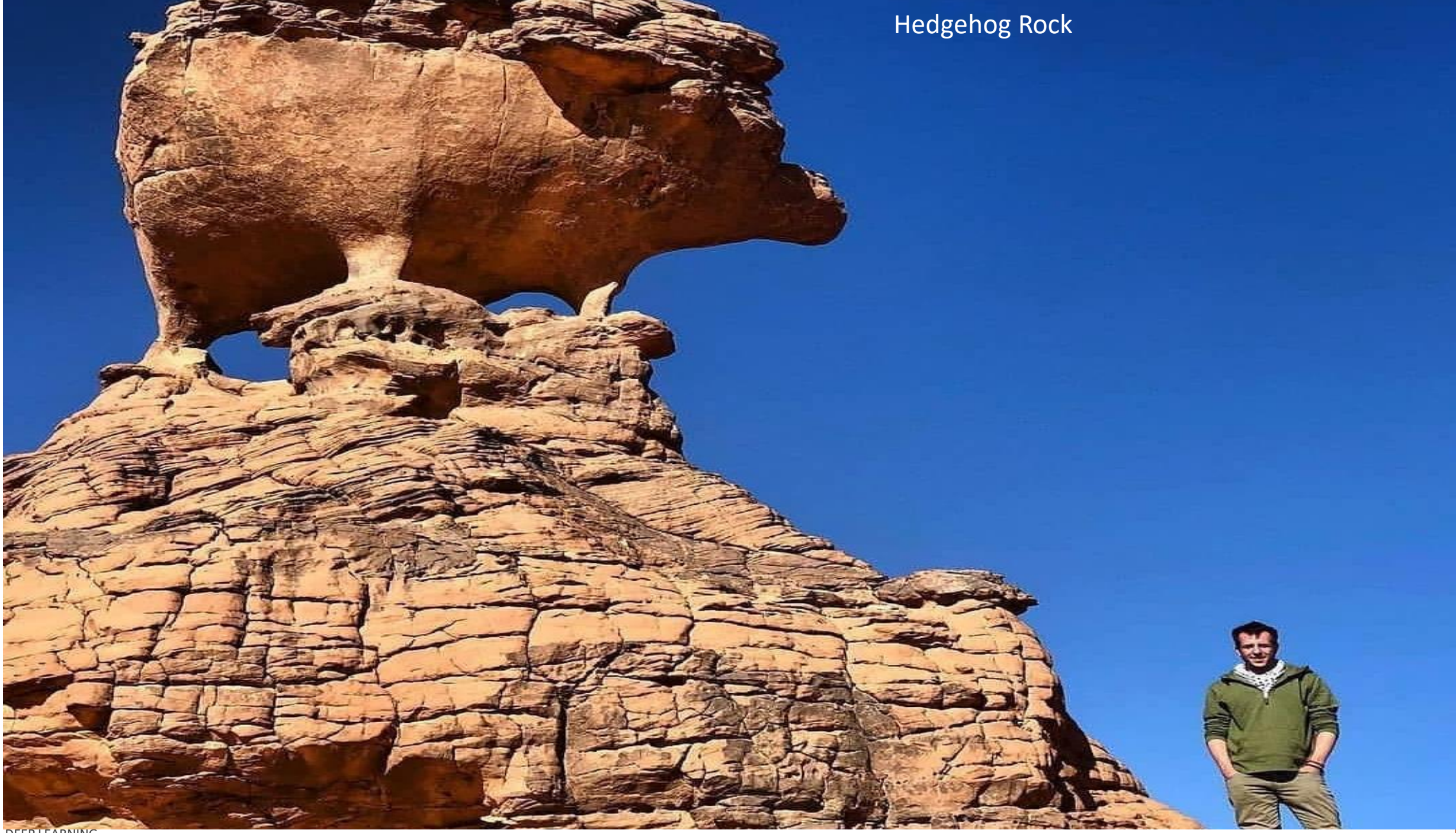

Hedgehog Rock

University of Science and Technology Bannu

Deep Learning

Lesson 1

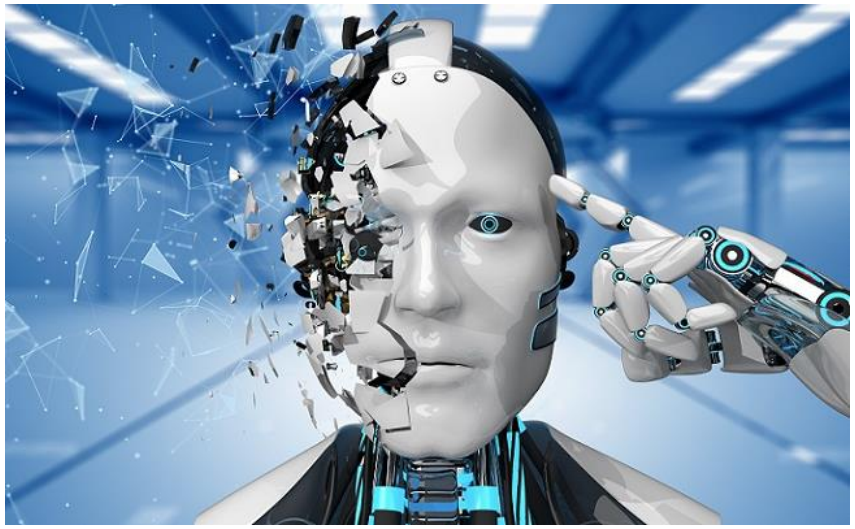
May 02, 2024

Learning Objectives

- Course Contents
- Recent Advances in AI
- Neural Networks

Welcome! In this course our focus will be:

1.Introduction to deep learning	2.Models	3. Python	4.DL for NLP
• Fundamentals of Deep learning • Neuron • Perceptron • Sigmoid Neuron • Structure of NN • Autoencoders • Restricted Boltzmann's Machine	• ANN (Artificial Neural Network) • CNN (Convolutional Neural Network) • RNN (Recurrent Neural Network) • LSTM (Long-Short Term Memory) • GRU (Gated Recurrent Unit) • Transformers	• DL with PyTorch • Numpy • Pandas • NLTK	• NER • POS



Recent Success with deep learning

GPT(Generative Pre-trained Transformer)

GPT-3, or the third-generation Generative Pre-trained Transformer, is a neural network machine learning model trained using internet data to generate any type of text. Developed by OpenAI, it requires a small amount of input text to generate large volumes of relevant and sophisticated machine-generated text.

GPT-3: OpenAI unveiled GPT-3 in June 2020, which is a new language model that surpasses GPT-2 in size by more than 100 times, with **175 billion** parameters and **96 layers**. Through training on a dataset comprising **499 billion tokens** extracted from web content

Google computer teaches itself to play Go and beats world's best within days

CLIVE COOKSON — LONDON

A self-taught computer has become the world's best player of Go, the fiendishly complex board game, without any input from human experts.

DeepMind, Google's artificial intelligence subsidiary in London, announced the milestone in AI less than two years after the highly publicised unveiling of AlphaGo, the first machine to beat human champions at the ancient Asian game. Details are published in the scientific journal Nature.

Previous versions of AlphaGo learned by analysing thousands of games between top human players. The new development, called AlphaGo Zero, starts just by knowing the rules and objective of the game.

"It learns to play simply by playing games against itself, starting from completely random play," said Demis Hassabis, DeepMind chief executive. "In doing so, it quickly surpassed human level of play and defeated the previously published version of AlphaGo by 100 games to zero."

David Silver, AlphaGo project leader, added: "By not using human data in any fashion, we can create knowledge by itself from a blank slate." Within a few days, the computer had not only learned Go from scratch but surpassed thousands of years of accumulated human wisdom about the game.

The team developed a new form of "reinforcement learning" to create AlphaGo Zero, combining search-based simulations of future moves with a neural network that decides which moves give the highest probability of winning. The network is constantly updated over

Although Go is in one sense extremely complex, with far more potential moves than there are atoms in the universe, in another sense it is simple because it is a "game of perfect information" — chance plays no part, as with cards or dice, and the state of play is defined entirely by the position of stones on the board.

The game involves surrounding more territory than your opponent. This aspect of Go makes it particularly susceptible to the computer simulations on which AlphaGo depends. DeepMind is now examining real-life problems that can be structured in a similar way, to apply the technology.

Mr Hassabis identified predicting the shape of protein molecules — an important issue in drug discovery — as a promising candidate. Other likely scientific applications include designing

Austria victor drives home 'political Islam' message

Interview ▸ PAGE 2

Google DeepMind Challenge Match

It was a five-game Go match between 18-time world champion Lee Sedol and AlphaGo, a computer Go program developed by Google DeepMind, played in Seoul, South Korea between 9 and 15 March 2016. AlphaGo won

Save

Found in translation: More accurate, fluent sentences in Google Translate

In 10 years, Google Translate has gone from supporting just a few languages to 103, connecting strangers, reaching across language barriers and even helping people find love. At the start, we pioneered large-scale statistical machine translation, which uses statistical models to translate text. Today, we're introducing the next step in making Google Translate even better: Neural Machine Translation. Neural Machine Translation has been generating exciting research results for a few years and in September, our researchers announced Google's version of this technique. At a high level, the Neural system translates whole sentences at a time, rather than just piece by piece. It uses this broader context to help it figure out the most relevant translation, which it then rearranges and adjusts to be more like a human speaking with proper grammar. Since it's easier to understand each sentence, translated paragraphs and articles are a lot smoother and easier to read. And this is all possible because of end-to-end learning system built on Neural Machine Translation, which basically means that the system learns over time to create better, more natural translations

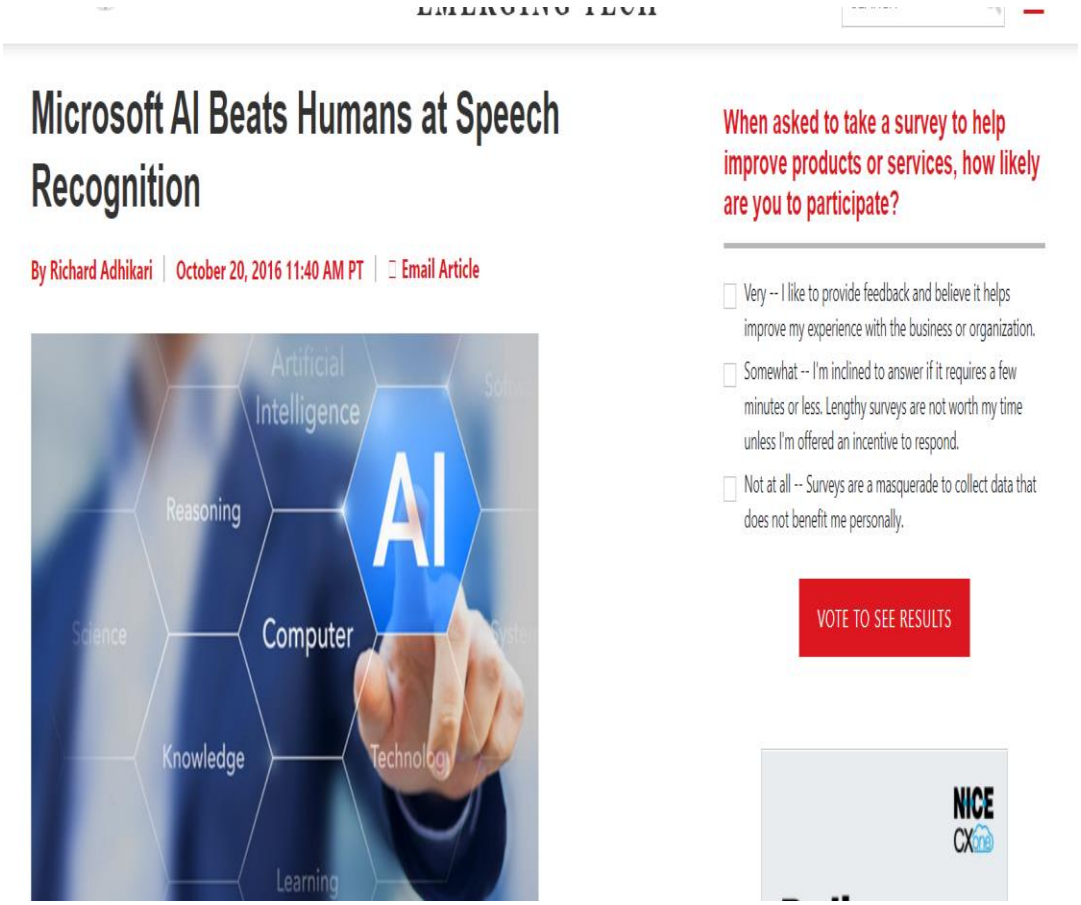
Recent Advances in AI

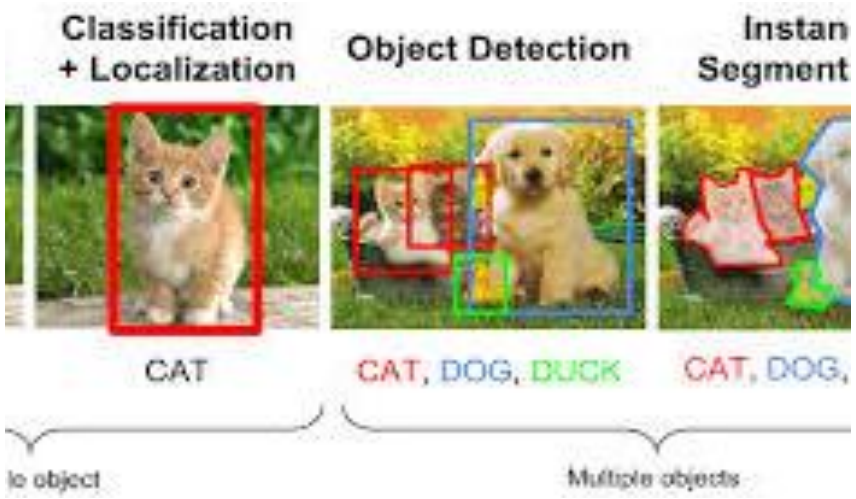
Microsoft's Artificial Intelligence and Research Unit earlier this week reported that its speech recognition technology had surpassed the performance of human transcriptionists.

The team last month published a paper describing its system's accuracy, said to be superior to that of IBM's famed Watson artificial intelligence.

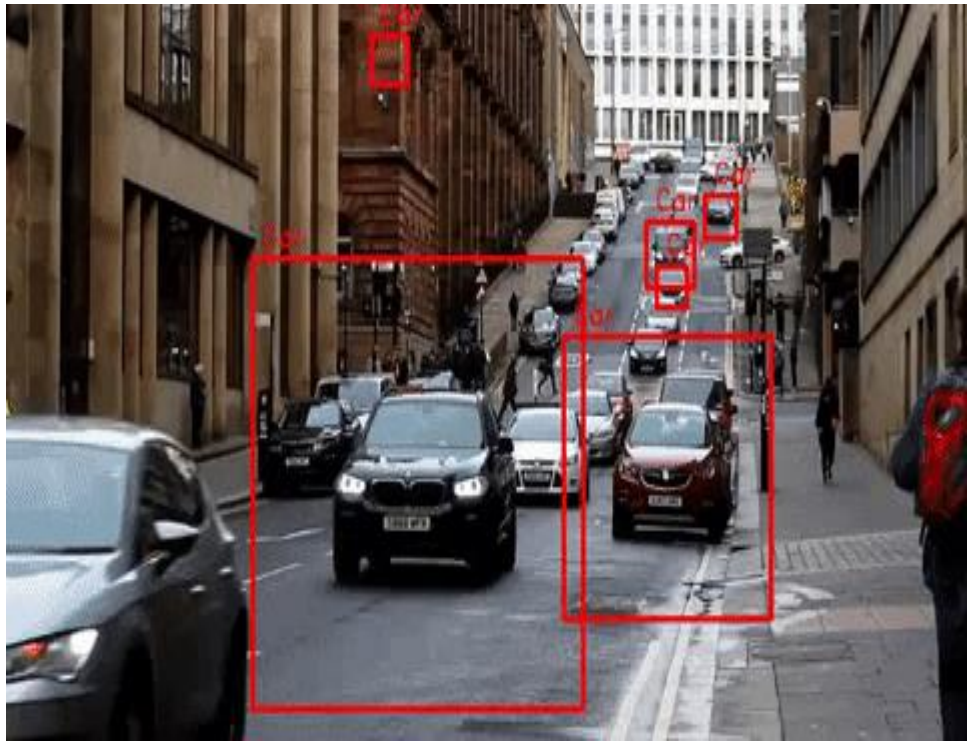
The error rate for humans on the widely used NIST 2000 test set is 5.9 percent for the Switchboard portion of the data, and 11.3 percent for the CallHome portion, the team said.

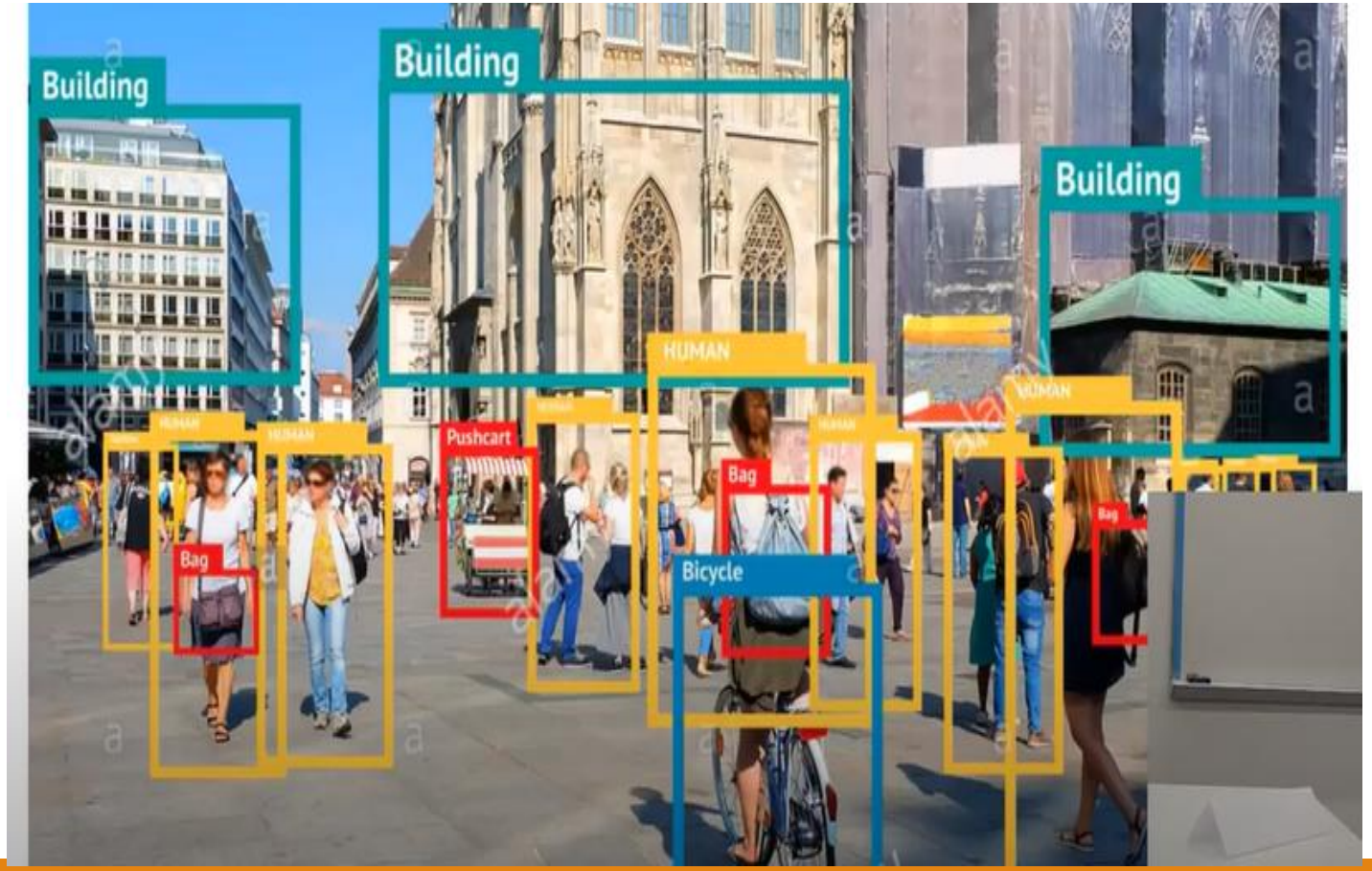
The team improved on the conversational recognition system that outperformed IBM's by about 0.4 percent, it reported.





Images segmentation and Recognition





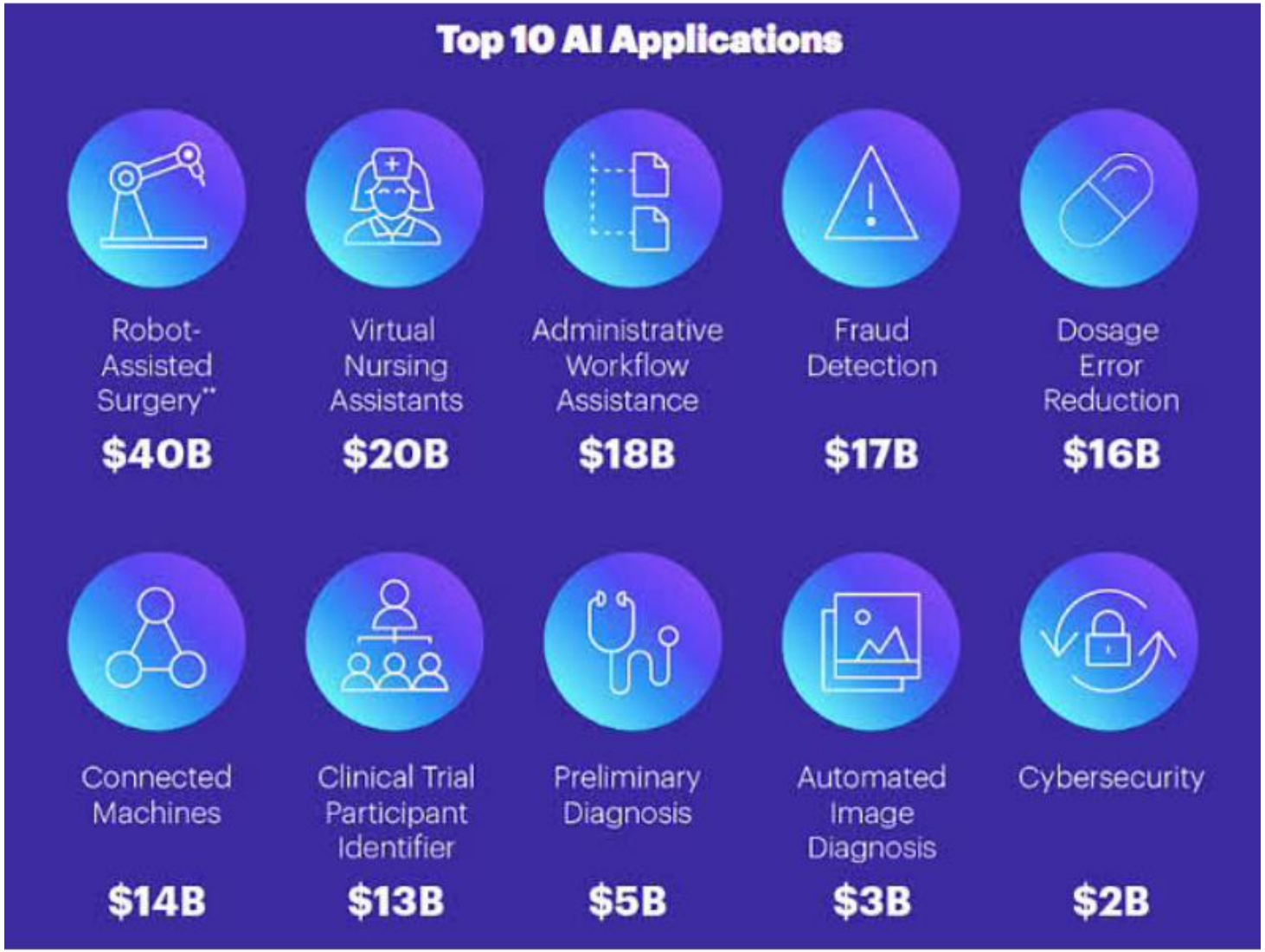
"man in black shirt is playing guitar."

"construction worker in orange safety vest is working on road."

"girl in pink dress is jumping in

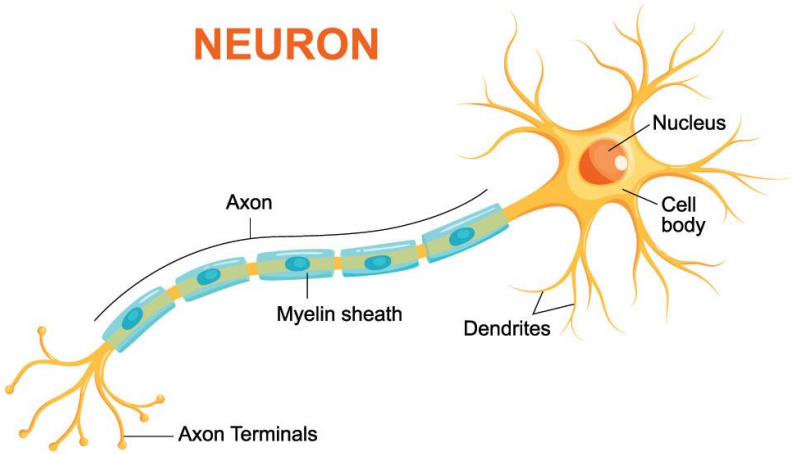
"black and white dog jumps over

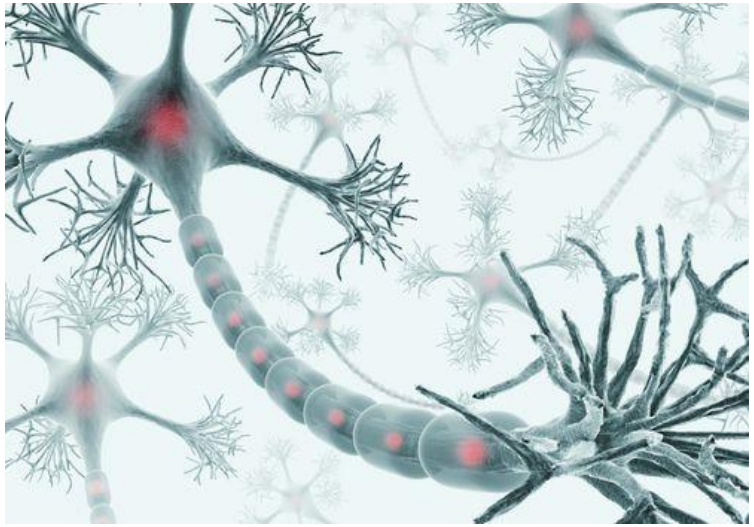
Auto image caption



Top AI Application

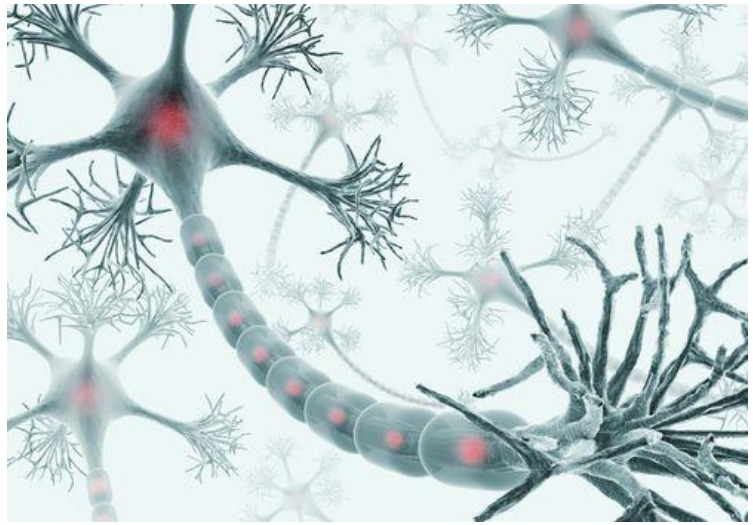

UK gives support to 1,000 PhD on AI
USA has been investing $10 billion of venture capital into AI
French government is planning to invest $1.8 billion in AI by 2022
China AI patent on AI grew by 190% in last 5 years
Cananda is building a reserach facility for AI for $127 million

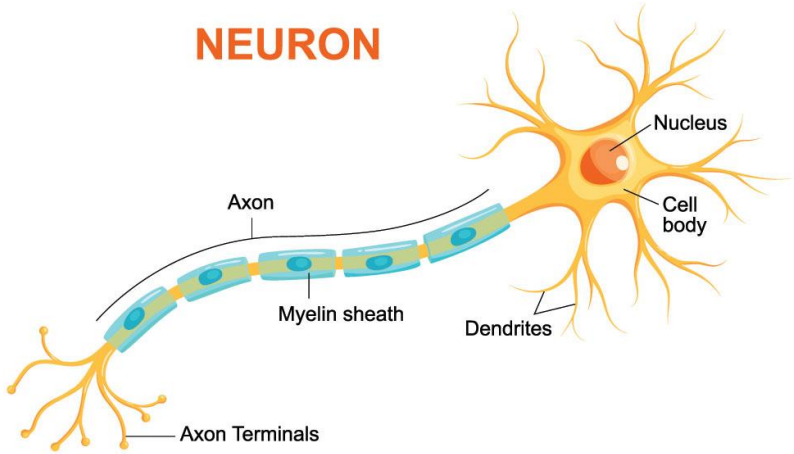




Neural Network

- In literature, neural networks and artificial neural networks (ANN) are the two words that are used interchangeably
- It is an advanced information processing models proposed after inspiration from natural human nervous systems
- It processes information alike human brains does.

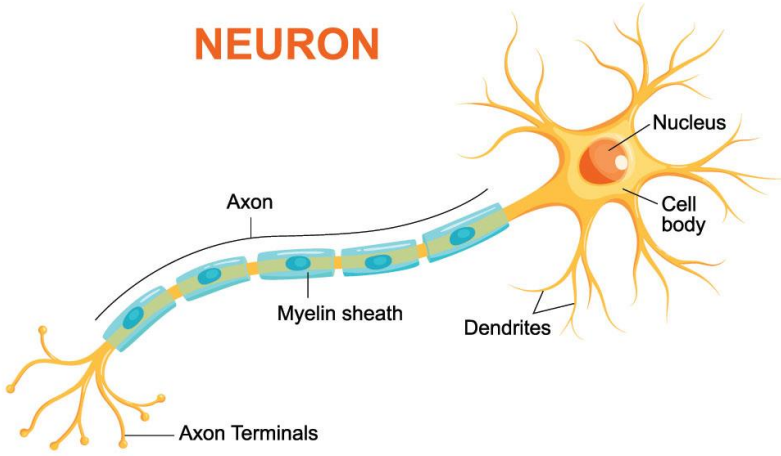




Biological Neuron

ABOUT 100 BILLION NEURON IN HUMAN BRAIN

Biological Neuron



The main body of the cell collect the incoming signals from the other neurons through its dendrites

The incoming signals are constantly being summed in the cell body.

If the result of the summation crosses as certain threshold, the cell body emits a signal of its own

This signal passes through the neuron’s axon, from where the dendrites of the other neurons pick it up

Biological Neuron

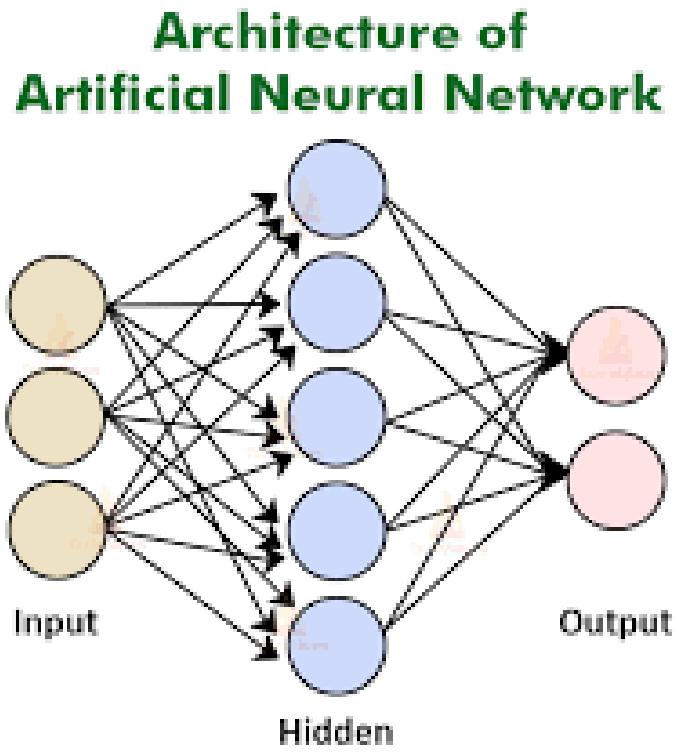
There are 1000 to 10000 dendrites in each neuron(few millimeters long)

There are only one Axon(several centimeter long)

The connection between dendrites and axon as electromagnetically and it is called synapses

The synapses modify the signal while passing it on to dendrites

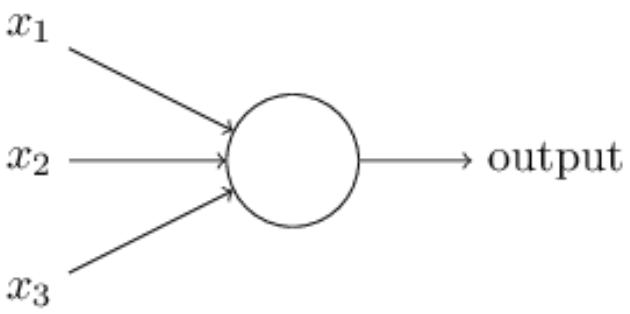
The human learning is stored in these synapses, and the connection of neuron with other neurons



Neural Network

- Its structure composed of galore interlinked information processing units termed as neurons functioning collectively to figure out particular problems
- ANNs, like humans neural system, learn by illustration.

Neural Network



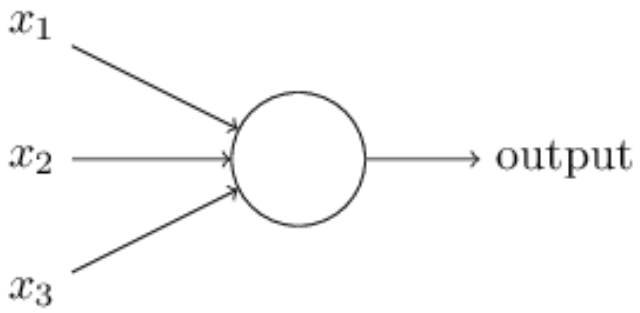
The two important types of artificial neuron are:

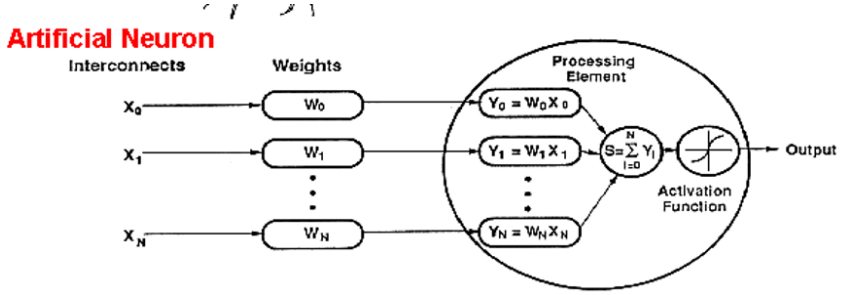
1. Perceptron

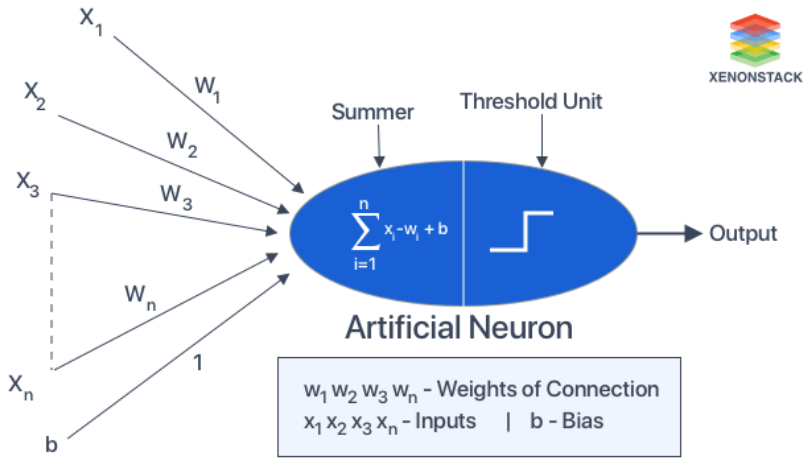
2. Sigmoid Neuron

Perceptron

- Perceptron is initially introduced by Frank Rosenblatt in 1950s and 1960s.
- Nowadays, it is a common practice to exercise new models of artificial neurons.
- Input to a general neuron consisted of several binary values $x_1,x_2,\ldots x_1,x_2,\ldots$, while the output is a single binary value







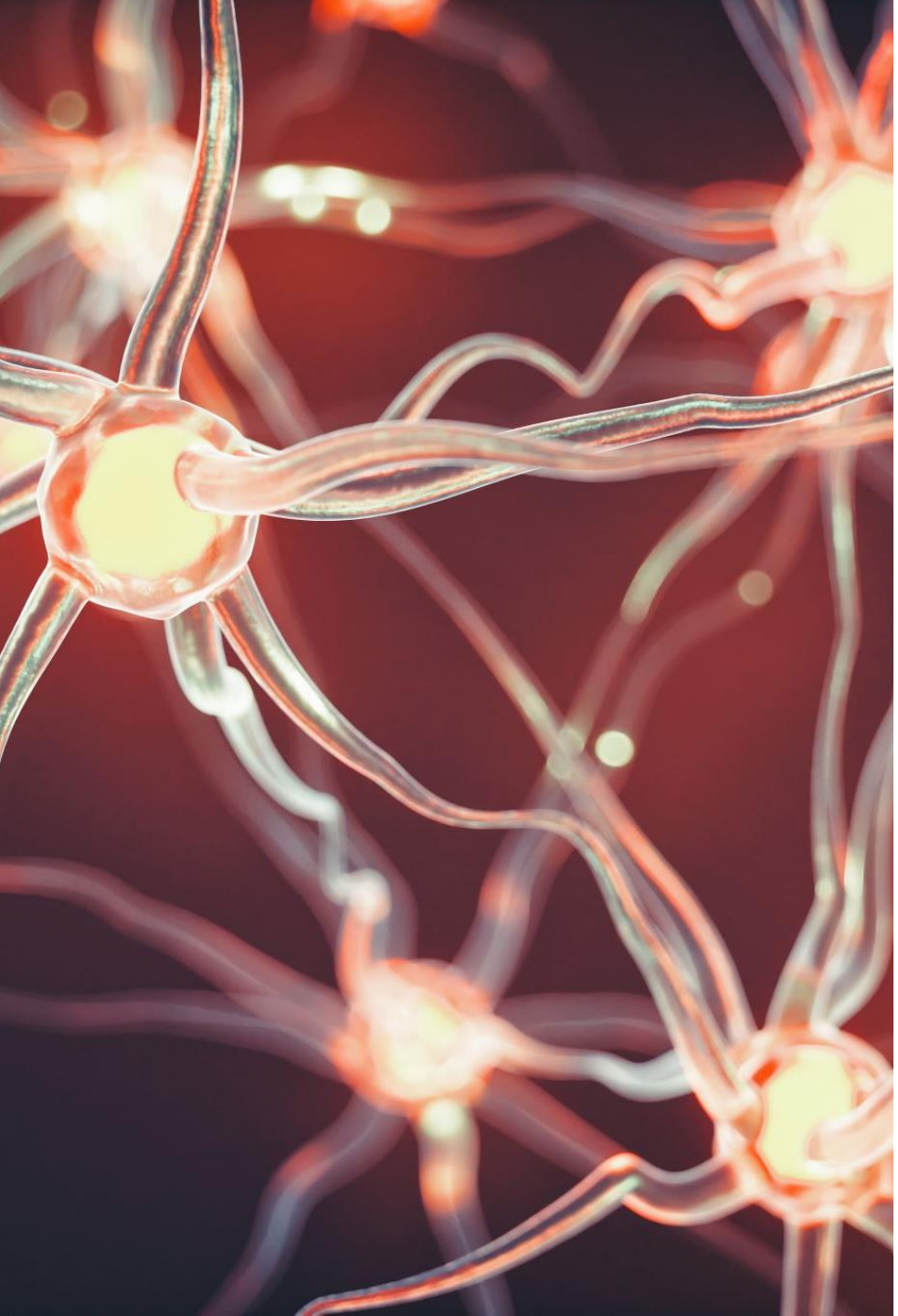
Perceptron Output

- Based on **Unit step activation function**
- $output = \begin{cases} 0, & if \ \sum_j x_j w_j \ \leq \ threshold \\ 1, & if \ \sum_j x_j w_j \ > threshold \end{cases}$

Limitations of Perceptron's

Binary input and output

The weights and bias are not adjustable



Sigmoid Neuron

- Structure wise Sigmoid neurons and perceptron's are similar
- The weights and bias of sigmoid neurons are adjustable and are adjusted such as that a minor change in inputs may reflect a minor change in their corresponding outputs
- In case of sigmoid neuron outputs are calculated using **Sigmoid Function**

$$\sigma(Z) \equiv \frac{1}{1+e^{-z}}$$

$$e^{-z} = \frac{1}{1+\exp(-\sum_j w_j x_j - b)}$$

Deep Learning, ML, AI, . . . oh my!

Mentimeter

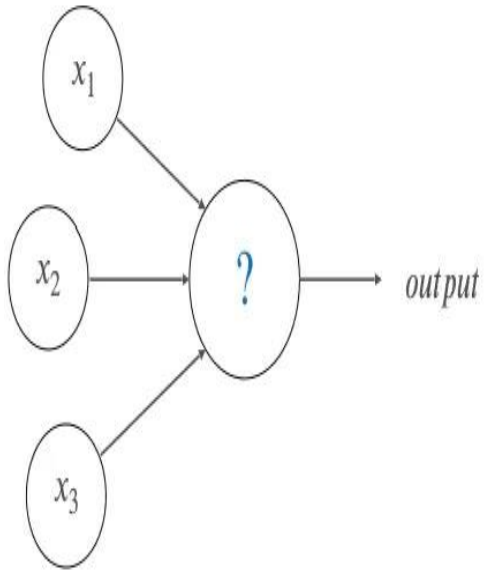
Mentimeter

Mentimeter

Go to https://www.menti.com/ and enter Gam Pin 39 81 34 9

Perceptron

Simplified (binary) artificial neuron



Do I snowboard this weekend?

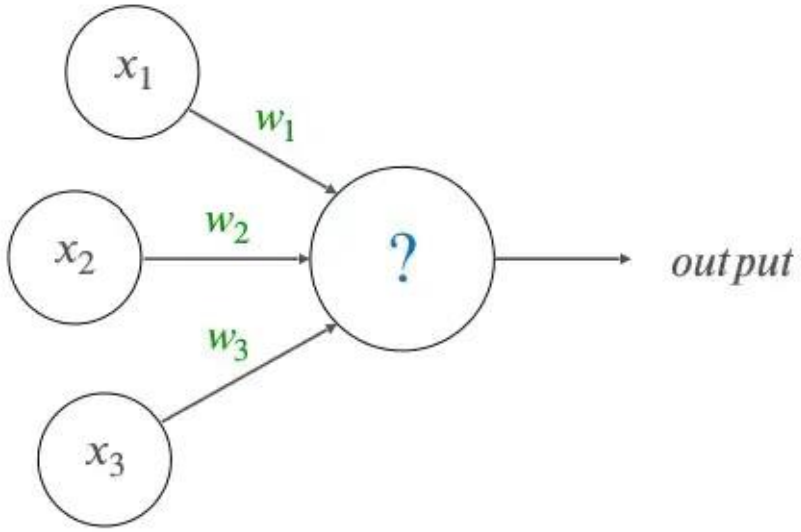
$x_1 \rightarrow$ *Is the weather good?*

$x_2 \rightarrow$ *Is the powder good?*

$x_3 \rightarrow$ *Am I in the mood to drive?*

Perceptron

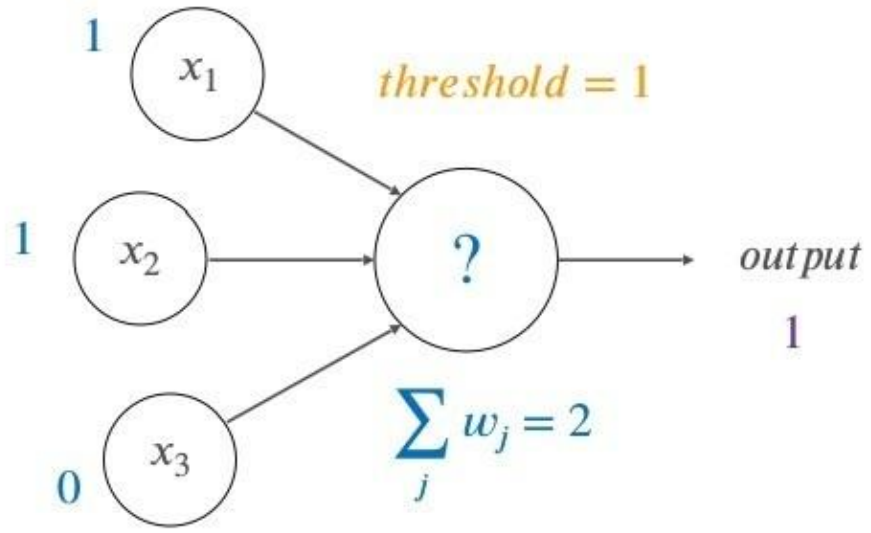
Simplified (binary) artificial neuron with weights



$$output = \begin{cases} 0, & \sum_{j=0}^{n} w_j x_j \leq threshold \\ 1, & \sum_{j=0}^{n} w_j x_j > threshold \end{cases}$$

Perceptron

Simplified (binary) artificial neuron; *no weights*



Do I snowboard this weekend?

$x_1 = 1$ (*good weather*)

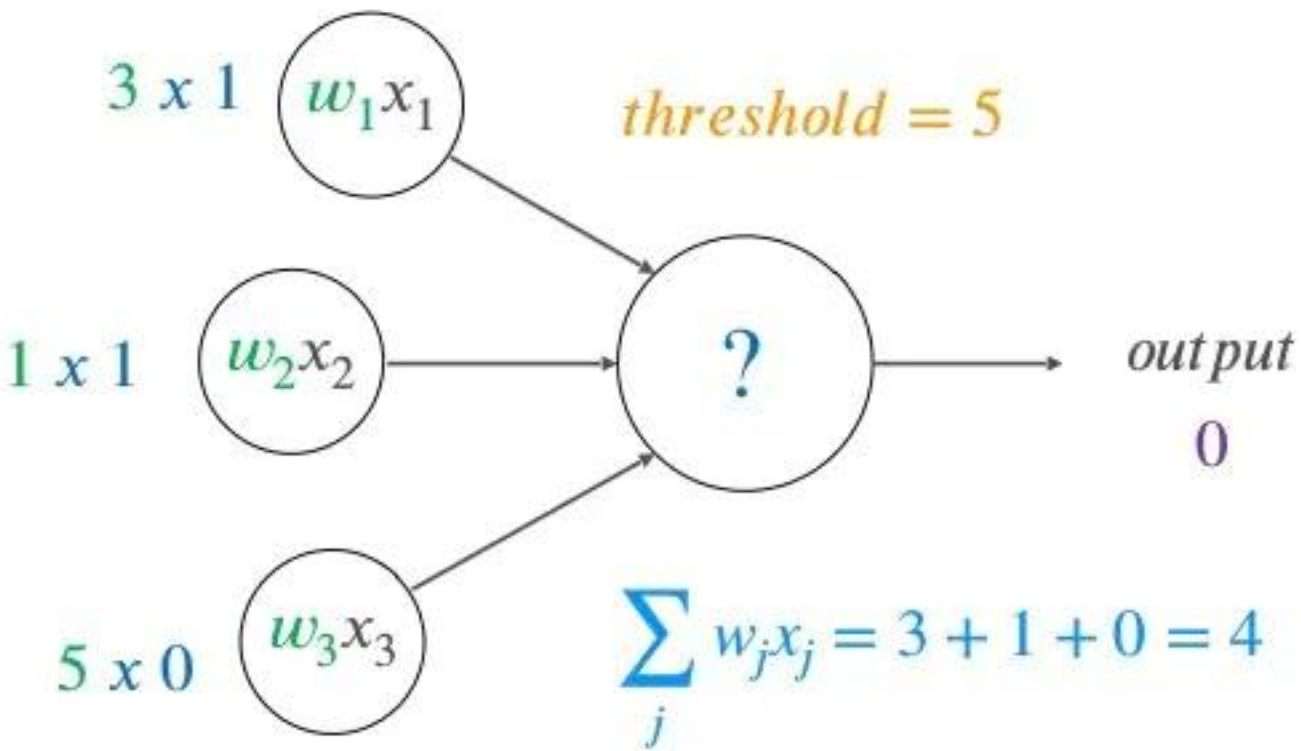
$x_2 = 1$ (*a lot of powder*)

$x_3 = 0$ (*driving sucks*)

Perceptron

Simplified (binary) artificial neuron; *add weights*

Persona: Après-ski'er



Do I snowboard this weekend?

$x_1 = 1$ (*good weather*)

$w_1 = 3$

$x_2 = 1$ (*a lot of powder*)

$w_2 = 1$

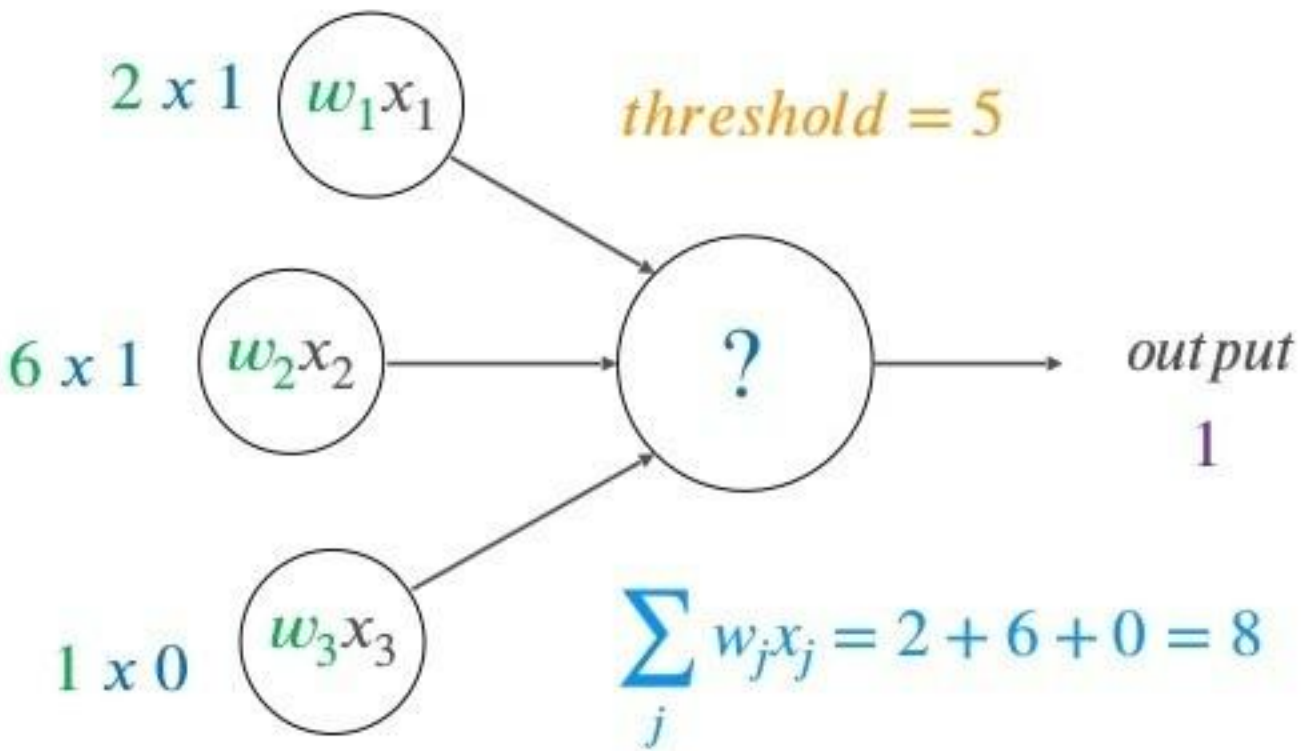
$x_3 = 0$ (*driving sucks*)

$w_3 = 5$

Perceptron

Simplified (binary) artificial neuron; *add weights*

Persona: Shredder



Do I snowboard this weekend?

$x_1 = 1$ (*good weather*)

$w_1 = 2$

$x_2 = 1$ (*a lot of powder*)

$w_2 = 6$

$x_3 = 0$ (*driving sucks*)

$w_3 = 1$

Introducing Bias

Perceptron needs to take into account the bias

$$output = \begin{cases} 0, & wx + b \leq 0 \\ 1, & wx + b > 0 \end{cases}$$

where b is how easy it is to get the perceptron to fire

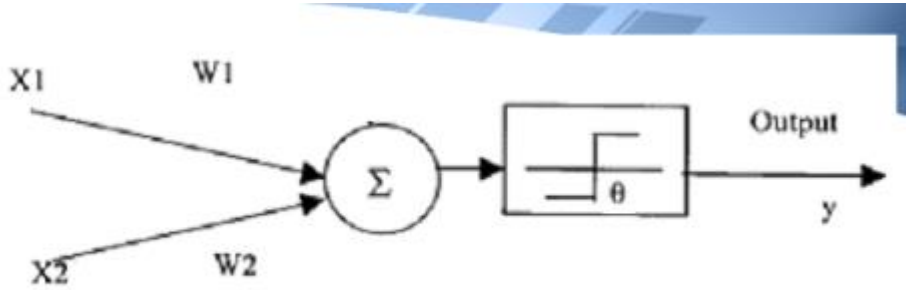
e.g. Shredder has a strong positive bias to go to Whistler while Après-Ski'er bias is not as strong

Artificial Neural Neuron

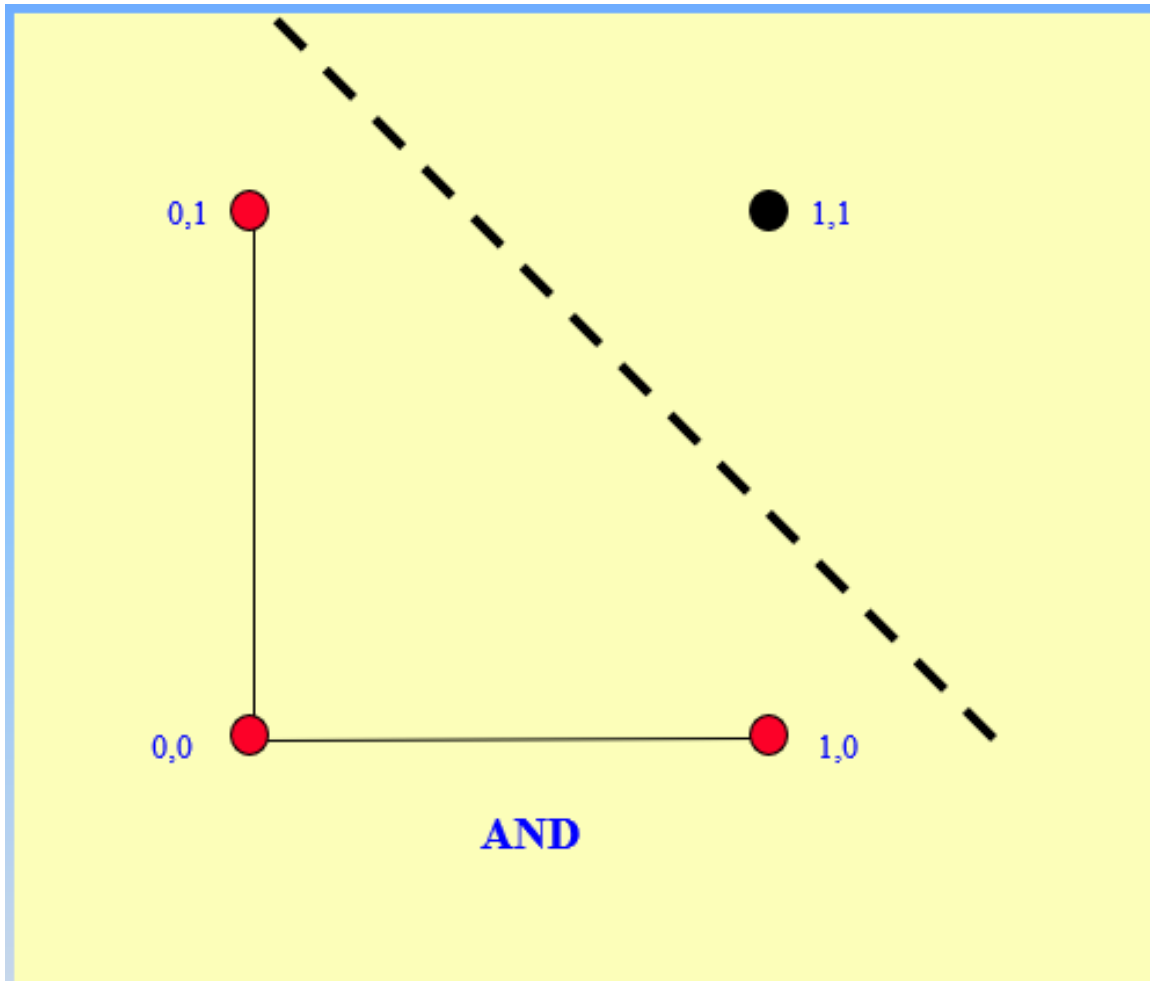
Implementation of AND Function

Let $W_1=W_2=1$

X_1	X_2	$X_1W_1 + X_2W_2$	Y
0	0	0	0
0	1	1	0
1	0	1	0
1	1	2	1



If we make Θ =2 (or any value >1 but <=2), we will get correct results with unit step activation function



Artificial Neural Neuron

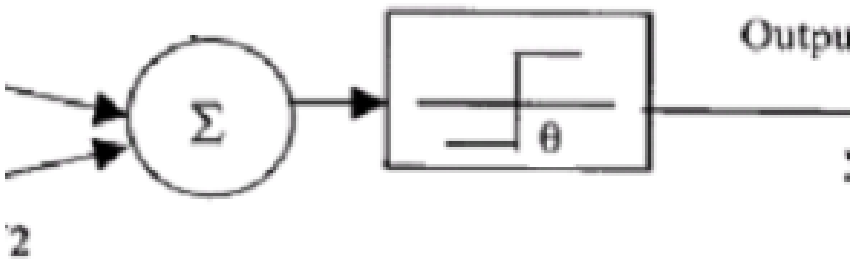
If we place the 4 points in a two coordinate system(X1 andX2),

Any new data falling on the left side of the line will give an output of zero and the data on the right side of the line will be classified as one

Epoch	Inputs		Desired output	Initial weights		Actual output	Error	Final weights	
	x_1	x_2	Y_d	w_1	w_2	Y	e	w_1	w_2
1	0	0	0	0.3	−0.1	0	0	0.3	−0.1
	0	1	0	0.3	−0.1	0	0	0.3	−0.1
	1	0	0	0.3	−0.1	1	−1	0.2	−0.1
	1	1	1	0.2	−0.1	0	1	0.3	0.0
2	0	0	0	0.3	0.0	0	0	0.3	0.0
	0	1	0	0.3	0.0	0	0	0.3	0.0
	1	0	0	0.3	0.0	1	−1	0.2	0.0
	1	1	1	0.2	0.0	1	0	0.2	0.0
3	0	0	0	0.2	0.0	0	0	0.2	0.0
	0	1	0	0.2	0.0	0	0	0.2	0.0
	1	0	0	0.2	0.0	1	−1	0.1	0.0
	1	1	1	0.1	0.0	0	1	0.2	0.1
4	0	0	0	0.2	0.1	0	0	0.2	0.1
	0	1	0	0.2	0.1	0	0	0.2	0.1
	1	0	0	0.2	0.1	1	−1	0.1	0.1
	1	1	1	0.1	0.1	1	0	0.1	0.1
5	0	0	0	0.1	0.1	0	0	0.1	0.1
	0	1	0	0.1	0.1	0	0	0.1	0.1
	1	0	0	0.1	0.1	0	0	0.1	0.1
	1	1	1	0.1	0.1	1	0	0.1	0.1

Threshold: $\theta = 0.2$; learning rate: $\alpha = 0.1$

Example of perceptron learning: the logical operation *AND*



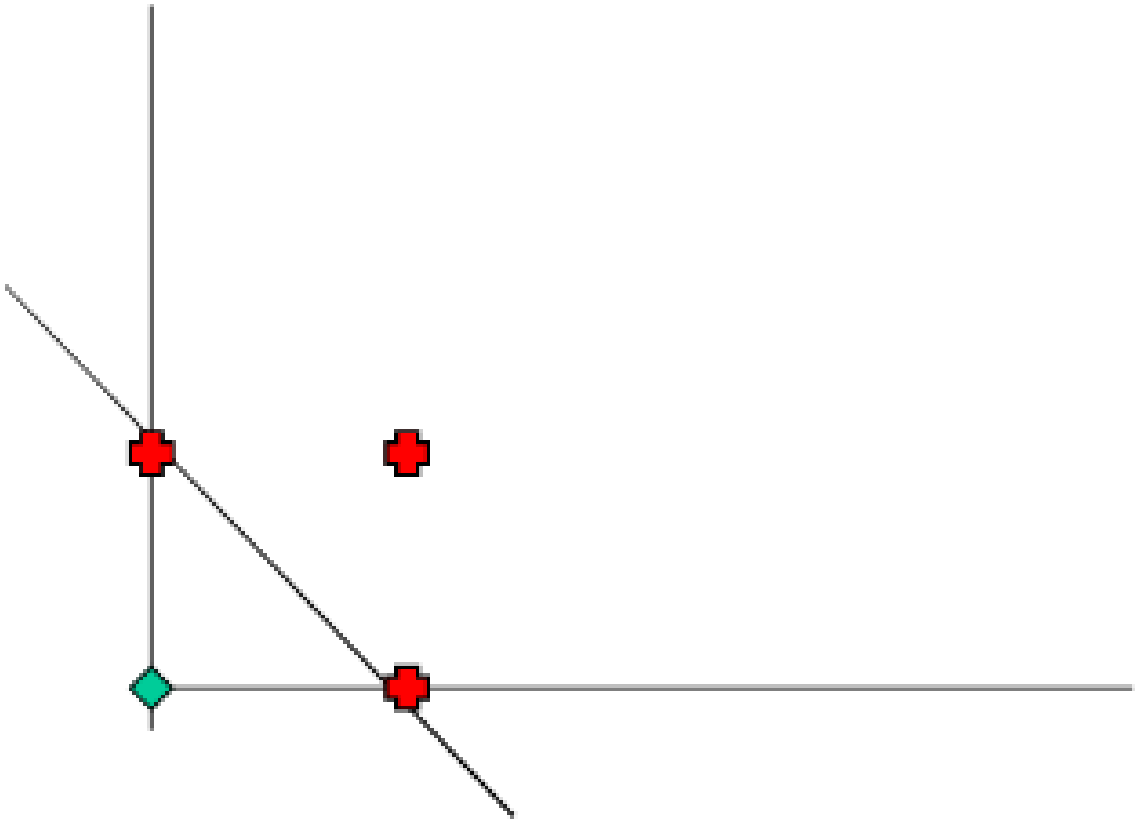
Artificial Neural Neuron

Implementation of OR Function

Let W1=W2=1

X_1	X_2	$X_1W_1 + X_2W_2$
0	0	0
0	1	1
1	0	1
1	1	2

If we make Θ =1 (or any value >0 but <=1), we will get correct results with unit step activation function



Artificial Neural Neuron

If we place the 4 points in a two-coordinate system(X1 andX2), and draw a line from points (0,1) to (1,0) then:

Any new data falling on the left side of the line will give an output of zero and the data on the right side of the line will be classified as one