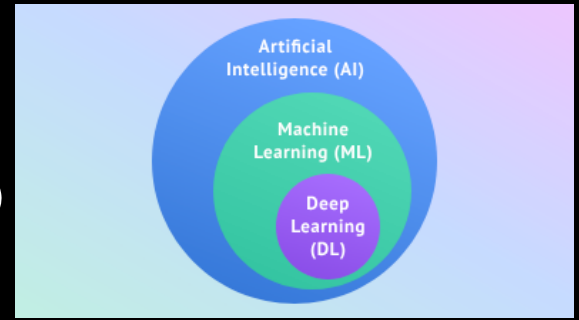


Deep Learning 101

2012 Subset of ML (Tabular Data)



Structured Unstructured
Images, Speech, Videos

SVM

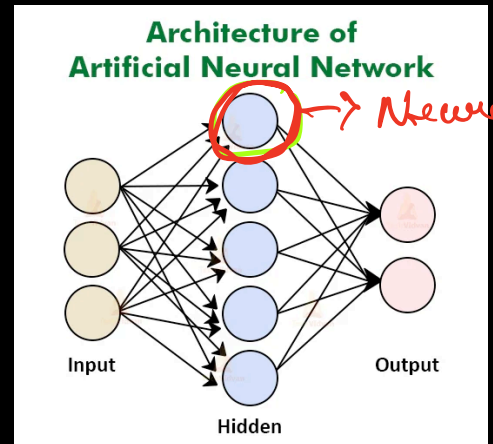
Representation Learning

Feature Engineering

Ex: - Dog Cat Classifier

Features / Attributes

x_1 x_2 x_3
10 12 SAT

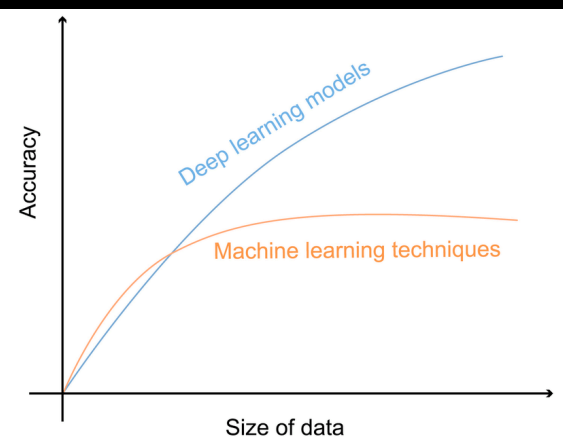


Input \rightarrow $\boxed{B.B}$ \rightarrow Output

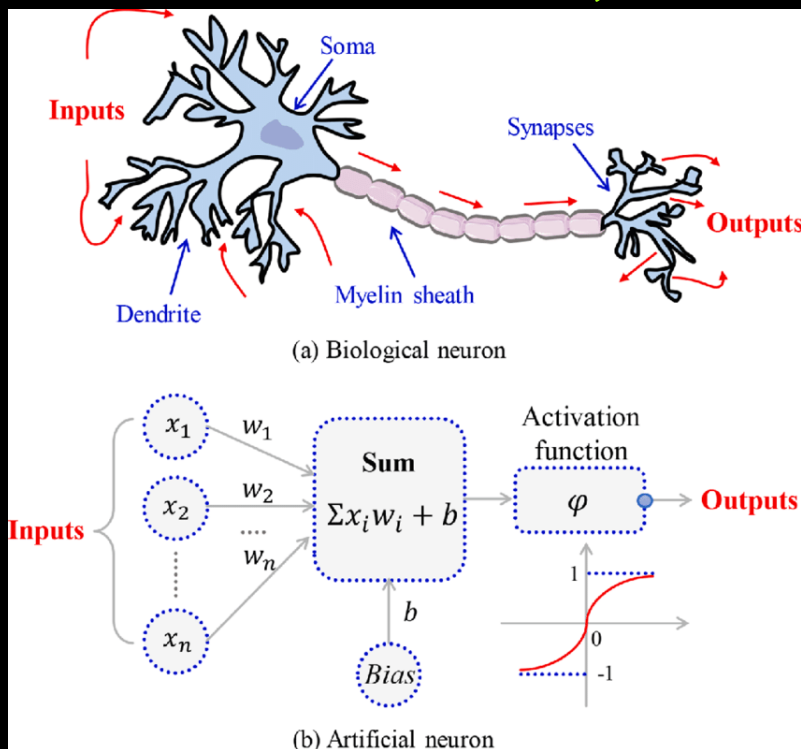
Neuron / Perception

Types of Neural Networks

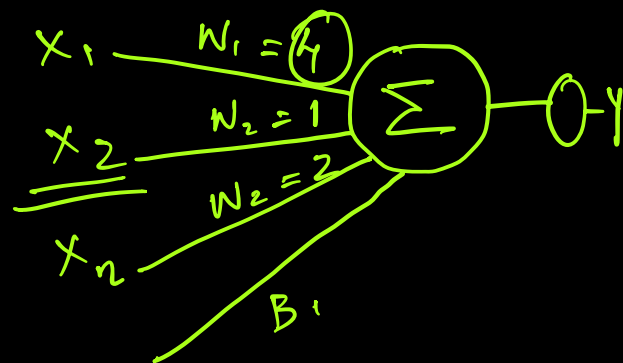
- 1) ANN (MLP) \rightarrow Structured / Unstructured
- 2) CNN \rightarrow Images / Videos
- 3) RNN \rightarrow Textual / Speech / Structured



4) GAN



Weights



Bias \rightarrow Constant
 $\textcircled{1}$

$$z = x_1 \cdot w_1 + x_2 \cdot w_2 + x_n \cdot w_n + b$$

$$z = \sum x_i w_i + b$$

$$g = f(z)$$

Sigmoid

Step function

Activation Functions

$$\left. \begin{array}{l} f = 8 \\ f = -3 \end{array} \right\}$$

$$\left. \begin{array}{l} f \geq 0 \quad , \quad 1 \\ f < 0 \quad , \quad 0 \end{array} \right\}$$

1) Data Dependency

DL \rightarrow Data Hungry

OOM

2) Hardware

\hookrightarrow GPU \rightarrow DL

\hookrightarrow CPU \rightarrow ML

218GB

3) Training Time

ML \rightarrow Minutes / Hours

DL \rightarrow Hours / Days / Weeks / Months

4) Feature Engineering

ML \rightarrow Manual

DL \rightarrow Automated

5) Interpretability

ML \rightarrow ✓

DL ✗

10^{th} 12^{th} SAT
= = =

Adoption of DL

1) Datasets

Lx Building Quality Datasets

Google, Microsoft, Facebook

Unlabelled
Labelled

img0001.jpg Dog

Annotation

Task

x ... 2015

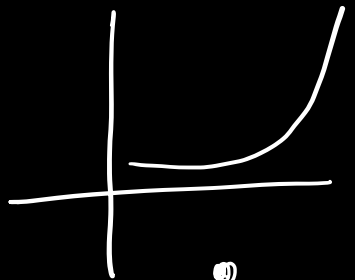
2x

16

4x

17

18



Hardware (GPUs)

Building Customized Hardware

GPU → Nvidia / AMD

CPU → Intel / AMD

Very Expensive

FP64A

Customized Sol

low, lower Xilinx

ASIC

Lx Bitcoin

Smart Watch

Gaming
Video Editing
DL

↳ Edge TPU

Crypto Mining

Macbooks

NPU → Neural Processing

↳ TPU - Tensor Processing

Simultaneous Instruction Multiple Data

SIMD Operation →

Frameworks / Libraries

Function

Top
Fastest

Keras

High level
wrapper

PyTorch ^{OOP}

↳ lua
↳ PyTorch
Meta

TensorFlow

Time
Money
Effort

Sonnet, Theano, Caffe, Chainer

Community

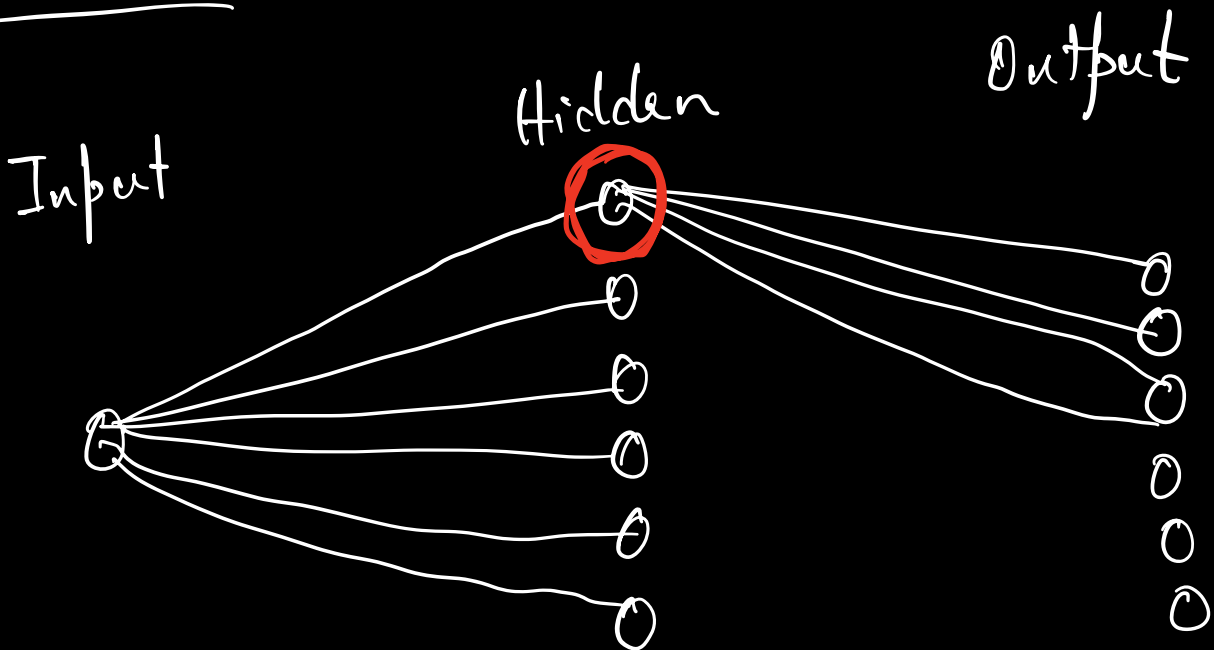
Adoption among researchers

Nobel Prize

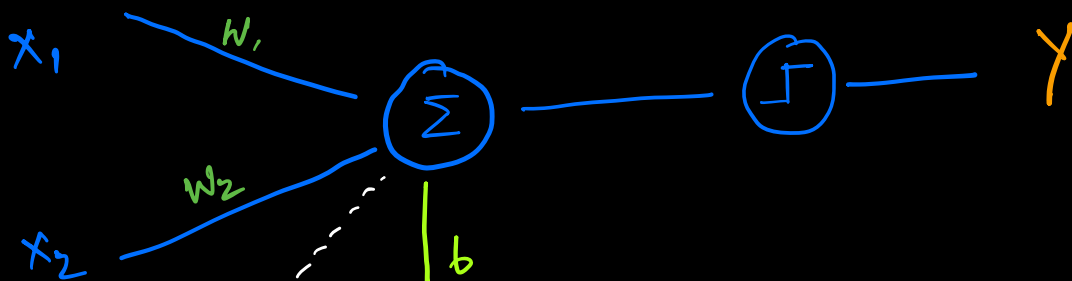
Physics

→ ~~James~~
Jeffrey Hinton

Neural Network



Neuron / Perceptron



X_n

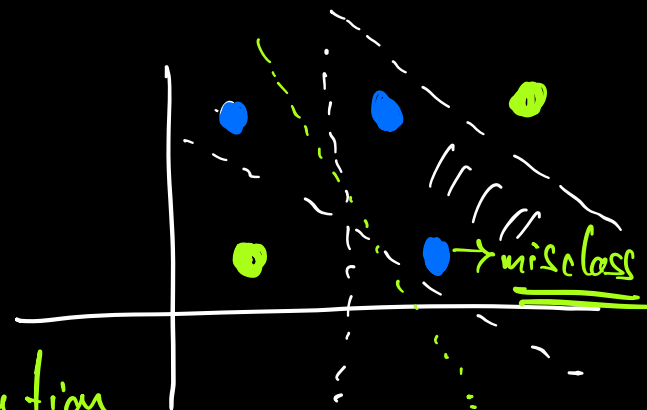
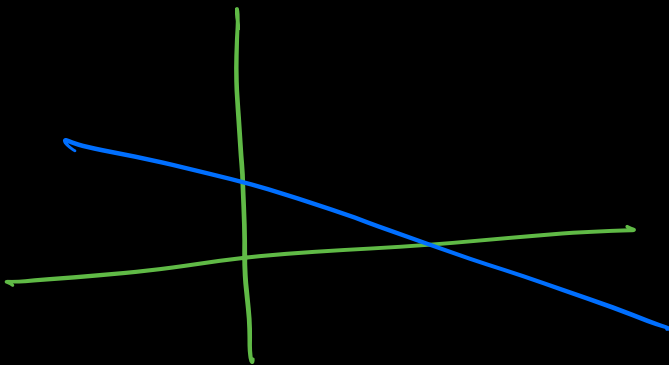
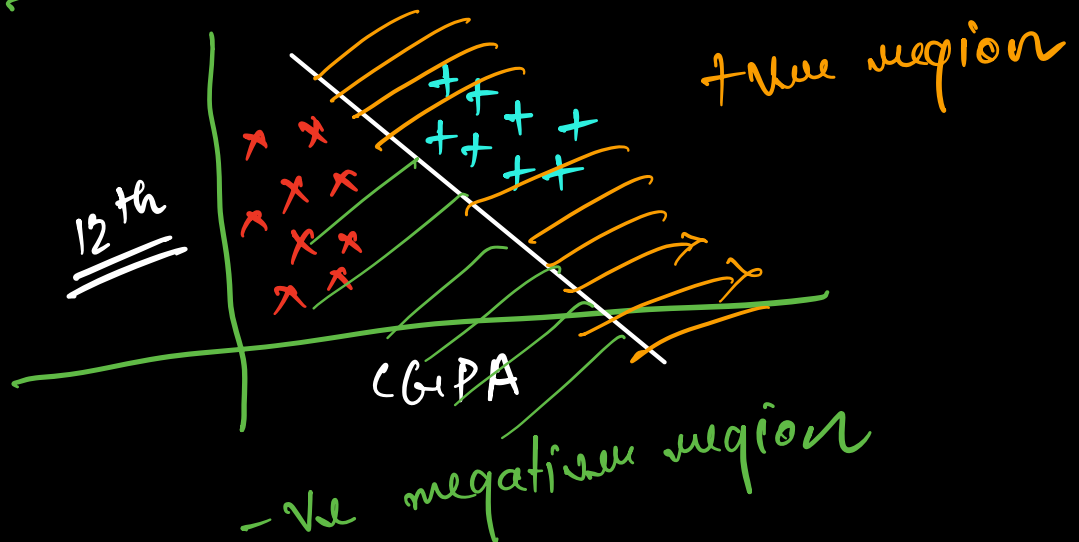
$$X_1 = A$$

$$X_2 = B, \text{ bias} = C$$

$$= X_1 W_1 + X_2 W_2 + b$$

$$= A W_1 + B W_2 + C$$

Line Equation



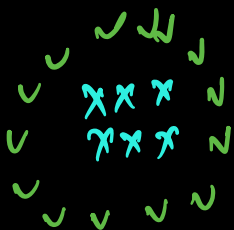
Problem of Misclassification

1) Manipulating the line

$$\textcircled{A}x + \textcircled{B}x + \textcircled{C}$$

I will try to change the values.

Linear Transformation



Non-linear
Data

fails with Perceptron

Perceptron works on linear Data

2D — line

3D — plane

more than D \rightarrow hyperplane

Misclassification

	10	12	SAT	γ
x_i	78	81	340	1
	64	61	240	0

if $x_i \in N$ and
 $\tilde{\gamma} \geq 0$ P

if $x_i \in P$ and
 $y \geq N$

y	\hat{y}	
1	0	
1	1	→ class
0	0	→ corr
0	1	→ misclassf.

Transformations

$$2x + 4y + 10 = 0$$

Very Big changes, $4x + 4y + 10 = 0$

Learning Rate (0.1 / 0.001 / 0.01)

Minor changes

$$w_{\text{new}} = w_{\text{old}} + \eta (y_i - \hat{y}_i) x_i$$

$$= w_{\text{old}} + 0.1 (y_i - \hat{y}_i) x_i$$

input → neuron

at times

Weight

What is Convergence?

Reaching Global Minimum
close to 0 \rightarrow 0 Loss
 \rightarrow 0 Errors
 \rightarrow 0 Misclassified
 \rightarrow 100% correct results

loop = 100

for i in range(100)

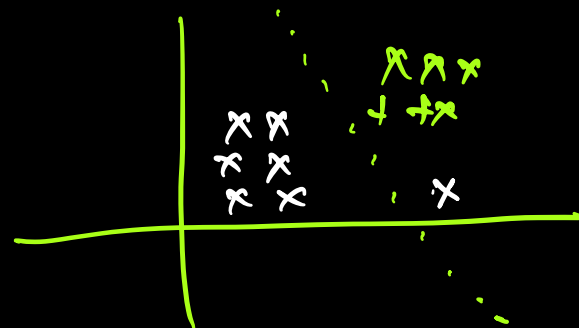
select random data point

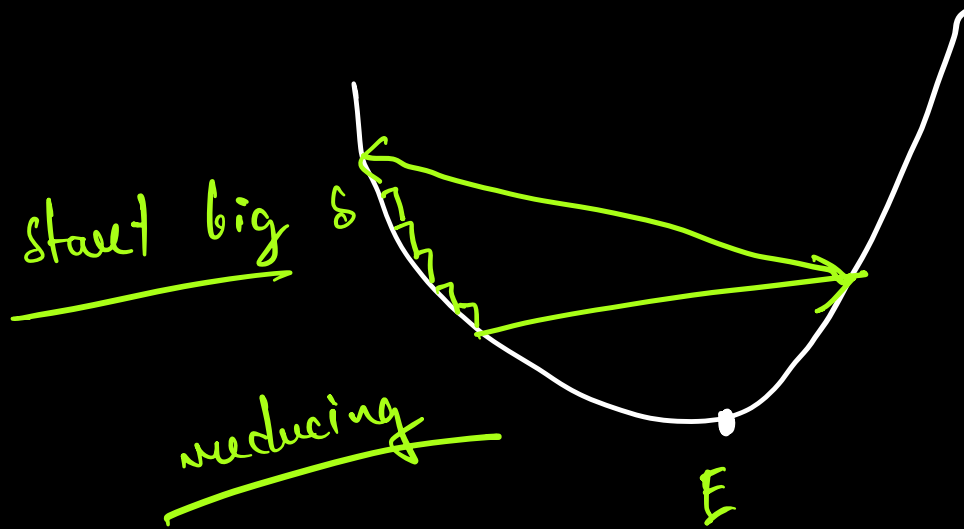
if $x_i \in N$ and $\sum w_i x_i > 0$

update the line parameters

A, B, C

1) learning rate





1000 steps

333 $\rightarrow 0.1$

333 $\rightarrow 0.01$

334 $\rightarrow 0.001$