Artificial Intelligence (AI) - 1st June 2024

- ANN/DNN/MLP - Tabular
- CNN/ConvNets - Image
- RNN/LSTM/GRU - Text/Translation/Audio to Text/Text to Audio
- NLP
- Transformer, BERT, GPT
- Computer vision/opencv - Face recog, Self driv
- Re-Inforcement learning -
- GANs & Autoencoder - unsup

* Gen AI - LLM

Basic to Adv

June, July, Aug, Sep - 1st Week

Industrial Revolution

(1) Steam-based Machine - 18th century.

(2) 19th - 20th century — 2nd revolution
→ Electrical energy-based mass Prodn

(3) late 20th century — computer & Internet based knowledge

(4) 4th Revolution - 21st century — 2012

AI & IT

$2 + 2 \div 2 = 3$

Perceive the world → Analyse & understand → Make informed Decision → AI

Act → Learn → Sense

(1) Rule Based AI

→ if, else condition

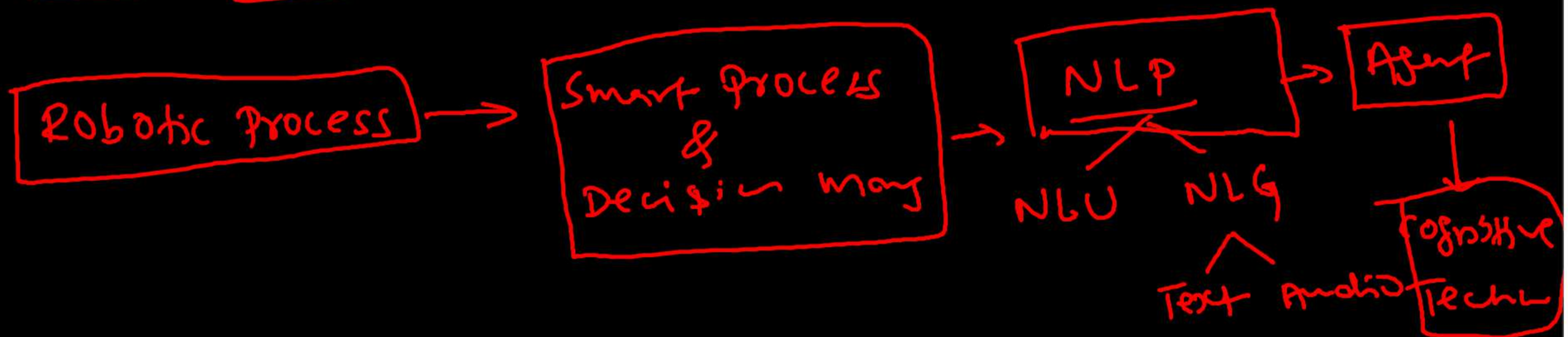
x_1	x_2	x_3	y
1	1	2	0
3	4	5	1
0	1	2	0

$x_1+x_2+x_3 > 5, 1$
otherwise zero

(2) Machine / Deep learning / Statical Based AI

→ Rules are very hard or we don't know rules - Text

AI Solution

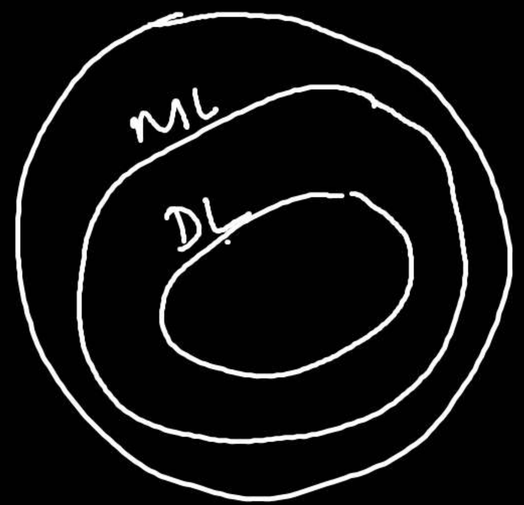


Game Theory

Markov Decision Process –

ML

Deep Learning



DL VS ML

1. Hardware dependency
2. Software dependency
3. Data dependency

4. Training & Testing Time
5. feature Selection
6. Interpretability
7. Algorithms

ML ↑
DL ↓

→ Hardware

2012 - Deep Learning by Gaming company

Software

ML

Numpy, pandas, SKlearn, plotly etc —

DL

Tensorflow, Keras, Pytorch, Theano, NLTK, Spacy, Textblob, Gensim etc.

Data Dependency

ML - less data / less variable

↓

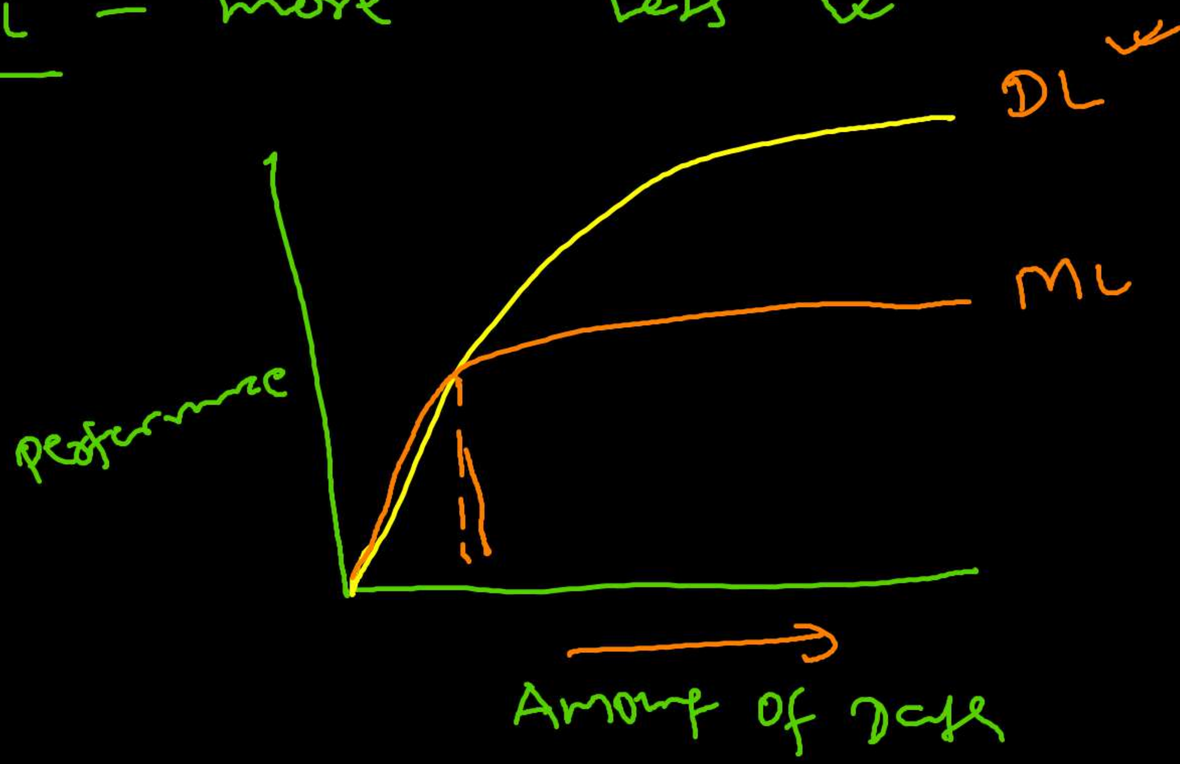
DL X

- DL
  - more / huge dataset - 50K / 5Million / bil~
  - Data hungry

Training & Test Time

ML — less more

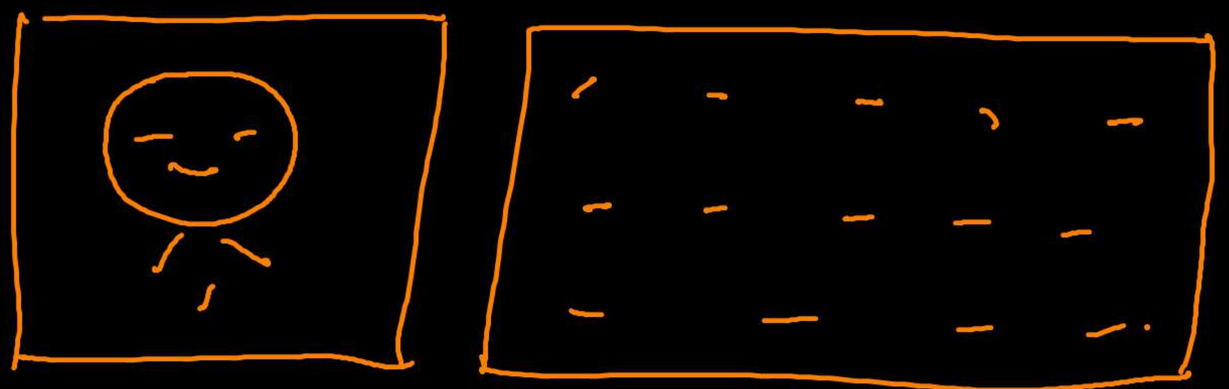
DL — more less

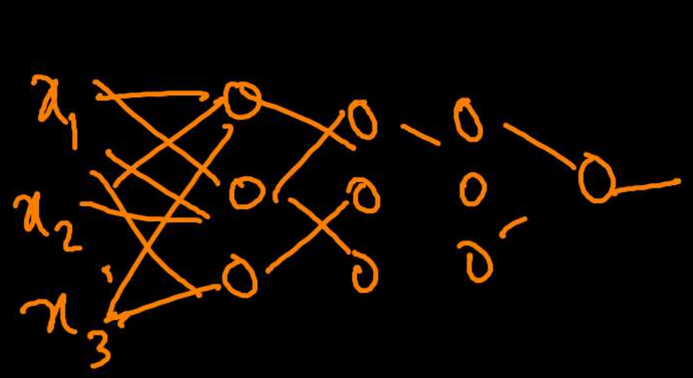


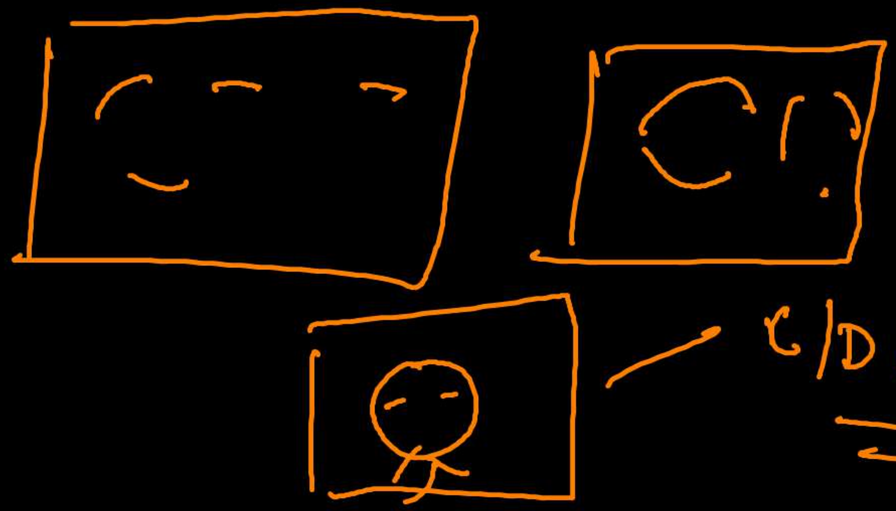
Feature selection

ML - Manual work

DL - Automation approach







C/D/h/E

Algorithms

Time series

ML - LR, LOGIT, DT, RT, Ada, XGB, GD, SV... KNN, SVM, NBT etc

DL - MLP, CNN, RNN LSTM

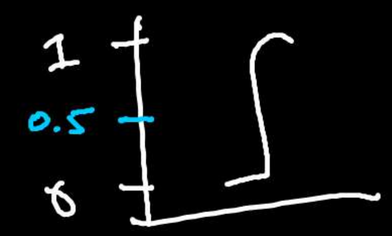
History Of Neural Network & Deep Learning

⇒ NN are an important ML Modelling framework

→ first simplest NN model built was a "perceptron".
  ↳ Built by Rosenblatt - 1957

⇒ With small changes a Perceptron can become a
  ✓ Logistic Regression - Sigmoid curve

MLE



x_0	x_1	x_2	y
5	10	20	0
50	100	zero	1
50	60	20	1
10	20	30	0

$-\infty$ to $+\infty$

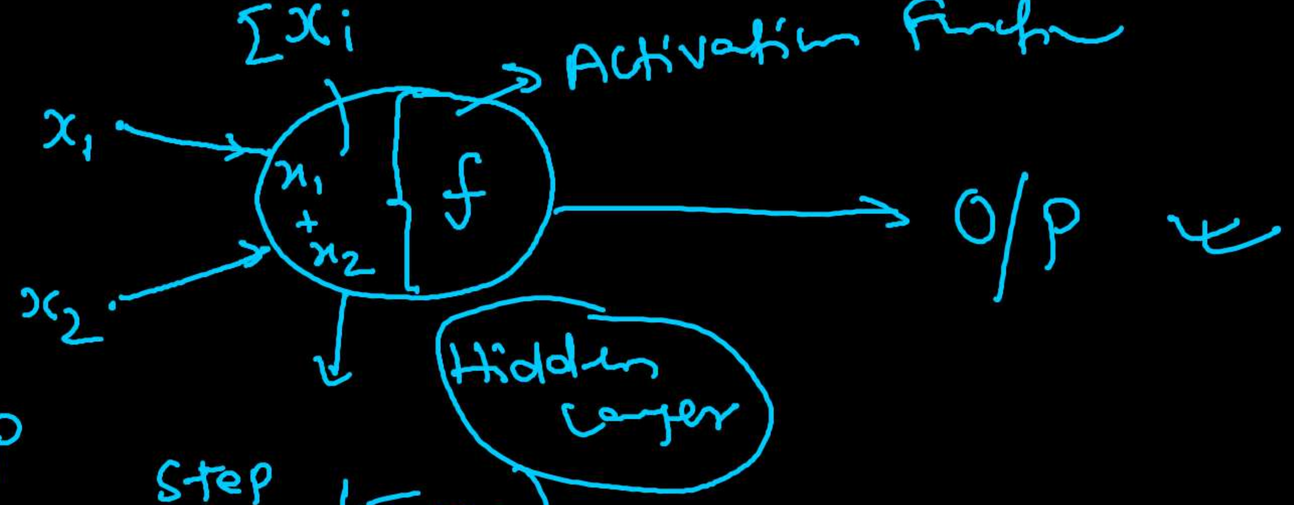
$x_0 + x_1 + x_2 > 100$, 1
otherwise, 0

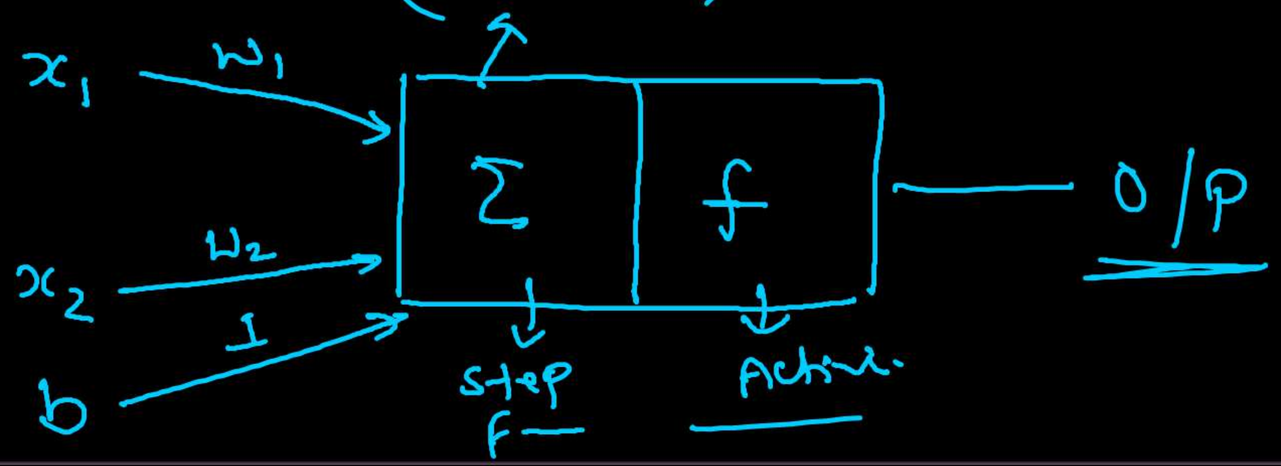
And Gate

I/P data

x_1	x_2	x_1+x_2	And O/P
1	0	1	0
0	1	1	0
1	1	2	1
0	0	0	0

if $x_1+x_2 < 2$, 0
otherwise 1





$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots$

$Y = \theta + w_1 x_1 + w_2 x_2 +$

Alan Turing (father of Modern Computing)

⇒ curiosity raise these questions

→ What is intelligence?

→ How should we build it artificially?

brilliant creatures

⇒ There was a biologically inspired way

"loosely" inspired from biology

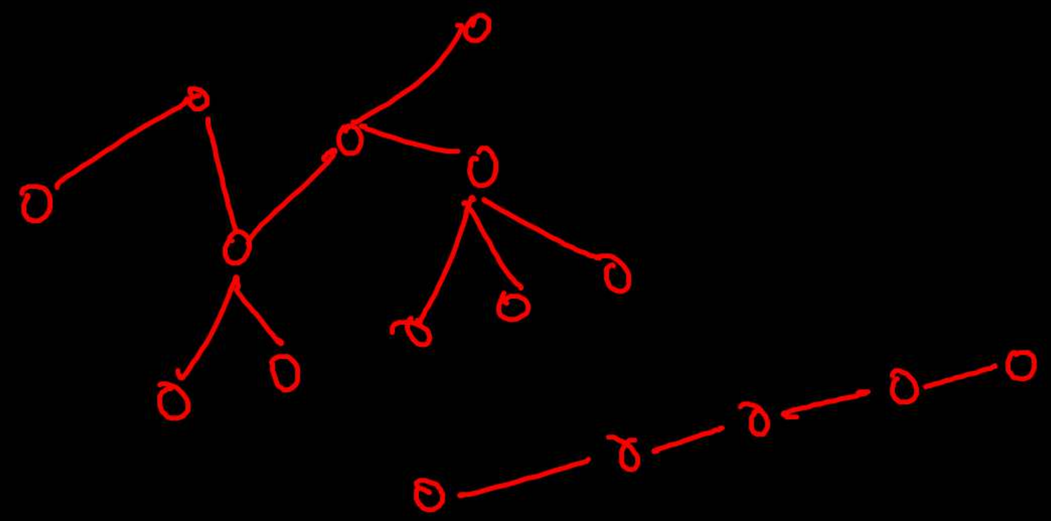
human, Monkey, Elephant, ants

⇒ Due to work in Neuroscience, a vague understanding

Of the working of brain led to developed

Biological Neuron VS Artificial Neuron

Perceptron → Single Neuron 1st AI wi



1986 (Mathematicians — Hinton & others)

2006 ↓ 20 years

→ Backpropogation

father of modern Deep Learning

code - Perceptron model

2012

ImageNet

Gradient desc

SVM RF GD
1990 - 2005

AI winter year - 1990-

⇒ insufficient computation power
→ " data
→ " algorithm

Activation — Sigmoid & tanh

optimizer — GD

↳ Perceptron is a linear model

→ Decision boundary / Decision surface

→ OVR - one vs Rest

→ AND Gate / OR Gate - Classification.

$x_1 + x_2 > 2$, Green
$x_1 + x_2 = 2$, Red
$x_1 + x_2 < 2$, blue

→ XOR Gate - failed perceptron - Single Neuron model

x_1	x_2	$x_1 + x_2$	O/P
1	0	1	
0	1	1	
1	1	2	
0	0	0	

↓

Sum / RF / GD

Geoffrey Hinton — Multilayer perceptron

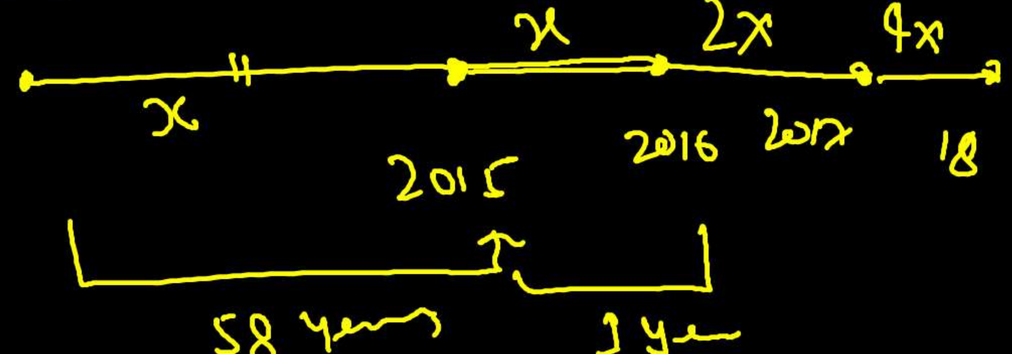
$x_1 + x_2 \geq 2$, 1 else 0
Red Green

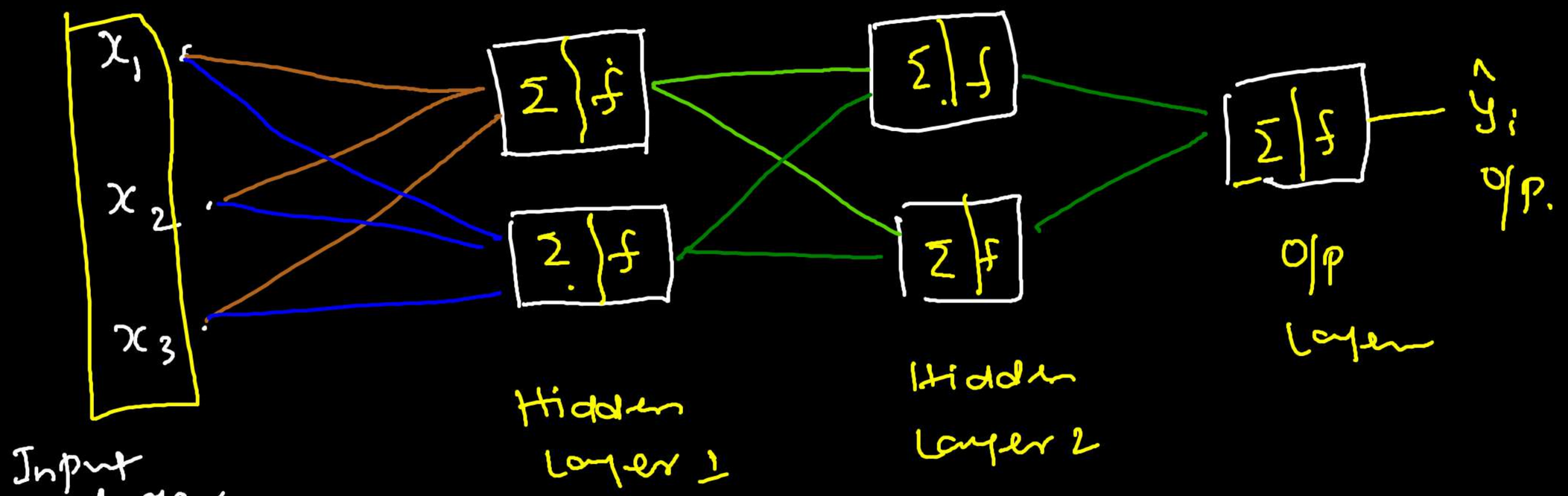
Why now ?

- → Hardware - GPU / TPU
- → Dataset
- → Architecture
- → Framework
- → Community

Perceptron
Rosenblatt

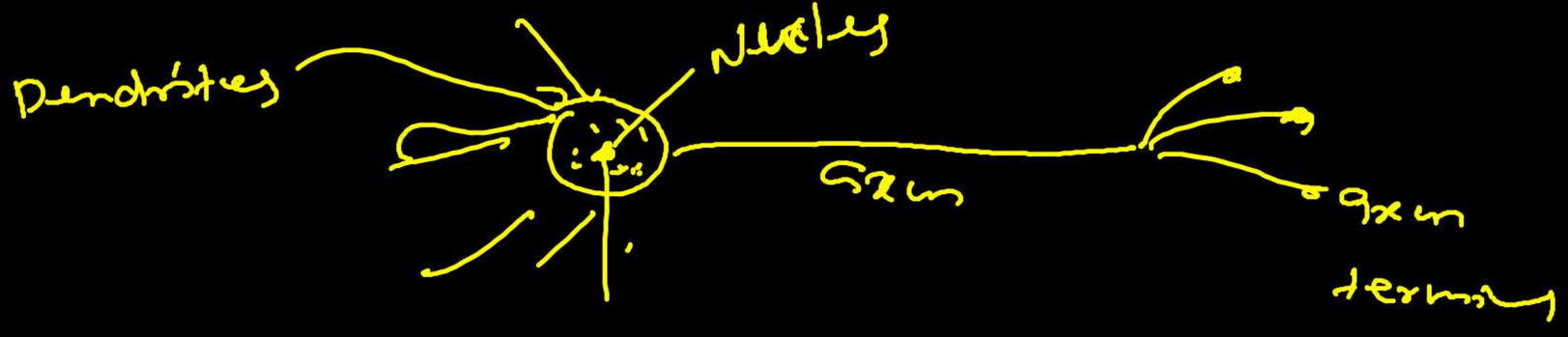
1957





How Biological Neurons Work

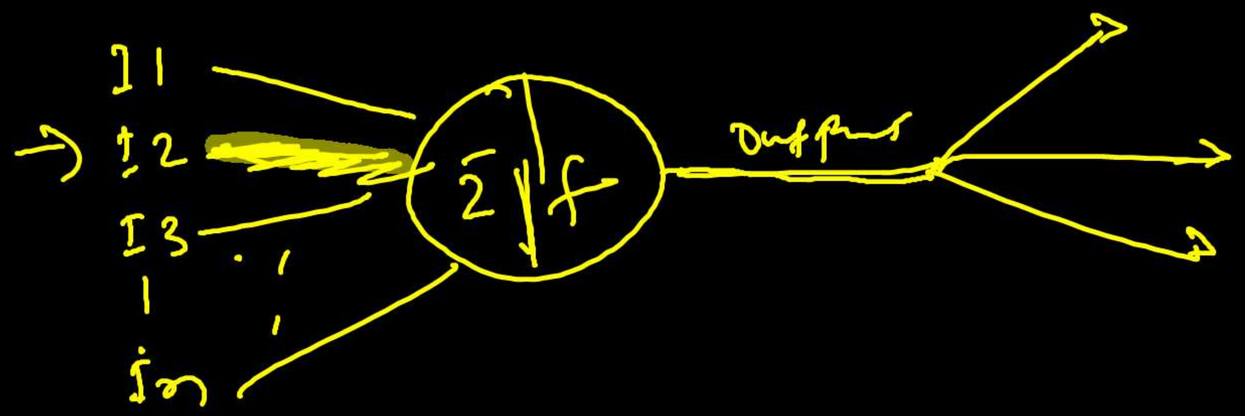
Biological Neuron



Simple

- Salary
- Cibil
- Family me

Simplified view of Neuron.



DNN / MLP

- → Perceptron
- → MLP
- → BackPropogation
- → Memoization
- → Chain Rule — calculus
- → Regularization
- → Dropout
- → Batch Normalization

- → Normalization
- → Weight Calculation
- → Activation Function
- → Optimization method
- → Early Stopping
- → Hyperparameter

-) each part - case Study
-) Real time..