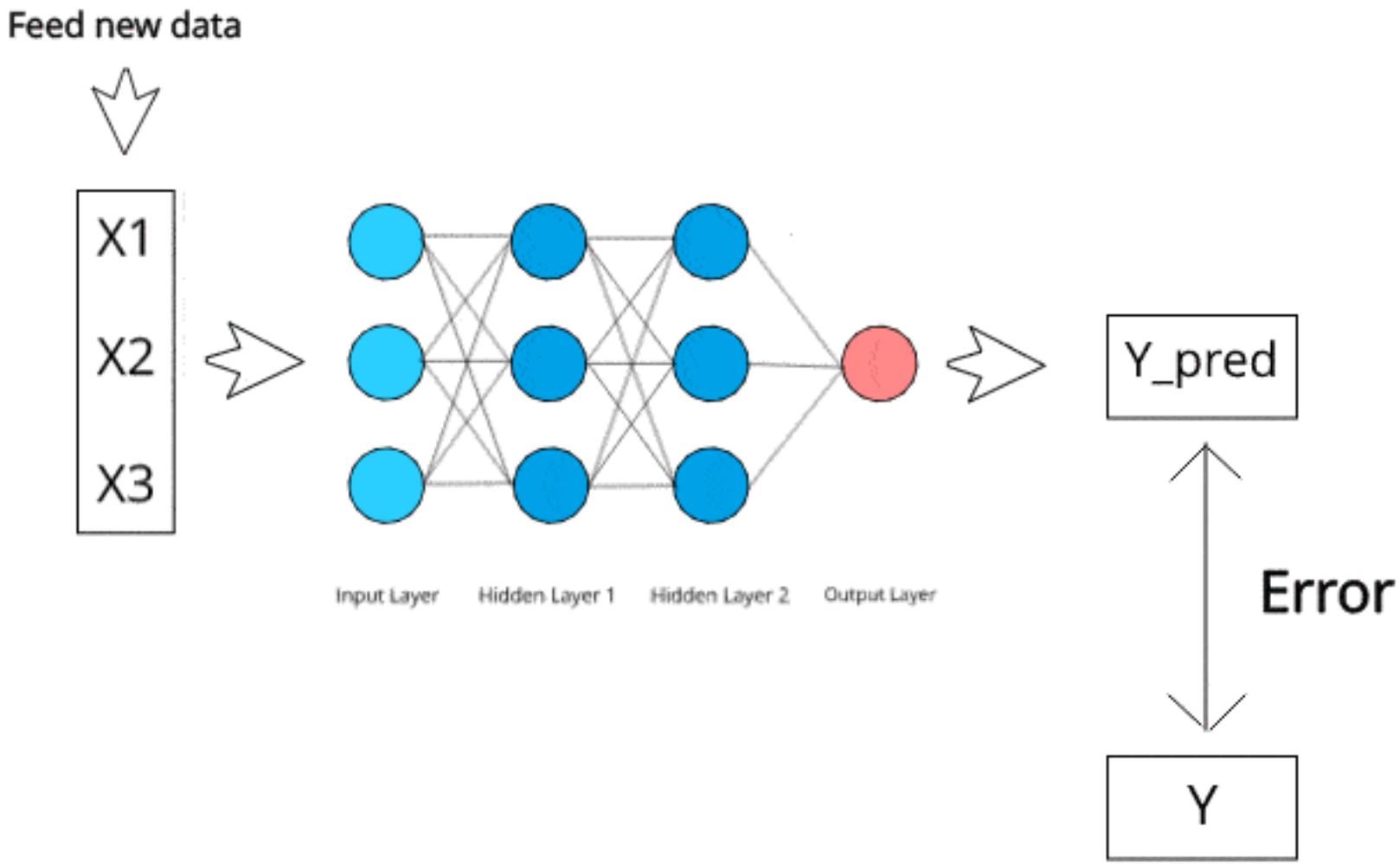
3 ←——— BACKWARD PROPAGATION ——— {W1 = 0,2, W2 = 0,1, W3 = 0,8, W4 = 0,4, ...}

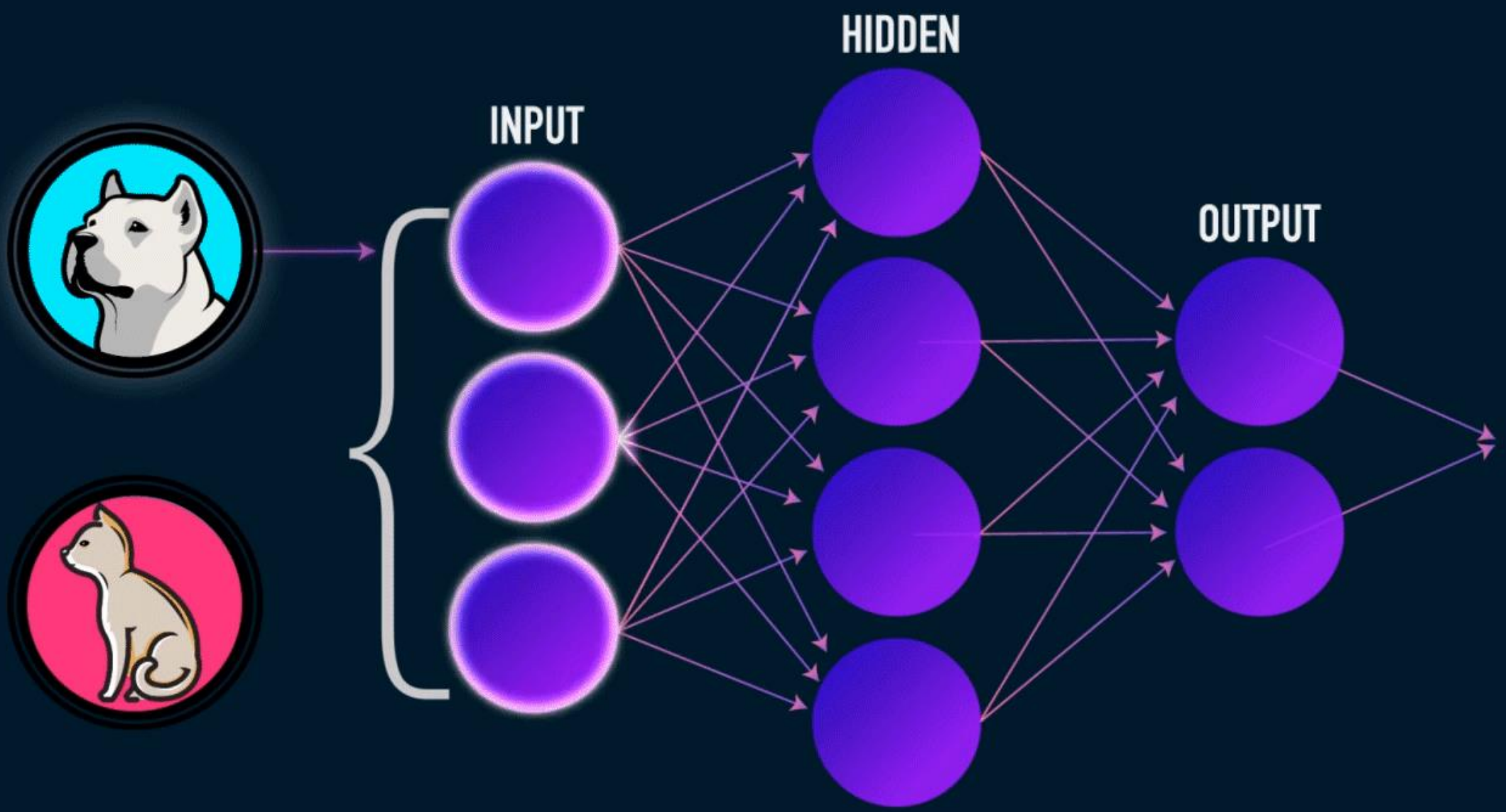
	1	0
1		
0		

4 ——— FORWARD ———→ EVALUATE (5)

6 ←——— BACKWARD PROPAGATION ——— {W1 = 0,3, W2 = 0,1, W3 = 0,8, W4 = 0,7, ...}

	1	0
1		
0		

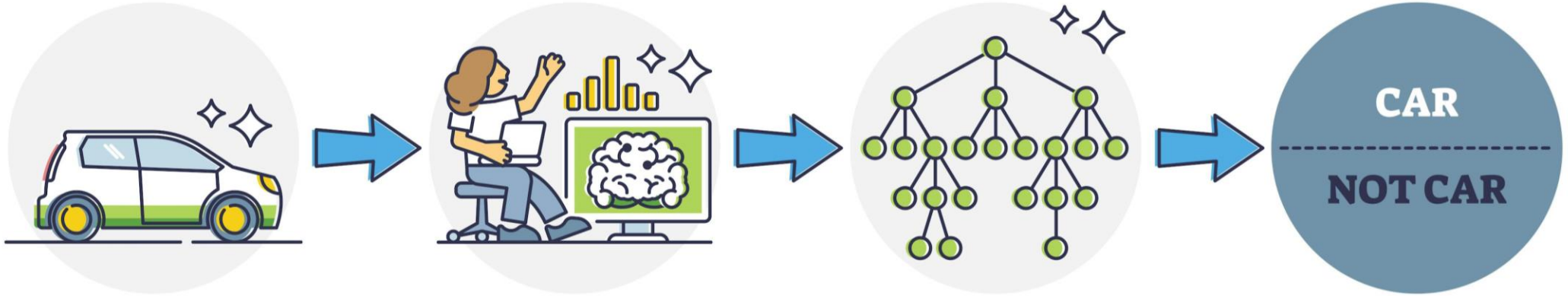

Feed new data
X1
X2
X3
Input Layer
Hidden Layer 1
Hidden Layer 2
Output Layer
Y_pred
Error
Y


HIDDEN
INPUT
OUTPUT

Example: **Diabetes Prediction**

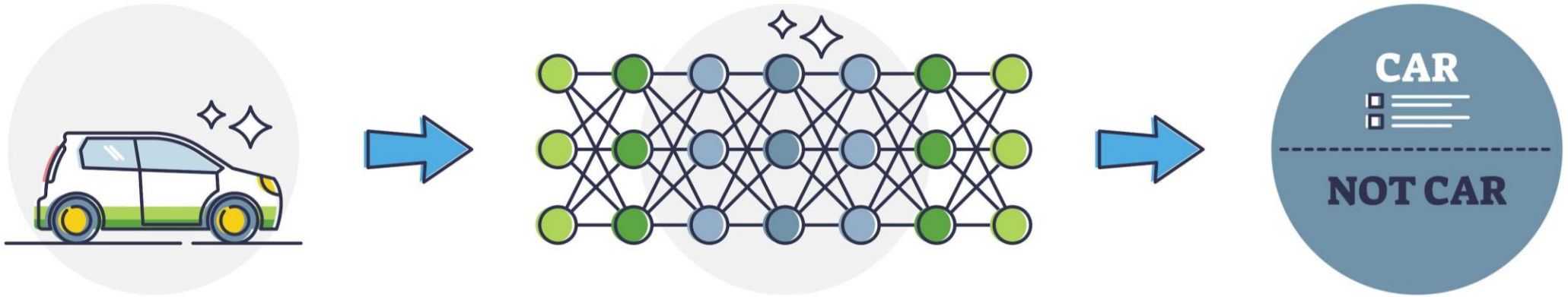
	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1
5	5	116	74	0	0	25.6	0.201	30	0
6	3	78	50	32	88	31.0	0.248	26	1
7	10	115	0	0	0	35.3	0.134	29	0
8	2	197	70	45	543	30.5	0.158	53	1
9	8	125	96	0	0	0.0	0.232	54	1

MACHINE LEARNING

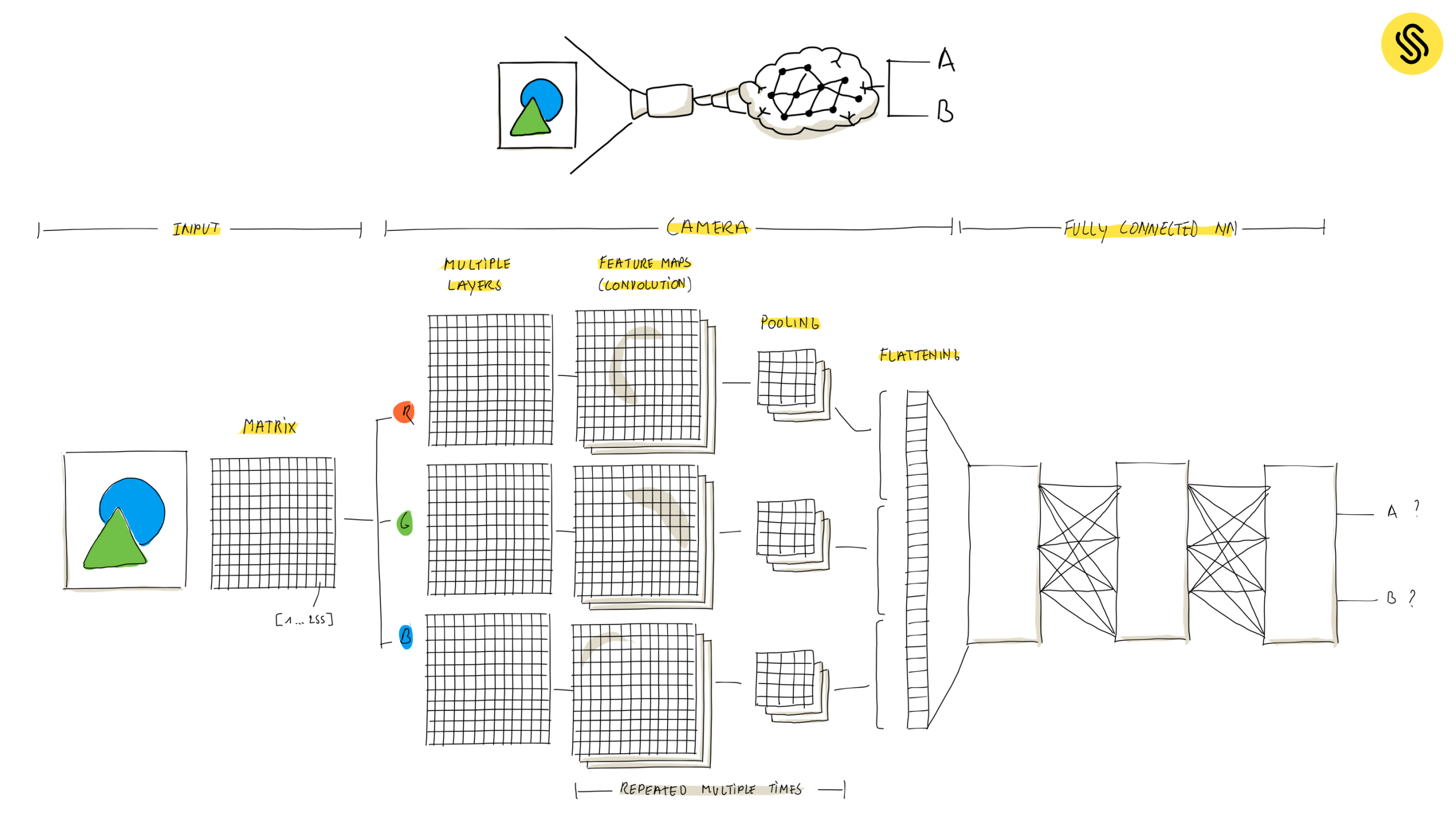


INPUT | FEATURE EXTRACTION | CLASSIFICATION | OUTPUT

DEEP LEARNING



INPUT | FEATURE EXTRACTION + CLASSIFICATION | OUTPUT

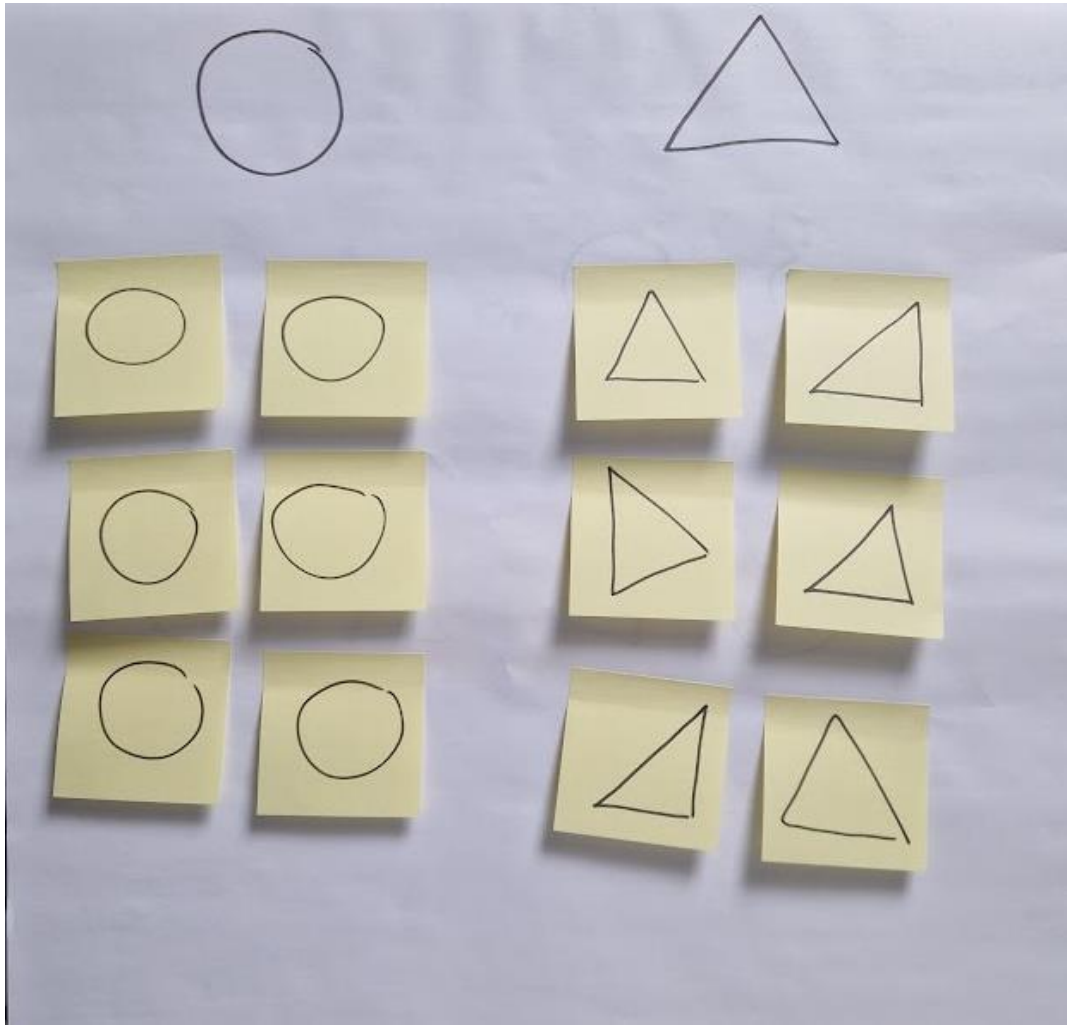

A
B
INPUT
CAMERA
FULLY CONNECTED NN
MULTIPLE LAYERS
FEATURE MAPS (CONVOLUTION)
POOLING
FLATTENING
MATRIX
R
G
B
[1 ... 255]
A ?
B ?
REPEATED MULTIPLE TIMES

2D CNN ANIMATION

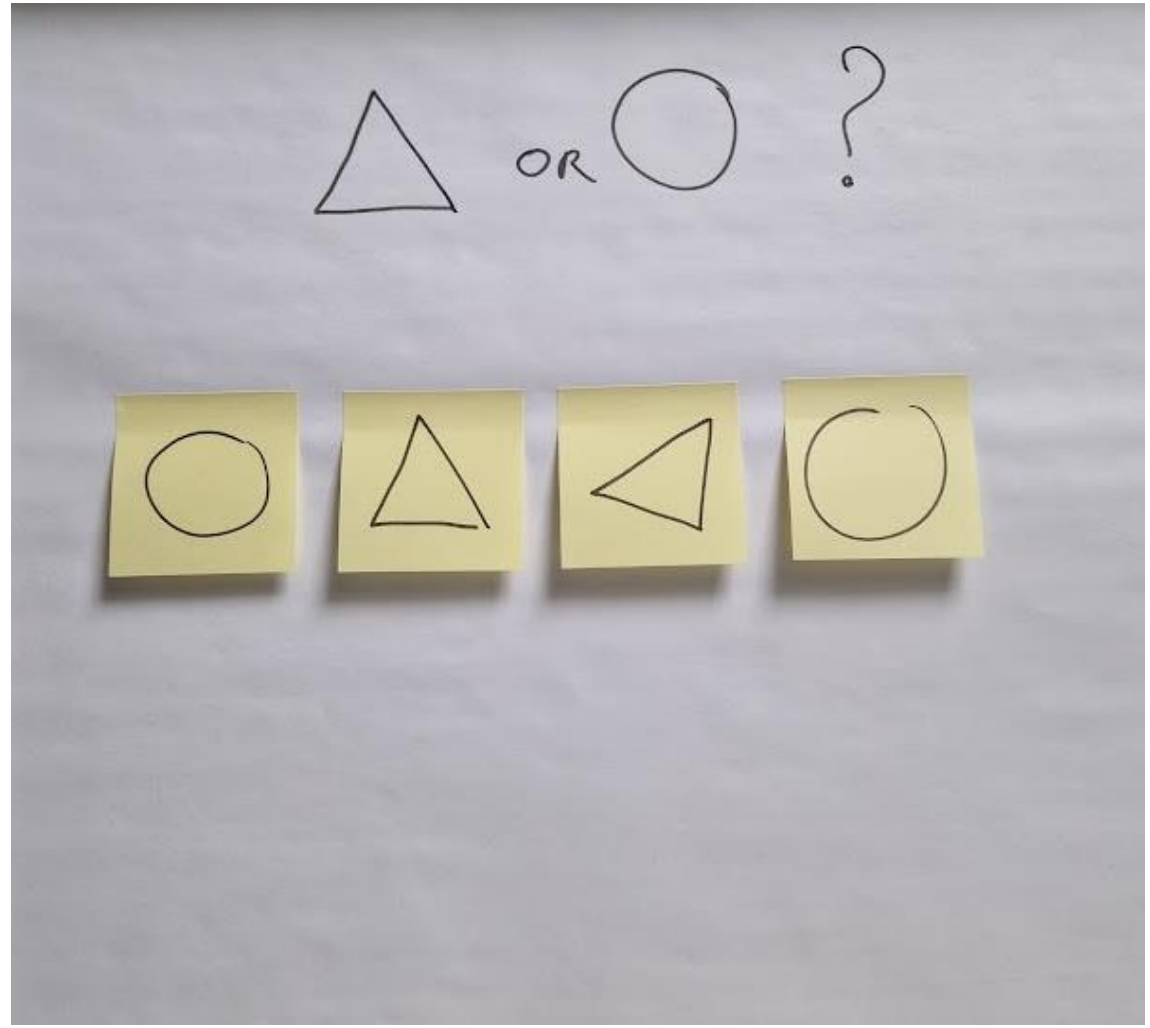
BASIC ARCHITECTURE

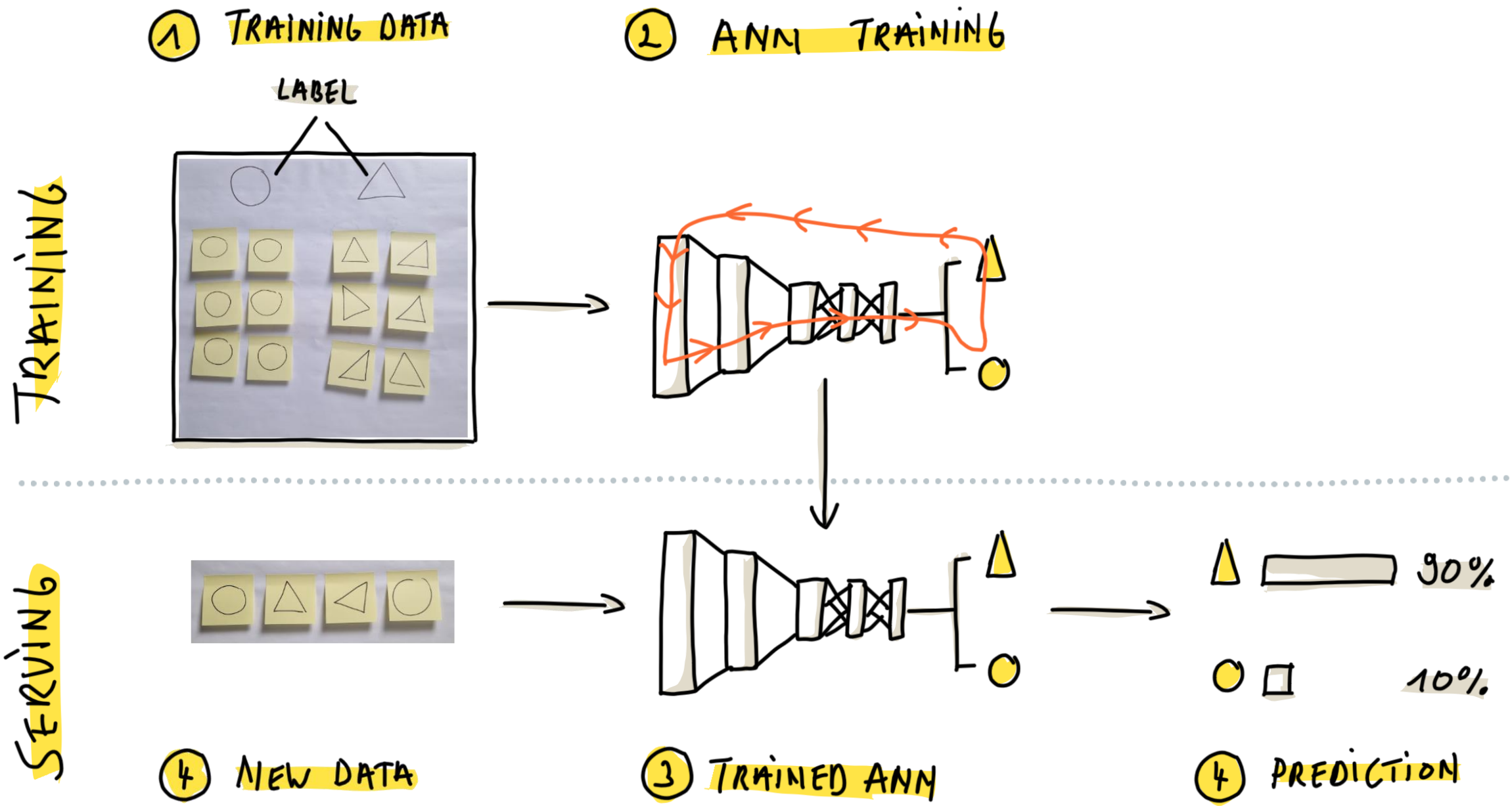
Source: https://www.youtube.com/watch?v=CXOGvCMLrkA

TRAINING



SERVING



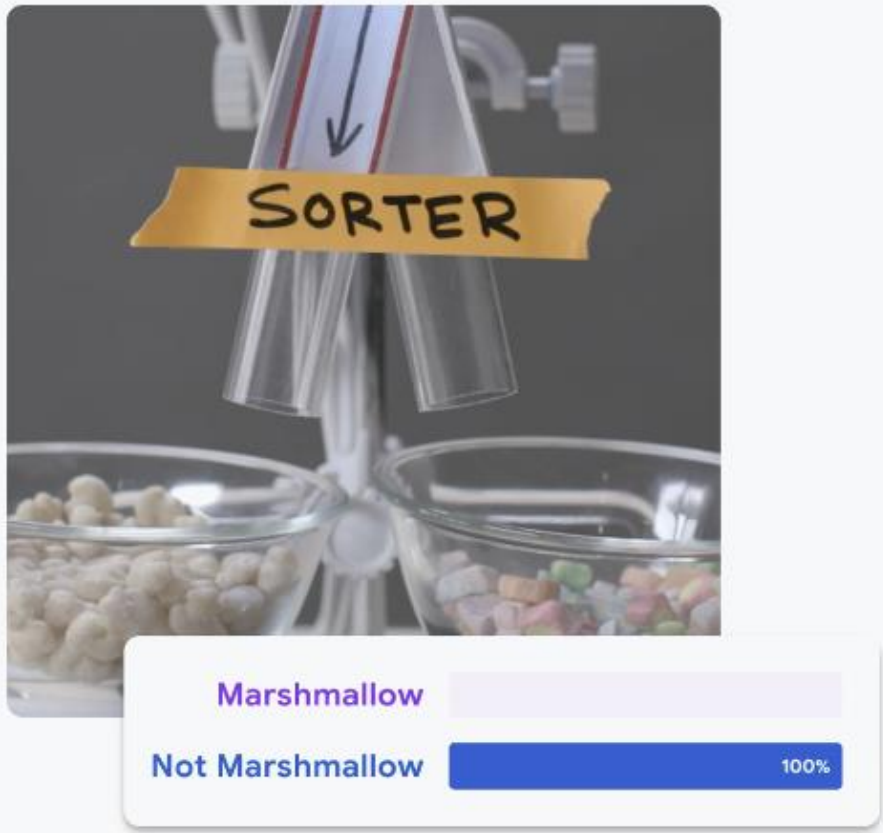

1 TRAINING DATA
LABEL
2 ANN TRAINING
TRAINING
SERVING
90%
10%
4 NEW DATA
3 TRAINED ANN
4 PREDICTION

Teachable Machine

Train a computer to recognize your own images, sounds, & poses.

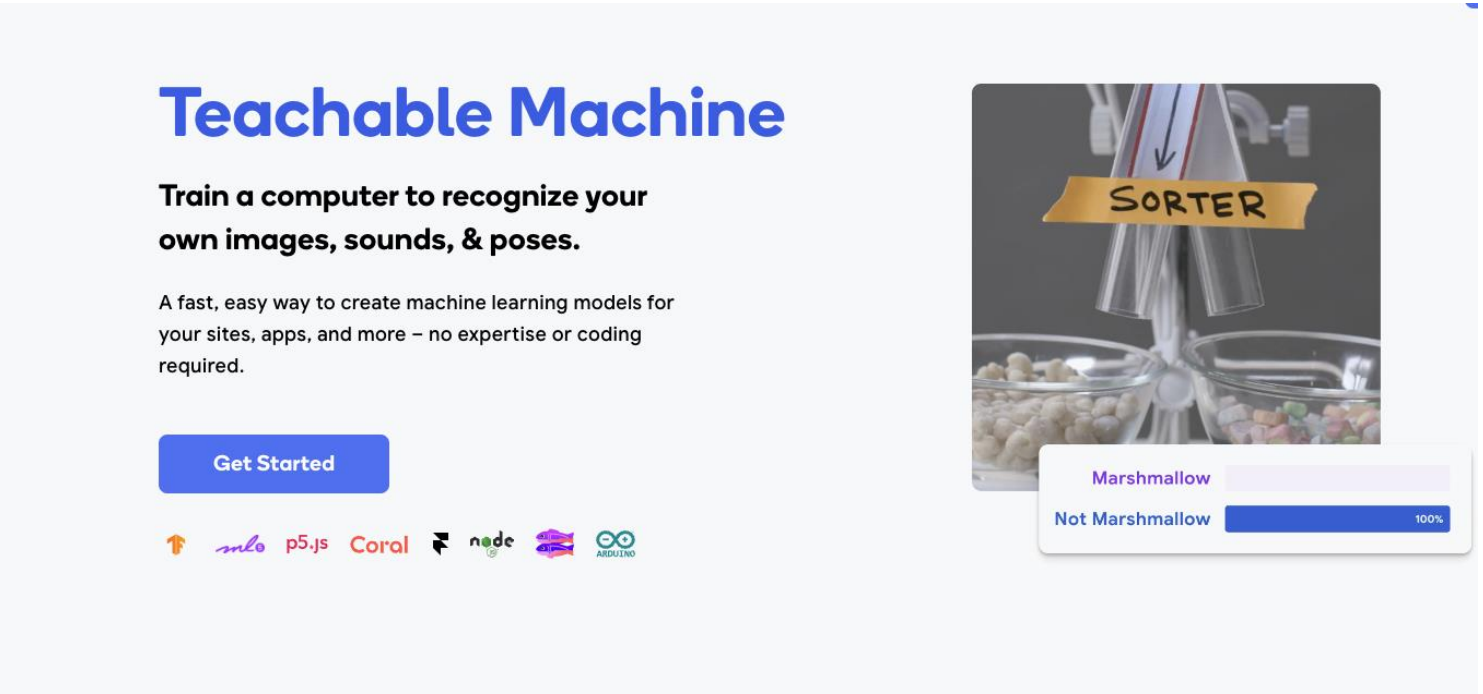
A fast, easy way to create machine learning models for your sites, apps, and more – no expertise or coding required.

Get Started



DEMO

OEFENING : Train je eigen **Deep Learning** model

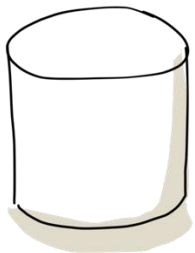


- Groepjes van 4
- Train het model om volgende bewegingen te herkennen:
  - Zwaaien
  - Drinken
  - Boksen
- Na het trainen probeer je modellen van andere groepen uit

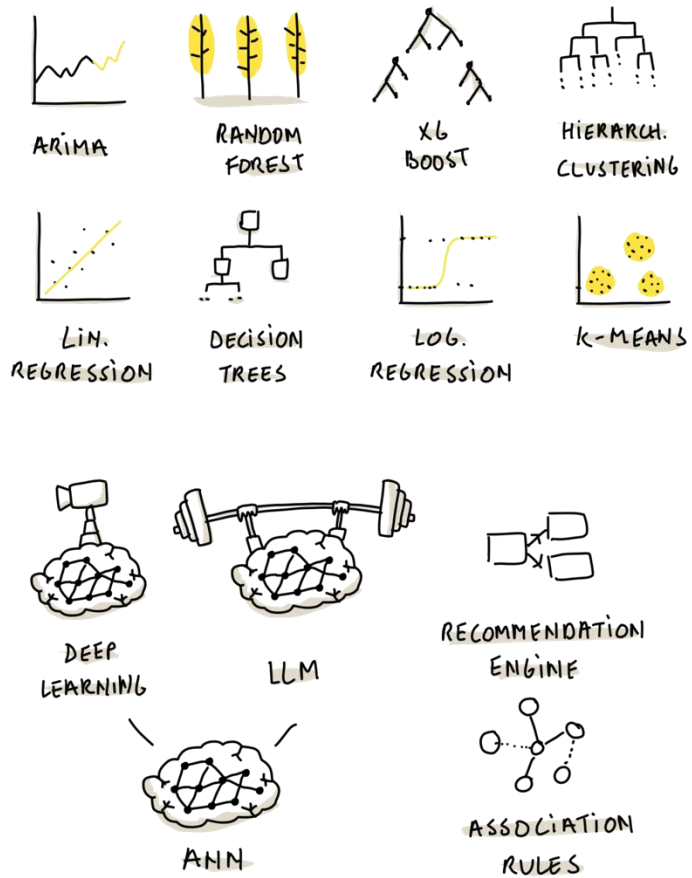
DEEL 4

AI MODEL QUALITY

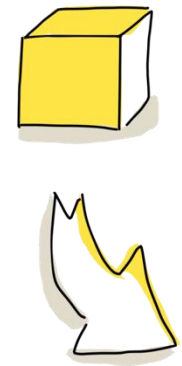
DATA



TRAINING



AI MODEL



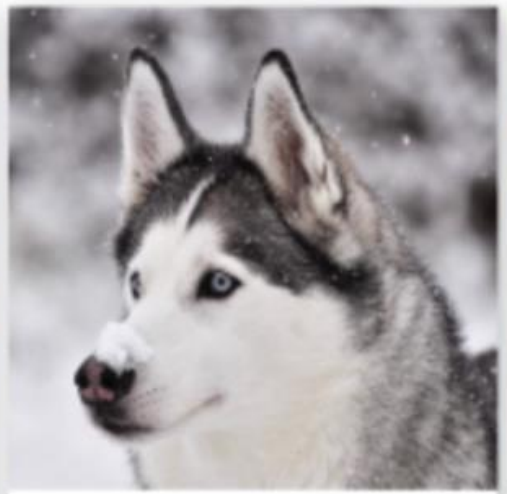
WHAT'S THE QUALITY OF THIS MODEL

Husky vs Wolf

Husky vs Wolf

Predicted: Husky
True: Wolf

Predicted: Wolf
True: Wolf



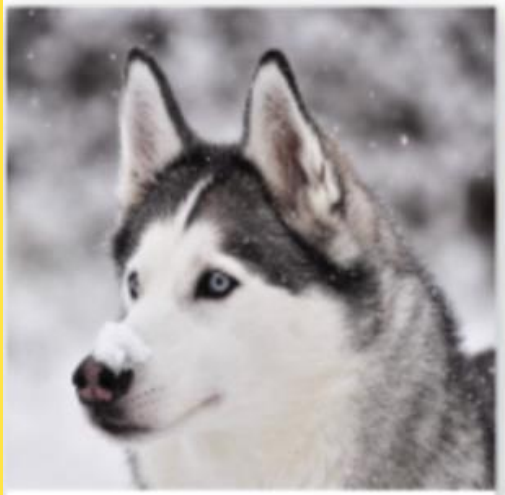
Predicted: Wolf
True: Husky

Predicted: Husky
True: Husky

Husky vs Wolf

Predicted: Husky
True: Wolf

Predicted: Wolf
True: Wolf



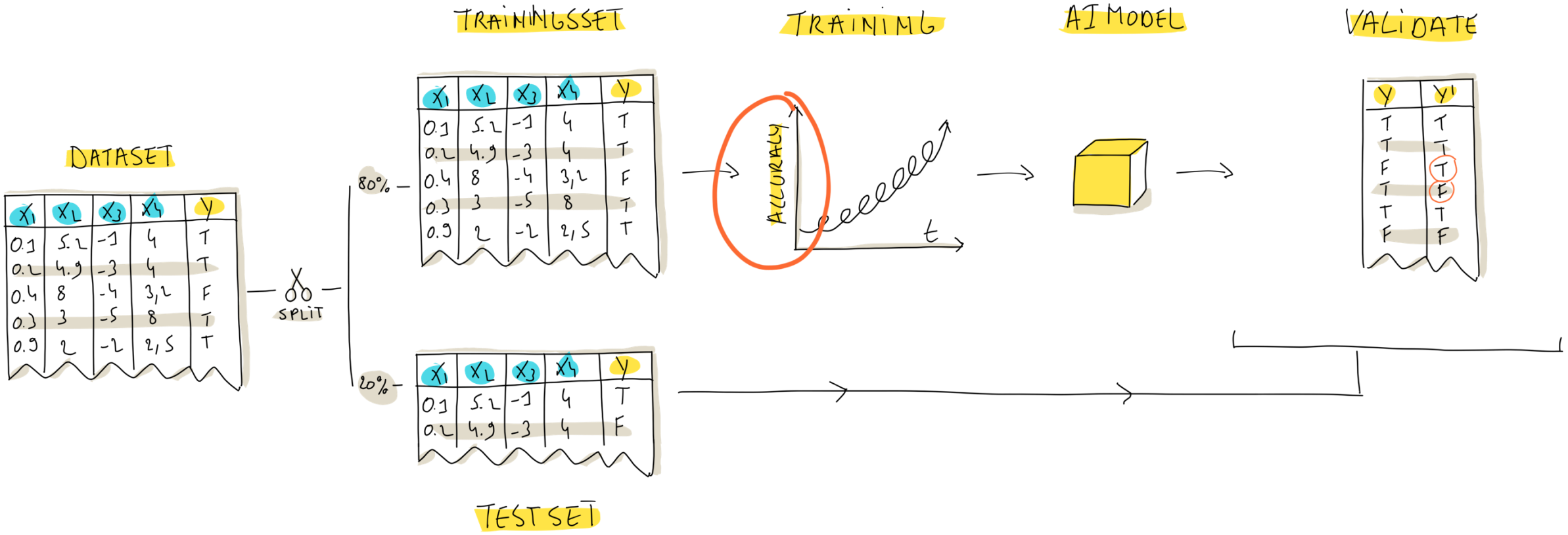
Predicted: Wolf
True: Husky

Predicted: Husky
True: Husky

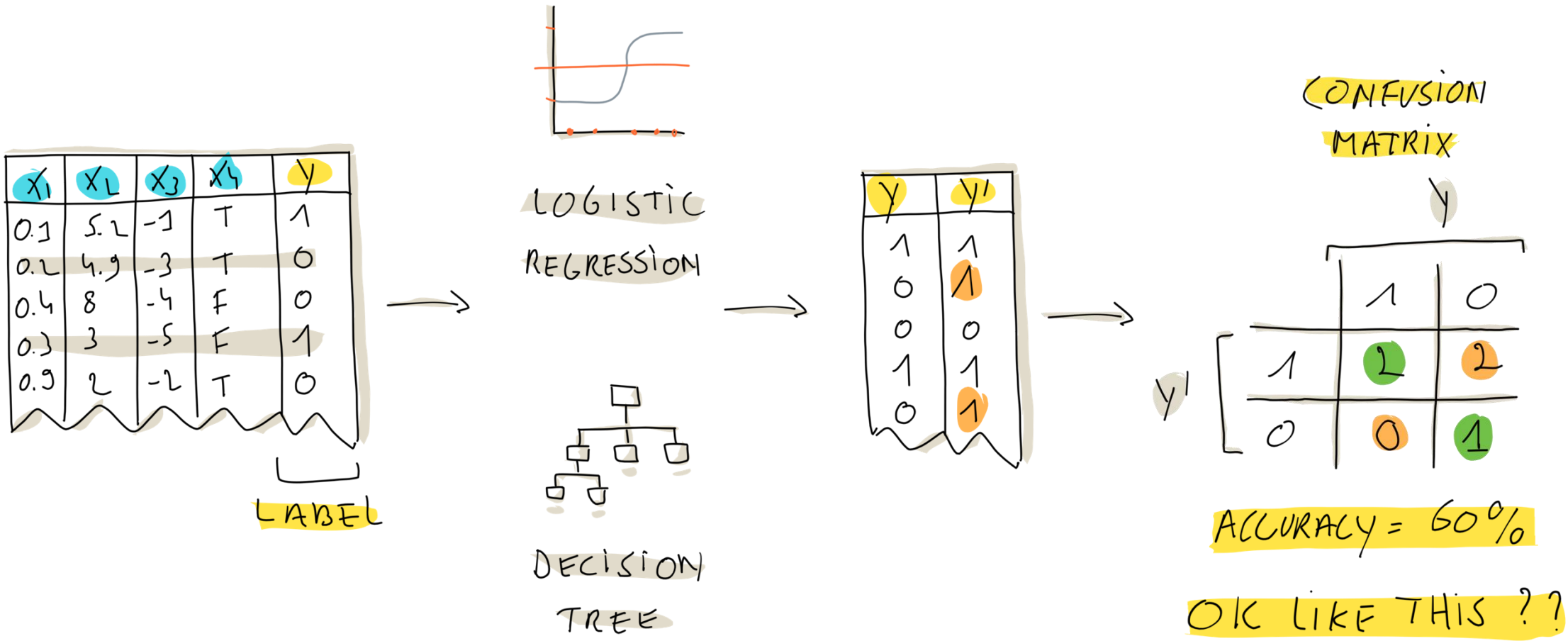
Husky vs Wolf

No Snow = Husky	Snow = Wolf	Snow = Wolf	No Snow = Wolf
			
Predicted: Husky True: Wolf	Predicted: Wolf True: Wolf	Predicted: Wolf True: Husky	Predicted: Husky True: Husky

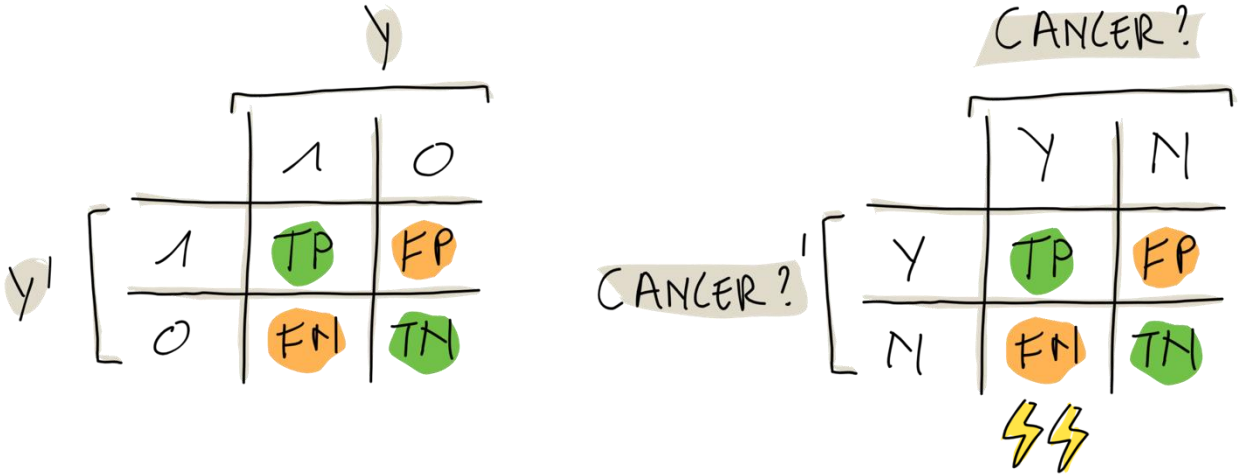
General Testing Approach

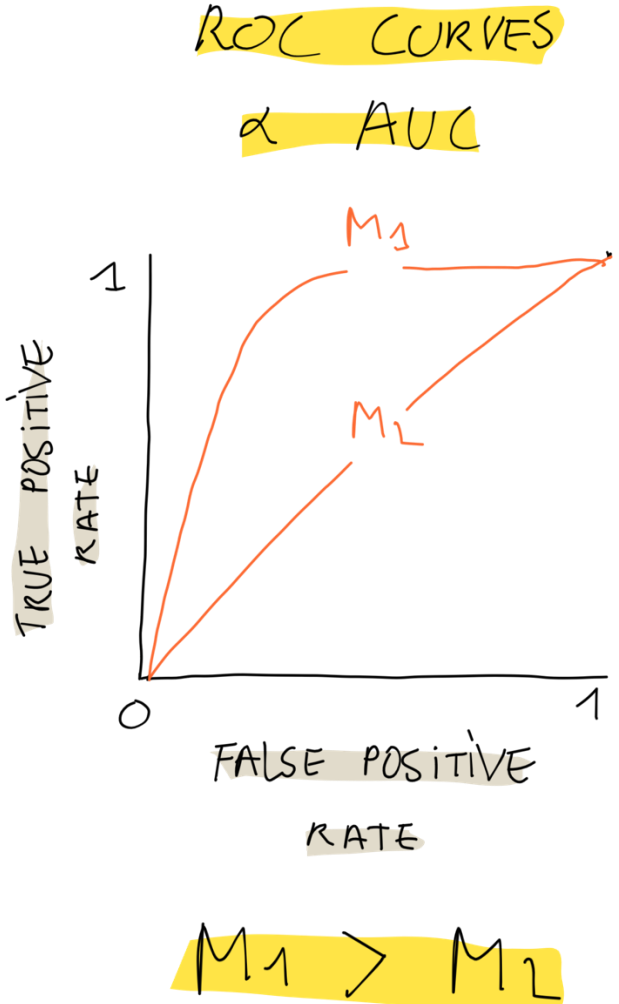


Accuracy: **Classification**



Accuracy: **Classification**





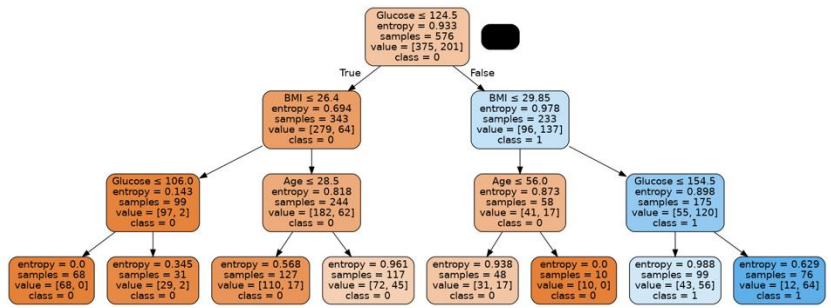
ROC: Receiver-operating characteristic
AUC: Area Under Curve

Example: **Diabetes Prediction**

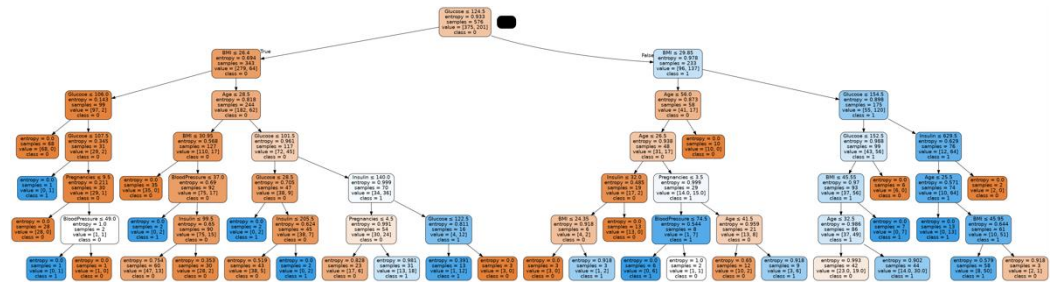
	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1
5	5	116	74	0	0	25.6	0.201	30	0
6	3	78	50	32	88	31.0	0.248	26	1
7	10	115	0	0	0	35.3	0.134	29	0
8	2	197	70	45	543	30.5	0.158	53	1
9	8	125	96	0	0	0.0	0.232	54	1

Logistic Regression

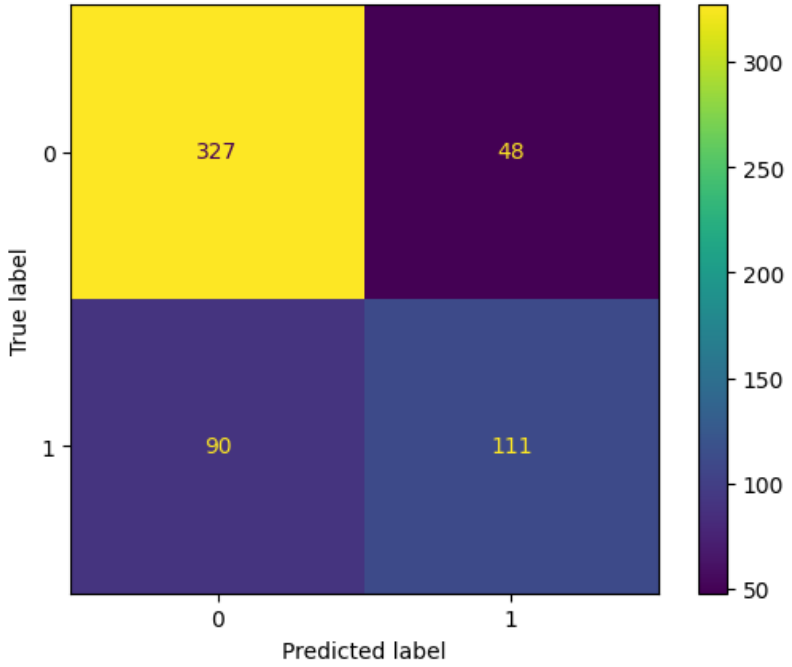
Decision tree - 3 levels



Decision tree - 6 levels

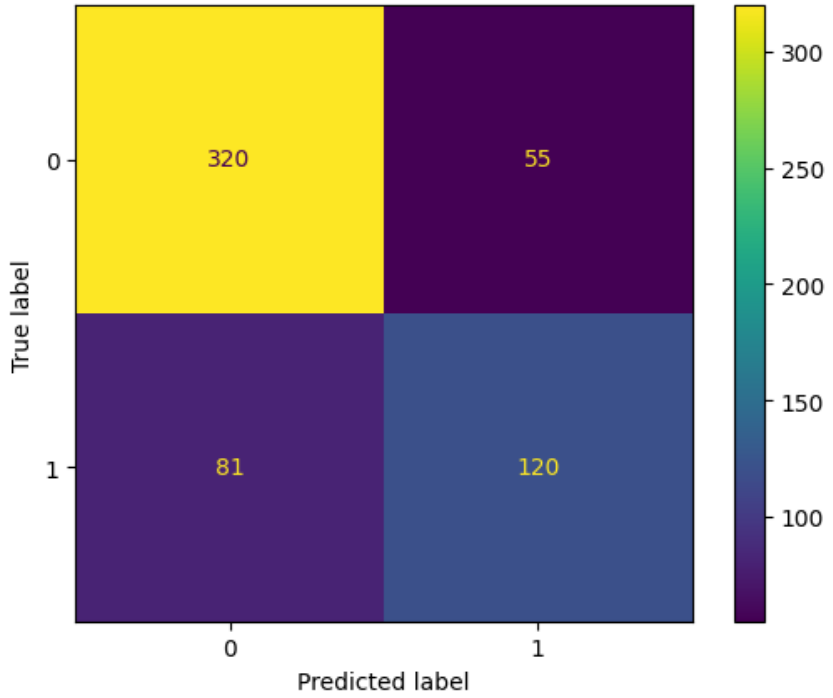


Logistic Regression



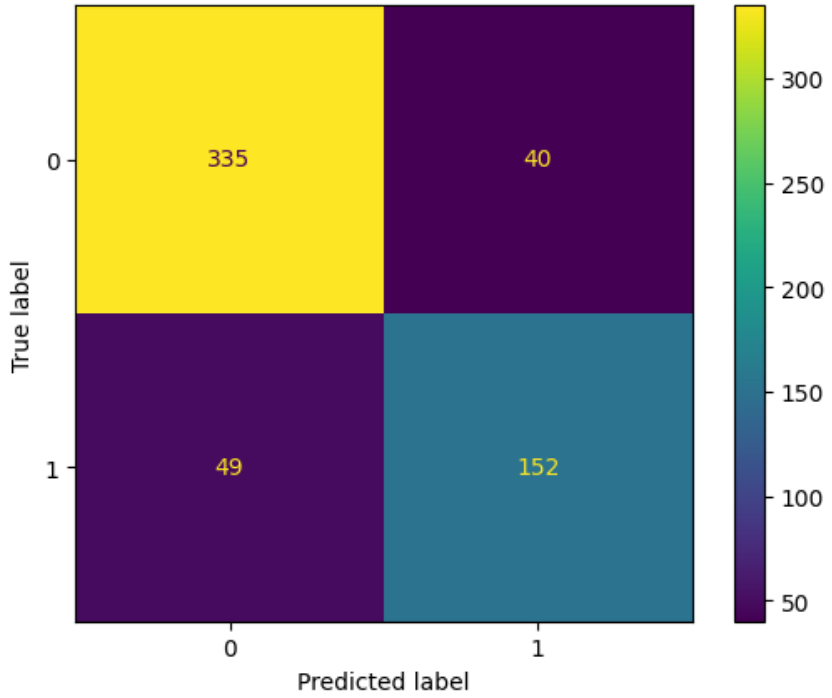
Acc = 76%

Decision tree - 3 levels

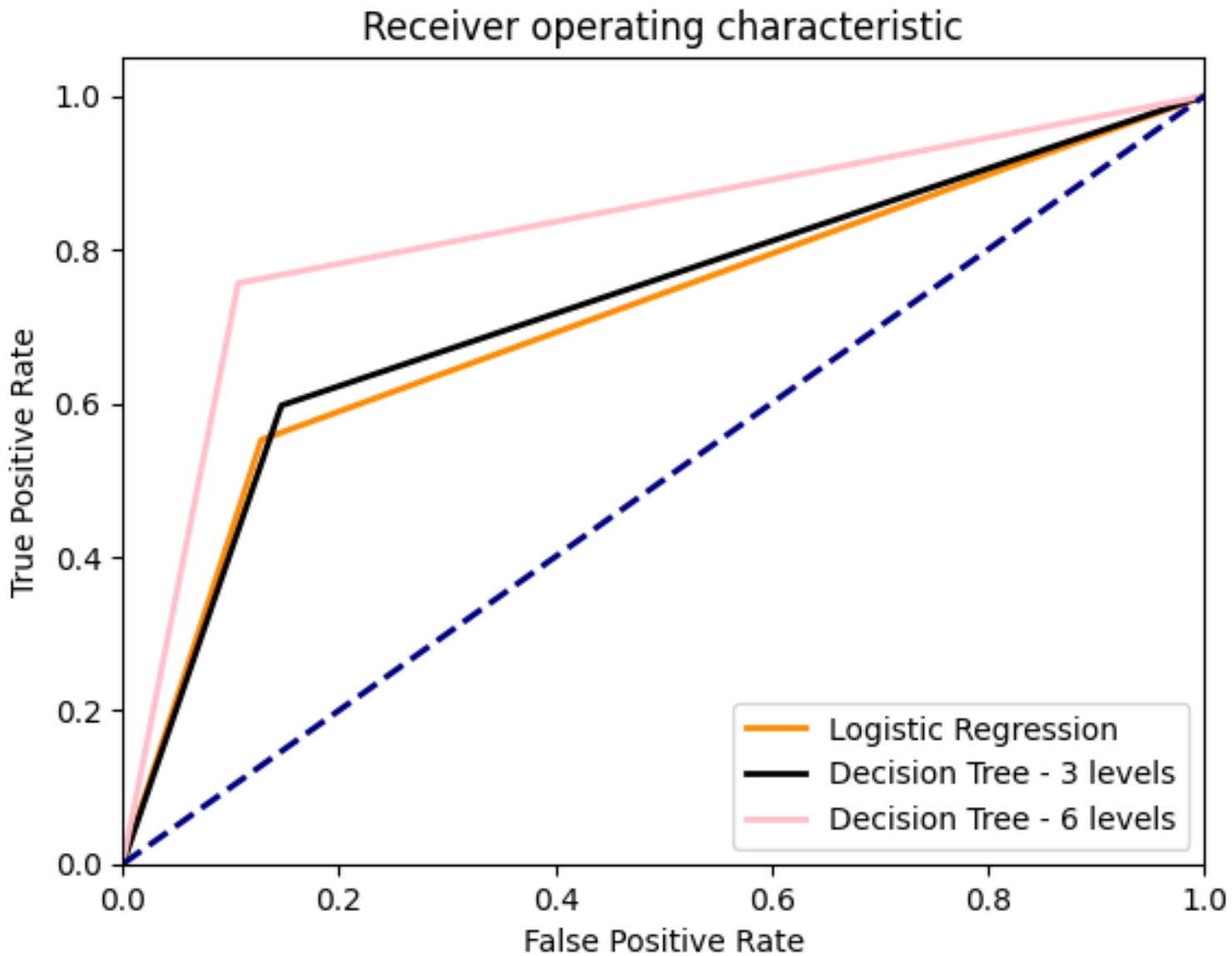


Acc = 76%

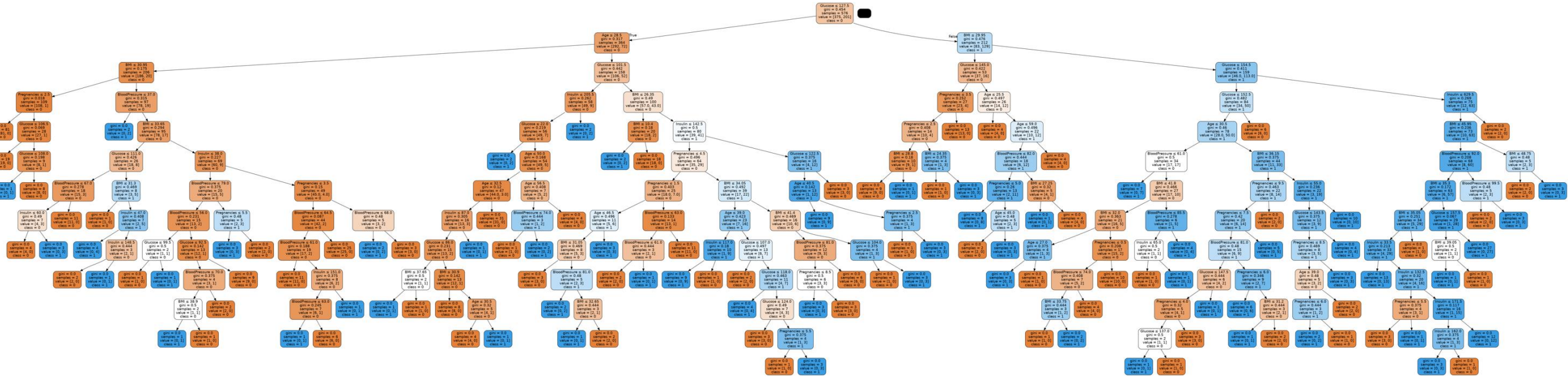
Decision tree - 6 levels

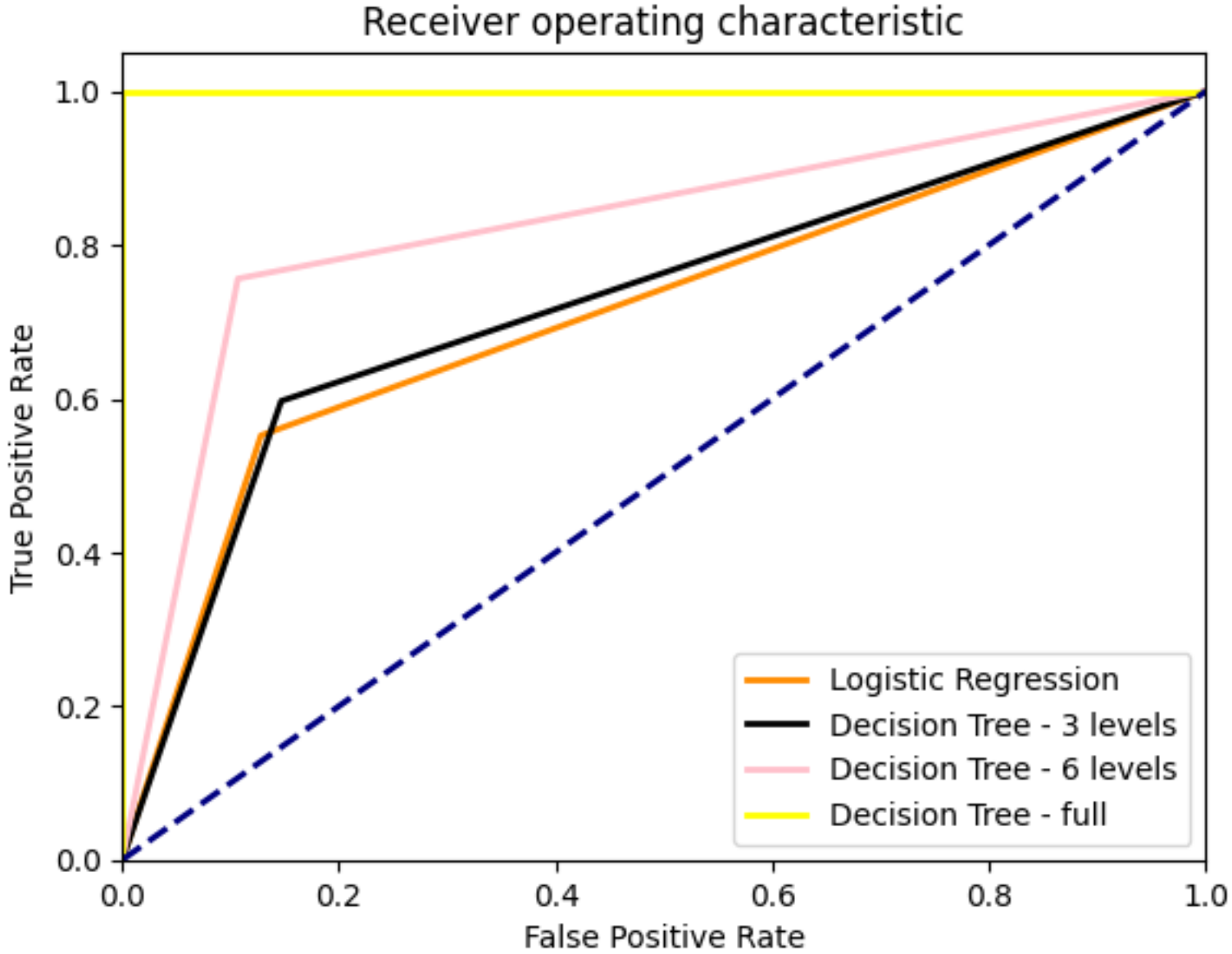


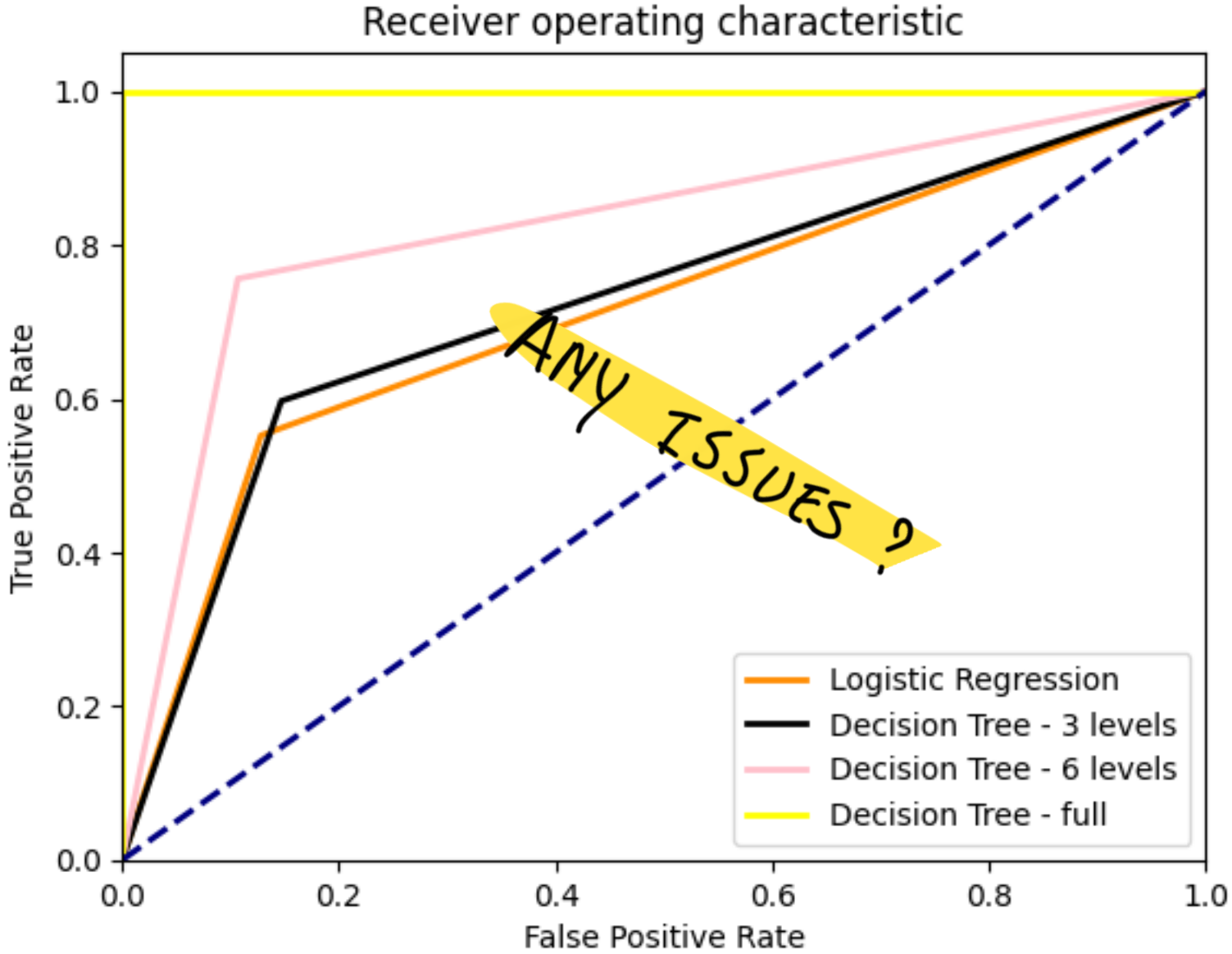
Acc = 85%

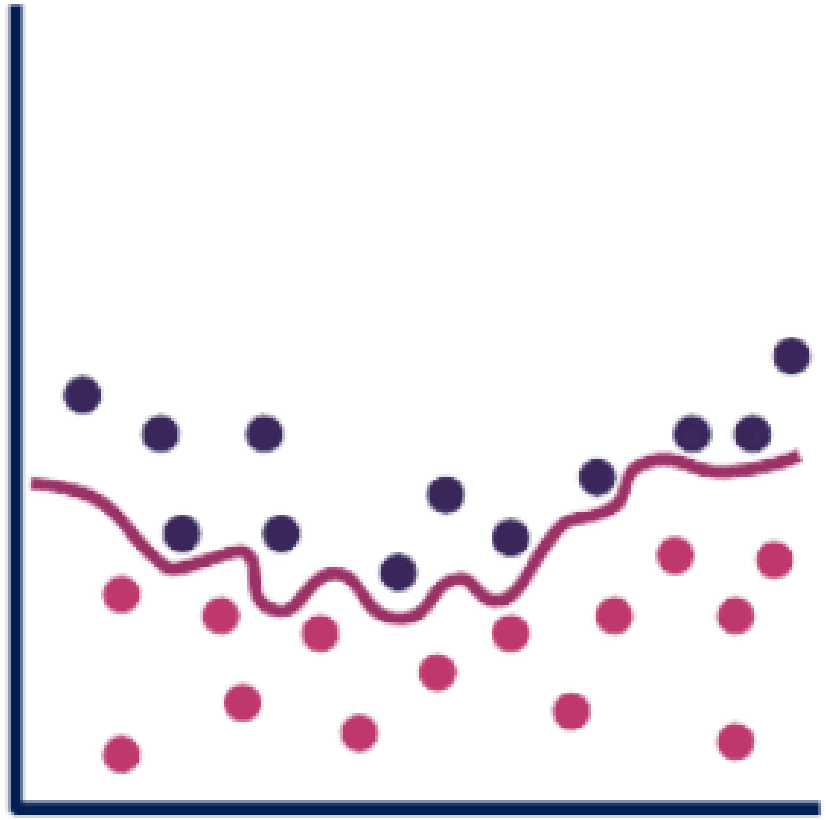

Receiver operating characteristic
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
Logistic Regression
Decision Tree - 3 levels
Decision Tree - 6 levels

We add one more model: **A “Full-Blown” Decision Tree**



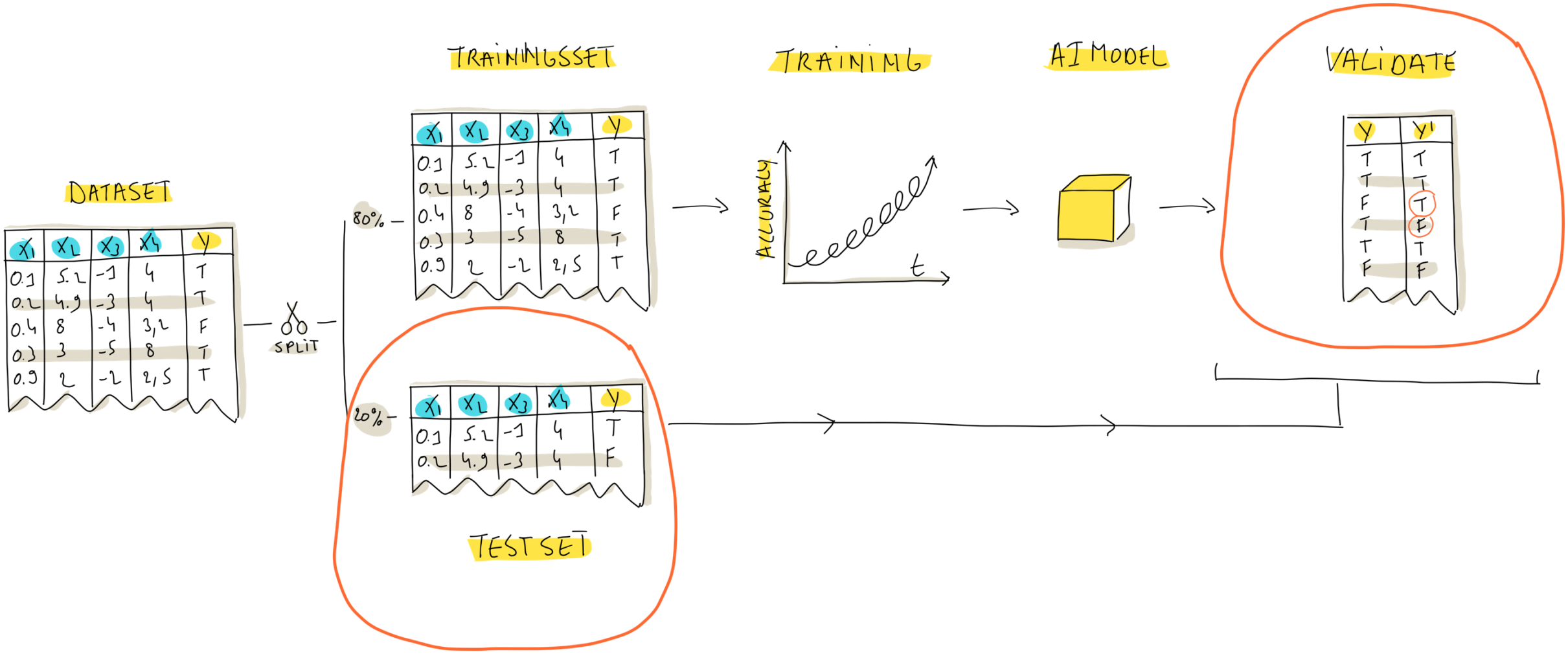

Receiver operating characteristic
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
Logistic Regression
Decision Tree - 3 levels
Decision Tree - 6 levels
Decision Tree - full

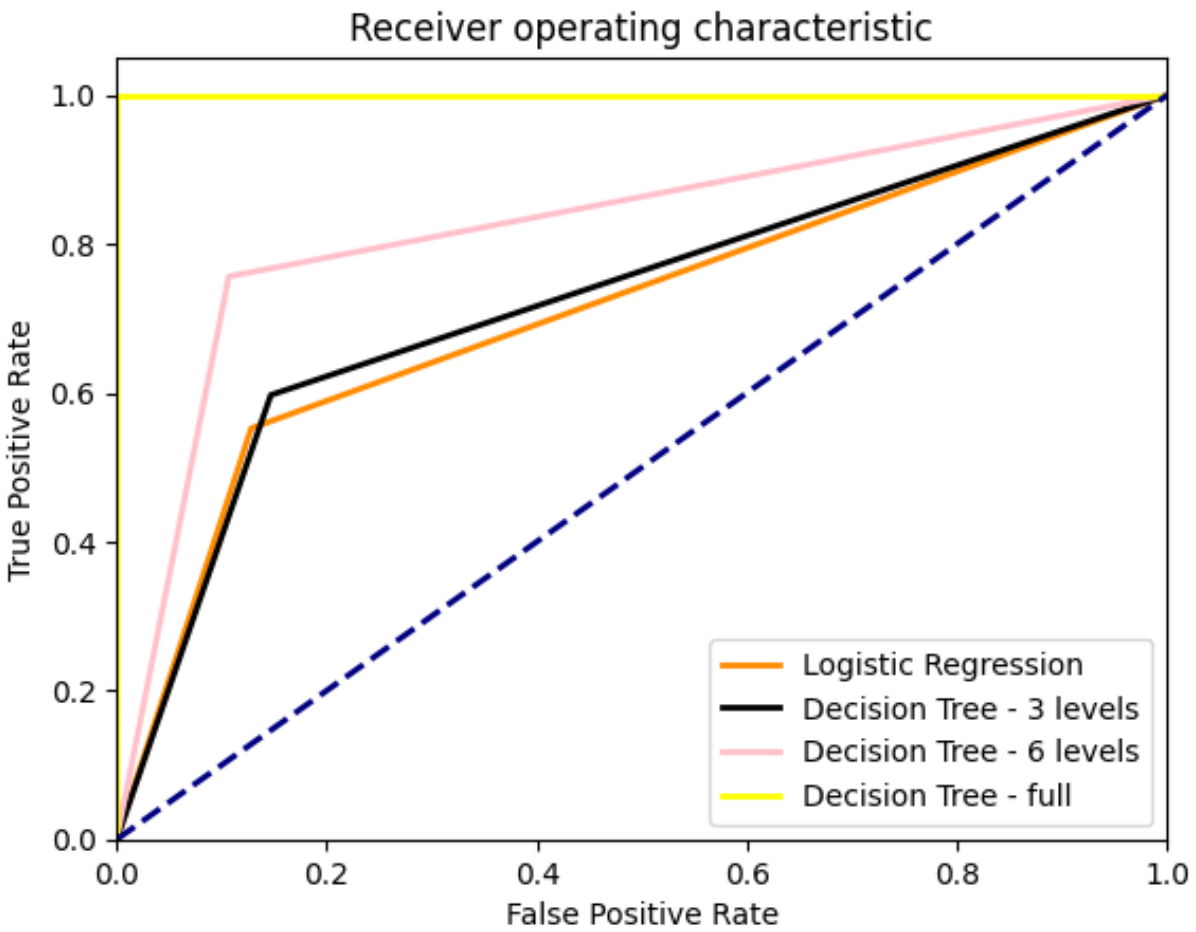

Receiver operating characteristic
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
ANY ISSUES ?
Logistic Regression
Decision Tree - 3 levels
Decision Tree - 6 levels
Decision Tree - full



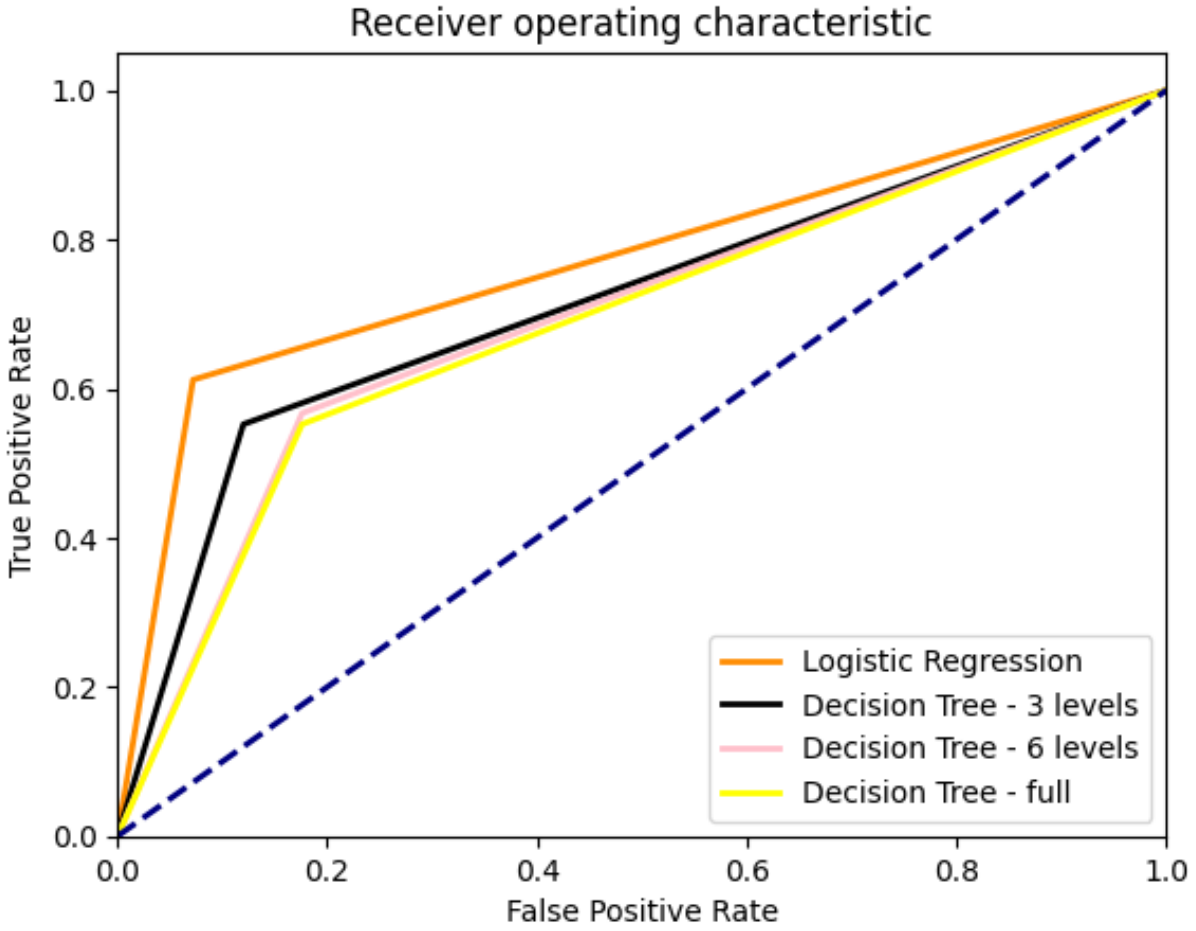
Overfitting

Validation?


DATASET
SPLIT
80%
20%
TRAININGSSET
TEST SET
TRAINIMG
ACCURACY
t
AI MODEL
VALIDATE



Trainingsdata



Testdata

Table of Contents