Workshop

Biologically Inspired Learning:

Neural Networks & Artificial Intelligence

Dr Amita Kapoor, Associate Professor, SRCASW

Venue: Hansraj College

dr.amita.Kapoor@ieee.org

Website: http://www.dramitakapoor.com

Purpose of the Talk

Basics of AI

- AI, ML, DL
- NN
- Model basic NN

Excite about AI

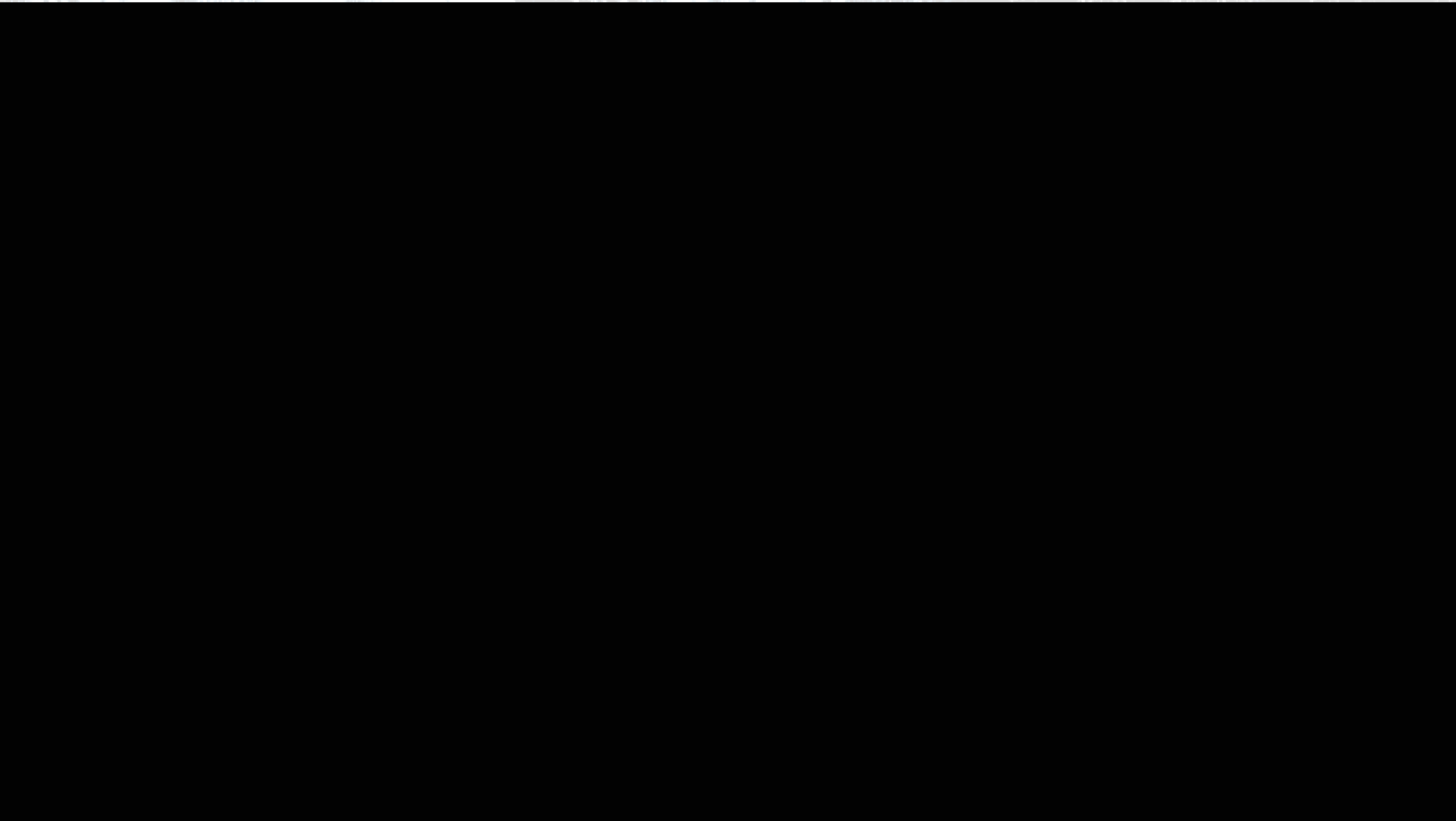
- Success Stories
- History of AI
- Great Personalities

Next Steps

- Online MOOCs
- Languages
- How to succeed!

Prof Hiroshi Ishiguro

Sophia- Hanson Robotics



Creating Art

Designing new Compoun ds

Gene Sequencing

Writing Books

Disease Detection

Grading answer Sheets

Face Recognition

Financial Market

Giving Legal advice

Music Generation

Machine Translation

Robotics

Drones

Self Driving Cars

Choose Any field and you will find application of Artificial Intelligence, Machine learning there

Artificial Intelligence

Emulate the intelligent Behaviour. Make machines do tasks, human are good at.

Machine Learning

Uses Statistical techniques that enable machines to improve performance with experience.

Deep Learning

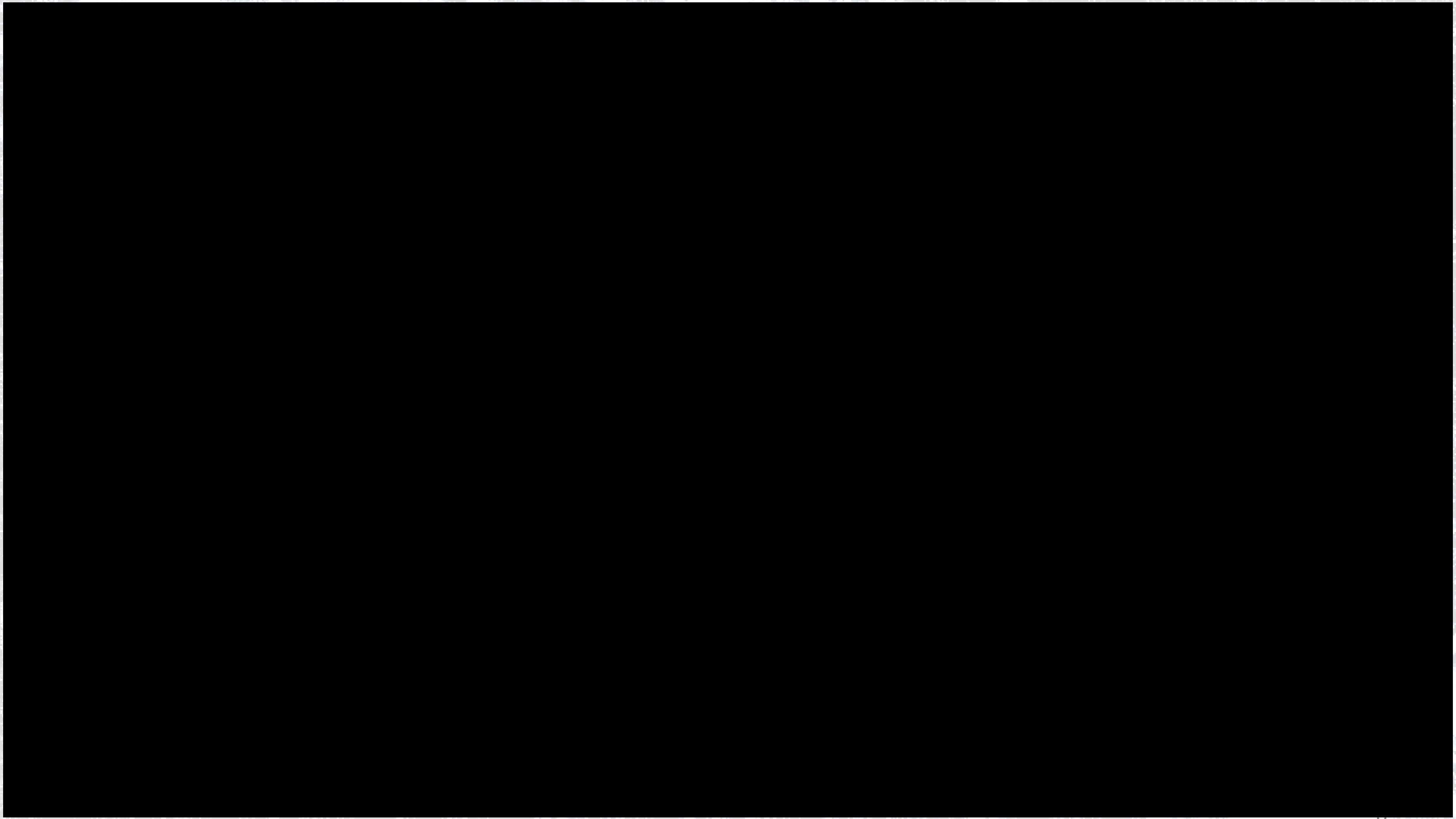
Multiple (Deep) layers of **Neural Networks**, that can be trained to perform task like speech and image recognition by learning through vast amounts of data.

AlphaGo

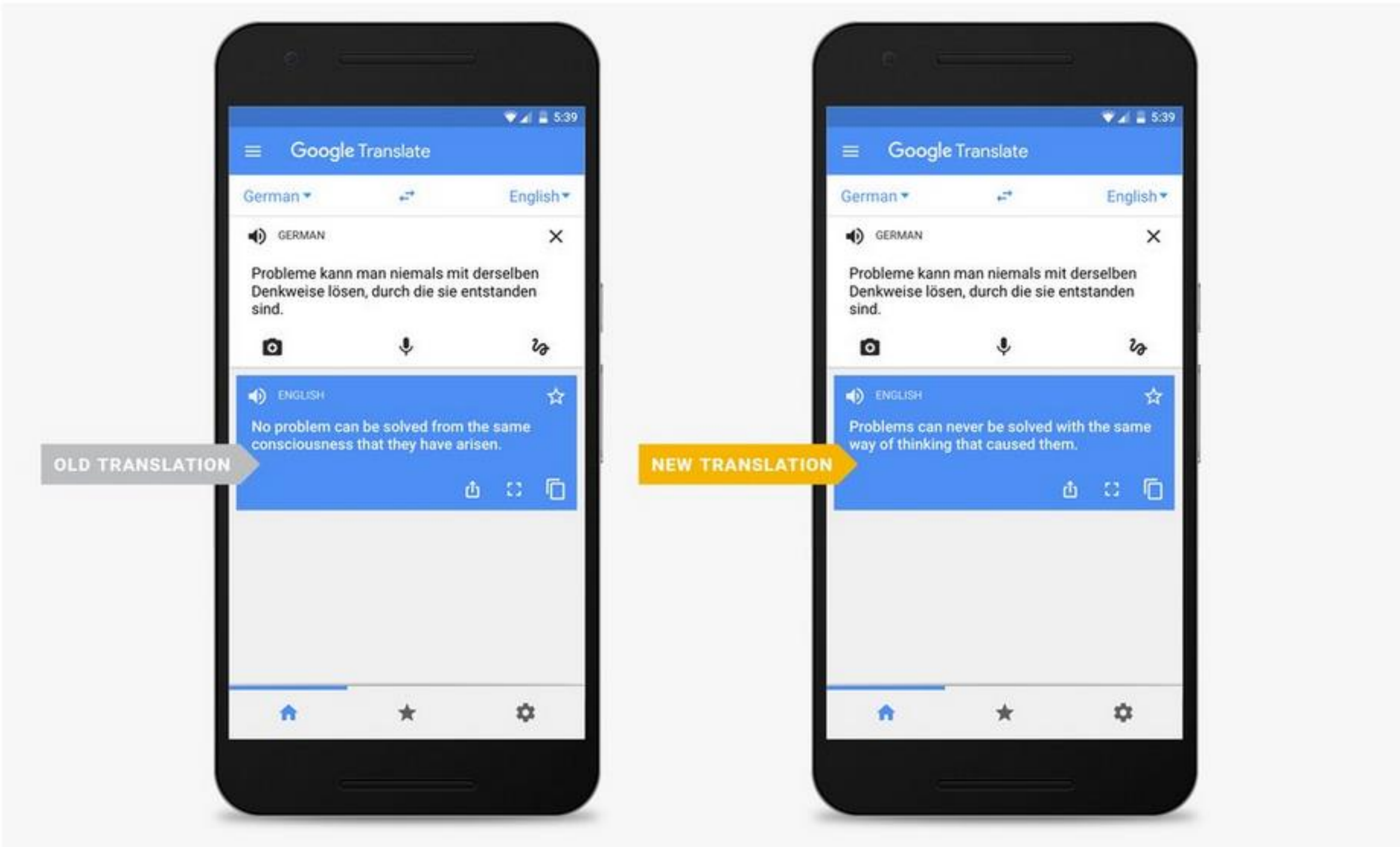
- The first computer program to defeat a professional human Go player, the first program to defeat a Go world champion, and arguably the strongest Go player in history.
- The game of Go originated in China 3,000 years ago.
- The rules of the game are simple: players take turns to place black or white stones on a board, trying to capture the opponent's stones or surround empty space to make points of territory.
- As simple as the rules are, Go is a game of profound complexity.
- There are an astonishing **10 to the power of 170 possible board configurations** - more than the number of atoms in the known universe

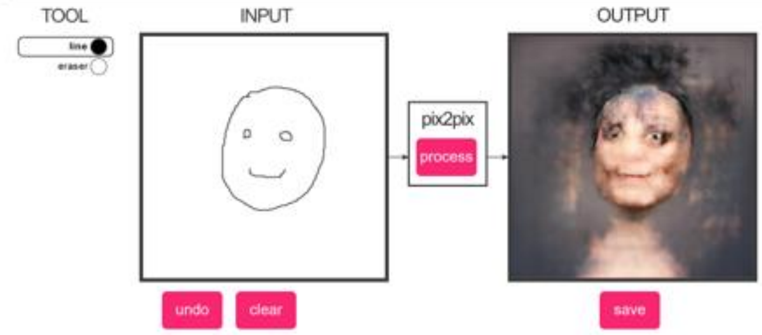
https://deepmind.com/research/alphago/

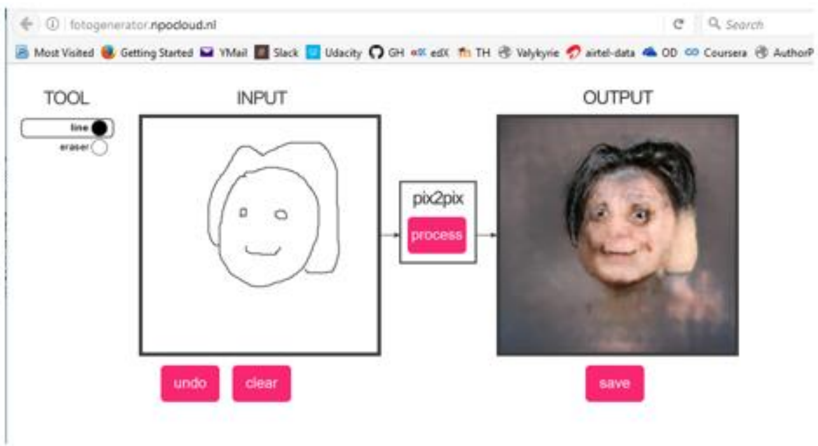
AlphaGo

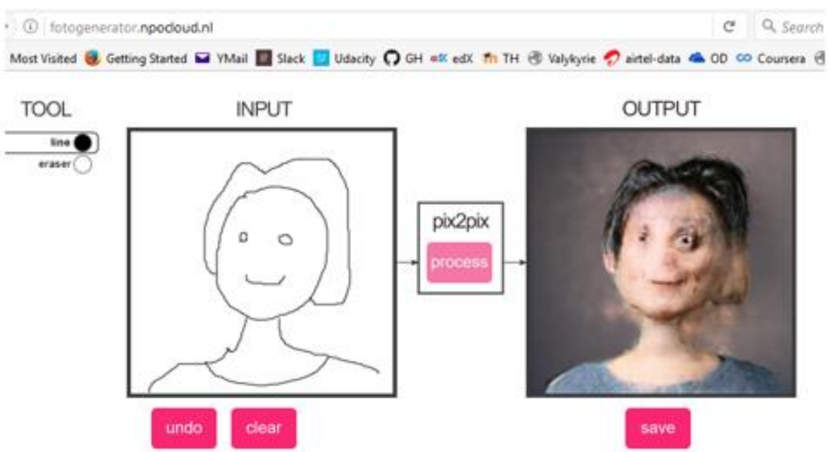


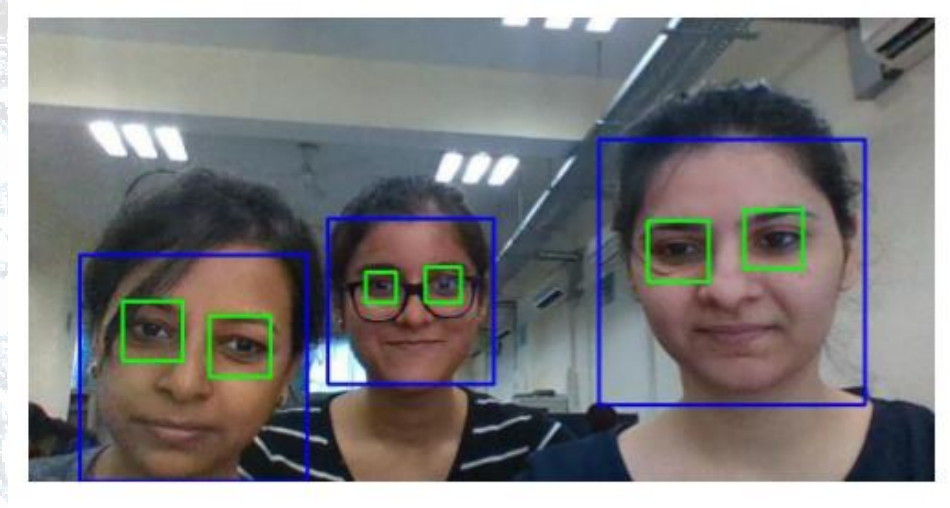
https://deepmind.com/research/alphago/

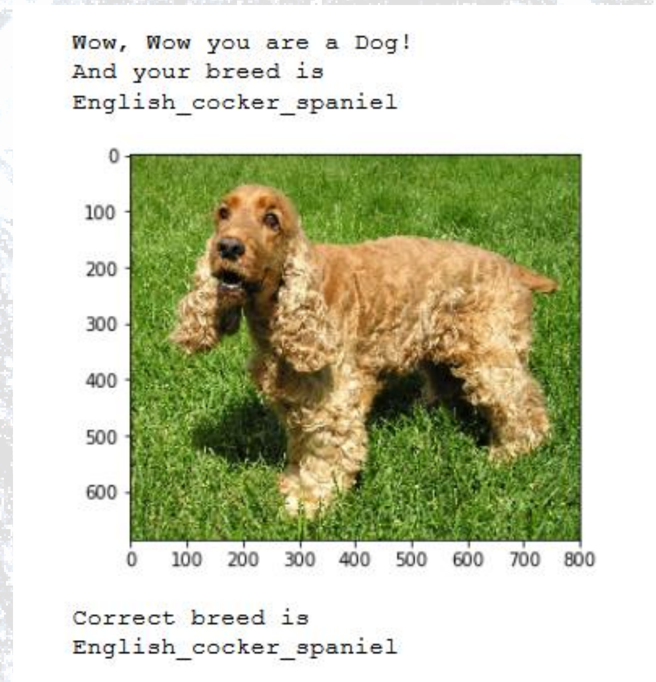

5:39
Google Translate
German
English
GERMAN
Probleme kann man niemals mit derselben Denkweise lösen, durch die sie entstanden sind.
ENGLISH
No problem can be solved from the same consciousness that they have arisen.
OLD TRANSLATION
5:39
Google Translate
German
English
GERMAN
Probleme kann man niemals mit derselben Denkweise lösen, durch die sie entstanden sind.
ENGLISH
Problems can never be solved with the same way of thinking that caused them.
NEW TRANSLATION

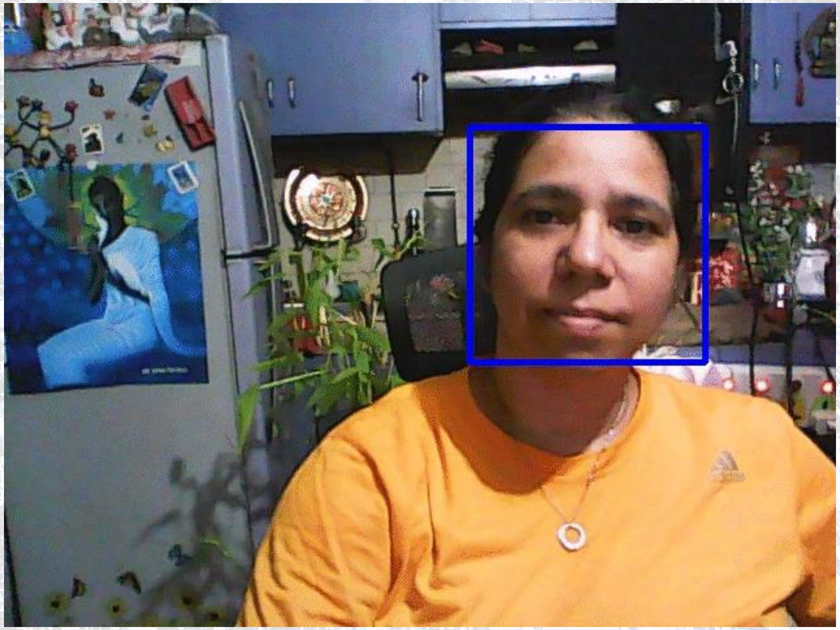

TOOL
line
eraser
INPUT
pix2pix
process
OUTPUT
undo
clear
save

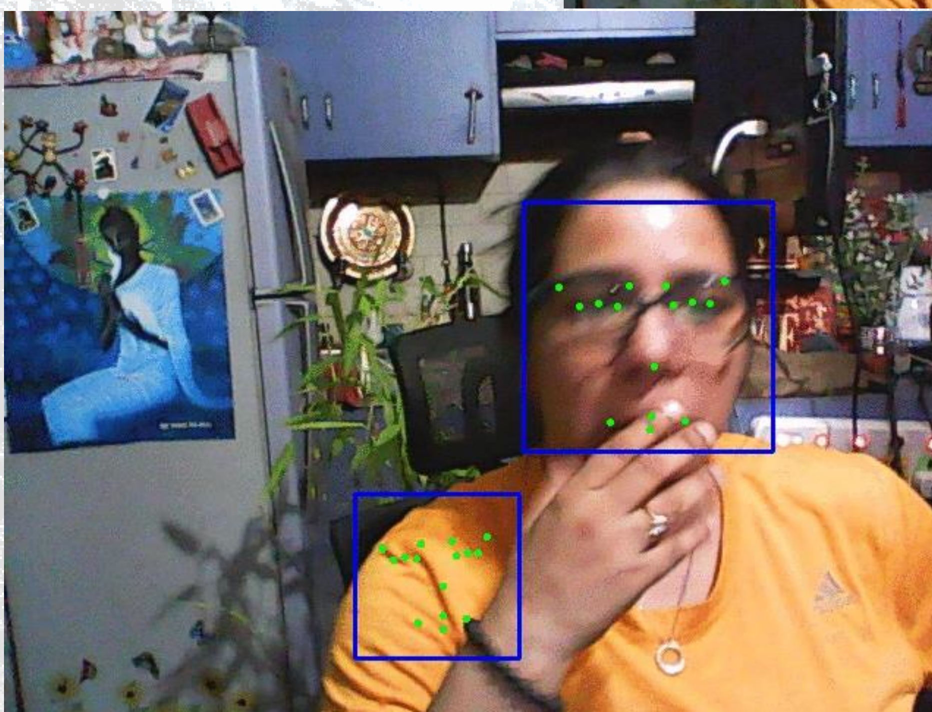

fotogenerator.npocloud.nl
Search
Most Visited Getting Started YMail Slack Udacity GH edX TH Valykyrie airtel-data OD Coursera AuthorP
TOOL
line
eraser
INPUT
pix2pix
process
OUTPUT
undo
clear
save

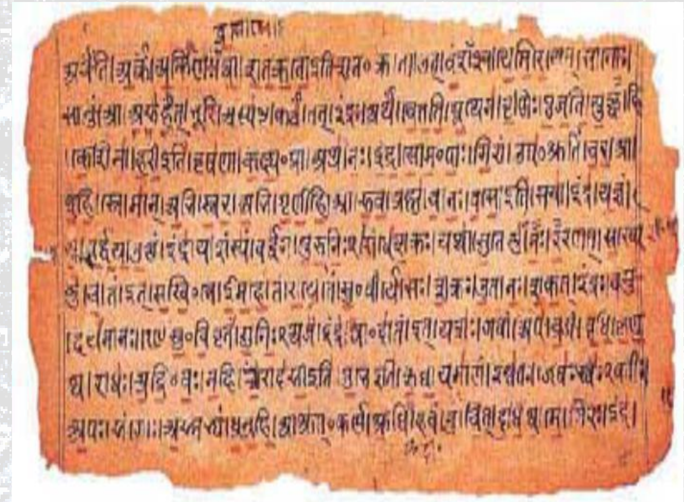

fotogenerator.npocloud.nl
Search
Most Visited Getting Started YMail Slack Udacity GH edX TH Valykyrie airtel-data OD Coursera
TOOL
line
eraser
INPUT
pix2pix
process
OUTPUT
undo
clear
save

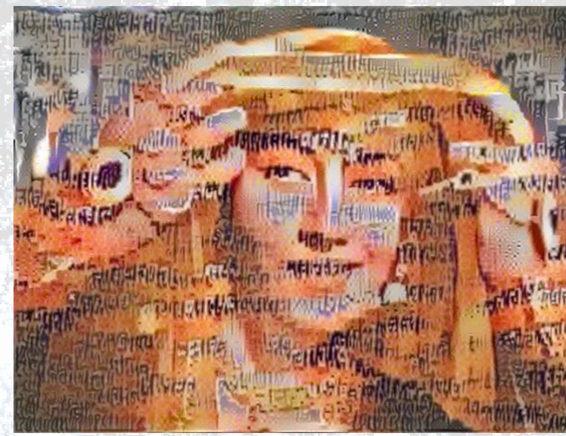


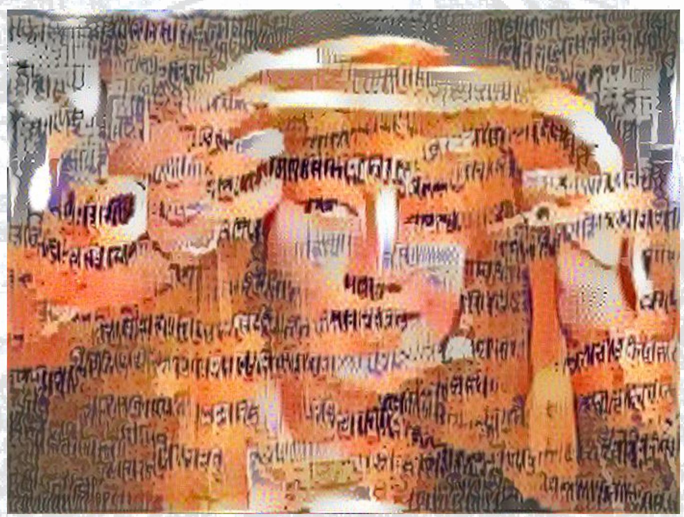

Wow, Wow you are a Dog!
And your breed is
English_cocker_spaniel
0
100
200
300
400
500
600
0
100
200
300
400
500
600
700
800
Correct breed is
English_cocker_spaniel











Biological Inspiration

To make the computers more robust and intelligent.

We take inspiration from the intelligent machine ever made

Human Brain

https://www.dreamstime.com/royalty-free-stock-images-human-brain-close-up-active-neurons-image18466049#

Features of the Brain

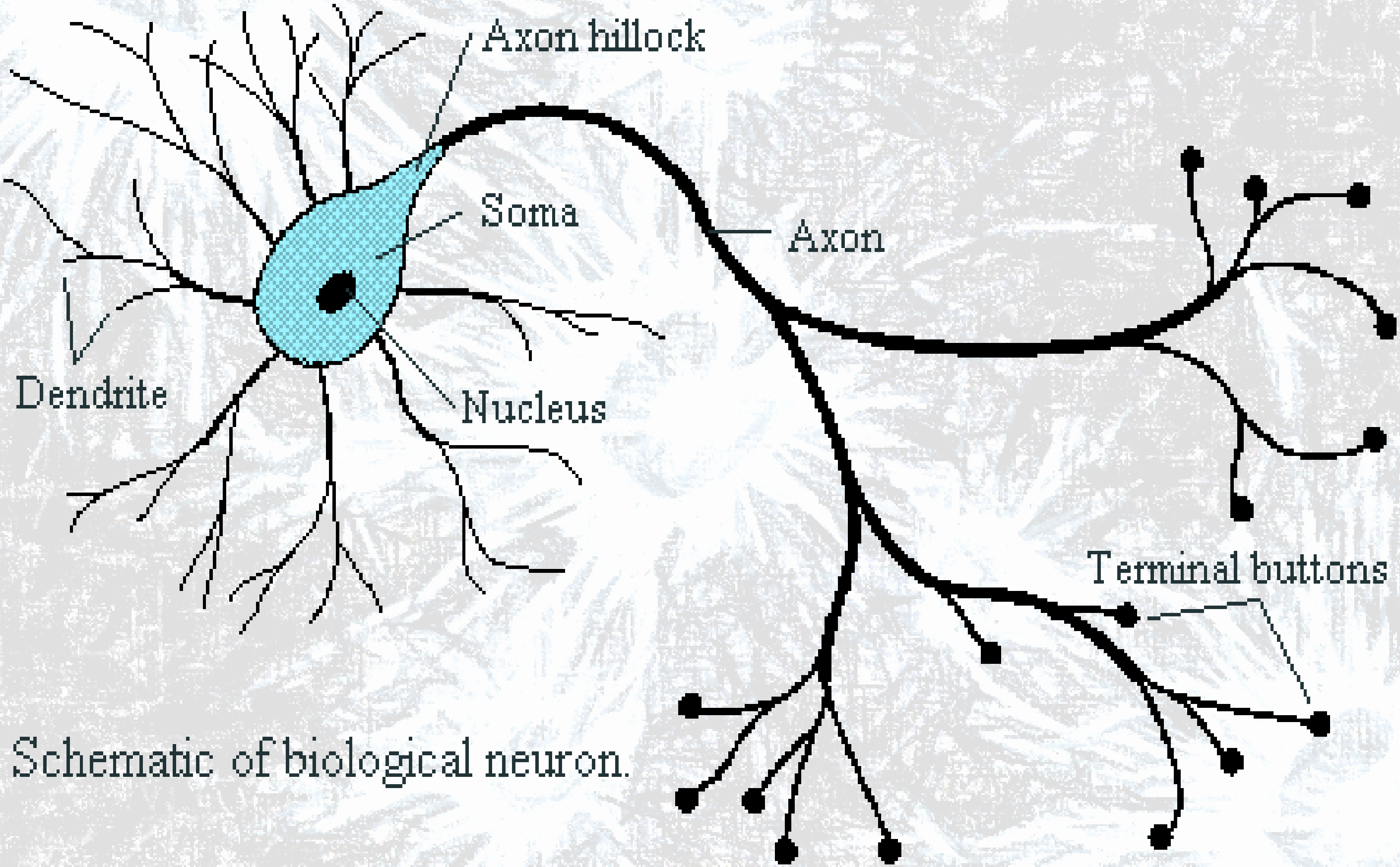
- Ten billion (10^{10}) neurons
- Neuron switching time $>10^{-3}$secs
- Face Recognition ~0.1secs
- On average, each neuron has several thousand connections
- Hundreds of operations per second
- High degree of parallel computation
- Distributed representations
- Compensated for problems by massive parallelism
- Graceful Degradation and Robust

How do we do it?

- The brain is a collection of about **10 billion** interconnected neurons.
- Each neuron is a cell that uses biochemical reactions to **receive**, **process** and **transmit** information.

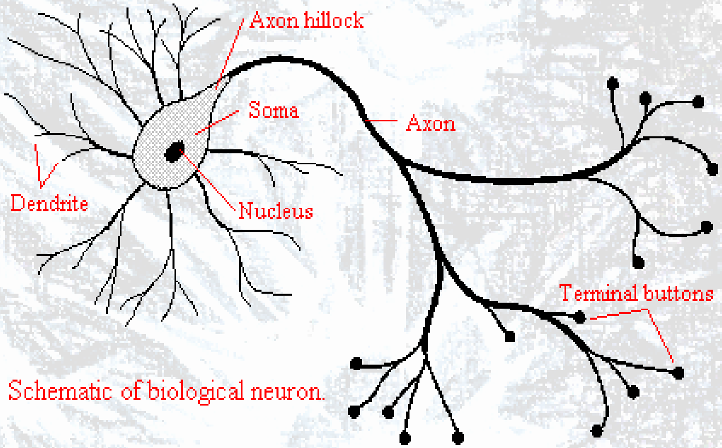
https://giphy.com/gifs/uofcalifornia-brain-neuroscience-neurons-xT0BKr4MvHdohFTe6s/download

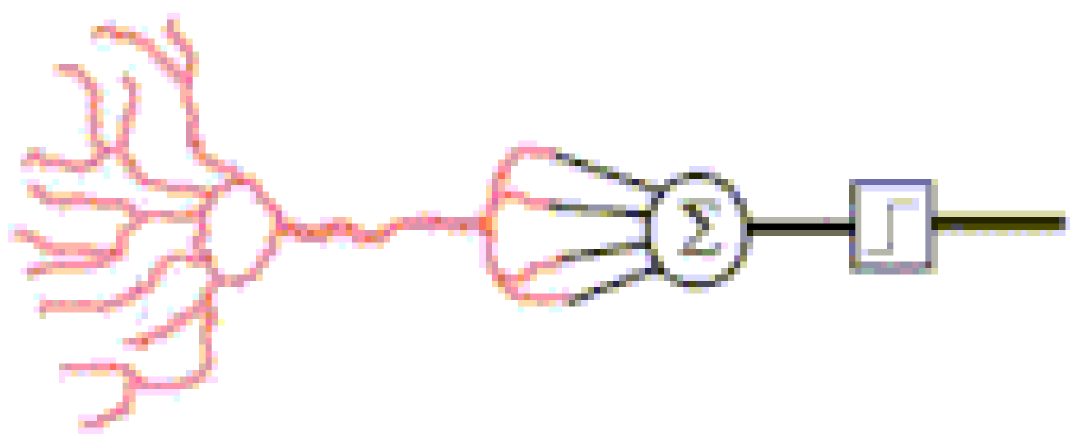
How do we do it?



http://serendip.brynmawr.edu/exchange/brains/definition/def-neuron

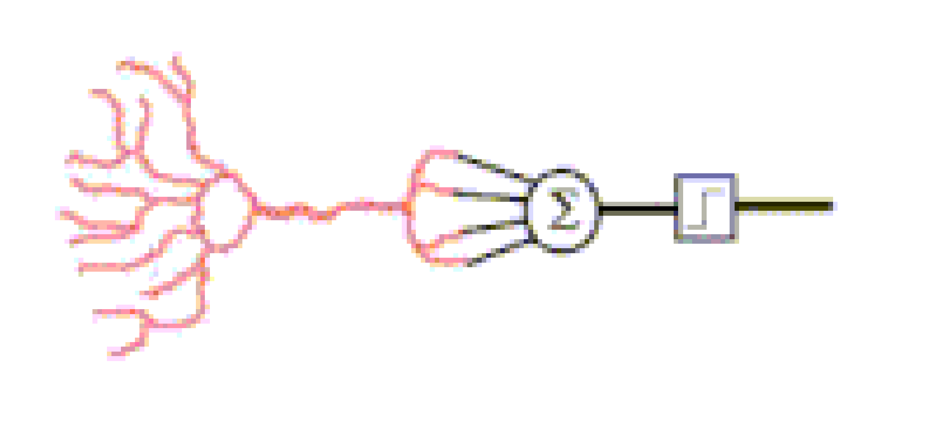
Can we make an Artificial Neuron?





Artificial Neuron

- Inputs I
- Weights W
- Activity $h = I_1 W_1 + I_2 W_2 + \ldots + I_N W_N - \theta = \sum_{i=1}^{N+1} I_i W_i$
- Activation function $\Theta(h) = \Theta\left(\sum_{i=1}^{N} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$



What can this single Neuron Do?

Can it think logically?

Logic?

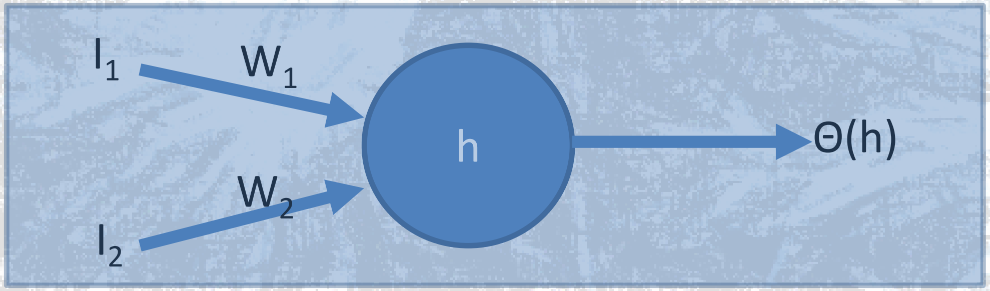
- Should I go to college?
  - If it is raining **OR** Ranbir Kapoor Movie Released?
  - If there is an important lecture **AND** I have done the assignment?
  - If my best friend is **NOT** coming

AND **OR** **NOT**

AND Logic

W_1=1, W_2 = 1, θ=2	$\Theta(h) = \Theta\left(\sum_{i=1}^{N} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$

I_1	I_2	$h=I_1W_1 + I_2W_2-\vartheta$	Θ(h)
1	1		
1	0		
0	1		
0	0		

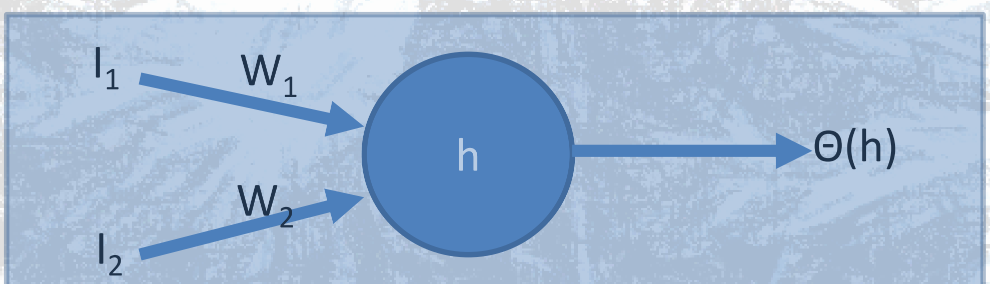


AND Logic

W_1=1, W_2 = 1, θ=2

$$\boldsymbol{\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}}$$

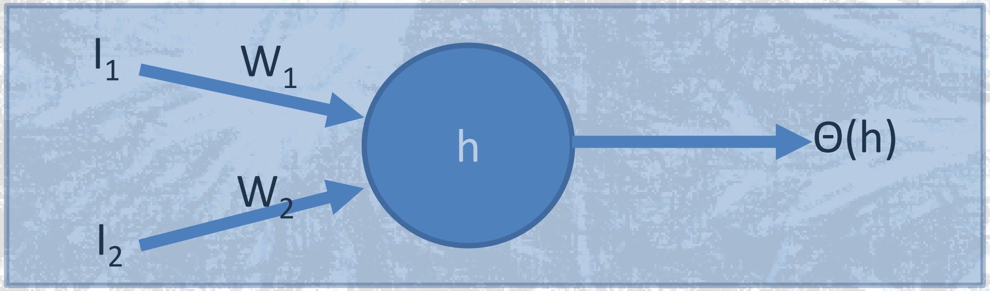
I_1	I_2	$h=I_1W_1 + I_2W_2-\vartheta$	Θ(h)
1	1	1 + 1 - 2 = **0**	
1	0		
0	1		
0	0		



AND Logic

$W_1=1, W_2 = 1, \theta=2$	$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$

I_1	I_2	$h=I_1W_1 + I_2W_2-\vartheta$	Θ(h)
1	1	1 + 1 - 2 = **0**	1
1	0		
0	1		
0	0		

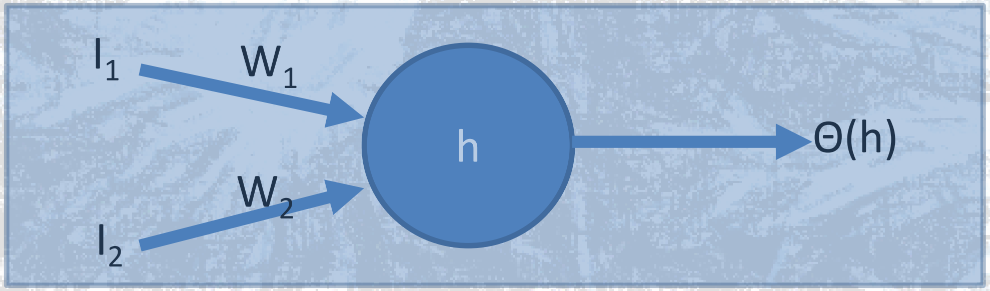


AND Logic

W_1=1, W_2 = 1, θ=2

$$\boldsymbol{\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}}$$

I_1	I_2	$h=I_1W_1 + I_2W_2-\vartheta$	Θ(h)
1	1	1 + 1 - 2 = **0**	1
1	0	1 + 0 – 2 = **-1**	
0	1		
0	0		

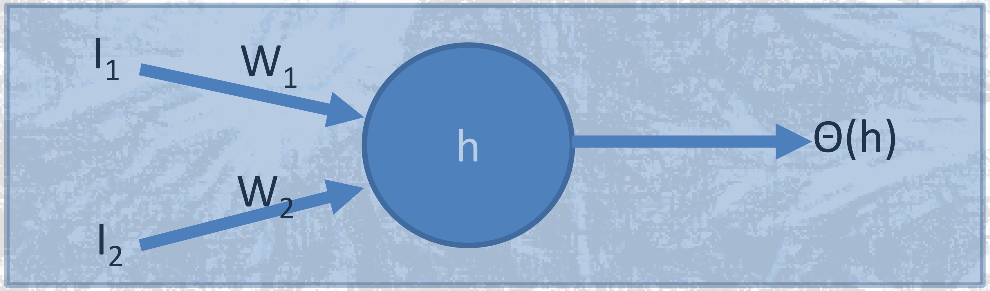


AND Logic

W_1=1, W_2 = 1, θ=2

$$\boldsymbol{\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}}$$

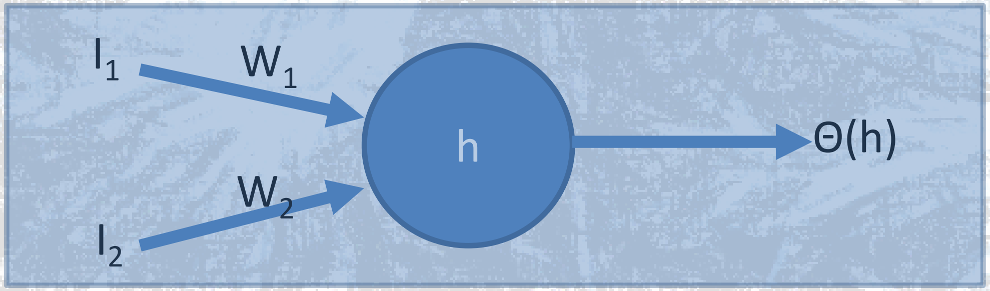
I_1	I_2	$h=I_1W_1 + I_2W_2 - \vartheta$	Θ(h)
1	1	1 + 1 - 2 = **0**	1
1	0	1 + 0 – 2 = **-1**	0
0	1		
0	0		



AND Logic

W_1=1, W_2 = 1, θ=2	$$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$$

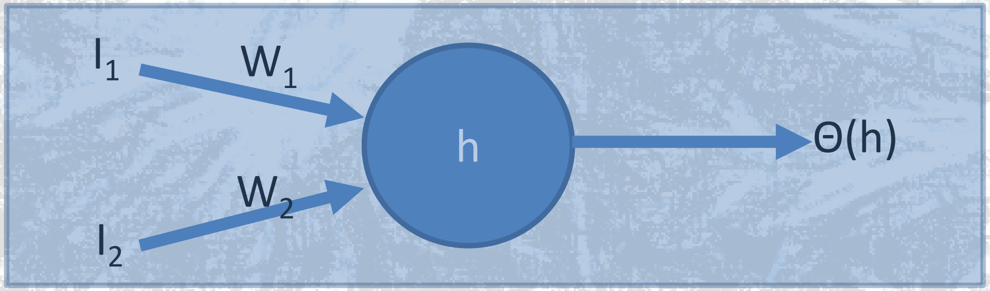
I_1	I_2	$h=I_1W_1 + I_2W_2-\vartheta$	Θ(h)
1	1	1 + 1 - 2 = **0**	1
1	0	1 + 0 – 2 = **-1**	0
0	1	0 + 1 – 2 = **-1**	
0	0		



AND Logic

$W_1=1, W_2 = 1, \theta=2$	$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$

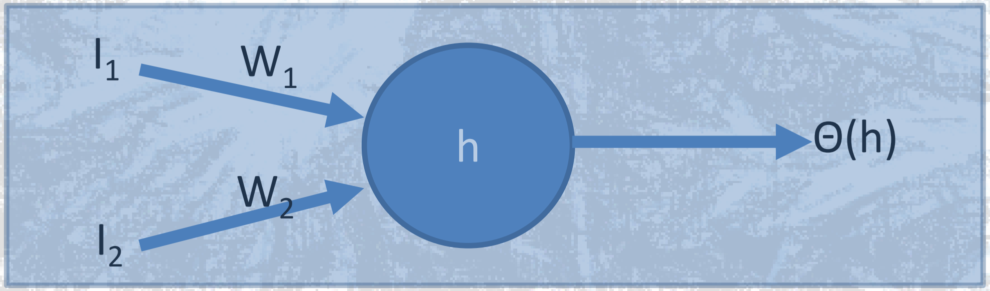
I_1	I_2	$h=I_1W_1 + I_2W_2-\vartheta$	Θ(h)
1	1	1 + 1 - 2 = **0**	1
1	0	1 + 0 – 2 = **-1**	0
0	1	0 + 1 – 2 = **-1**	0
0	0		



AND Logic

$W_1=1, W_2 = 1, \theta=2$	$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$

I_1	I_2	$h=I_1W_1 + I_2W_2-\vartheta$	Θ(h)
1	1	1 + 1 - 2 = **0**	1
1	0	1 + 0 – 2 = **-1**	0
0	1	0 + 1 – 2 = **-1**	0
0	0	0 + 0 -2 = **-2**	

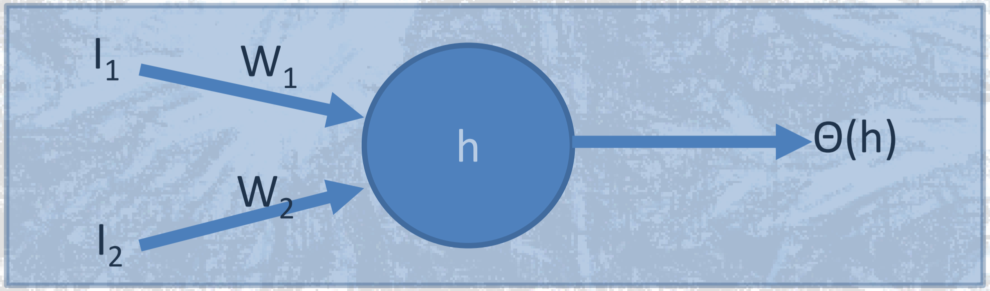


AND Logic

W_1=1, W_2 = 1, θ=2

$$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$$

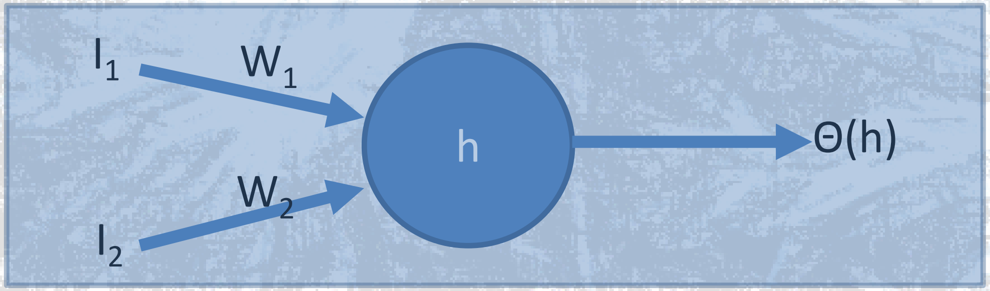
I_1	I_2	$h=I_1W_1 + I_2W_2-\vartheta$	Θ(h)
1	1	1 + 1 - 2 = **0**	1
1	0	1 + 0 – 2 = **-1**	0
0	1	0 + 1 – 2 = **-1**	0
0	0	0 + 0 -2 = **-2**	0



OR Logic

$W_1=1, W_2 = 1, \theta=1$	$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & \text{if } h \geq 0 \\ 0 & \text{otherwise} \end{matrix}$

I_1	I_2	$h=I_1W_1 + I_2W_2 - \vartheta$	$\Theta(h)$
1	1		
1	0		
0	1		
0	0		

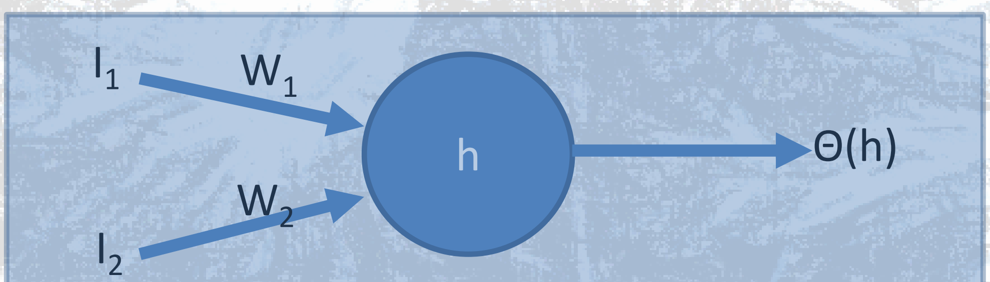


OR Logic

W_1=1, W_2 = 1, θ=1

$$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$$

I_1	I_2	$h=I_1W_1 + I_2W_2 -\vartheta$	Θ(h)
1	1	1 + 1 - 1 = **1**	
1	0		
0	1		
0	0		

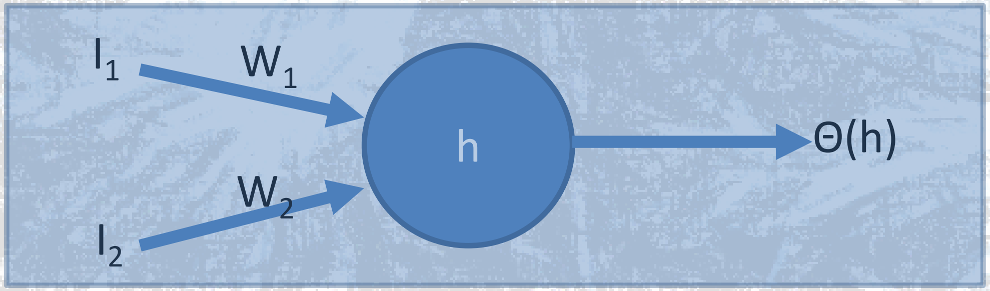


OR Logic

W_1=1, W_2 = 1, θ=1

$$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$$

I_1	I_2	$h=I_1W_1 + I_2W_2 - \vartheta$	Θ(h)
1	1	1 + 1 - 1 = **1**	1
1	0		
0	1		
0	0		

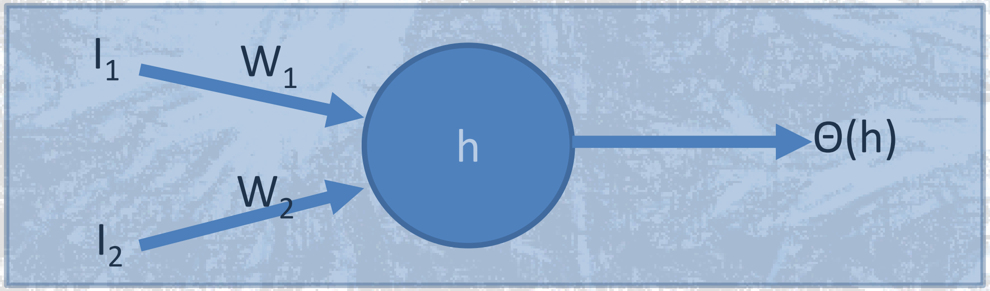


OR Logic

W_1=1, W_2 = 1, θ=1

$$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$$

I_1	I_2	$h=I_1W_1 + I_2W_2 - \vartheta$	Θ(h)
1	1	1 + 1 - 1 = **1**	1
1	0	1 + 0 – 1 = **0**	
0	1		
0	0		

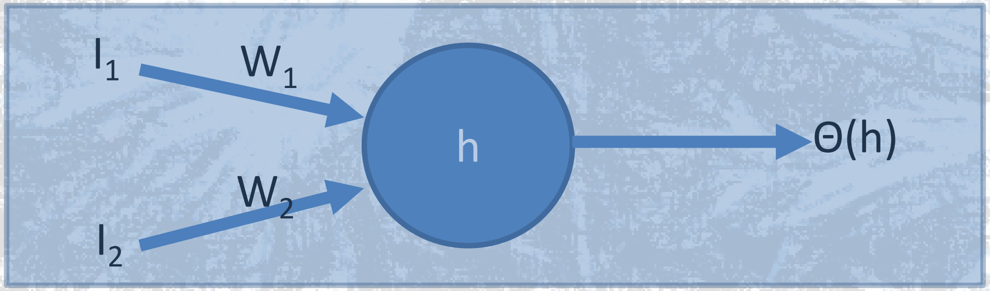


OR Logic

W_1=1, W_2 = 1, θ=1

$$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$$

I_1	I_2	$h=I_1W_1 + I_2W_2 - \vartheta$	Θ(h)
1	1	1 + 1 - 1 = **1**	1
1	0	1 + 0 – 1 = **0**	1
0	1		
0	0		

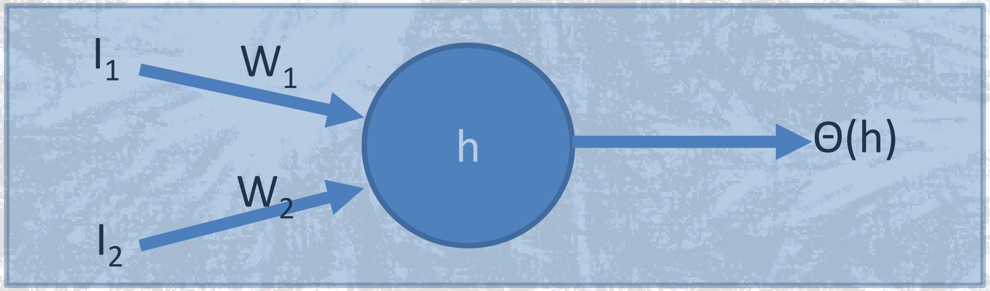


OR Logic

W_1=1, W_2 = 1, θ=1

$$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$$

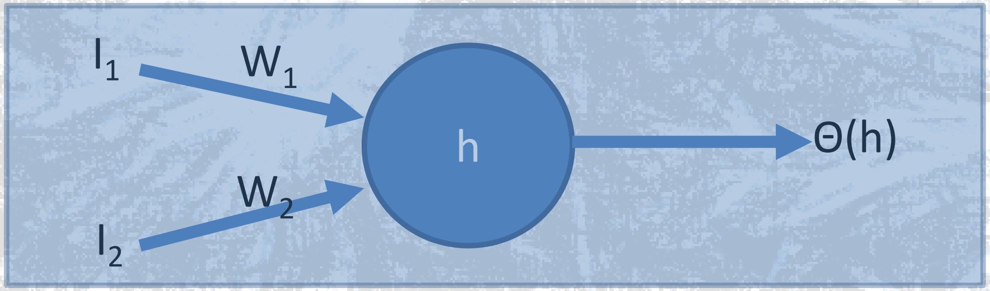
I_1	I_2	$h=I_1W_1 + I_2W_2 - \vartheta$	Θ(h)
1	1	1 + 1 - 1 = **1**	1
1	0	1 + 0 – 1 = **0**	1
0	1	0 + 1 – 1 = **0**	
0	0		



OR Logic

$W_1=1, W_2 = 1, \theta=1$	$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$

I_1	I_2	$h=I_1W_1 + I_2W_2 - \vartheta$	Θ(h)
1	1	1 + 1 - 1 = **1**	1
1	0	1 + 0 – 1 = **0**	1
0	1	0 + 1 – 1 = **0**	1
0	0		

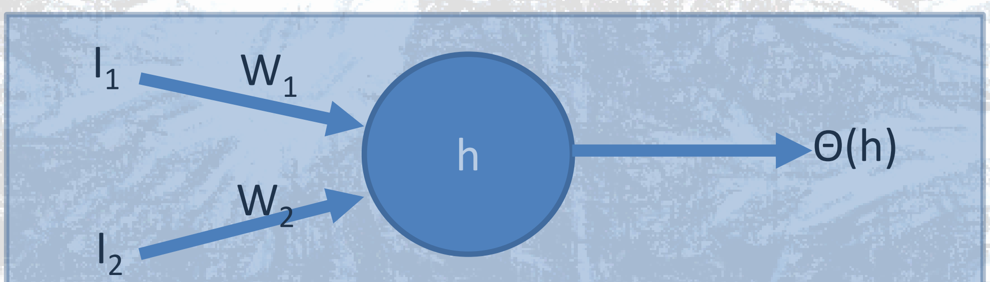


OR Logic

W_1=1, W_2 = 1, θ=1

$$\boldsymbol{\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}}$$

I_1	I_2	$h=I_1W_1 + I_2W_2 - \vartheta$	Θ(h)
1	1	1 + 1 - 1 = **1**	1
1	0	1 + 0 – 1 = **0**	1
0	1	0 + 1 – 1 = **0**	1
0	0	0 + 0 -1 = **-1**	

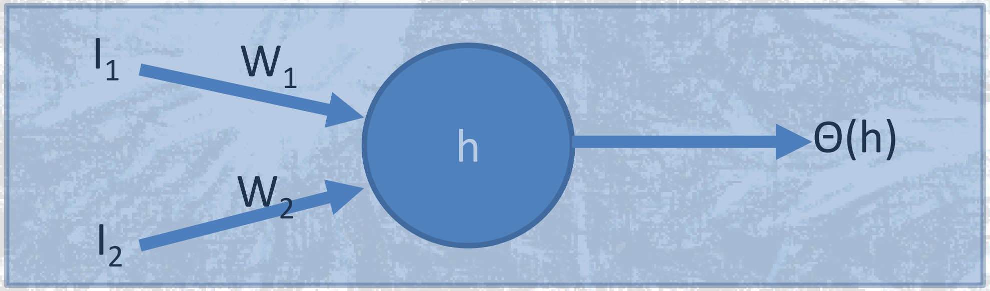


OR Logic

W_1=1, W_2 = 1, θ=1

$$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$$

I_1	I_2	$h=I_1W_1 + I_2W_2 - \vartheta$	Θ(h)
1	1	1 + 1 - 1 = **1**	1
1	0	1 + 0 – 1 = **0**	1
0	1	0 + 1 – 1 = **0**	1
0	0	0 + 0 -1 = **-1**	0

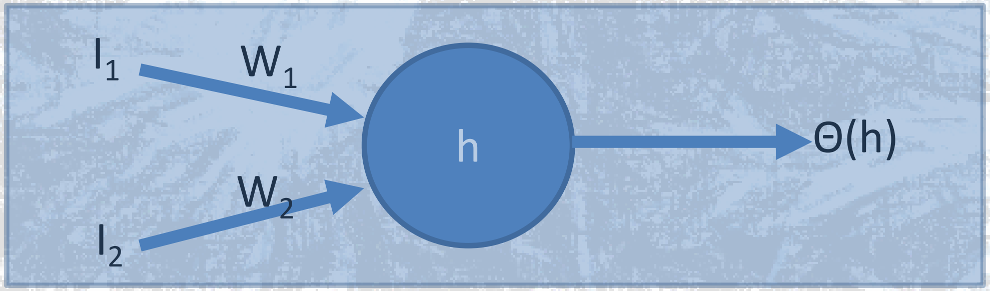


OR Logic

W_1=1, W_2 = 1, θ=1

$$\boldsymbol{\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}}$$

I_1	I_2	$h=I_1W_1 + I_2W_2 - \vartheta$	Θ(h)
1	1	1 + 1 - 1 = **1**	1
1	0	1 + 0 – 1 = **0**	1
0	1	0 + 1 – 1 = **0**	1
0	0	0 + 0 -1 = **-1**	0

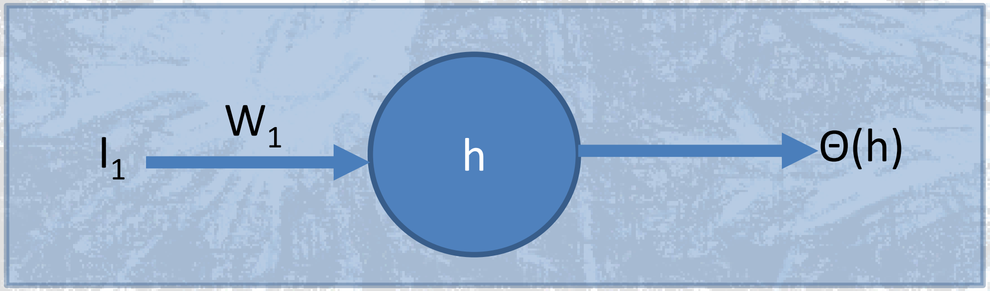


NOT Logic

W_1=-1, θ=0

$$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$$

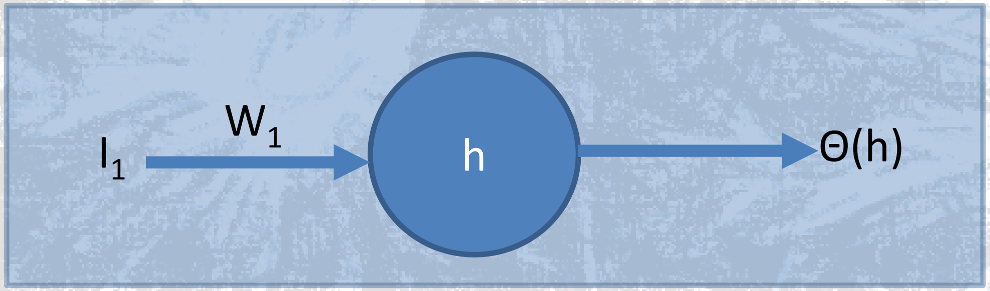
I_1	$h=I_1W_1 - \vartheta$	Θ(h)
1	-1 - 0 = -**1**	
0		



NOT Logic

W_1=-1, θ=0	$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$

I_1	$h=I_1W_1 - \vartheta$	Θ(h)
1	-1 - 0 = -**1**	0
0		

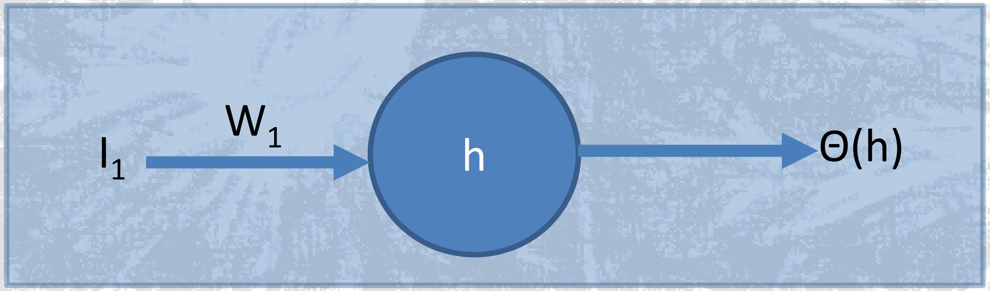


NOT Logic

W_1=-1, θ=0

$$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$$

I_1	$h=I_1W_1 - \vartheta$	Θ(h)
1	-1 - 0 = **-1**	0
0	0 - 0 = **0**	

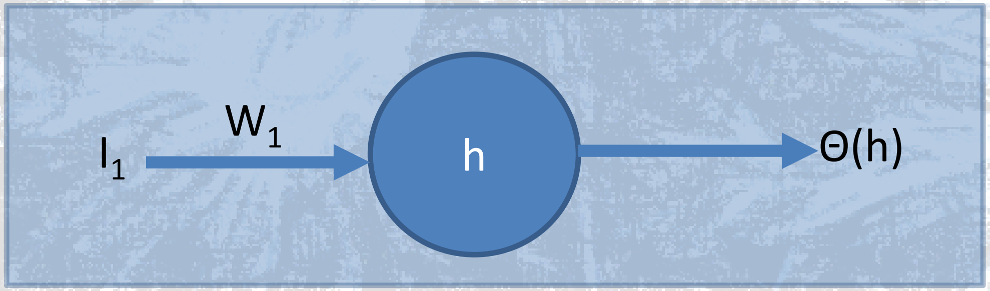


NOT Logic

W_1=-1, θ=0

$$\boldsymbol{\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}}$$

I_1	$h=I_1W_1 - \vartheta$	Θ(h)
1	-1 - 0 = **-1**	0
0	0 - 0 = **0**	1

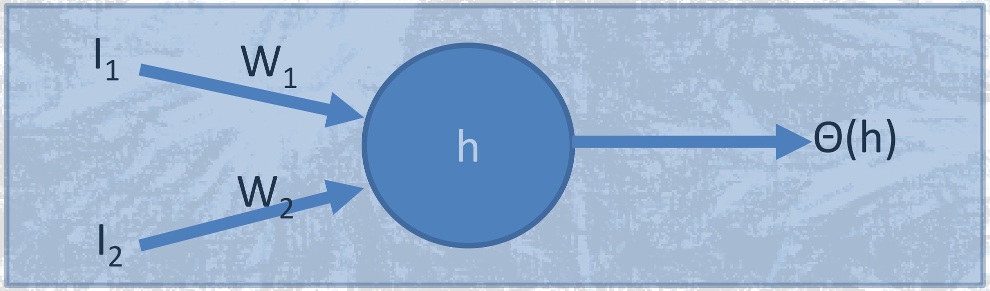


NOT AND(NAND)Logic

W_1= , W_2 = , θ=

$$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$$

I_1	I_2	$h=I_1W_1 + I_2W_2-\vartheta$	Θ(h)
1	1		0
1	0		1
0	1		1
0	0		1

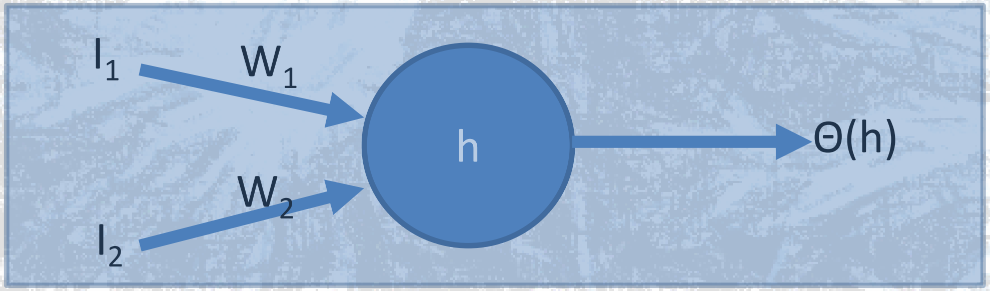


NOT AND(NAND)Logic

W_1= -1 , W_2 = -1 , θ= -1.5

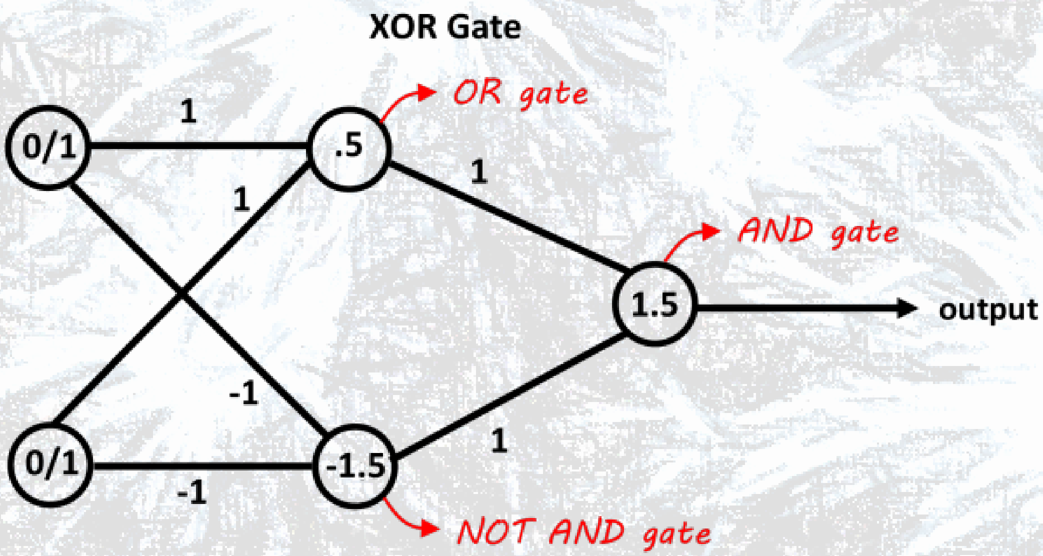
$$\Theta(h) = \Theta\left(\sum_{i=1}^{N+1} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$$

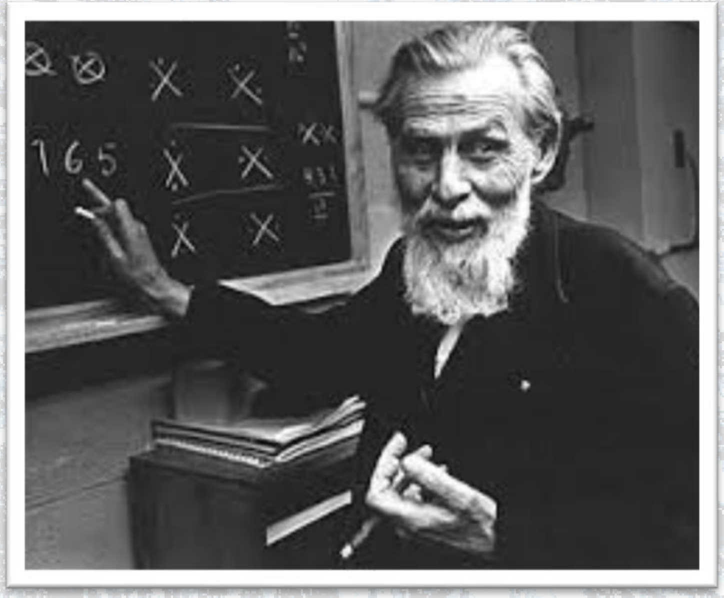
I_1	I_2	$h=I_1W_1 + I_2W_2-\vartheta$	Θ(h)
1	1	**-1 -1 +1.5 = -0.5**	0
1	0	**-1+0+1.5= 0.5**	1
0	1	**0-1+1.5=0.5**	1
0	0	**0+0+1.5=1.5**	1



XOR Logic

I_1	I_2	$h=I_1W_1 + I_2W_2 - \vartheta$	$\Theta(h)$
1	1	?	0
1	0	?	1
0	1	?	1
0	0	?	0





Warren McCulloch

Walter Pitts

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 5, 1943

A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY

WARREN S. MCCULLOCH AND WALTER PITTS

FROM THE UNIVERSITY OF ILLINOIS, COLLEGE OF MEDICINE, DEPARTMENT OF PSYCHIATRY AT THE ILLINOIS NEUROPSYCHIATRIC INSTITUTE, AND THE UNIVERSITY OF CHICAGO

Because of the "all-or-none" character of nervous activity, neural events and the relations among them can be treated by means of propositional logic. It is found that the behavior of every net can be described in these terms, with the addition of more complicated logical means for nets containing circles; and that for any logical expression satisfying certain conditions, one can find a net behaving in the fashion it describes. It is shown that many particular choices among possible neurophysiological assumptions are equivalent, in the sense that for every net behaving under one assumption, there exists another net which behaves under the other and gives the same results, although perhaps not in the same time. Various applications of the calculus are discussed.

I. *Introduction*

Theoretical neurophysiology rests on certain cardinal assumptions. The nervous system is a net of neurons, each having a soma and an axon. Their adjunctions, or synapses, are always between the axon of one neuron and the soma of another. At any instant a neuron has some threshold, which excitation must exceed to initiate an im-

McCulloch Pitts Model

https://link.springer.com/article/10.1007%2FBF02478259
https://chatbotslife.com/keras-in-a-single-mcculloch-pitts-neuron-317397cccd45
http://nautil.us/issue/21/information/the-man-who-tried-to-redeem-the-world-with-logic

Can it Learn?

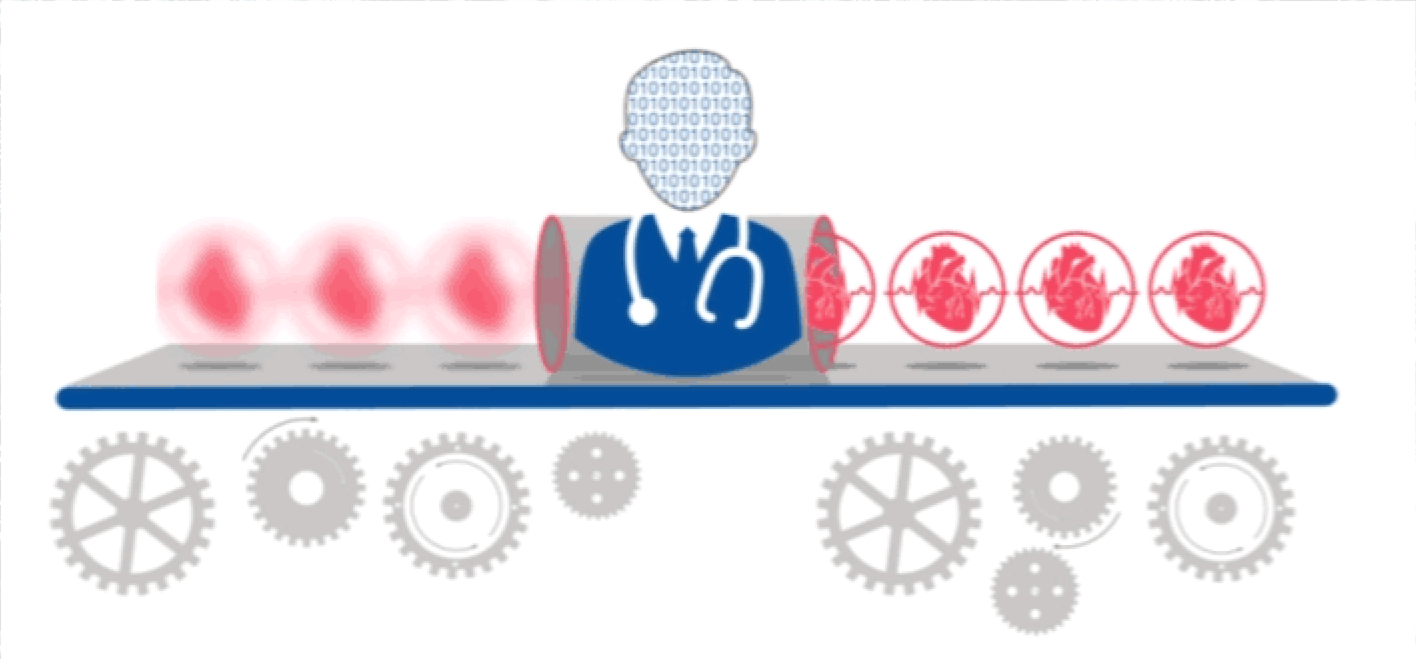
- How does human learn?
  - Parents and teachers teach us
  - We read books and introspect what they mean
  - Motivated by a goal: Good job, praise from peers, desire to excel=> Force us to learn.

Can it Learn?

- How does human learn?
  - Parents and teachers teach us:
    - **Supervised Learning**
  - We read books and introspect what they mean
    - **Unsupervised Learning**
  - Motivated by a goal: Good job, praise from peers, desire to excel=> Force us to learn
    - **Reinforcement Learning**

The most understood learning

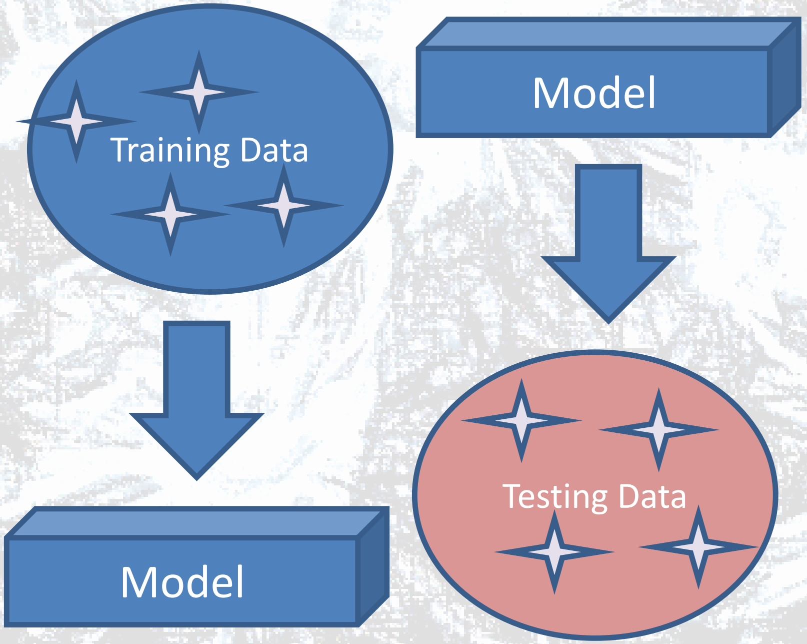
Supervised Learning



The most understood learning

Supervised Learning

- Input, Output pair is given as the training data (x, y).
- With large data we divide it into **Training**, **Testing** and **Cross Validation**



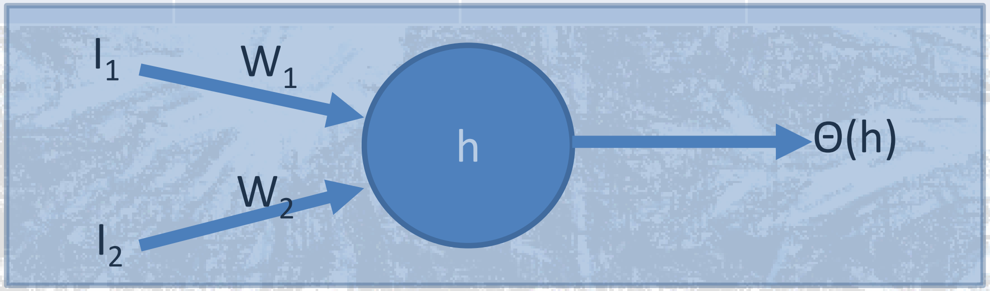
Learn AND Logic?

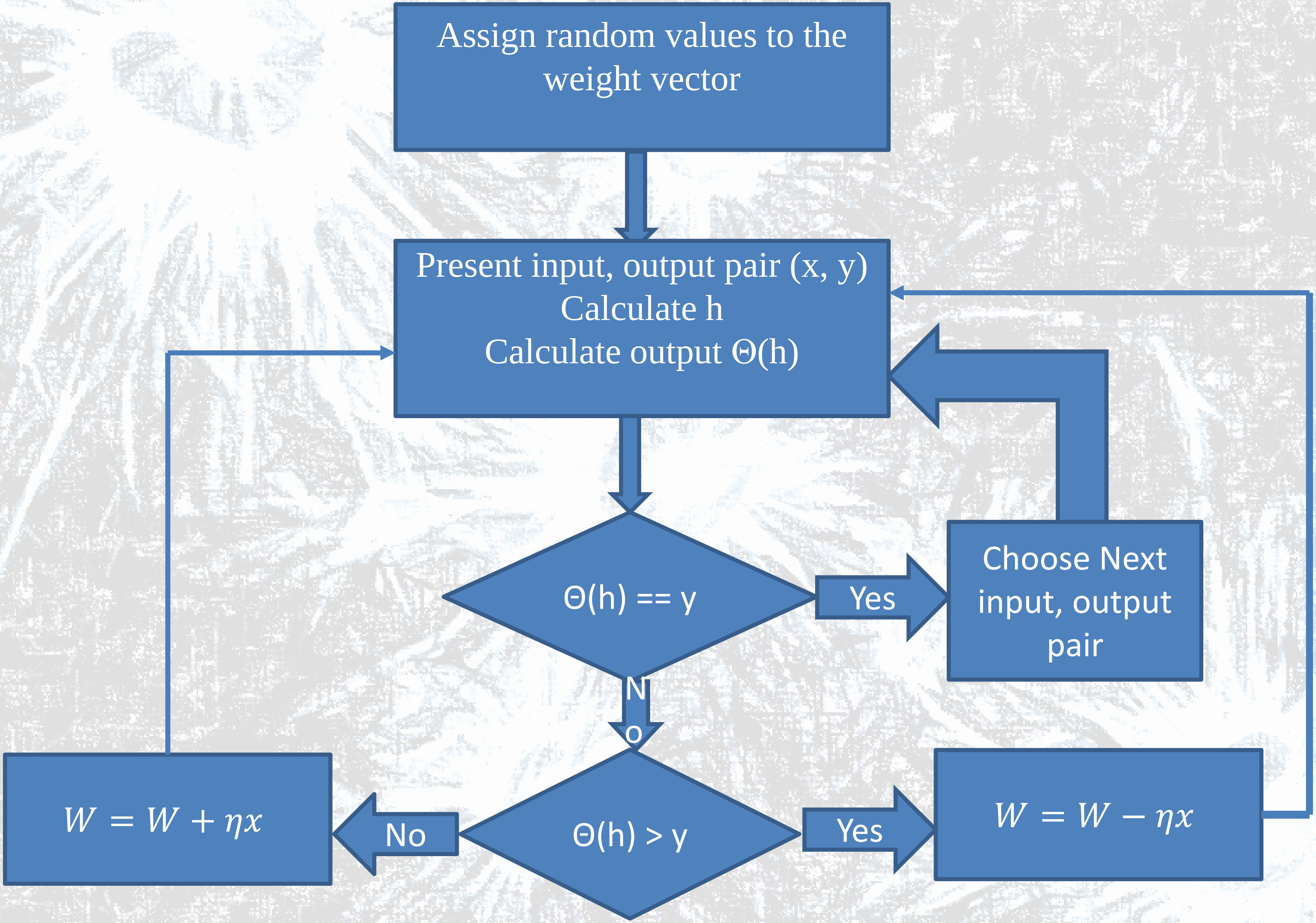
$\theta = W_3 I_3$

W_1=0, W_2 = 0, W_3=0 ; I_3 = -1

$$\boldsymbol{\Theta(h) = \Theta\left(\sum_{i=1}^{N} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}}$$

I_1	I_2	$h=I_1W_1 + I_2W_2+I_3W_3$	Θ(h)	Desired	Weights
1	1	0 + 0 - 0 = **0**	1	1	Unchanged
1	0	0 + 0 - 0 = **0**	1	0	Decrease
0	1	0 + 0 - 0 = **0**	1	0	Decrease
0	0	0 + 0 - 0 = **0**	1	0	Decrease



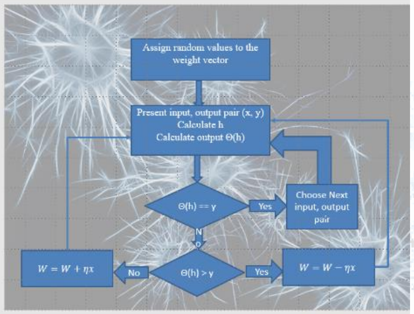


Assign random values to the weight vector
Present input, output pair (x, y)
Calculate h
Calculate output Θ(h)
Θ(h) == y
Yes
Choose Next input, output pair
No
Θ(h) > y
No
$W = W + \eta x$
Yes
$W = W - \eta x$

Learn AND Logic?

W_1=0, W_2 = 0, W_3=0 ; I_3 = -1	$$\Theta(h) = \Theta\left(\sum_{i=1}^{N} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$$

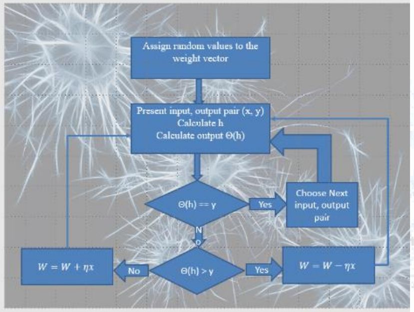
I_1	I_2	$h=I_1W_1 + I_2W_2+I_3W_3$	Θ(h)	Desired	Weights	New Weights
1	1	0 + 0 - 0 = **0**	1	1	Unchanged	**W_1=0, W_2 = 0, W_3=0**



Learn AND Logic?

W_1=0, W_2 = 0, W_3=0 ; I_3 = -1	$$\Theta(h) = \Theta\left(\sum_{i=1}^{N} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$$

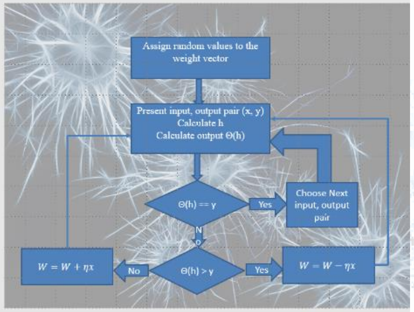
I_1	I_2	$h=I_1W_1+I_2W_2+I_3W_3$	Θ(h)	Desired	Weights	New Weights
1	1	0 + 0 - 0 = **0**	1	1	Unchanged	**W_1=0, W_2 = 0, W_3=0**
1	0	0 + 0 - 0 = **0**	1	0	Decrease	W1 = W1-I1 = -1 W2= W2-I2 = 0 W3 = W3-I3= +1



Learn AND Logic?

W_1=-1, W_2 = 0, W_3=1 ; I_3 = -1	$$\Theta(h) = \Theta\left(\sum_{i=1}^{N} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}$$

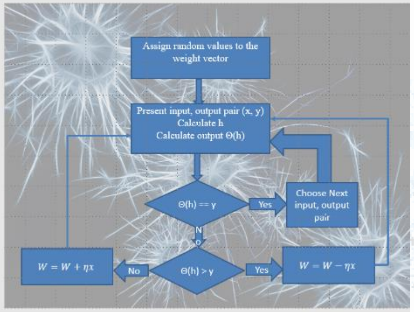
I_1	I_2	$h=I_1W_1 + I_2W_2+I_3W_3$	Θ(h)	Desired	Weights	New Weights
1	1	0 + 0 - 0 = **0**	1	1	Unchanged	**W_1=0, W_2 = 0, W_3=0**
1	0	0 + 0 - 0 = **0**	1	0	Decrease	W1 = W1-I1 = -1 W2= W2-I2 = 0 W3 = W3-I3= +1
1	0					



Learn AND Logic?

W_1=-1, W_2 = 0, W_3=1 ; I_3 = -1	$$\boldsymbol{\Theta(h) = \Theta\left(\sum_{i=1}^{N} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}}$$

I_1	I_2	$h=I_1W_1 + I_2W_2+I_3W_3$	Θ(h)	Desired	Weights	New Weights
1	1	0 + 0 - 0 = **0**	1	1	Unchanged	**W_1=0, W_2 = 0, W_3=0**
1	0	0 + 0 - 0 = **0**	1	0	Decrease	W1 = W1-I1 = -1 W2= W2-I2 = 0 W3 = W3-I3= +1
1	0	-1 + 0 - 1 = 0	1	1	Unchanged	**W_1=-1, W_2 = 0, W_3=1**

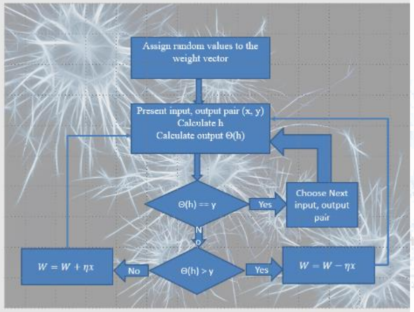


Learn AND Logic?

W_1=-1, W_2 = 0, W_3=1 ; I_3 = -1

$$\boldsymbol{\Theta(h) = \Theta\left(\sum_{i=1}^{N} I_i W_i\right) = \begin{matrix} 1 & if\ h \geq 0 \\ 0 & otherwise \end{matrix}}$$

I_1	I_2	$h=I_1W_1+I_2W_2+I_3W_3$	Θ(h)	Desired	Weights	New Weights
1	1	0 + 0 - 0 = **0**	1	1	Unchanged	W_1=0, W_2 = 0, W_3=0
1	0	0 + 0 - 0 = **0**	1	0	Decrease	W1 = W1-I1 = -1 W2= W2-I2 = 0 W3 = W3-I3= +1
1	0	-1 + 0 - 1 = -2	0	0	Unchanged	W_1=-1, W_2 = 0, W_3=1
0	1	0 + 0 – 1 = -1	0	0	Unchanged	W_1=-1, W_2 = 0, W_3=1



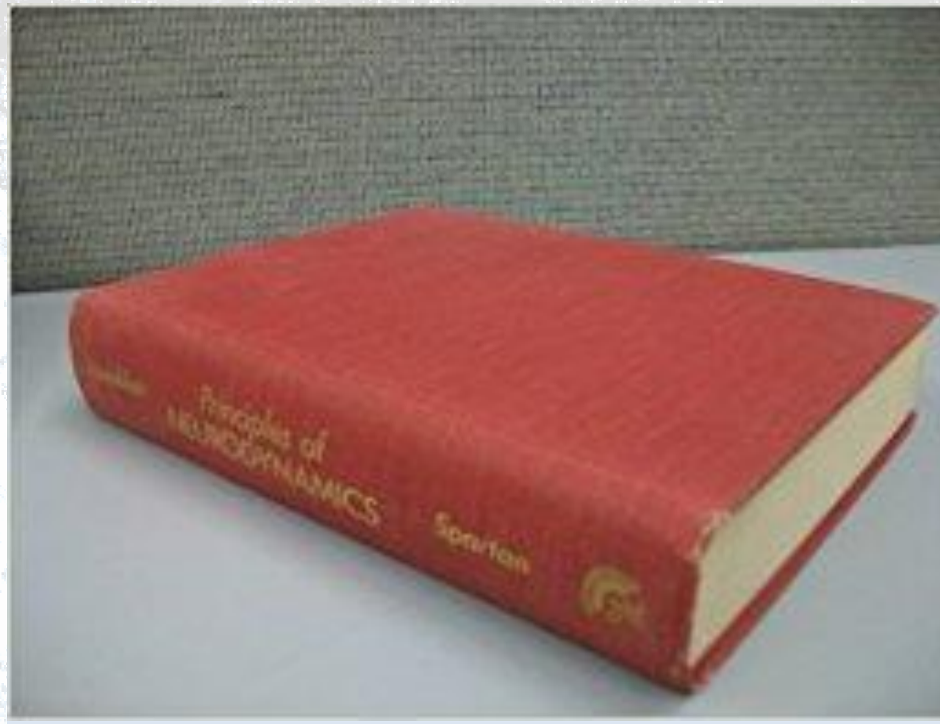
Learn AND Logic?

W_1=-1, W_2 = 0, W_3=1 ; I_3 = -1	$$\boldsymbol{\Theta(h) = \Theta\left(\sum_{i=1}^{N} I_i W_i\right) = \begin{matrix} 1 & \text{if } h \geq 0 \\ 0 & \text{otherwise} \end{matrix}}$$

I_1	I_2	$h=I_1W_1 + I_2W_2+I_3W_3$	Θ(h)	Desired	Weights	New Weights
1	1	0 + 0 - 0 = **0**	1	1	Unchanged	W_1=0, W_2 = 0, W_3=0
1	0	0 + 0 - 0 = **0**	1	0	Decrease	W1 = W1-I1 = -1 W2= W2-I2 = 0 W3 = W3-I3= +1
1	0	-1 + 0 - 1 = -2	0	0	Unchanged	W_1=-1, W_2 = 0, W_3=1
0	1	0 + 0 – 1 = -1	0	0	Unchanged	W_1=-1, W_2 = 0, W_3=1
0	0	0 + 0 – 1 = -1	0	0	Unchanged	W_1=-1, W_2 = 0, W_3=1

Frank Rosenblatt

Perceptron

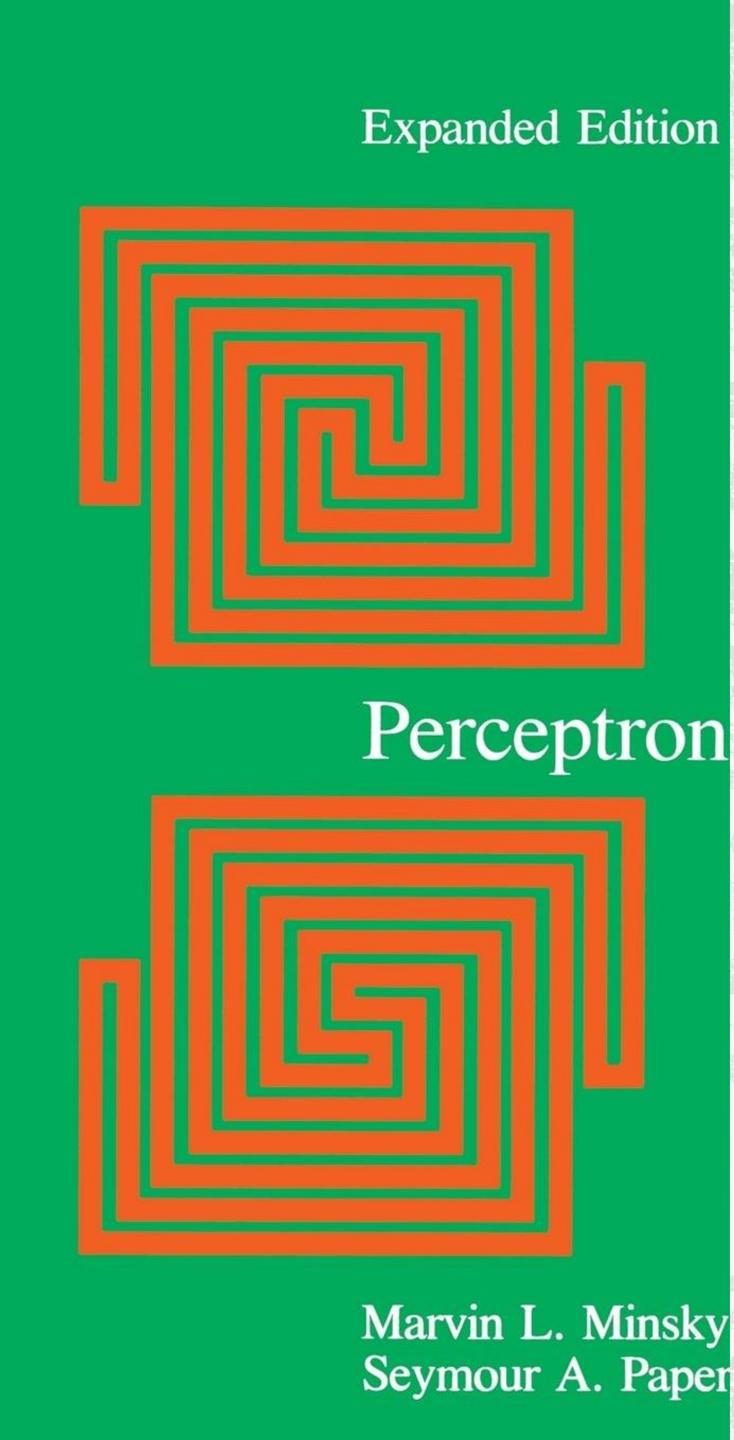


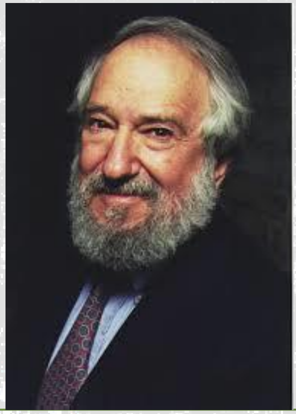
Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms

"[The perceptron] is the embryo of an electronic computer that [The Navy] expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence."

-Frank Rosenblatt

But!!



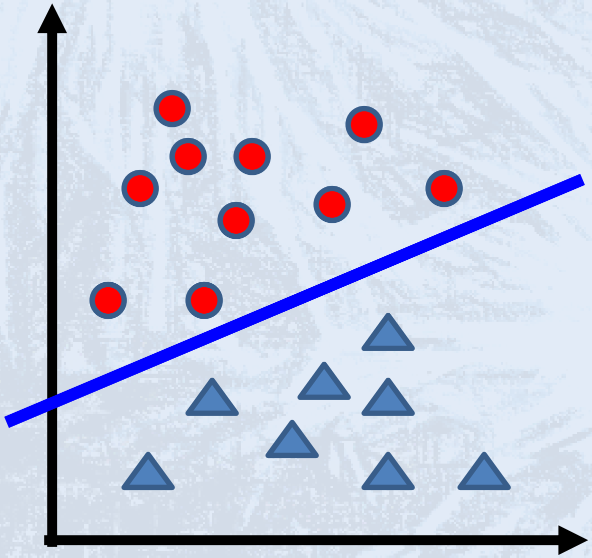


Marvin Minsky Seymour Papert

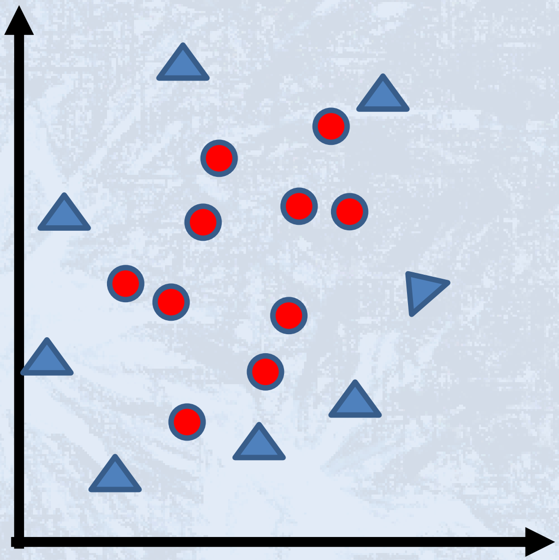
Detailed Study of Perceptrons and their capabilities.

- **Proved Perceptron cannot solve non-linearly separable problems. Eg. XOR.**
- **In addition, training time grows exponentially with the size of the input.**

Linearly vs Non-linearly Separable

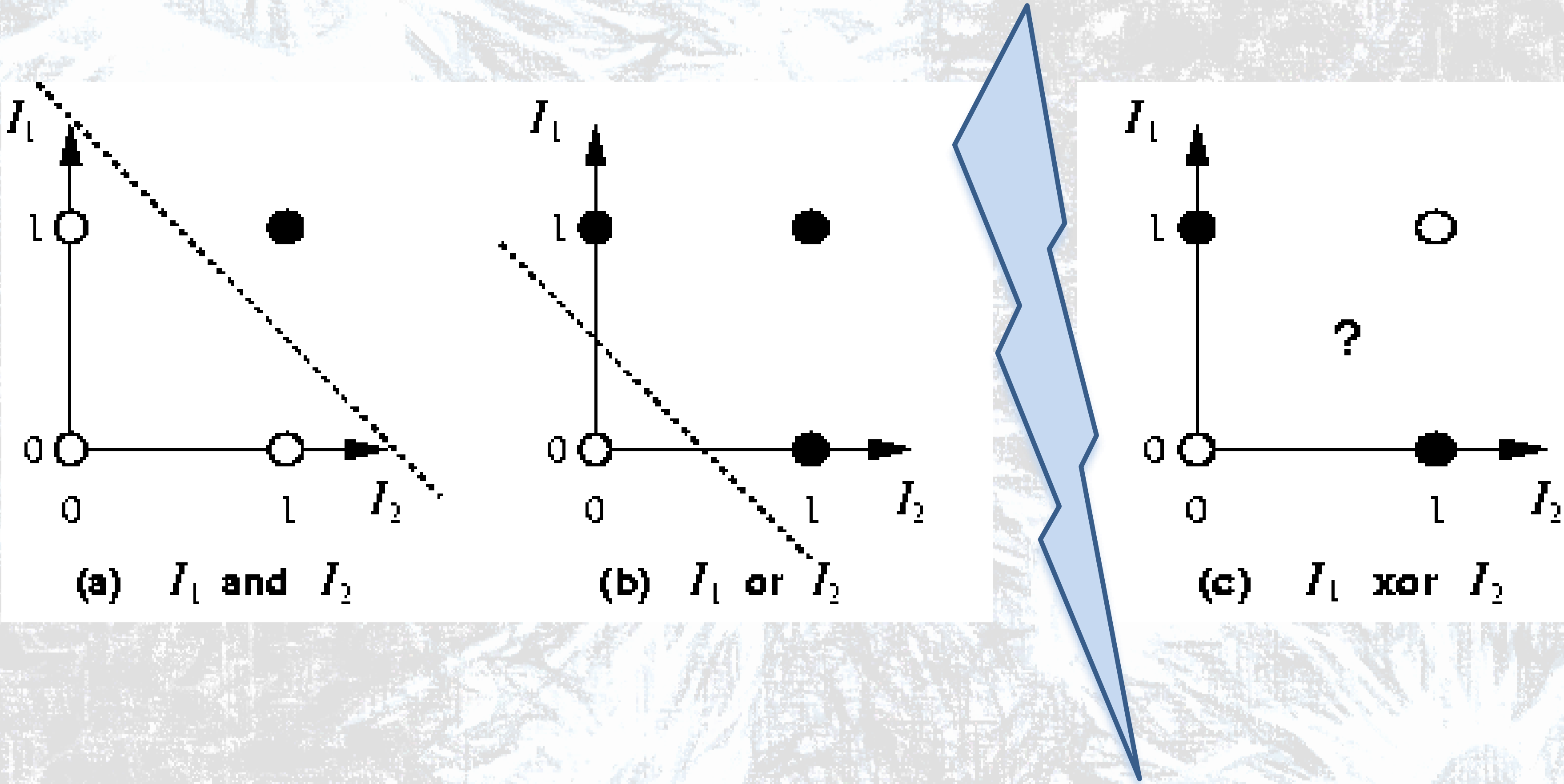


Linearly separable



Non-linearly separable

AND/OR/ XOR



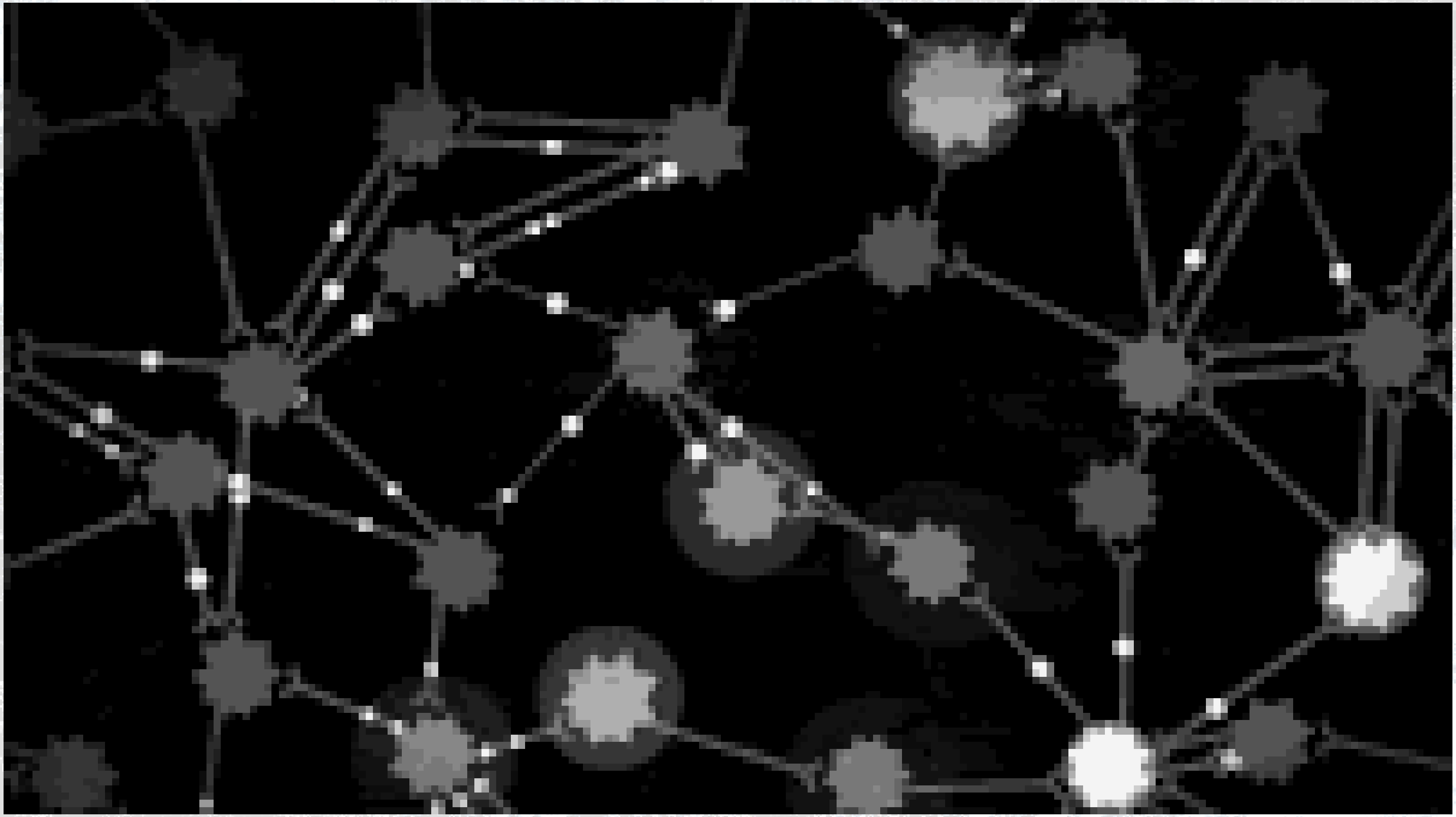
The First AI Winter
1974-1980

AI Failure’s

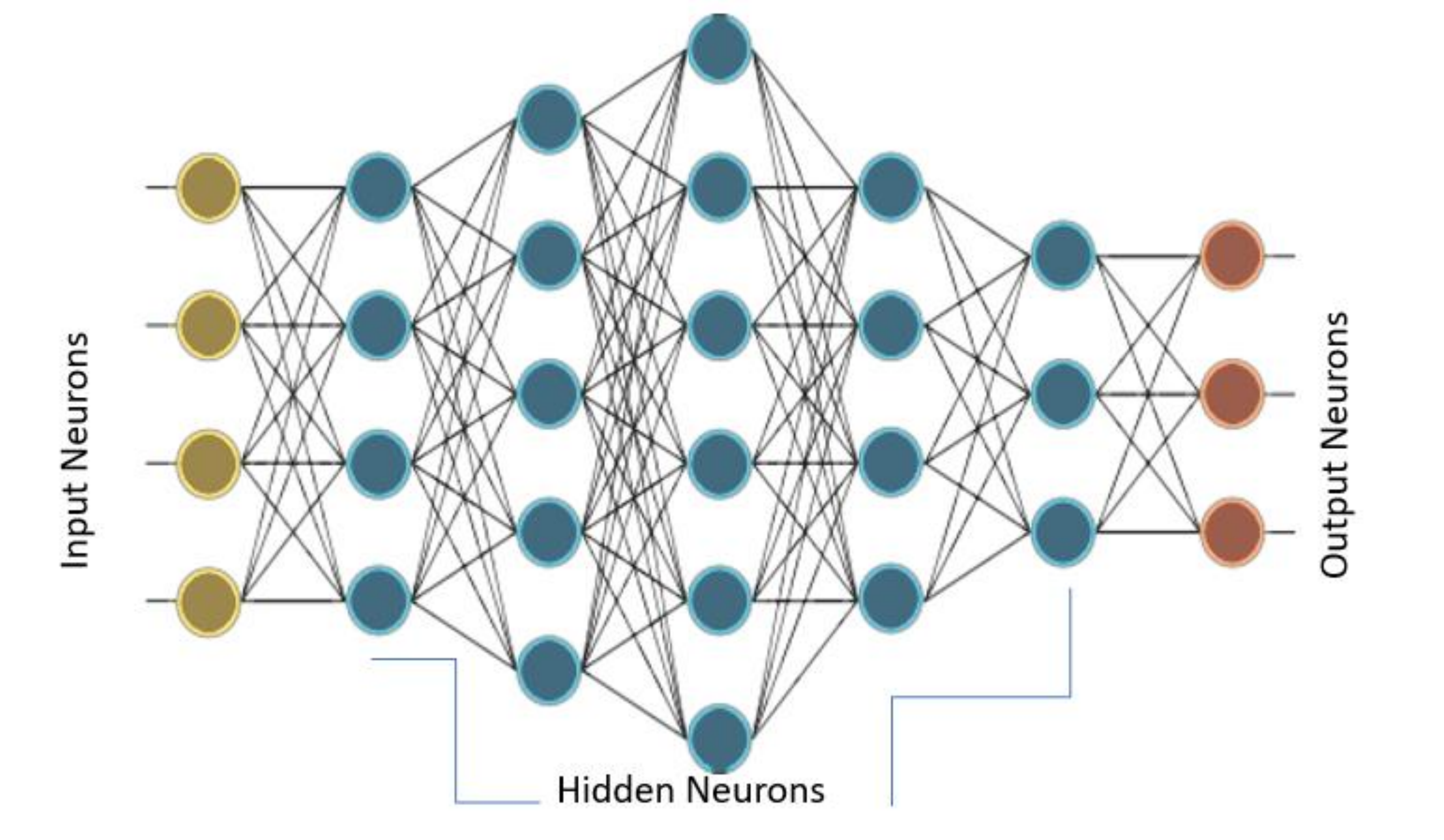
- 1969: Non linear problems.
- 1966:Machine Translation
- 1971-1975: Speech Understanding
- James Lighthill paper “**Artificial Intelligence: A General Survey**”
  - AI researchers had failed to address the issue of combinatorial explosion when solving problems within real world domains.
  - That is, the report states that AI techniques may work within the scope of small problem domains, but the techniques would not scale up well to solve more realistic problems.
  - The report represents a pessimistic view of AI that began after early excitement in the field.

What if we connect many such artificial neurons?

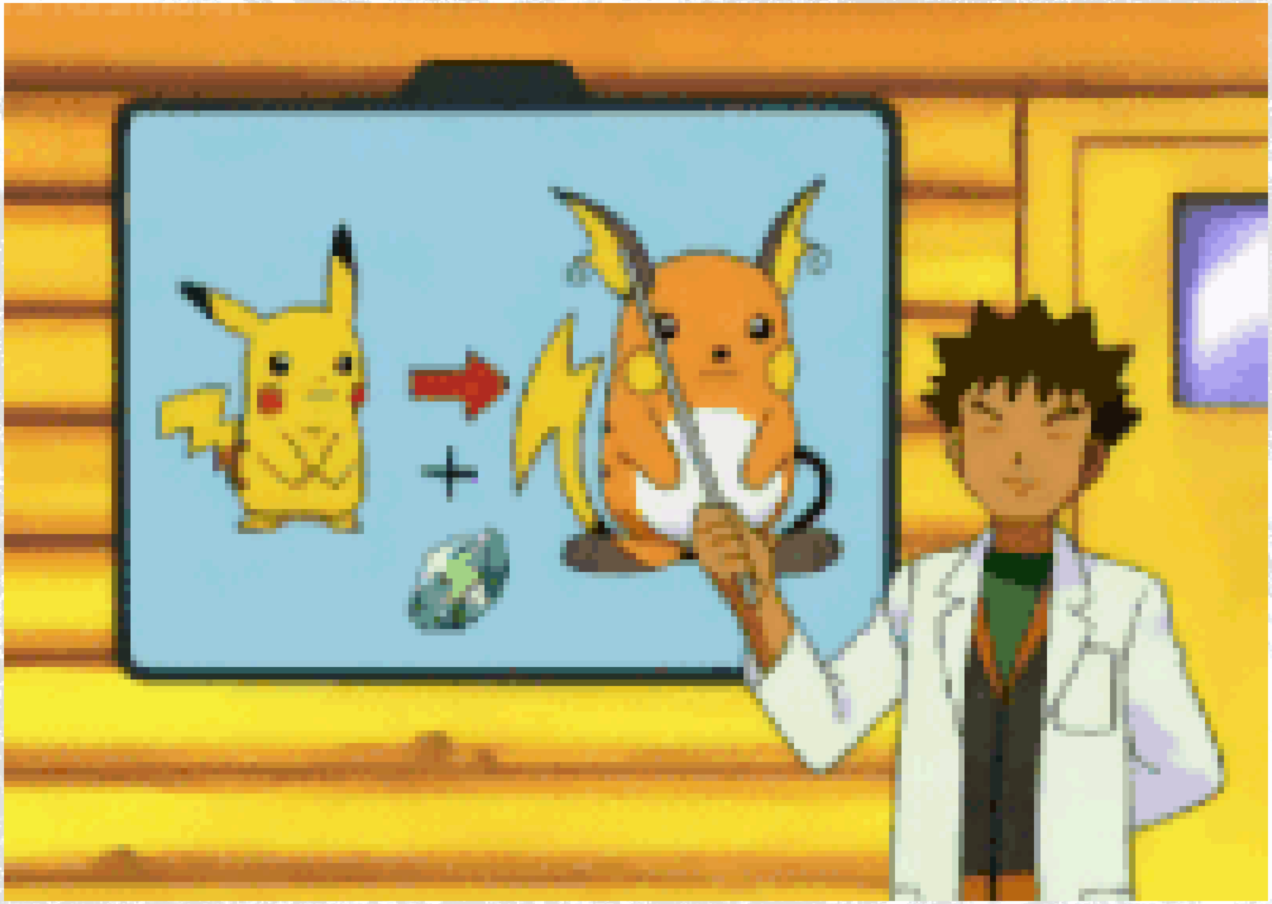
Neural Networks



Multilayer Perceptron (MLP)



But then how will we train them?



Back Propagation Algorithm

David Rumelhart

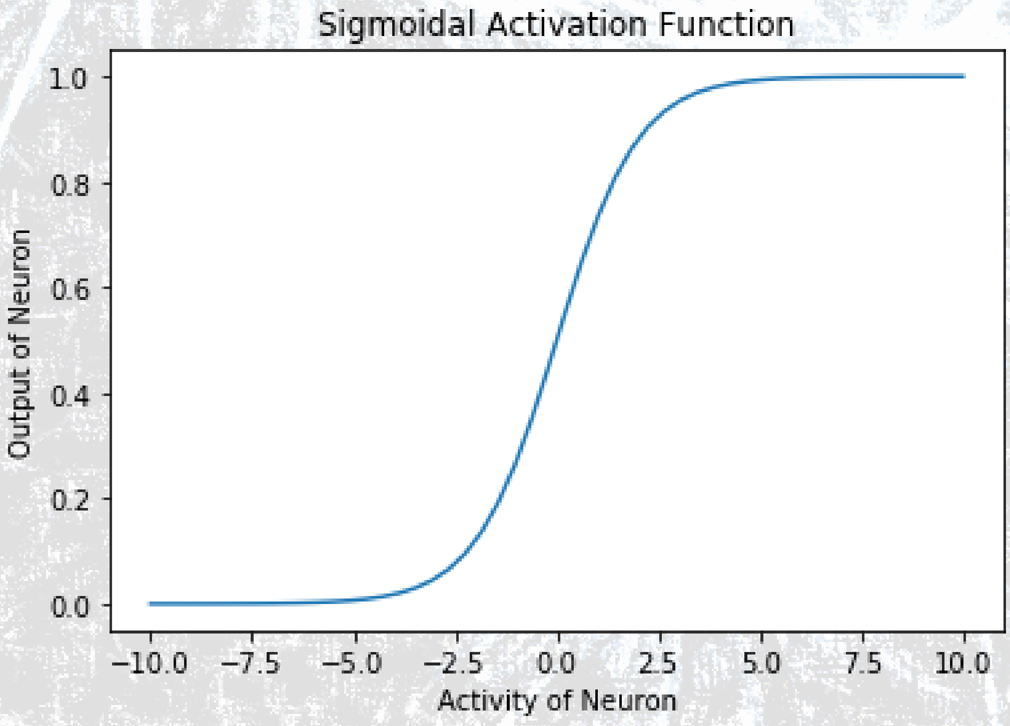
Geoffrey Hinton

Ronald Williams

- BPA was derived in 1960 by Henry J Kelley for Control theory
- In 1986 Rumelhart, Hinton and Williams showed experimentally that BPA can generate useful internal representation of incoming data in hidden layers.

Changes for BPA

- The activation function should be differentiable
- Sigmoid function $g(x) = \dfrac{1}{1+e^{-x}}$ $\quad g'^{(x)} = g(x)(1-g(x))$



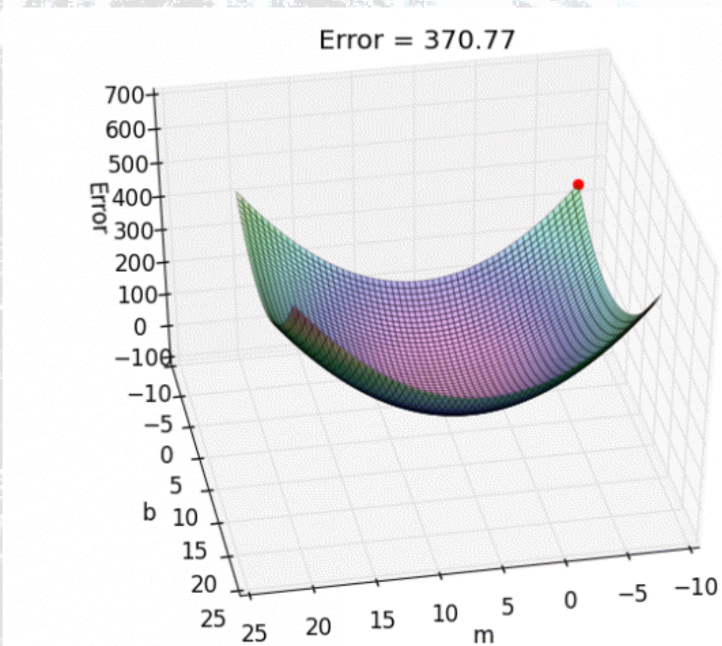
$$
\begin{aligned}
&= \frac{d}{dx}\left[\frac{1}{1+e^{-x}}\right] \\
&= \frac{d}{dx}\left(1+\mathrm{e}^{-x}\right)^{-1} \\
&= -(1+e^{-x})^{-2}(-e^{-x}) \\
&= \frac{e^{-x}}{(1+e^{-x})^2} \\
&= \frac{1}{1+e^{-x}} \cdot \frac{e^{-x}}{1+e^{-x}} \\
&= \frac{1}{1+e^{-x}} \cdot \frac{(1+e^{-x})-1}{1+e^{-x}} \\
&= \frac{1}{1+e^{-x}} \cdot \left(\frac{1+e^{-x}}{1+e^{-x}} - \frac{1}{1+e^{-x}}\right) \\
&= \frac{1}{1+e^{-x}} \cdot \left(1 - \frac{1}{1+e^{-x}}\right)
\end{aligned}
$$

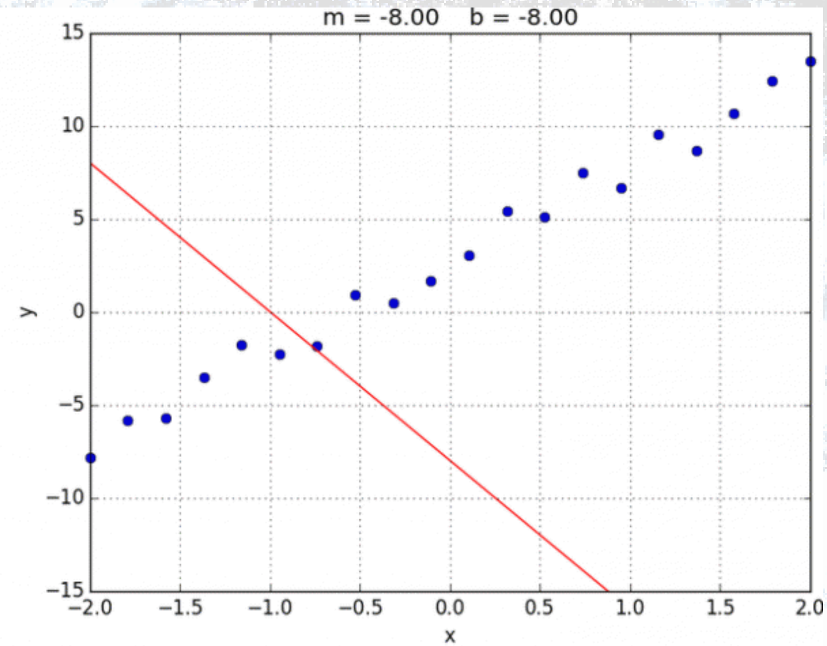
Gradient Descent Algorithm

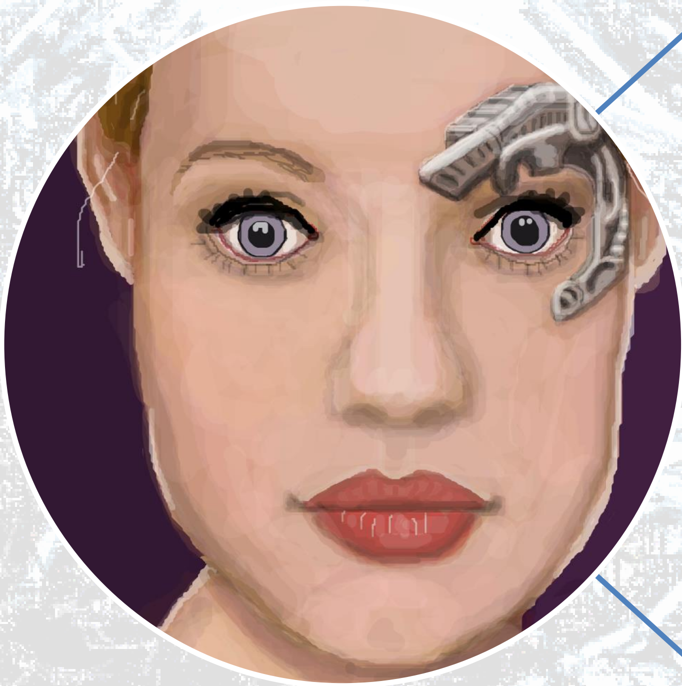
$$E = \frac{1}{2N}\sum_{x}\|y(x) - y'(x)\|^2$$

$$\Delta w_{ij} = -\eta\frac{\partial E}{\partial w_{ij}} = -\eta o_i \delta_j$$

$$\delta_j = \frac{\partial E}{\partial o_j}\frac{\partial o_j}{\partial \mathrm{net}_j} = \begin{cases} (o_j - t_j)o_j(1-o_j) & \text{if } j \text{ is an output neuron,} \\ (\sum_{\ell \in L}\delta_\ell w_{j\ell})o_j(1-o_j) & \text{if } j \text{ is an inner neuron.} \end{cases}$$







Supervised Learning:

Classification /Regression

- Multilayered Perceptrons (MLP)
- Convolutional Neural Networks (CNN)
- Recurrent Neural Networks (RNN)
- Deep Belief Networks

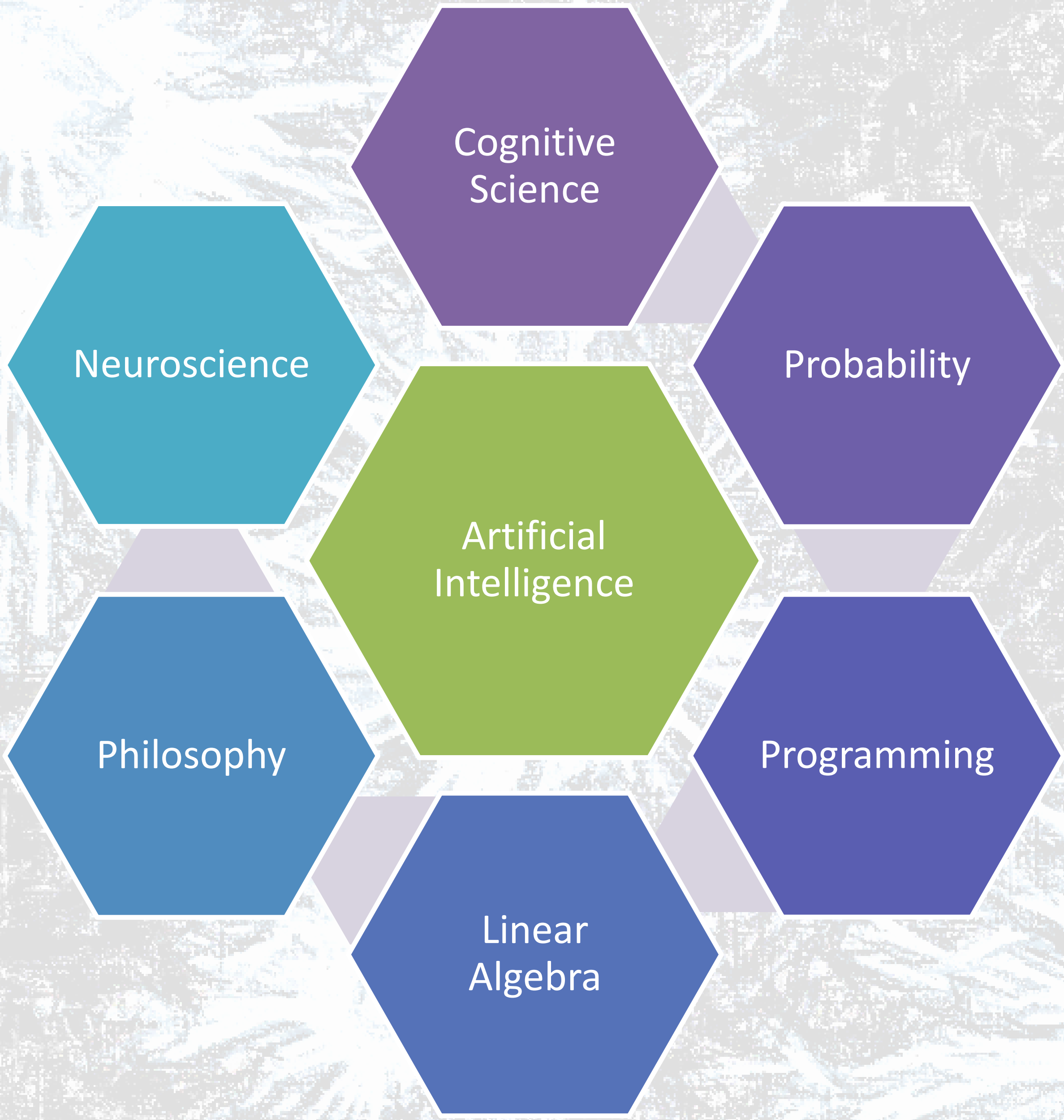
Unsupervised Learning:

Dimensionality Reduction/ Generating Data

- Self Organized Maps (SOM)
- Restricted Boltzmann Machines (RBM)
- Autoencoders
- Generative Adversial Networks (GANs)
- Deep Belief Networks

Reinforcement Learning: Robotics/ Inventory Management

- Q- learning
- Monte Carlo Methods
- Temporal Difference Methods


Cognitive Science
Neuroscience
Probability
Artificial Intelligence
Philosophy
Programming
Linear Algebra

Probability

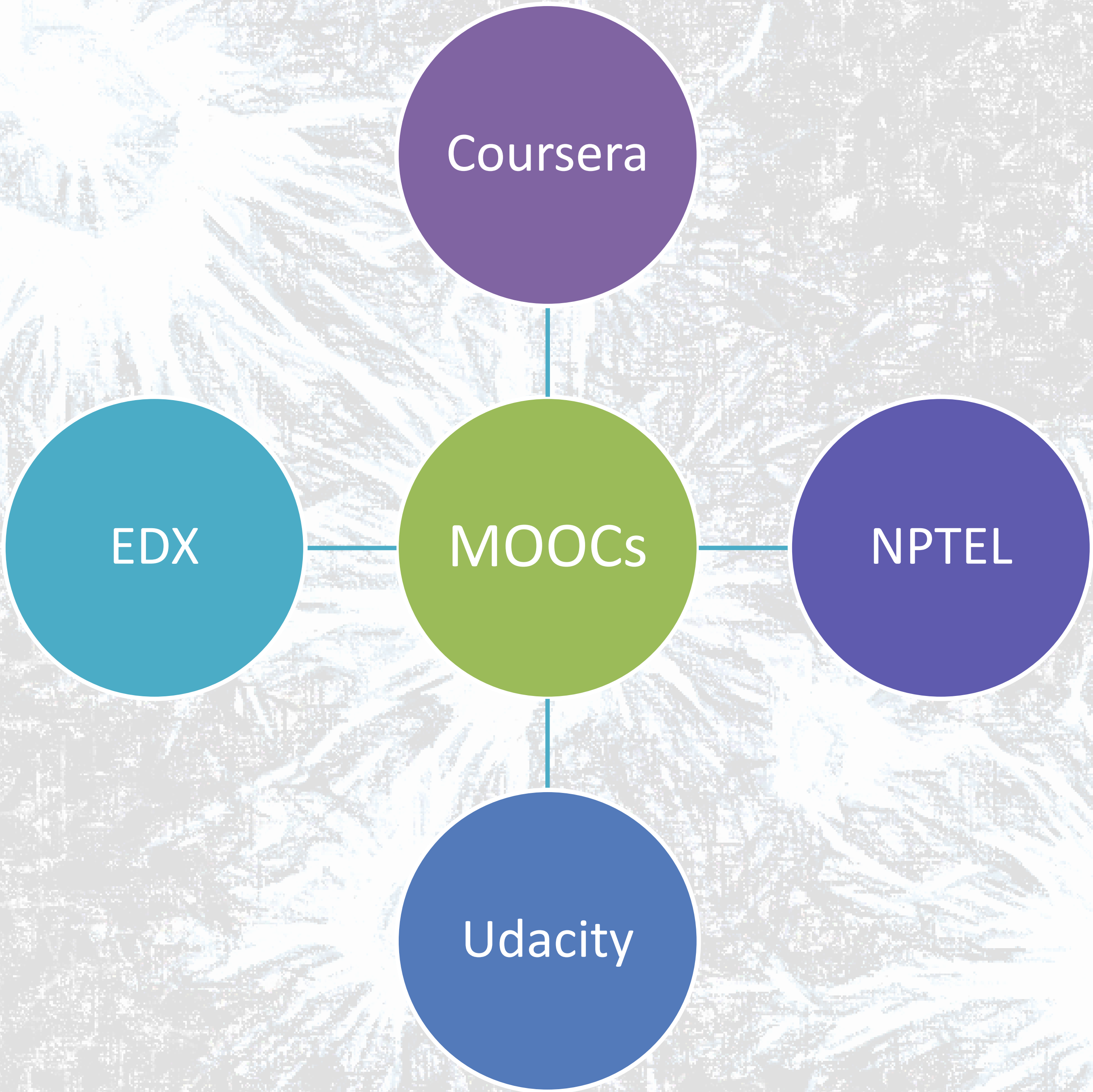
- https://courses.edx.org/courses/course-v1:MITx+6.041x_4+1T2017/course/
- https://www.coursera.org/specializations/probabilistic-graphical-models

Linear Algebra

- https://www.khanacademy.org/math/linear-algebra
- https://classroom.udacity.com/courses/ud953

Python

- https://www.coursera.org/specializations/python
- http://www.datasciencecentral.com/profiles/blogs/learn-python-in-3-days-step-by-step-guide


Coursera
EDX
MOOCs
NPTEL
Udacity