Till now

→ Statistics
→ Data toolkit
→ EDA.

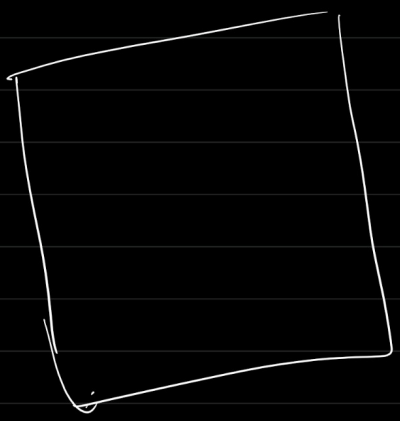
ML-1

Maths behind ML.

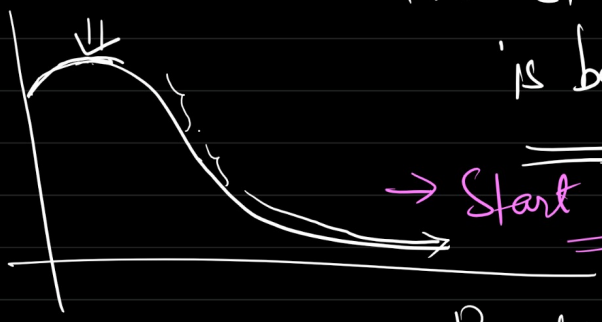
ML
→ intuition
→ implementation
→ Maths.

→ No fear from mathematics.

→ Be punctual.



ML



No student is bad

→ Start thinking

- Practice 50%

Statistics

* Todays

→ Intro to ML
→ Termino logies
→ Pre-process | feature engineering.
→ Matrices.

two types

Not in methodological (its not)

methodological → Successful

Motivation of ML

→ Cadbury has used customised using ML.
→ Predictive typing
→ Alexa, Siri or Voice Assistant
→ Autonomous vechicle
→ Credit fraud detection.
→ GPT, Bard

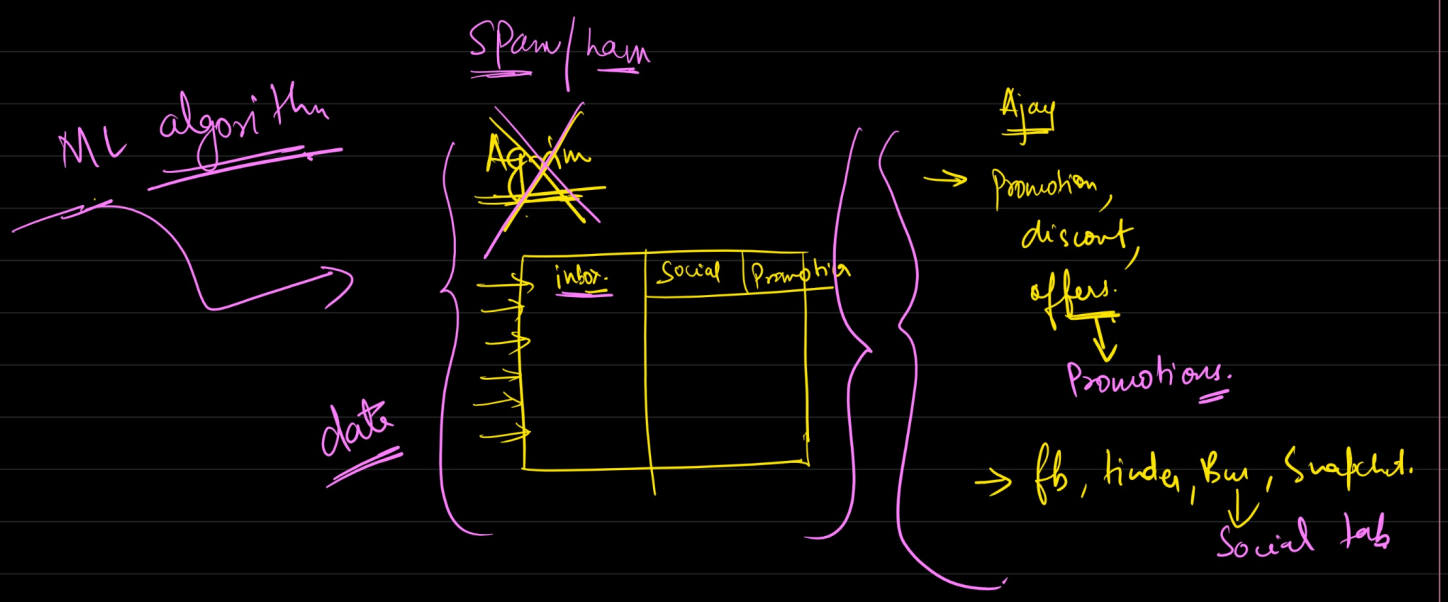
Stats > ML > DL > NLP > LLM's

Scen-1

'inbox'
=
promotions – 500
socials – 300
Spam

Scen-2

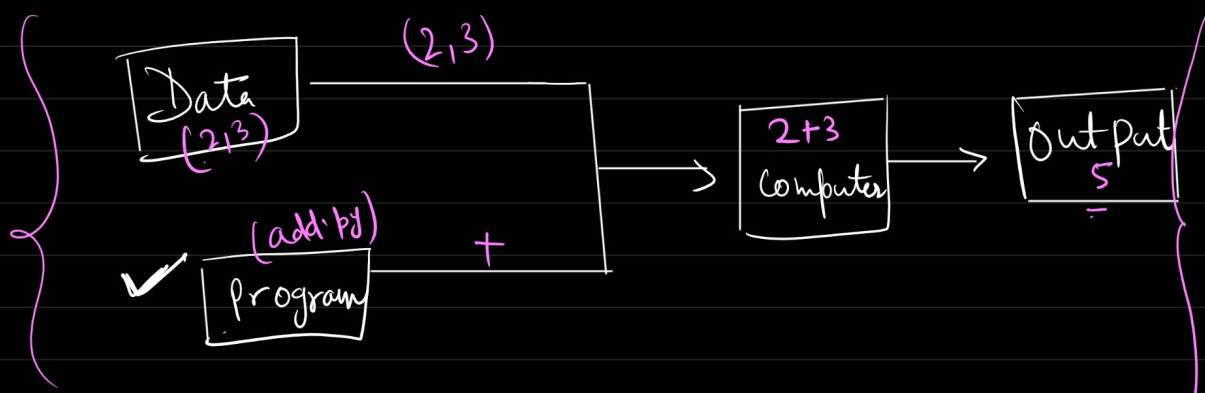
'inbox'
→
→



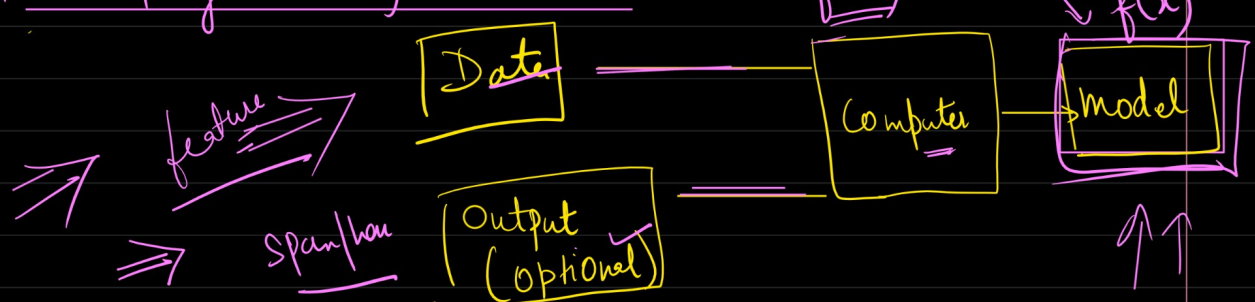
* What is ML ??
→ Machine learns patterns from the data and tries to replicate the same in future.

* Arthur Samuel (1959) :→ The subfield of Computer Science that gives computers the ability to learn without explicitly Programmed.

* Tradition Programming Paradigm



* ML programming Paradigm



Spam/ham

	feature	output
data (Agrim)	Email-1	Spam
	email-2	ham
	email-3	Spam
	email 4	ham
	email-5	Spam

Tom Mitchel (1997)

spam/ham watching Agrim's mail label.

A computer program is said to learn from experience E with respect to some class of Task T and Performance measure P if its performance measure P at task T improves with Experience E.

classify email spam/ham

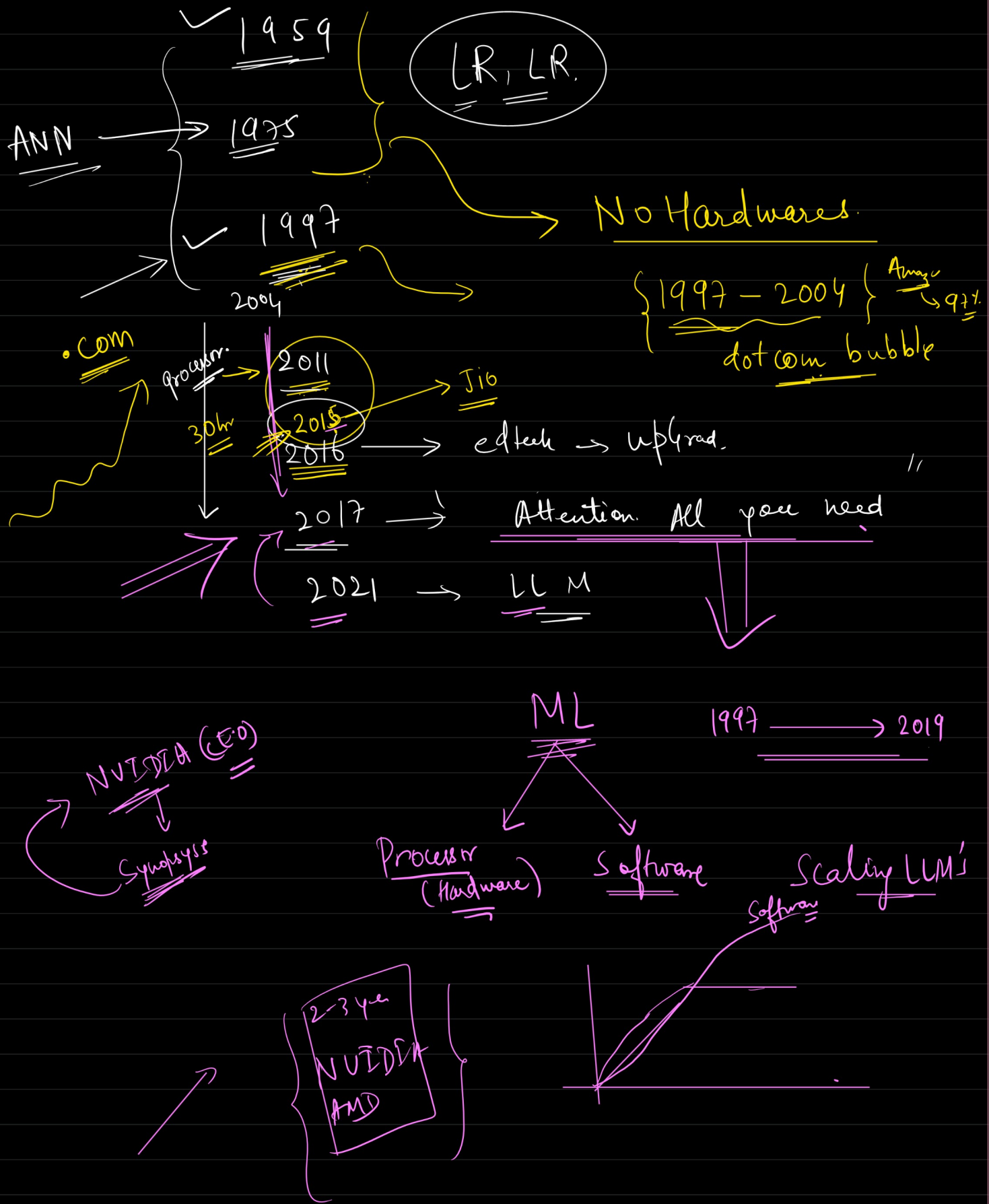
Performance measure P → Accuracy (success rate) → The no. of emails correctly classified

Ajay Defn

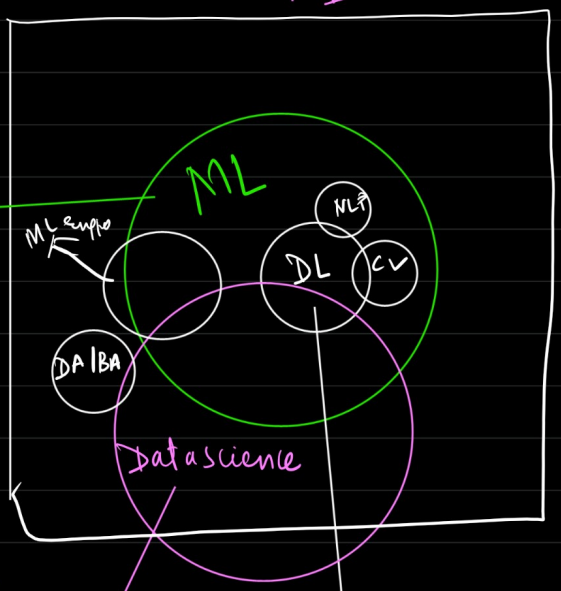
ML is a field of learning algorithms that:

→ improves their performance P
→ at executing a task T
→ over Experience E

Spam/ham


1959
ANN
1975
1997
2004
LR, LR.
No Hardwares.
{1997 – 2004} Amazon 97%
dot com bubble
.com
processor
30hr
2011
2015
2016
Jio
edtech → upGrad.
2017 → Attention All your need
2021 → LLM
NVIDIA (CEO)
Synopsys
ML
1997 → 2019
Processor (Hardware)
Software
Scaling LLM's
Software
2-3 yrs
NVIDIA
AMD

AI Vs ML Vs DL Vs DS



describes how computers and other tech mimics human intelligen

→ Smart Applications that can perform own task without human intervention

eg: Robots
eg Self driving car
eg Devin AI.

that focuses on creating algos & statistical models to let computers learn and make prediction without explicitly programmed from data.

eg Sloan/loan
eg Diabetic or not
eg. price of house

Specialise ML Algos that mimics the human brain

eg: Chatgpt
eg Devin AI
- eg object detection
- eg Image Recog
eg Recommendation

'interdisciplinary field comprised of CS, math, Stats, and domain knowledge

* Types of ML

1. Supervised ML
2. Unsupervised ML
3. Semi-supervised ML
4. Reinforcement learning.

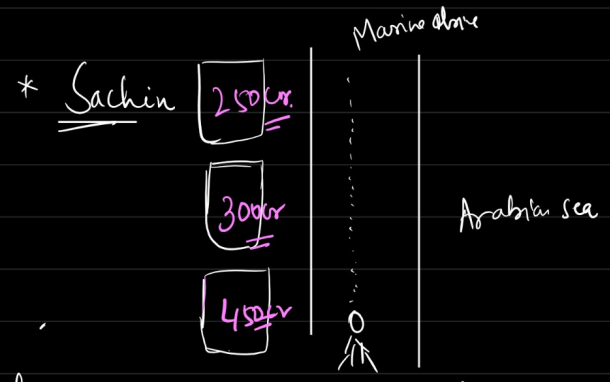
① Supervised ML

Supervisor

The ML style where dependent variable is present

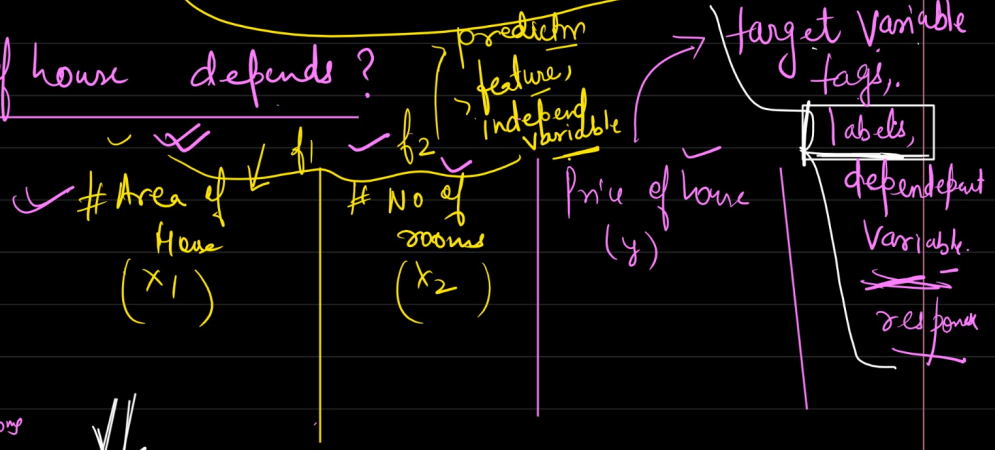
→ y is present (Historical data)

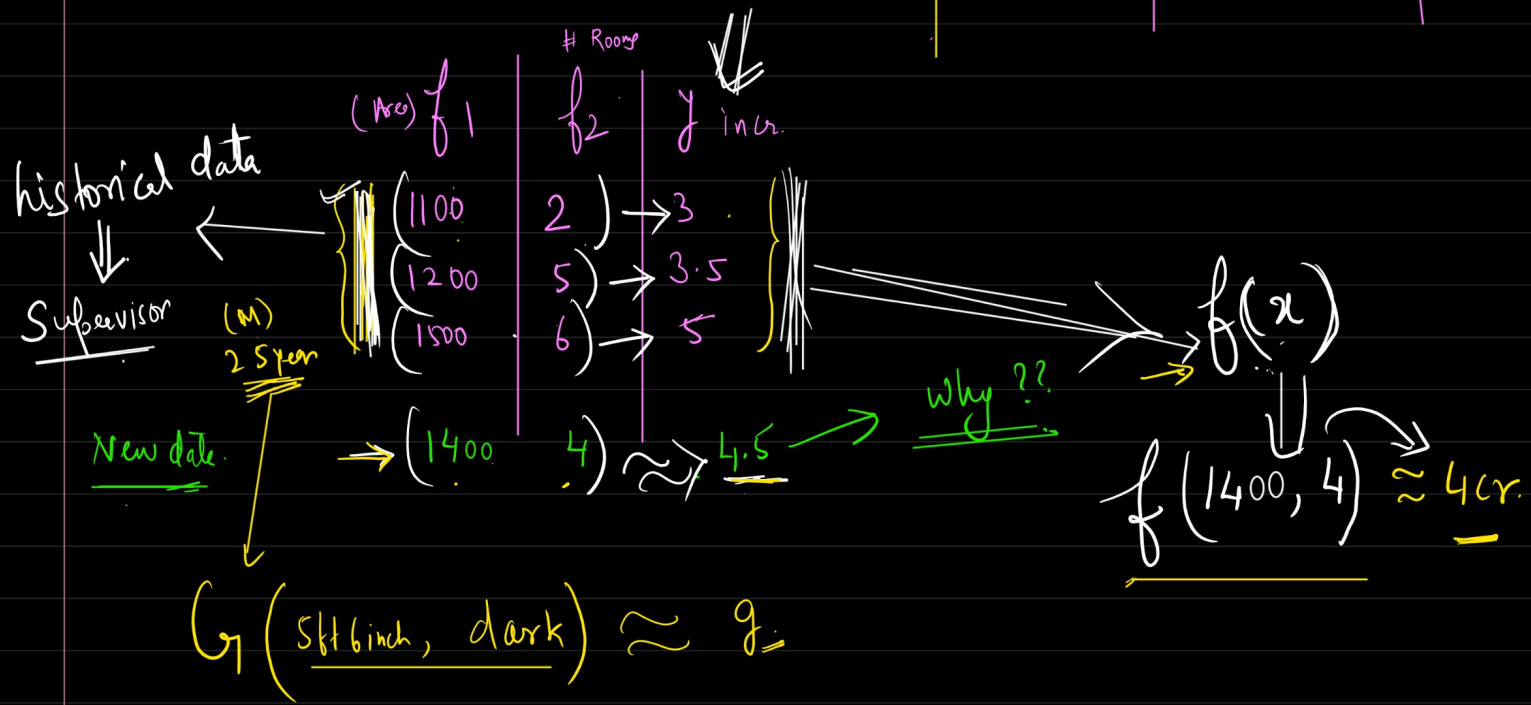
Q Predict Price of House



Price of house ∝ (# of rooms, Area, Distance from Sea beach, bathrooms.)

On what factors price of house depends?





G (5ft 6inch, dark) ≈ ?

✓ Target Variable ≈ SL

SL

Regression

→ y is continuous

Eg: Price of house

Price — 11 cr
11 cr 9 Paise
11 cr 2 Rs 3 paise

Eg: Score of student

Classification

→ y is discrete (finite)

Eg: Spam / ham.
Pass / fail
diabetic / No di
Cancer / N cant
Defa

SL

Reg

→ SLR
→ MLR
→ Polynomial.
→ Lasso
→ Ridge.

Classification

→ Logistic Regression (1,0)
→ Multi-class Classification
→ Naive Bayes Classifier

#hours Stud	Mark obta	Pass / Fa
8	80%	1
3	70%	1
1	20%	0

KNN
DT
RF ← Baggily vs Boss
XGR
SVMR

KNN
DTC
RF
XGB
SVMR

PCA

② Unsupervised learning (y is not Present)

Un - Supervised

No supervisor.

No Historical data.

Can you find Similar groups in that data ??

→ ajay is blind
→ on Earth

ajay (clothing)

Ruchika

Shirts → 10
Jackets → 15
Jean → 10

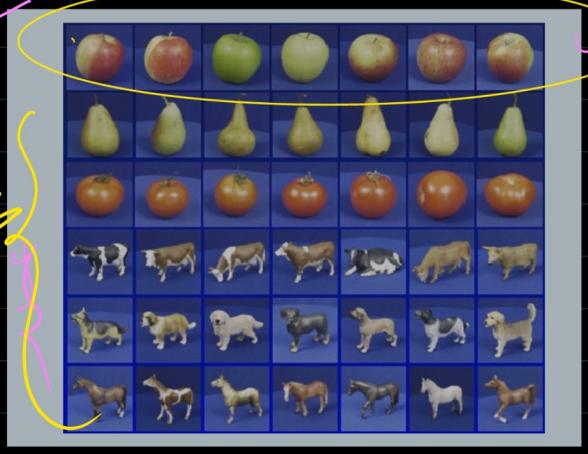
$J_1, J_{11}, S_1, J_1, J_2$

With similar features, shape, → Segregating

rack
→ Shirts
→ Jackets
→ Jeans

Rajiv -
→ fruits
→ Animals

data



Aman
Apple
Pear
Tomat
Hor

Lashita - different breeds

Ruchika → Brown Cow

Narration → Apple Pin - APPL

Sai → I am seen in different colour

data ⇒ Segregate in different groups.

2 S.

Lara

USL - No Wish

Expenditure	Salary (in Lakhs)
5000	20
6000	10
8000	35
—	—
—	—
—	—
—	—

Expenditure ⇒

LS-HE

LS-LE

HS-HE → loyalty → retaining members

HS-LE

→ {Parents, family, brothers,}

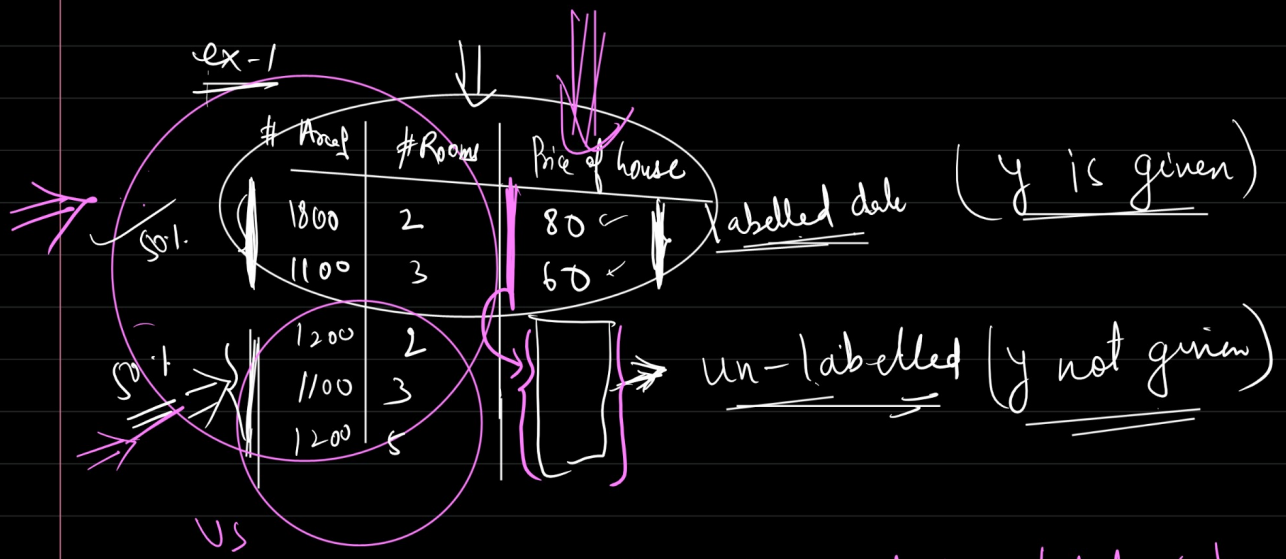
→ discount

Salary

③ Semi-Supervised

↳ Combination of Supervised & USL
↳ labelled & labelled data.

ex-1



* Either you drop label column and consider complete dat as USL problem Statement.
* drop 50% of row where target/label is not given & consider rest 50% as SL.

ex-2

Netflix recommendation.

Amazon Prime → Romatic, thriller, social cause, Action thriller → (USL) → different groups.

You went inside comedy

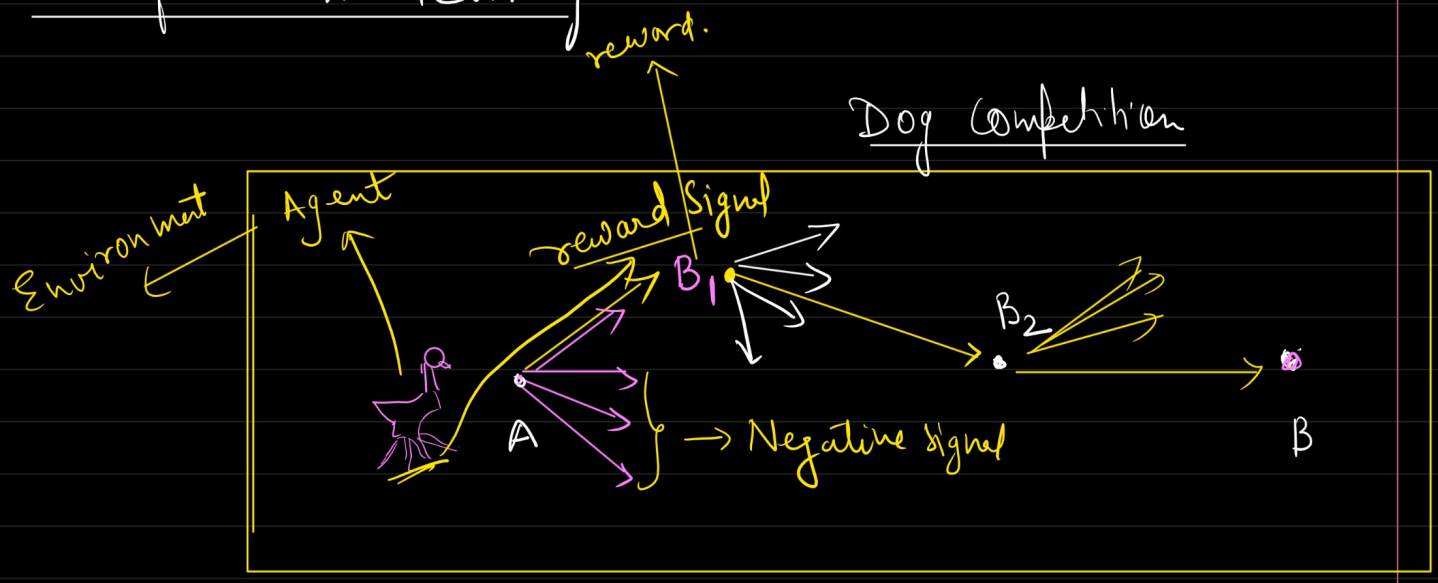
SL

Based on your watch

Since you have watched Golmaal, you can also watch

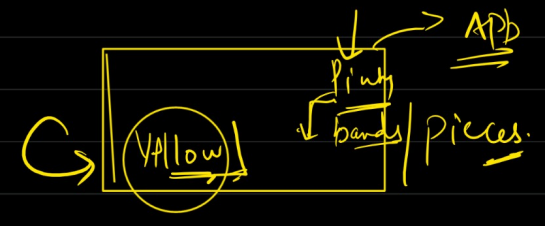
→ Hera pheri, Garam masala., Welcome

④ Reinforcement learning



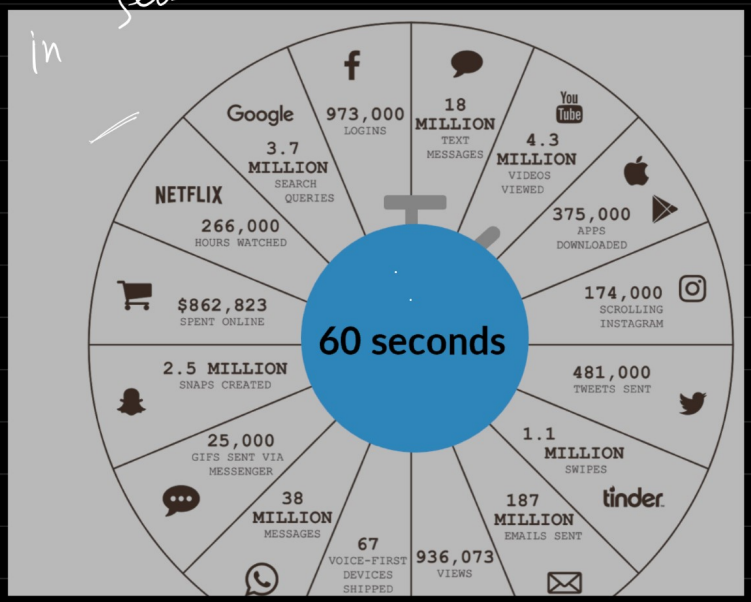
* An agent interacts with an environment and watches the result of interaction
* Environment gives feedback via positive or negative reward signal

* Ludo
* Chess



Chess

* Self driving car.



→ Use this data
↓
informed business decision