



AI – ML – DL



By
ADITYA PRABHAKARA



Introducing Myself

Aditya S P (sp.aditya@gmail.com)

Freelance trainer and technologist

Boring Stuff about me:

- 15+ years of experience in development and training
- Started with Java, moved to Android and now working on Big Data Technologies

Interesting Things about me:

- Actually Nothing !



My Current work

To create a real time score rendering system for a financial company

- ✓ Platform - AWS
- ✓ Architecture - Serverless
- ✓ AWS Services - Glue, Lambda, API Gateway, RDS, Amazon ML, SageMaker
- ✓ Programming Language - Python
- ✓ Algorithms - Regression

Agenda

- 1) Introduction to Machine learning and Deep Learning
- 2) AI
- 3) Developing AI solutions
- 4) Use cases
- 5) Societal impact of AI
- 6) Deep Learning and Neural Networks
- 7) Artificial Neural Networks
- 8) Reinforcement learning

Objective and expectation

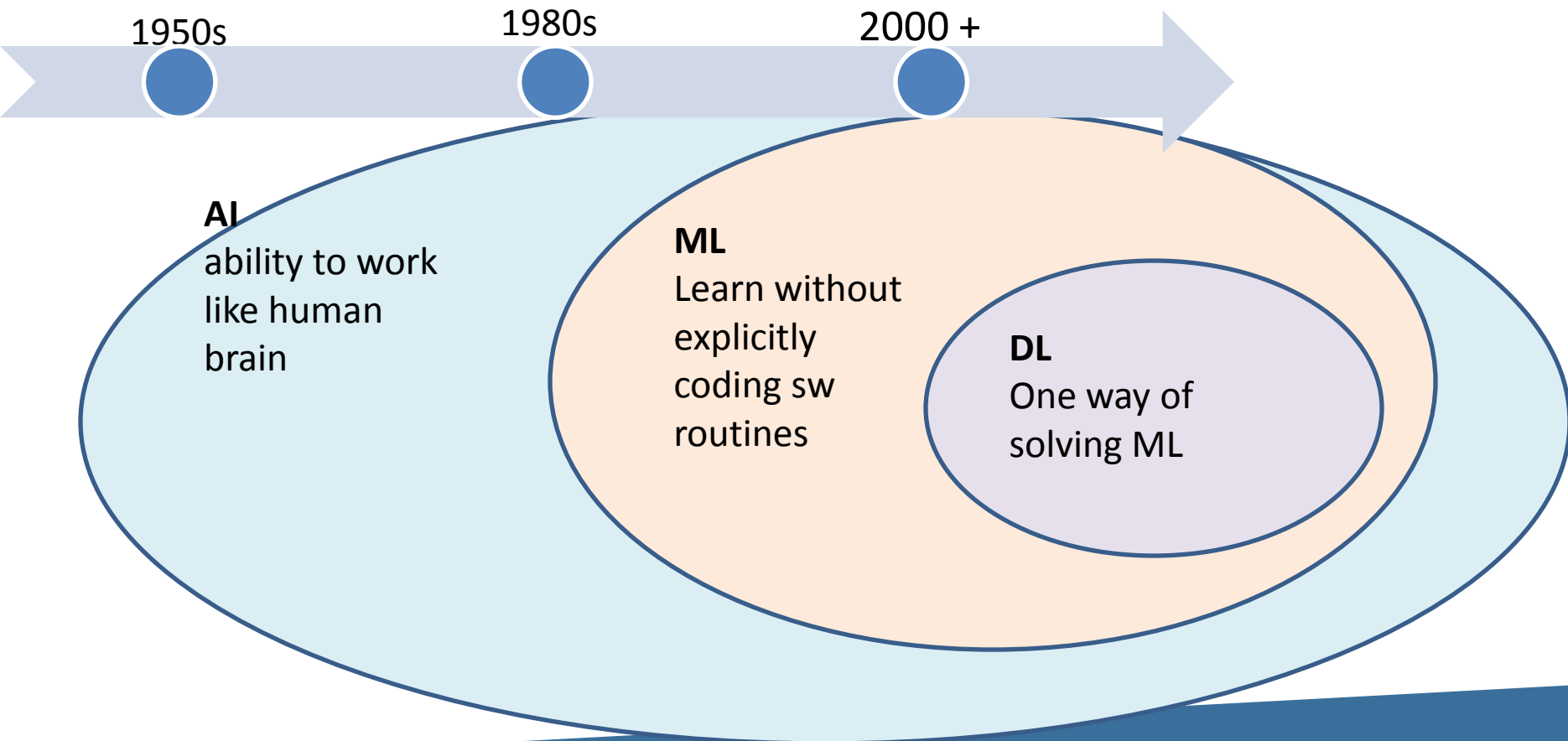
- 1) Very comfortable with the differences of AI-ML-DL
- 2) Hands on experience of executing a few ML/DL programs
- 3) Clean intuitive understanding of Algorithms and steps
- 4) Familiarity with a lot of jargon and terminology
- 5) Will I be ML/DL expert after this course? No, definitely not
- 6) Is this an advanced course? No, This is an introductory course
- 7) I already know these!
- 8) How much more time is required for me to become ML/DL expert? ~6 months of constant efforts



AI vs ML vs DL

By
ADITYA PRABHAKARA

❖ AI vs ML vs DL



❖ AI vs ML vs DL

AI

- Machines which can perform tasks like human
- Very generic and broad spectrum
- Narrow AI vs Generic AI

ML

- Core way of achieving ML
- Learn without being explicitly programmed
- So training is required
- Regression, Prediction,

DL

- Computation intensive Machine learning
- Inspired by how brain works like interconnecting neurons
- ANN

✦ Use cases

Smarter healthcare

Traffic control

JIT manufacturing

Trading analytics

Search quality

Customer behaviour

Fraud detection

Customer experience management

Predictive maintenance

Churn analysis

Security anomaly detection



An Intuitive Understanding of ML

By
ADITYA PRABHAKARA

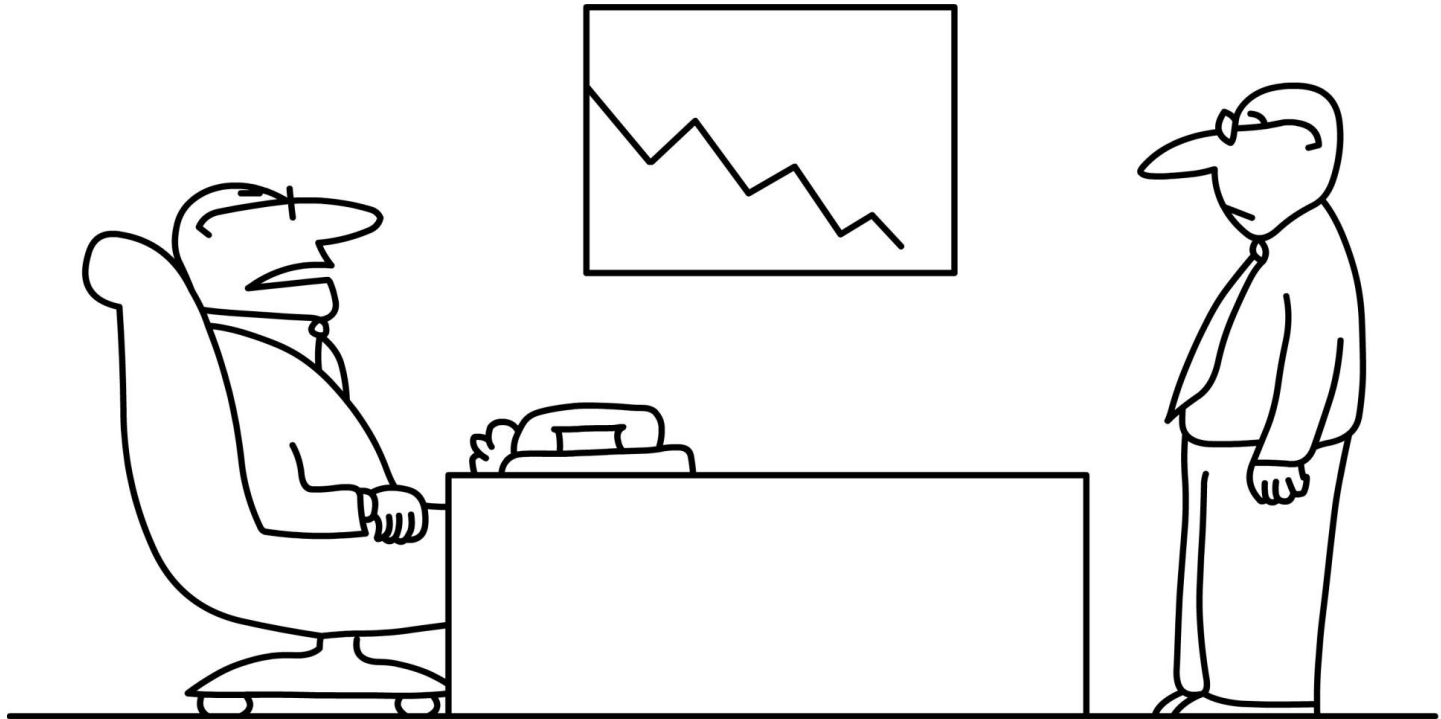


Chapter: Introduction

❖ Scene 1



❖ Scene 2



©2008 www.timoelliott.com

"It would appear, Hopkins, that your gut feel was only indigestion"

ML



A widely quoted definition of ML algorithm by Tom M Mitchell

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 at tasks in T , as measured by P , improves with experience E

❖ Step 1 - Intuition



Given the value of x to be 3,
what is the value of y ?

❖ Step 2 - Intuition



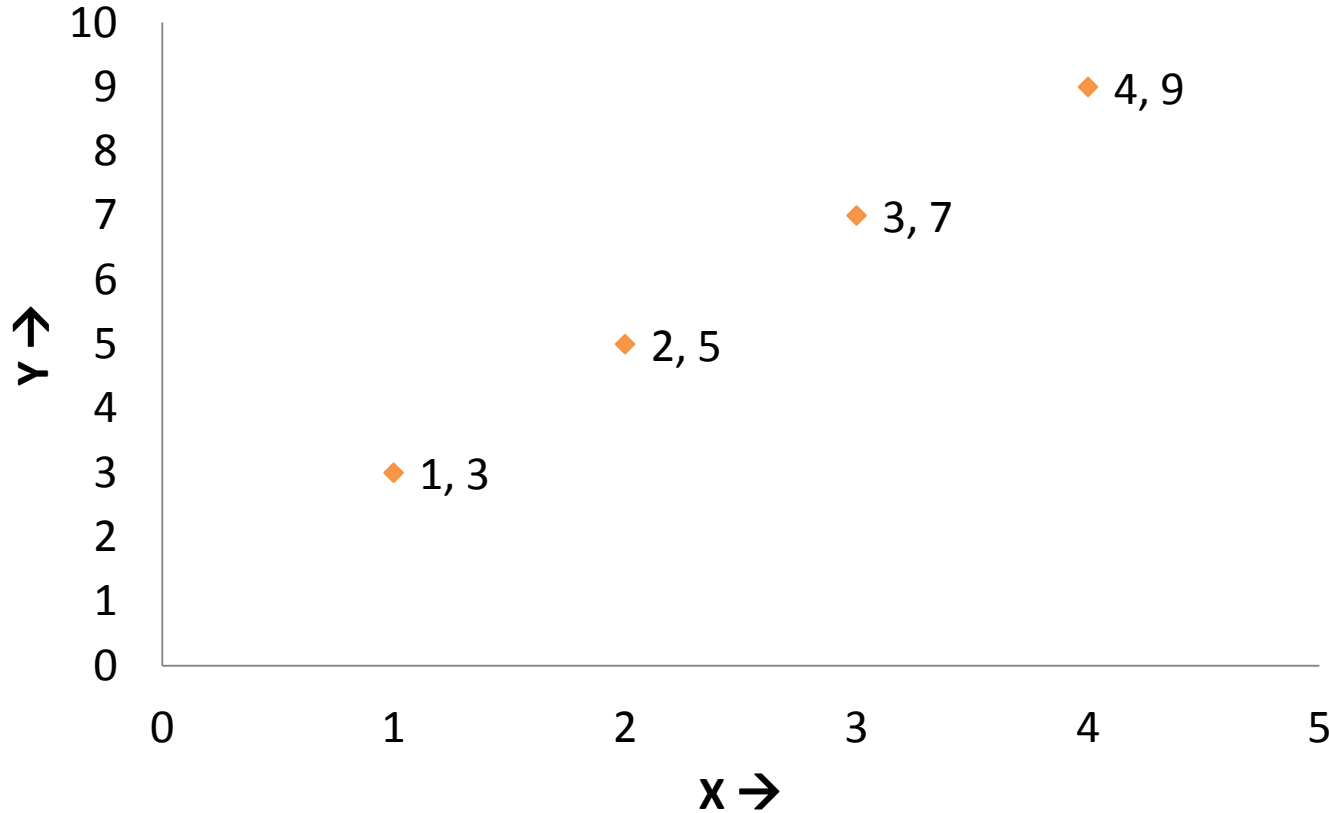
❖ Terminology

Given the value of x to be 3, what is the value of y ?

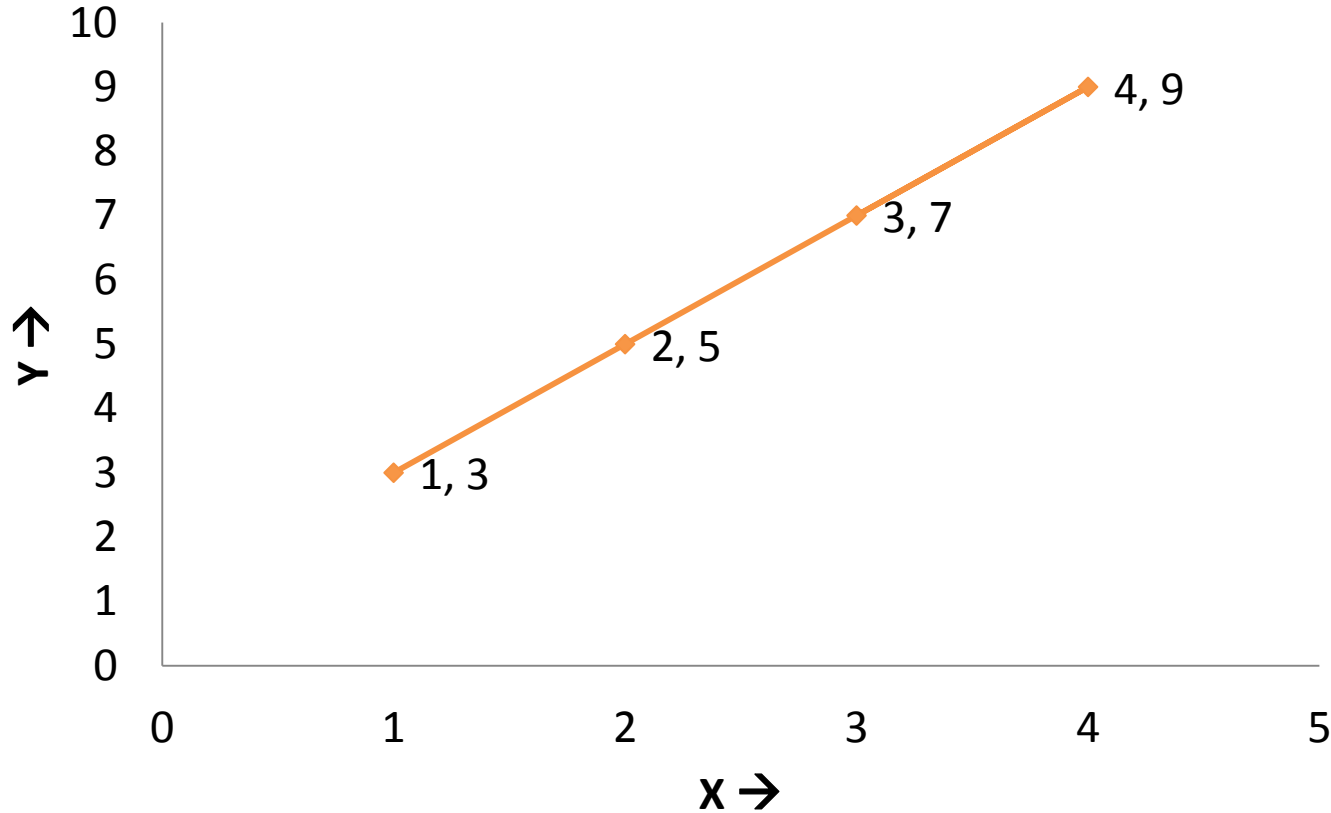
- ✓ Feature
- ✓ Target Value

Y should be 7 given x is 3
How did we make that decision?

❖ Step 3 - Intuition



❖ Step 4 - Intuition



❖ Step 4 - Intuition

Equation of a line – Linear Equation

$$y = mx + c$$

Y is the target

X is the feature

C is the intercept

M is the coefficient

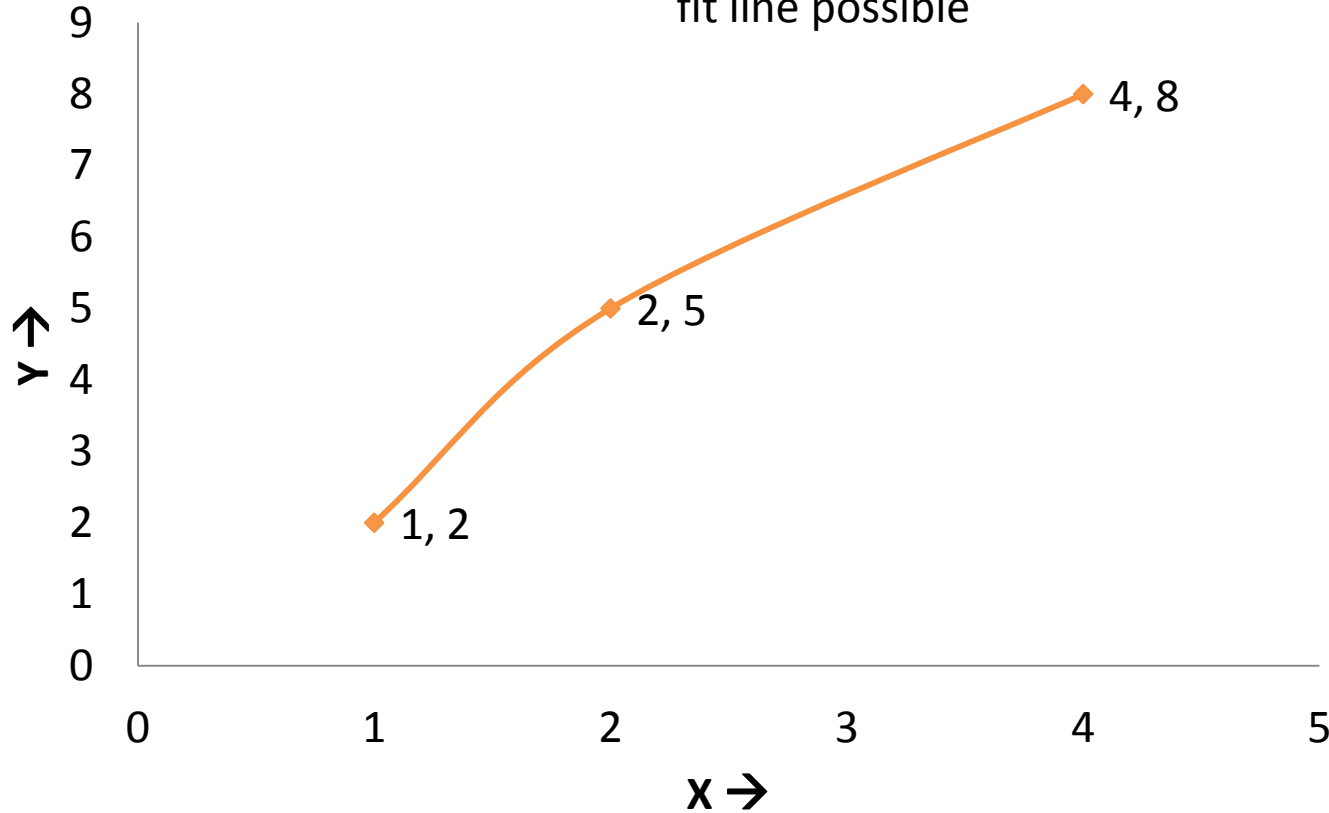
X-Values	Y-Values
1	3
2	5
4	9
3	7

$$y=2x+1$$

❖ Step 5 - Intuition

Add some confusion

Cannot create an equation for this kind of a line
So the best solution would be to find the best fit line possible

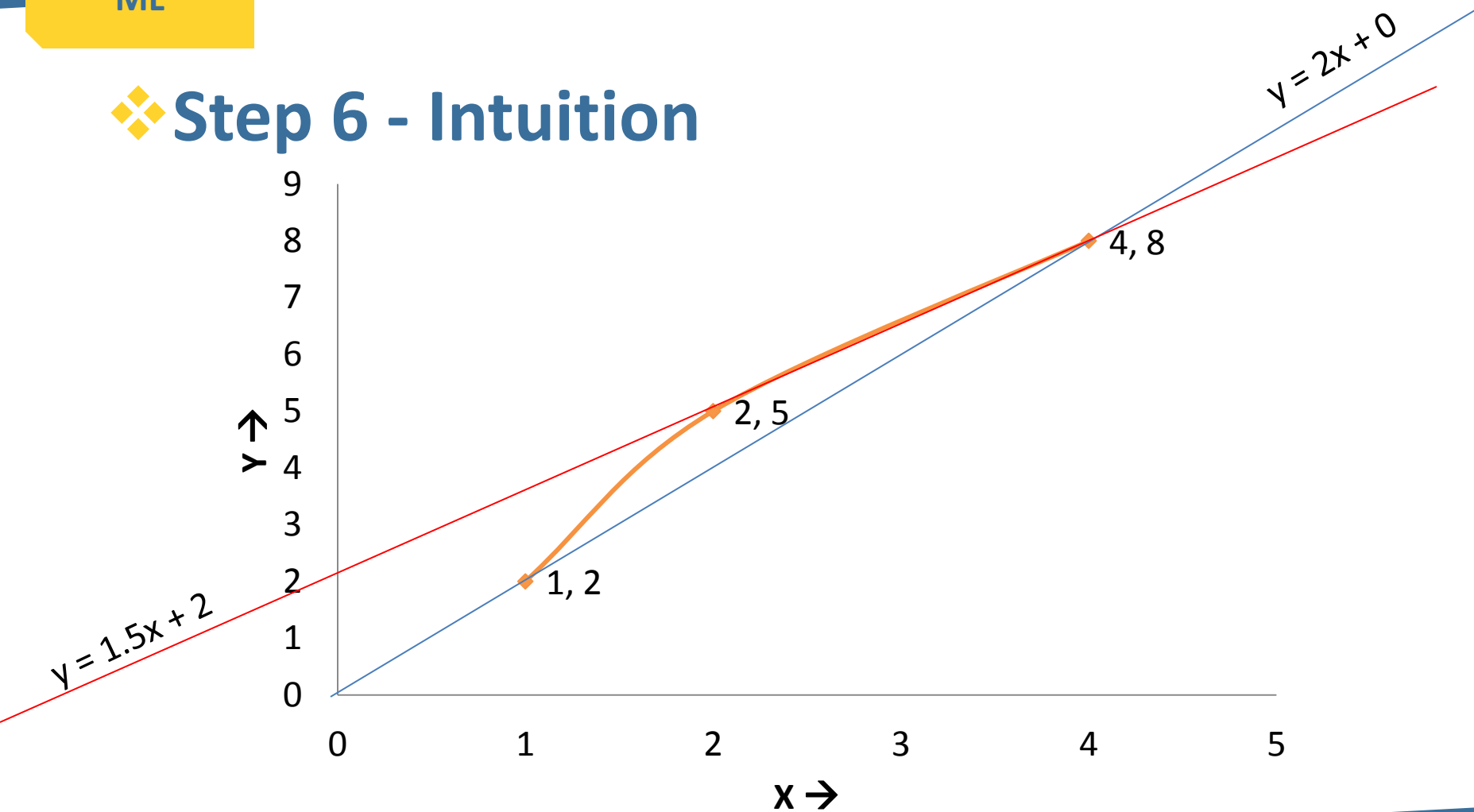


❖ Step 4 – Best fit line strategy

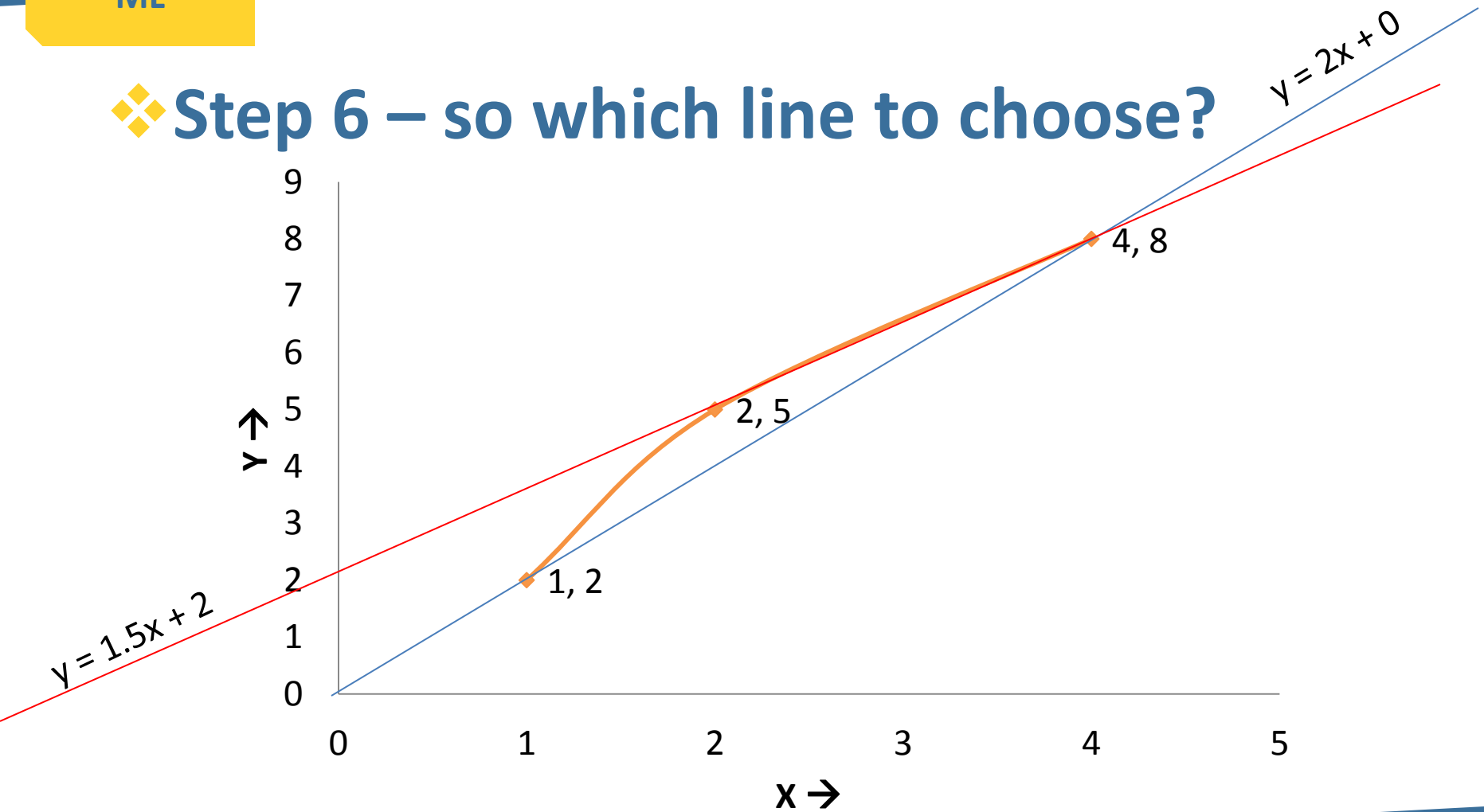
Would try to go through most number of points ?

If there are more than one line possible which to choose?

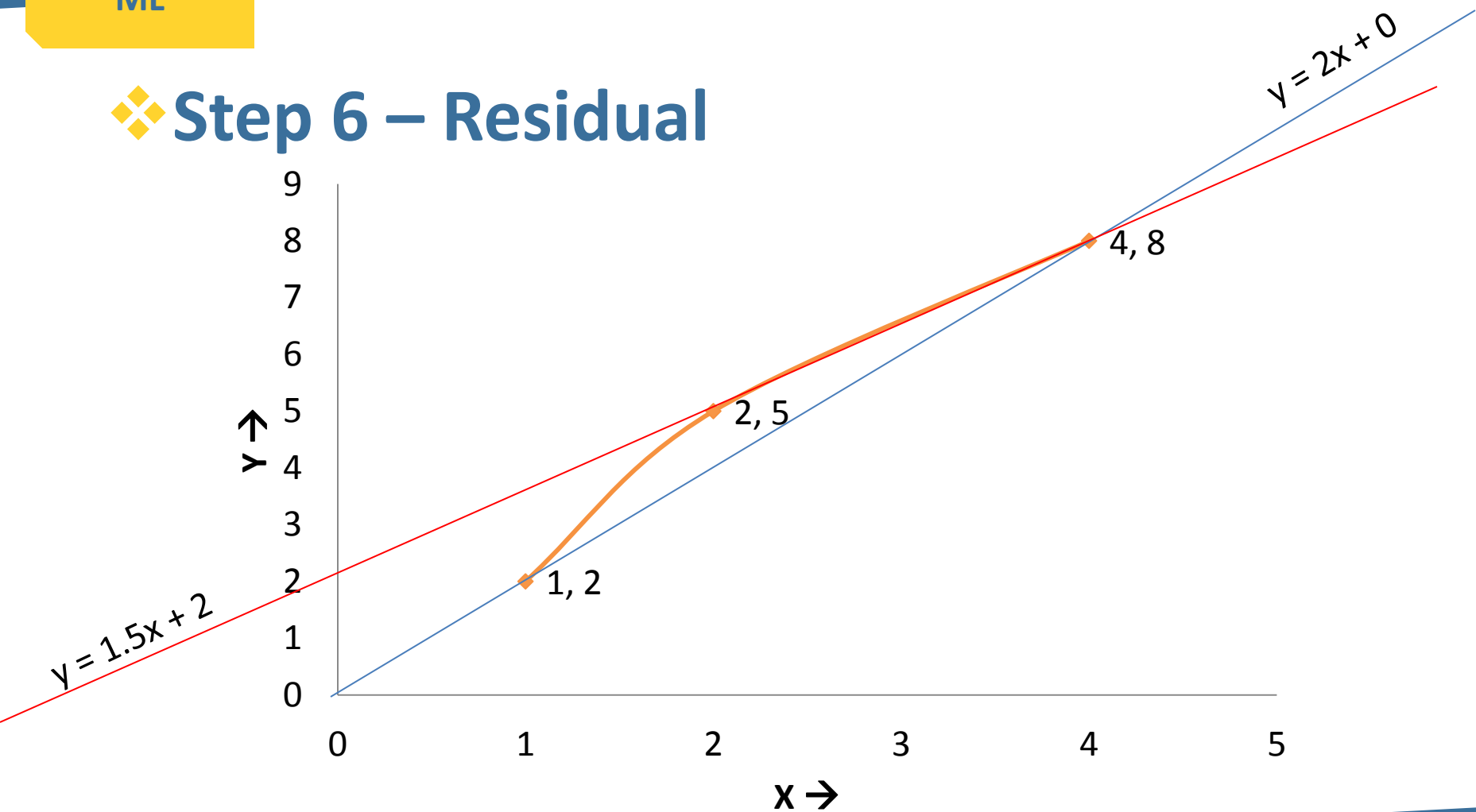
❖ Step 6 - Intuition



❖ Step 6 – so which line to choose?



❖ Step 6 – Residual



❖ Let us look at a real life example

House price prediction

The “Hello World” of prediction examples

See the example with rooms with prices comparison

✦ Features

It is not that only room numbers decide the price of the house. There can be many

For ex:

Total Area,

Construction material,

Number of schools near by,

Age of the house,

Number of ATMs near by

Correlation

Not all features affect the same way.

i.e. number of atms do not affect the price as total area does.

❖ Features

Considering all the features a generalized line equation will look like

$$y = m_1x_1 + m_2x_2 + m_3x_3 + m_4x_4 + \dots + c$$

$$2.15282755e+01 * \text{Area Income} + 1.64883282e+05 * \text{House Age} + \\ 1.22368678e+05 * \text{Number of Rooms} + 2.23380186e+03 * \text{Avg. Area Number of} \\ \text{Bedrooms} + 1.51504200e+01 * \text{Area Population} + \\ -2640159.79685$$

y is the target

$x_1, x_2, x_3, x_4 \dots$ are features

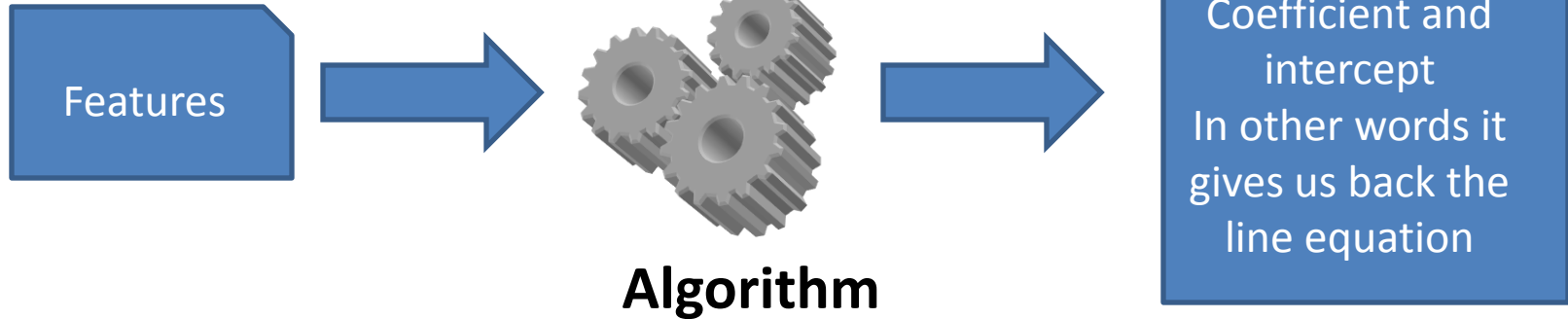
$m_1, m_2, m_3, m_4 \dots$ are coefficients

c ... which is the intercept

In this example:

❖ Features

- ✓ When multiple features start affecting the price, then just a graph plotting would not be enough.
- ✓ We have to feed this data through an algorithm that will help me decide the coefficients and intercept



❖ Algorithm

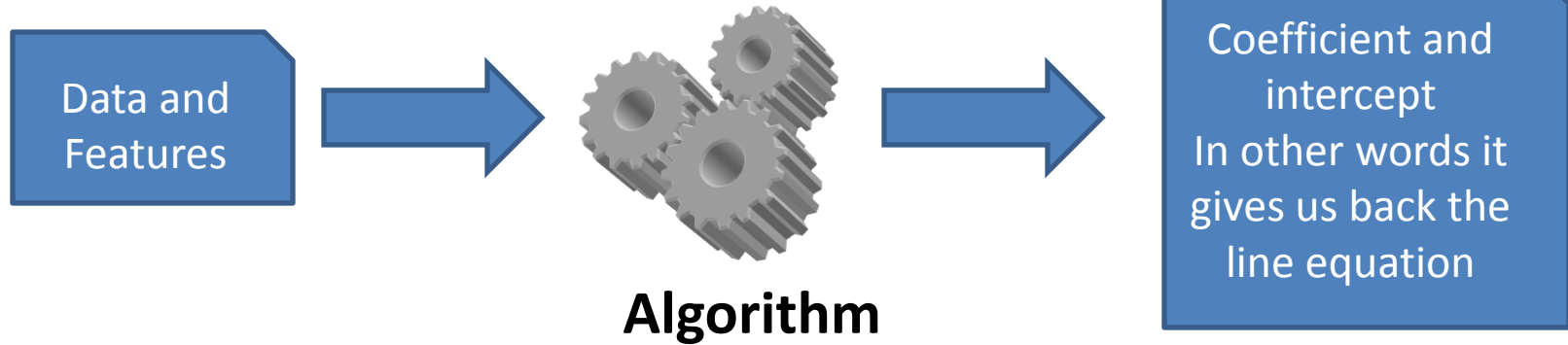
Current algorithm

1. Equation of line => Linear Equation
2. Class of algorithm
 1. Regression

In spite of a stellar performance of a single feature, other features will pull the final target towards the mean

=> The final value will regress towards the mean

❖ The question – where is the machine learning?



Every time we change the features or data, machine learns about it and adjusts its estimate of co-efficients and intercept. There by giving a brand new line equation!

❖ Now let us go back to the definition of ML

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 at tasks in T , as measured by P , improves with experience E

Experience $E \Rightarrow$ Data and features

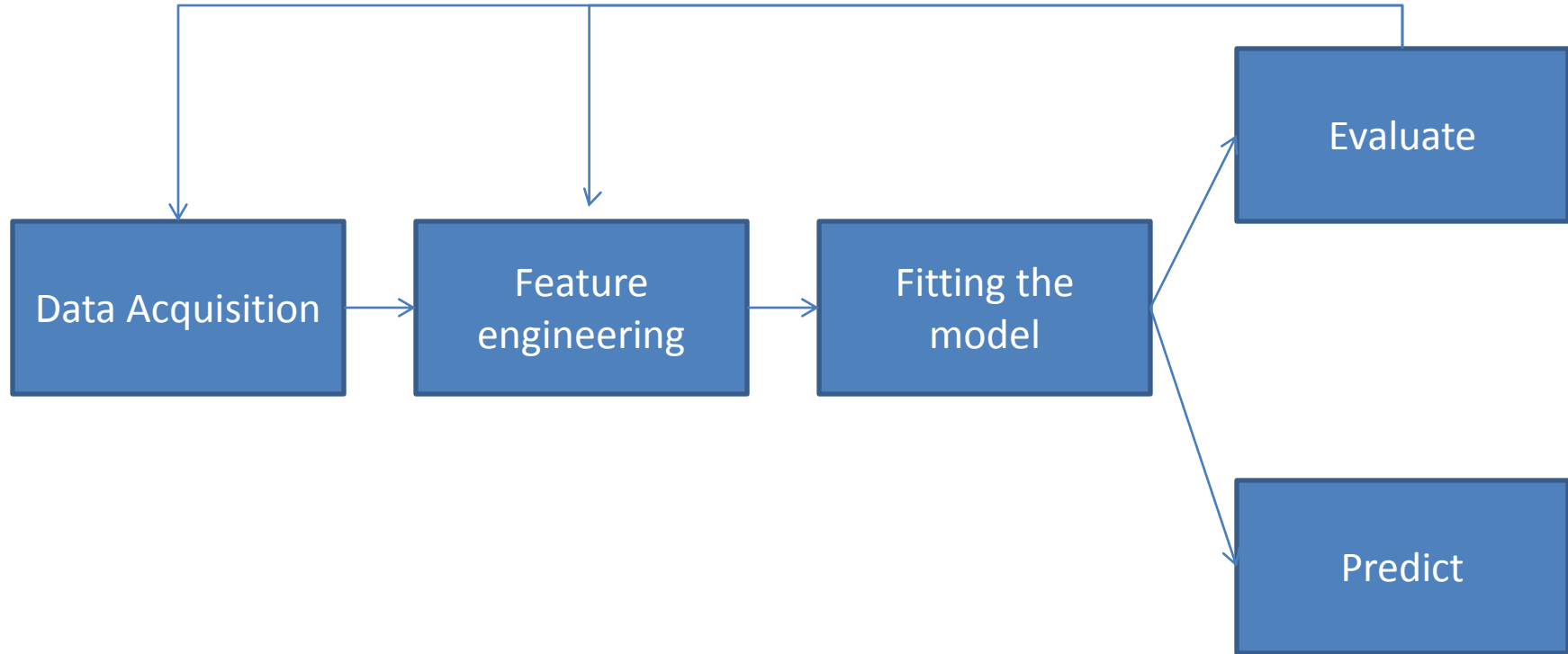
Class of task $T \Rightarrow$ predicting the price (in this case)

Performance measure $P \Rightarrow$ sum of residuals (Error)

Machine learns because

If its performance at tasks in T , as measured by P , improves with experience E

❖ High level ML process



ML

 Demo

ML

❖ Learning Mind Map



Chapter: Classification

❖ Use case

From Wikipedia

Probability of passing an exam versus hours of study

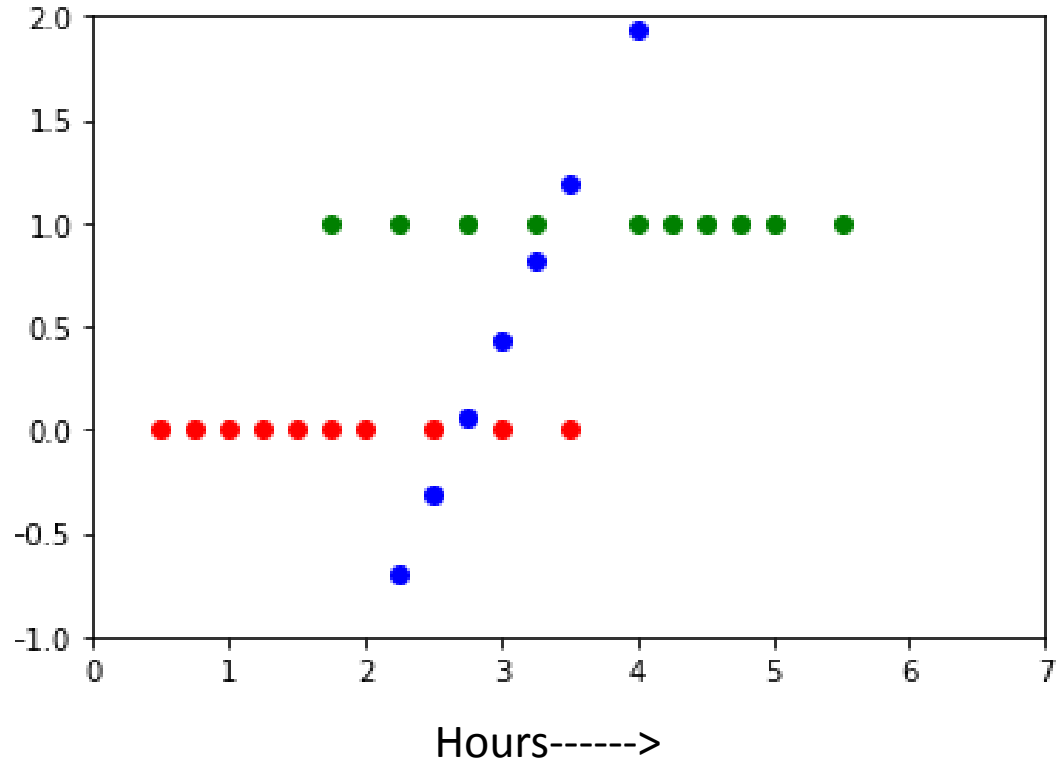
1. Let us start with what we know. Let us start to treat this as linear regression problem

	Coefficient
Intercept	-4.0777
Hours	1.5046

❖ Use case

Y values

-3.32540
-2.94925
-2.57310
-2.19695
-1.82080
-1.44465
-1.44465
-1.06850
-0.69235
-0.31620
0.05995
0.43610
0.81225
1.18840
1.94070
2.31685
2.69300
3.06915
3.44530
4.19760





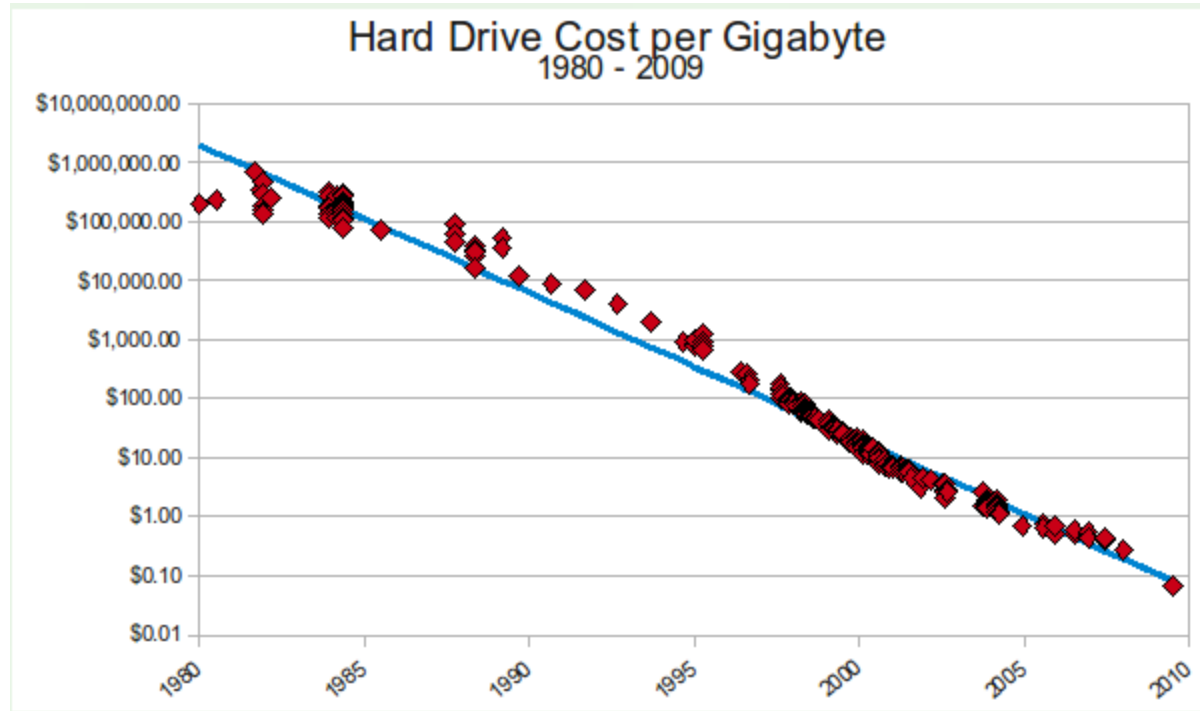
Chapter: Deep learning

❖ What is Deep Learning

1998 – Ad about what is internet !!

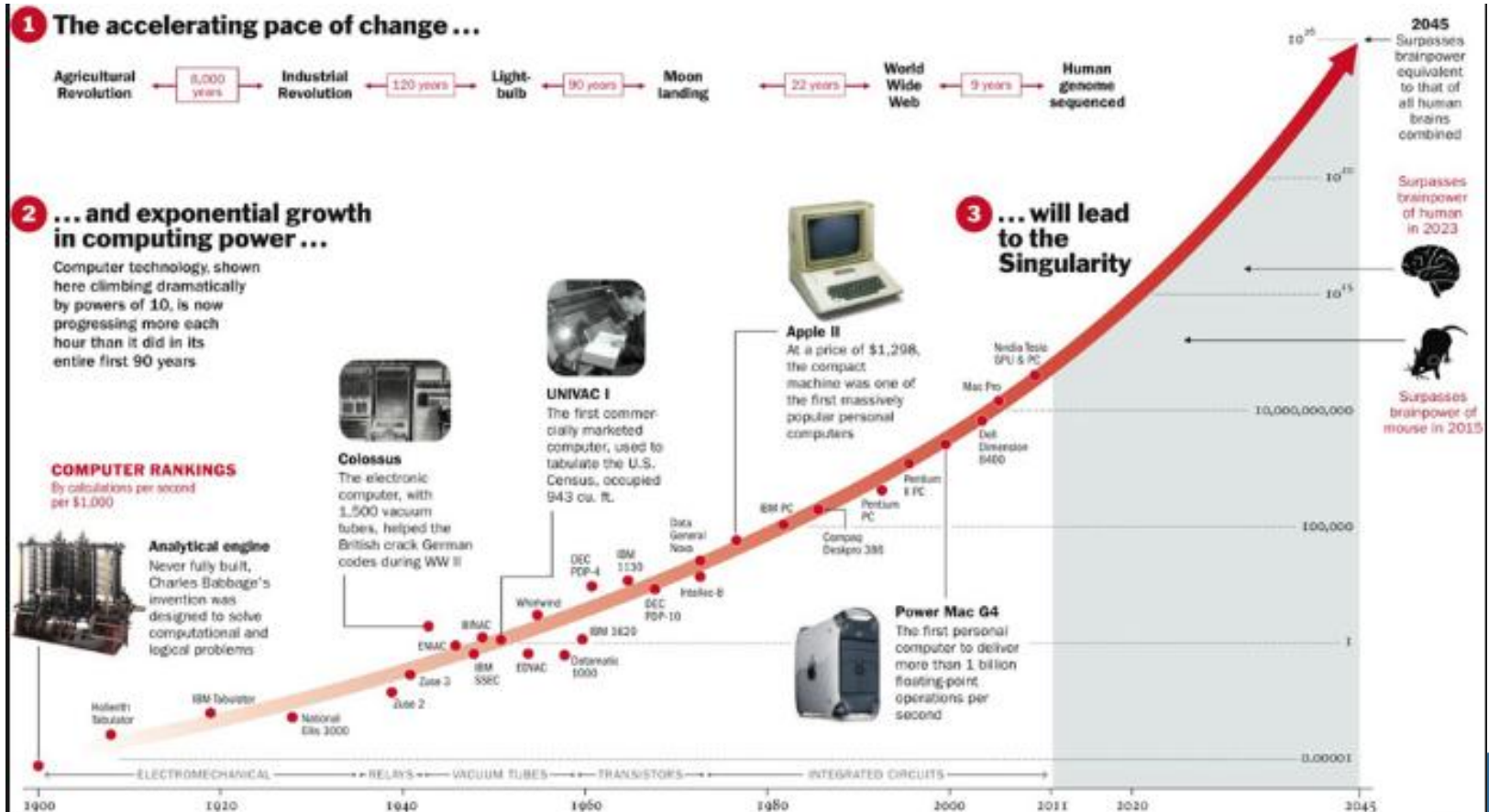
https://www.youtube.com/watch?v=qh_AZWq1lfo

❖ Hard drive prices



Source: <http://www.mkomo.com/cost-per-gigabyte>

❖ Processing capacity (content.time.com)



❖ Father of DL



Geoffrey Hinton

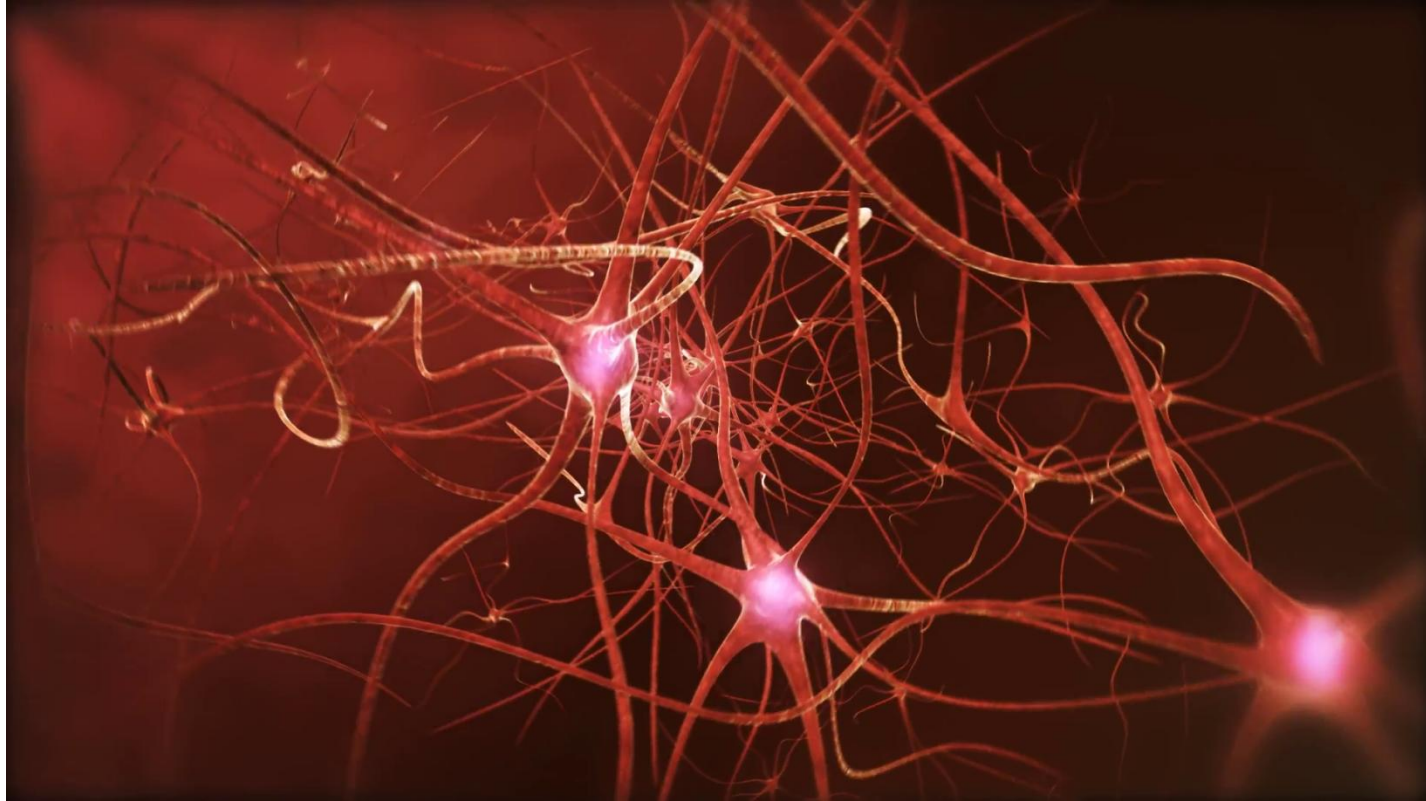
Currently works at Google

Computer algorithms mimicking human brain

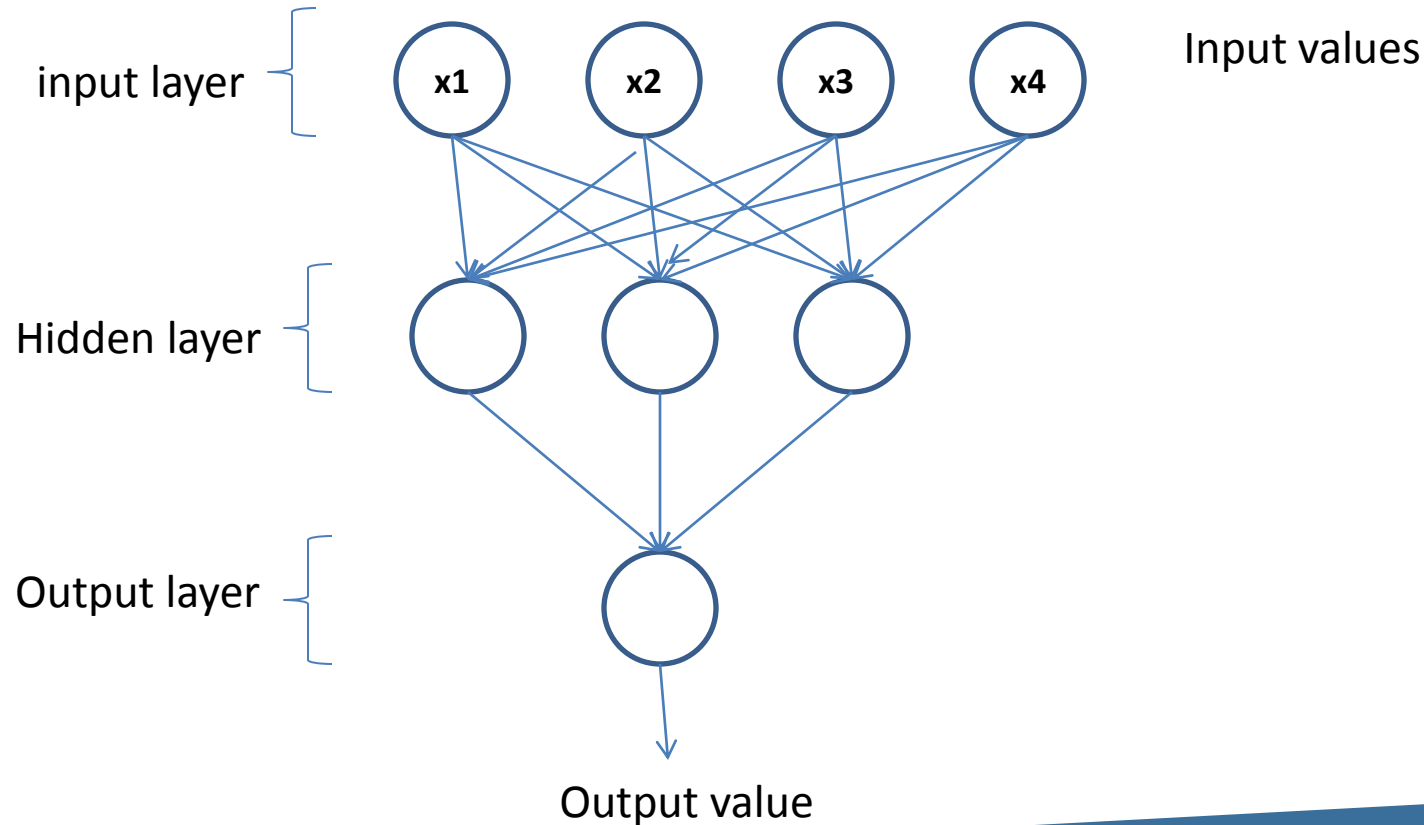
From Wikipedia,

- > Moved from USA to Canada mainly in disapproval of military funding of AI research
- > AI will be used to “terrorize people”

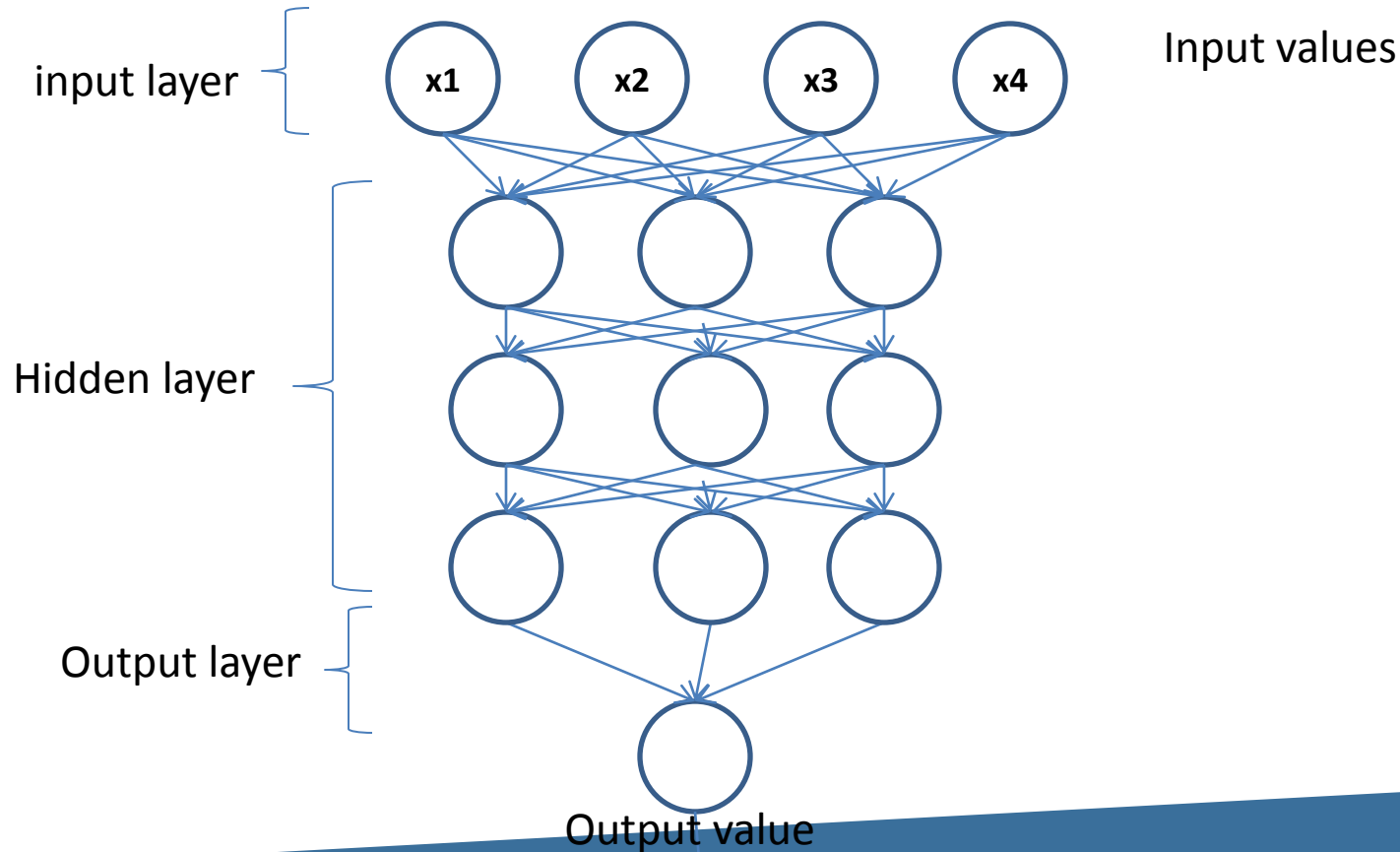
✦ Neuron



❖ DL computer program mimics the same



❖ DL computer program mimics the same



❖ Father of DL



Geoffrey Hinton

Currently works at Google

Computer algorithms mimicking human brain

From Wikipedia,

- > Moved from USA to Canada mainly in disapproval of military funding of AI research
- > AI will be used to “terrorize people”

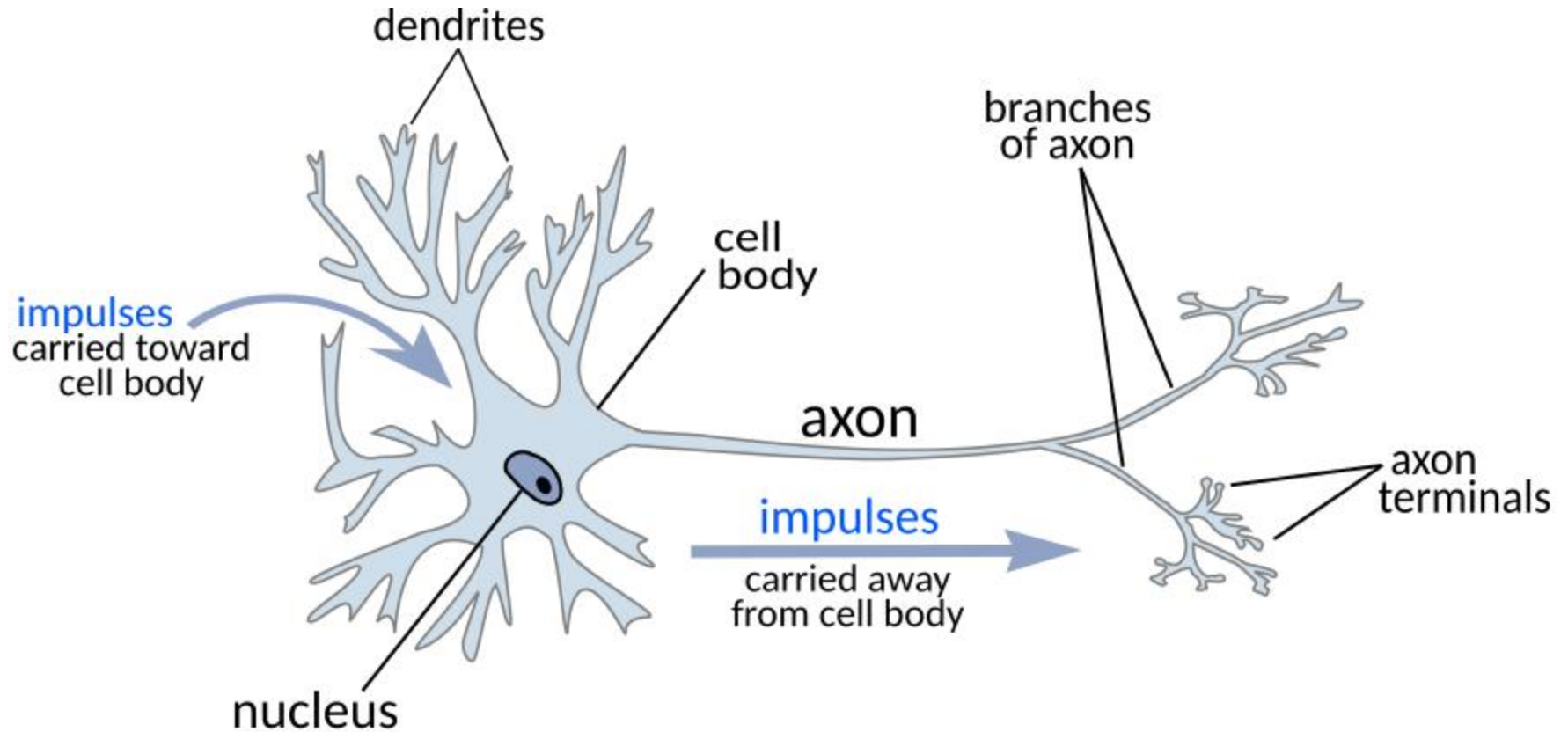


Chapter: Artificial Neural Networks

Agenda

- 1) The network of neurons
- 2) The activation function
- 3) Working of an ANN
- 4) How do they learn
- 5) Backpropagation

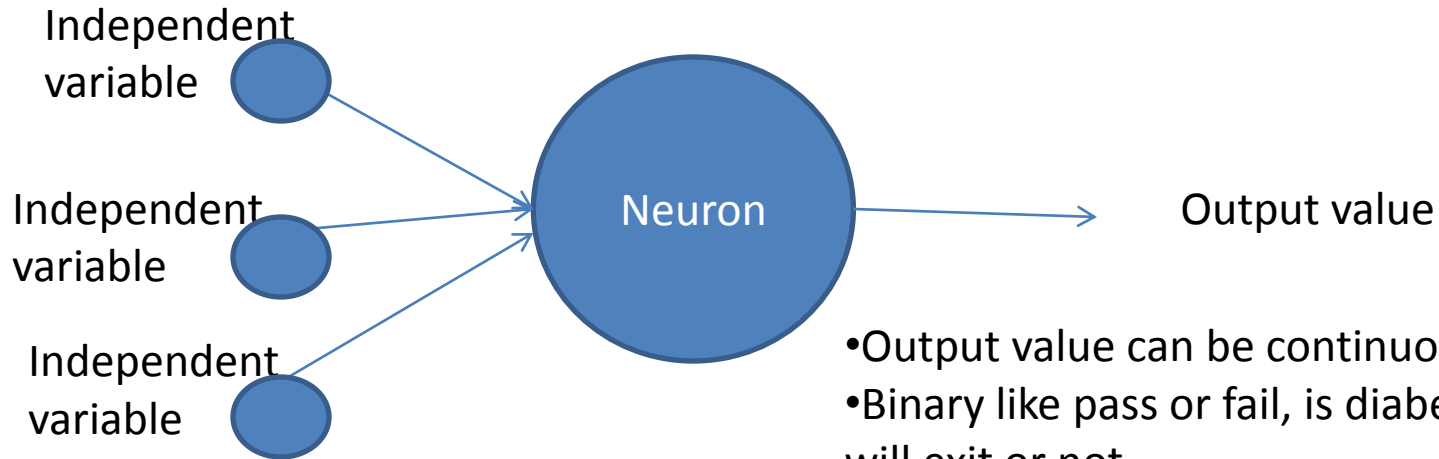
❖ Neuron



✦ Neuron

A basic building block of ANN
Dendrites are the input values

- Independent variables represent one observation or one row in a database
- Standardize or normalize the input variables

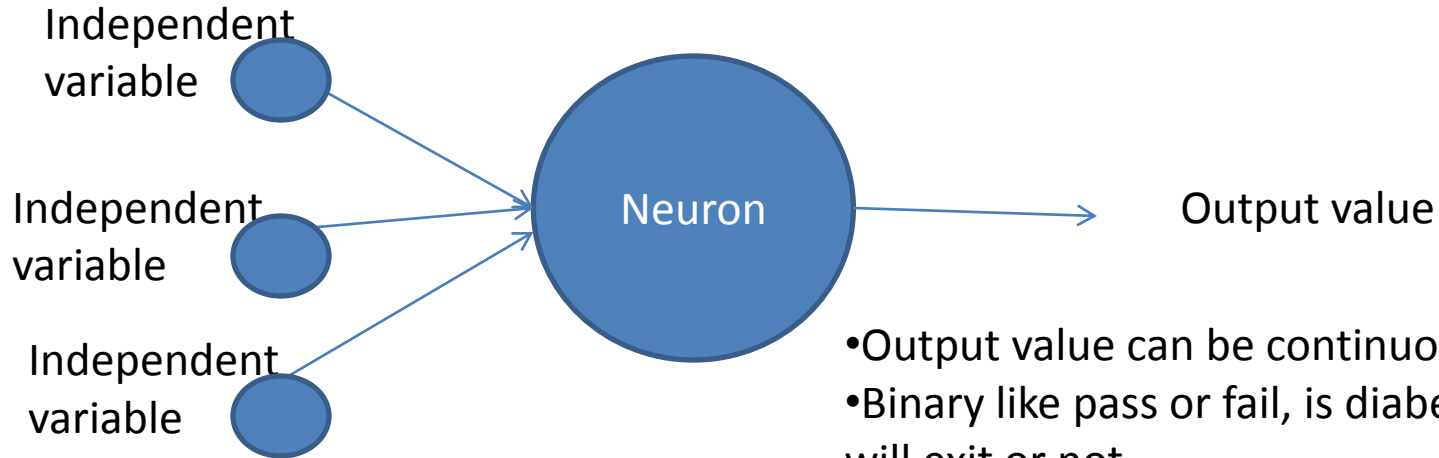


- Output value can be continuous like price
- Binary like pass or fail, is diabetic or not, will exit or not
- Categorical

✦ Neuron

A basic building block of ANN
Dendrites are the input values

- Independent variables represent one observation or one row in a database
- Standardize or normalize the input variables

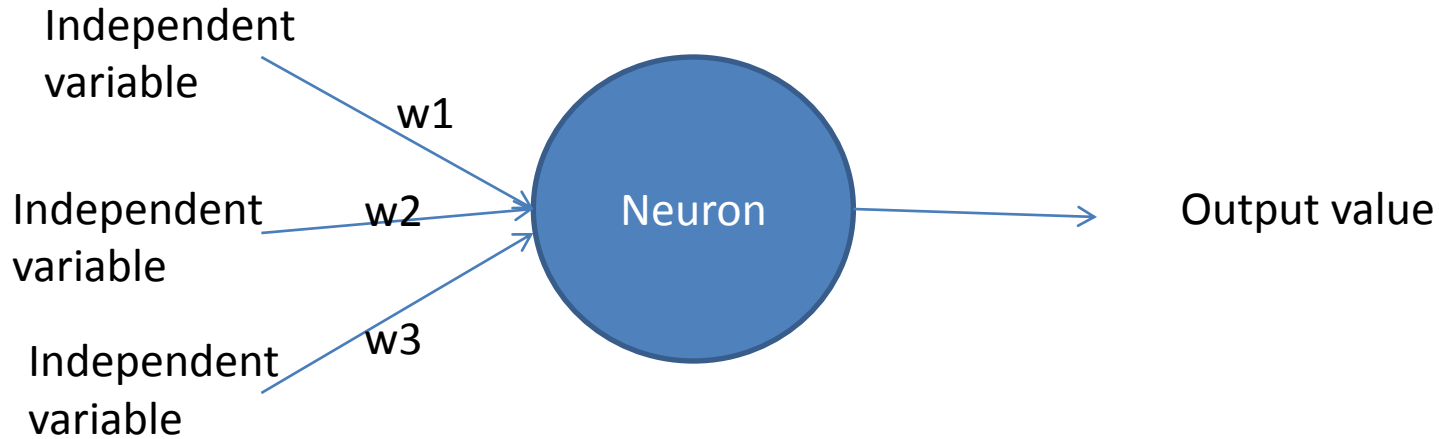


- Output value can be continuous like price
- Binary like pass or fail, is diabetic or not, will exit or not
- Categorical

✦ Neuron

A basic building block of ANN

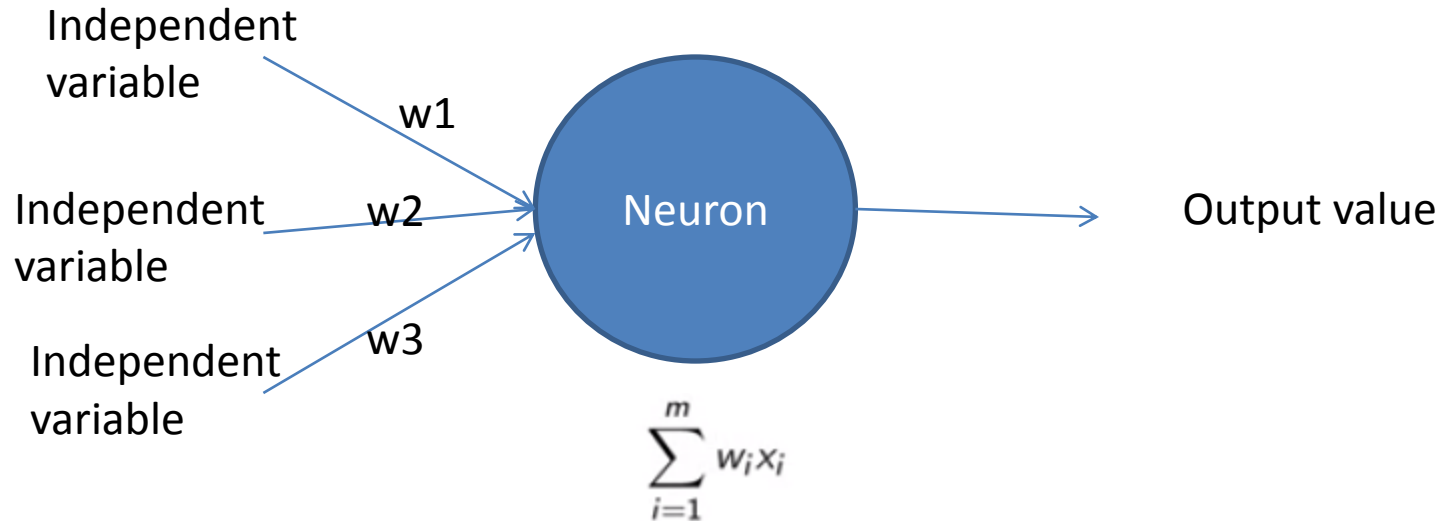
- Weights get attached to each input
- Implies how important one input is for the particular neuron



✦ Neuron

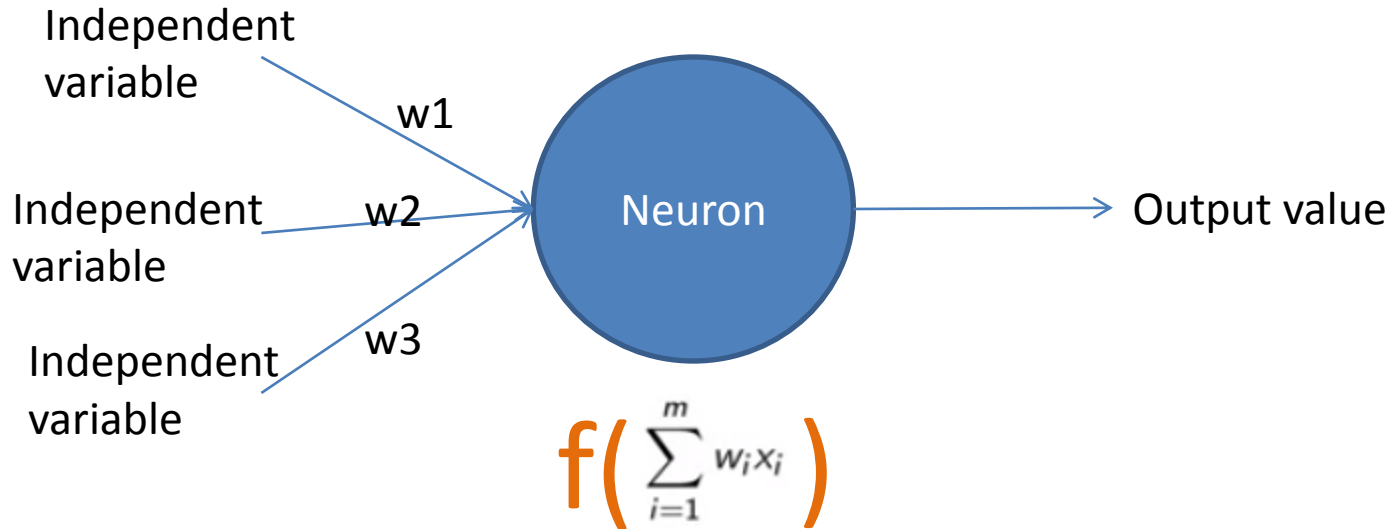
- In the neuron these are added to get a value

A basic building block of ANN

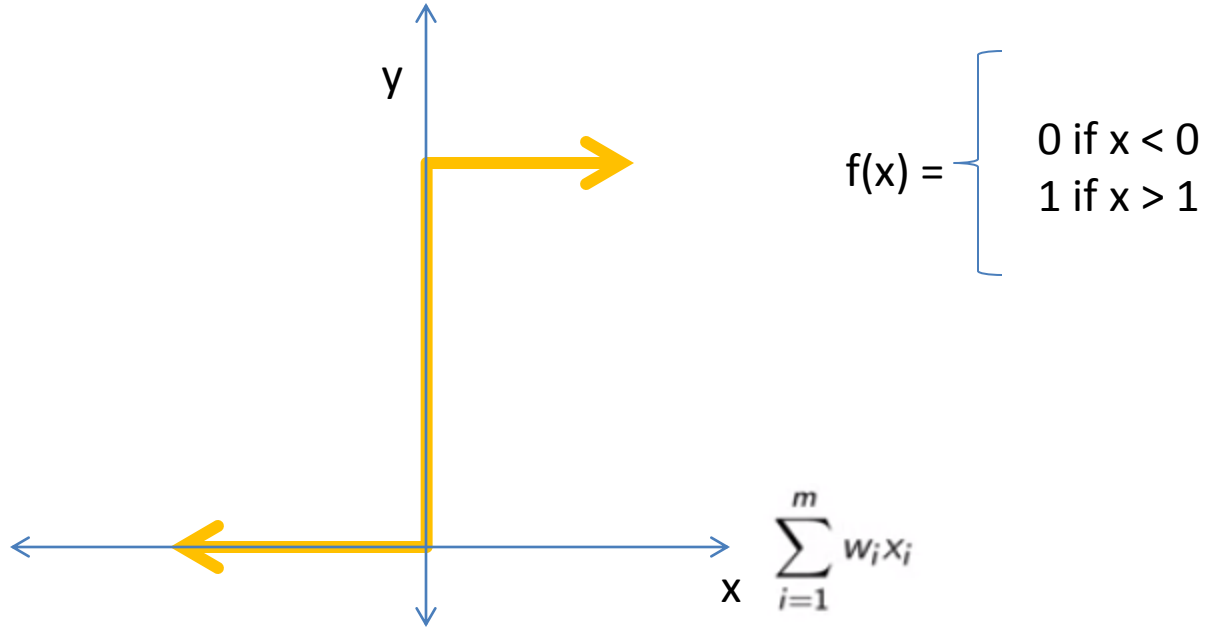


❖ Activation function

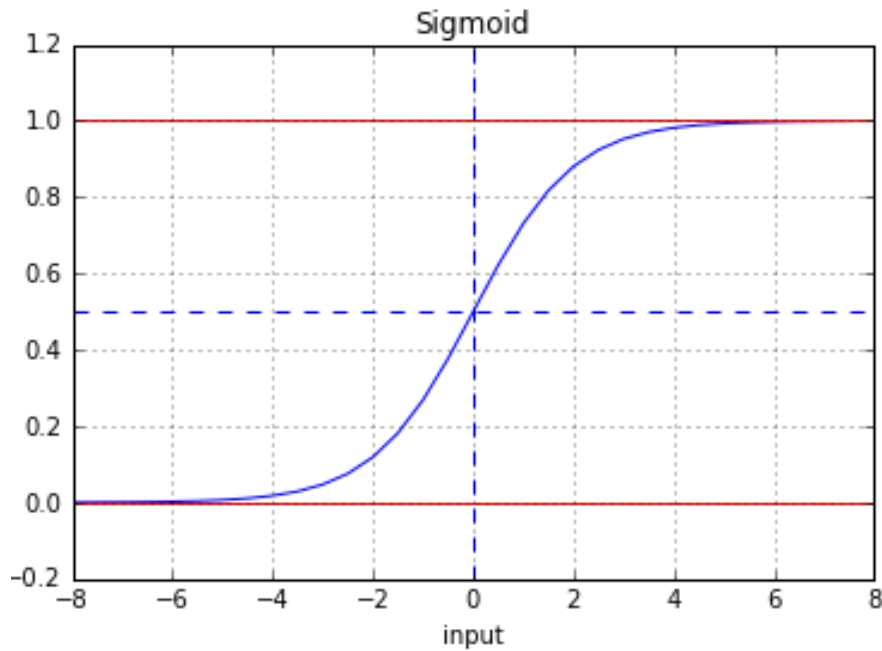
A basic building block of ANN



❖ Threshold function



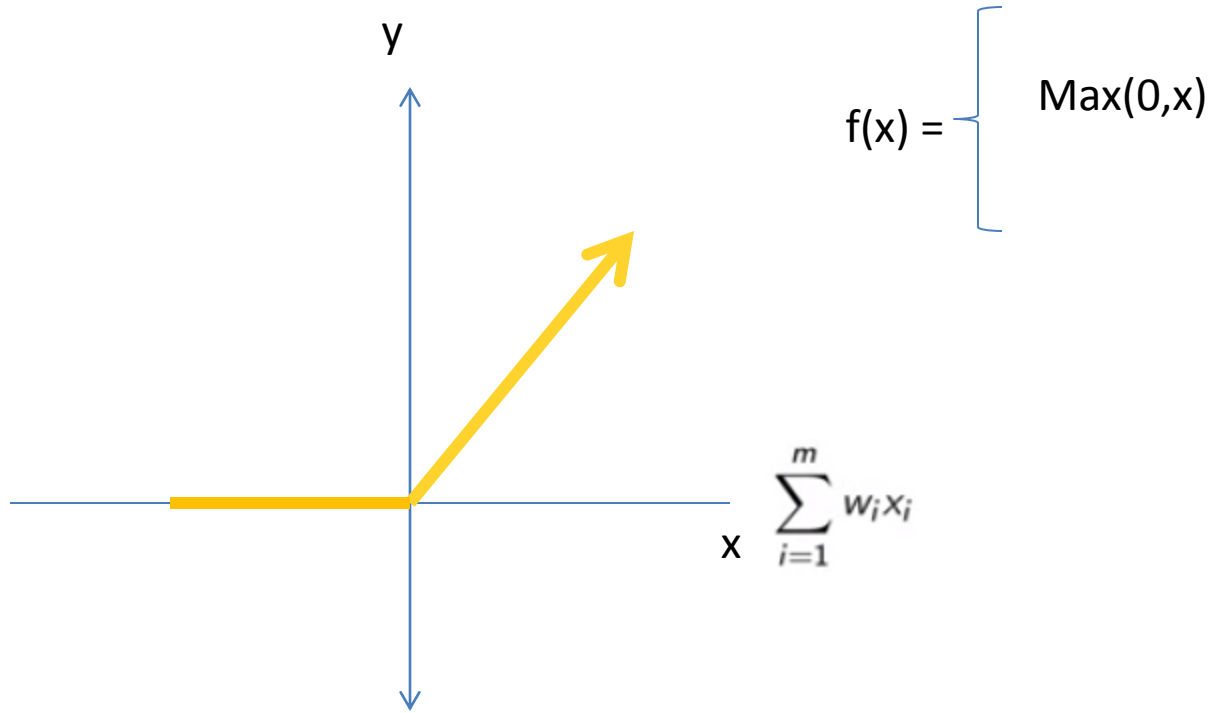
❖ Sigmoid function



$$f(x) = \frac{1}{1 + e^{-x}}$$

$$\sum_{i=1}^m w_i x_i$$

❖ Threshold function



❖ Application

