Lecture - 1

Introduction to Neural Networks

Who am I?

- Assistant Professor at Department of CSE
- **Research areas:** Computer Vision, Language, Knowledge Graphs.
- **PhD from:** IIIT Hyderabad
- **Post Doctorate from:** IISc Bangalore
- **Contact:** mishra@iitj.ac.in
- **Feel free to call me by my first name (Anand)**

In this fractal ...

- **Neural Networks:** From a single neuron to multi-layer perceptron. Learning algorithms, backpropagation, loss function, optimization techniques. (4 Lectures)
- **Deep Neural Network:** CNN and RNN (2 Lectures)
- **SVM** (2 Lecture)

Books: The course will not follow a specific book, but will draw from a number of sources. I will list them in the Google Classroom.

Major resources:

- 11-785 Introduction to Deep Learning (CMU): **Bhiksha Raj**
- CS7015: Deep Learning (IIT-M): **Mitesh Khapra**
- CS231n: CNNs for Visual Recognition: **Fei-Fei Lee et al.**
- http://neuralnetworksanddeeplearning.com
- https://www.deeplearningbook.org
- Duda, Hart, and Strok, **Pattern Classification**
- Neural Networks by Rojas
- …
- Few of my own examples …

In this fractal …

- Two announced short quizzes (Best will be considered) **(3-4% weight)**
- 1 Assignment: Release: **May 9, 2021**, Due: **May 30, 2021 (early bird), June 8, 2021 (with minor penalty). (10% weight)**
- **Major exam: June 5/6 (20% weight)**

Learning outcome

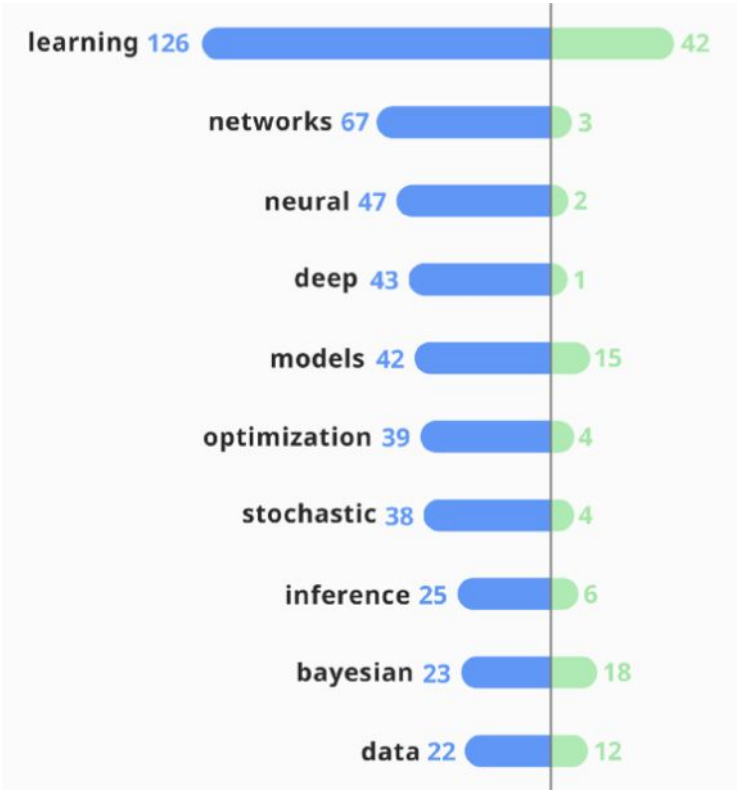
- Few lectures are not enough to be an expert! This is just a beginning!
- Learning to formulate ML problems
- Basic understanding of how Neural Networks work
- Implementing neural networks from scratch
- Familiarity with training

Overall: Set you up for further work in this area, next semester some of these concepts you will lean in much more detail

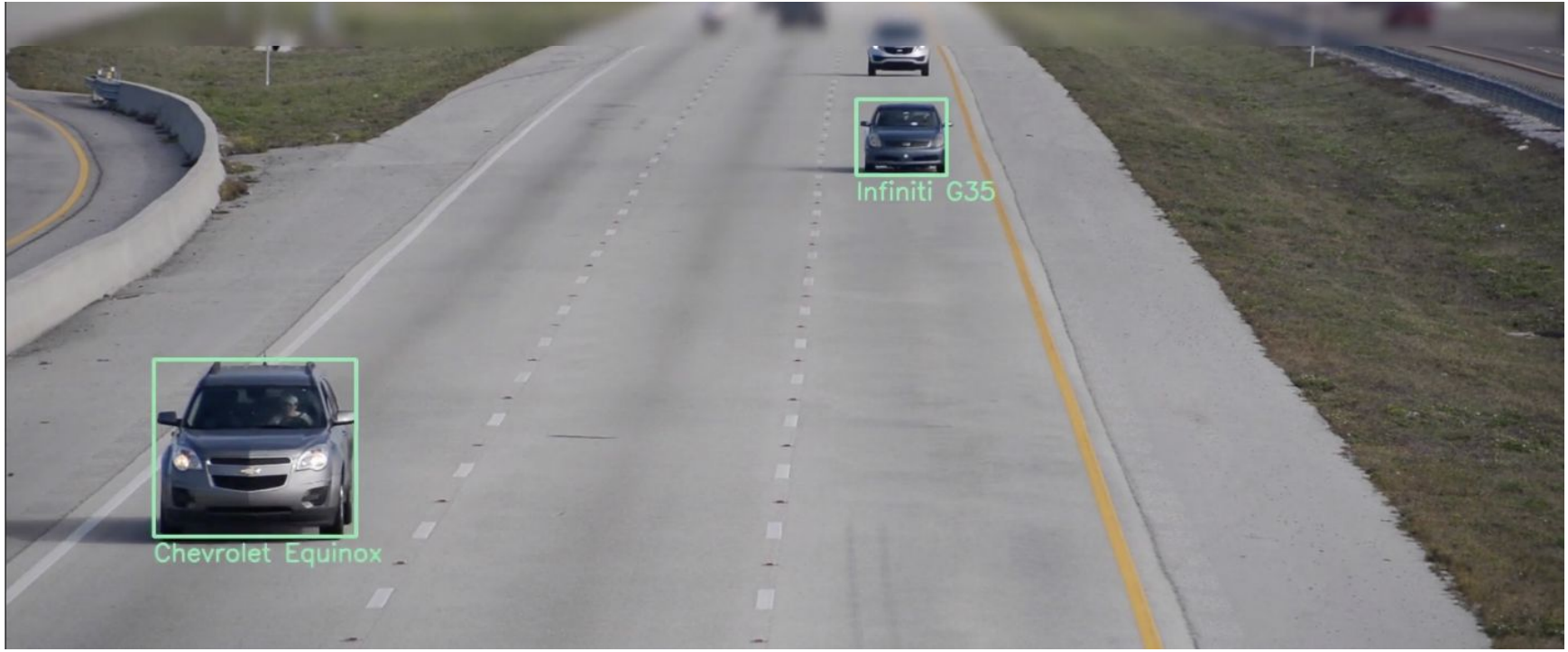
Neural Networks are taking over!

- Major thrust area
- New State of the art in many ML tasks

Neural Networks are taking over!

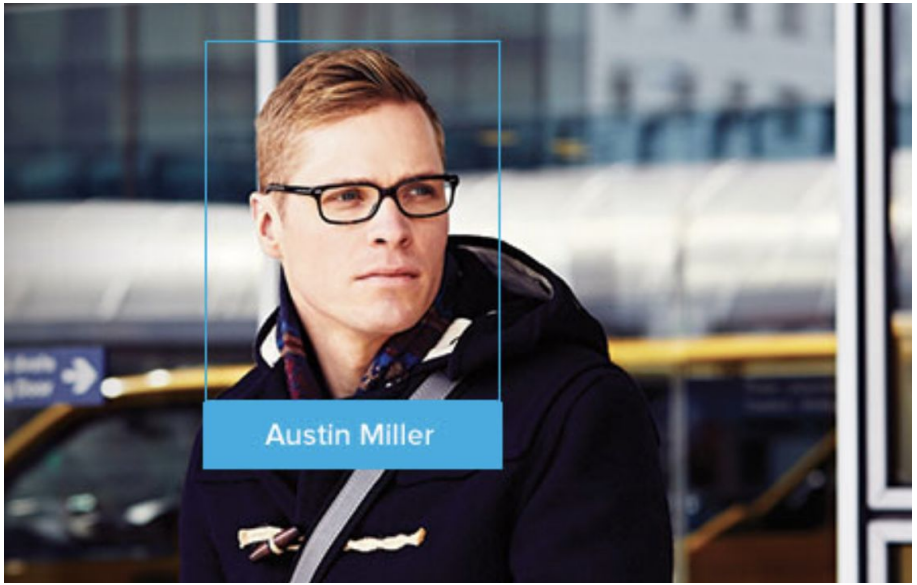


Visual Recognition



https://www.sighthound.com/technology/

Visual Recognition



https://www.sighthound.com/technology/

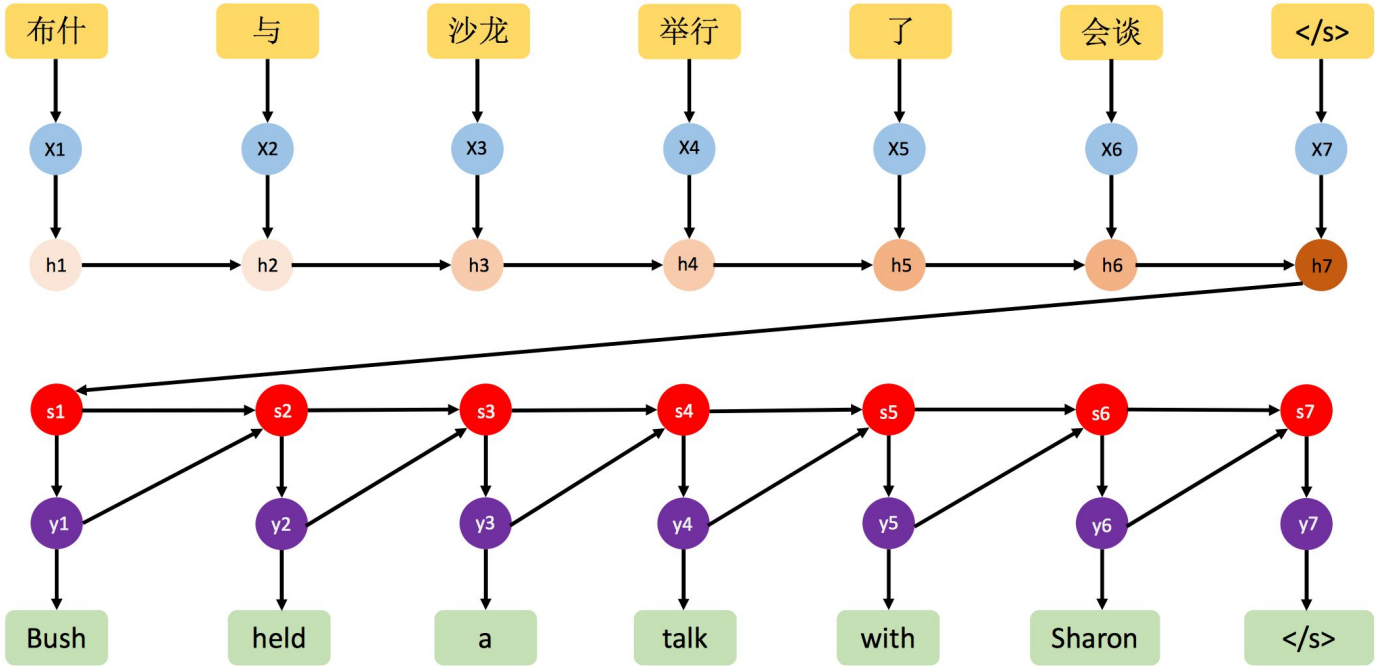
Describing Images

A child in a helmet is riding a bike.

A group of people are walking on a busy street.

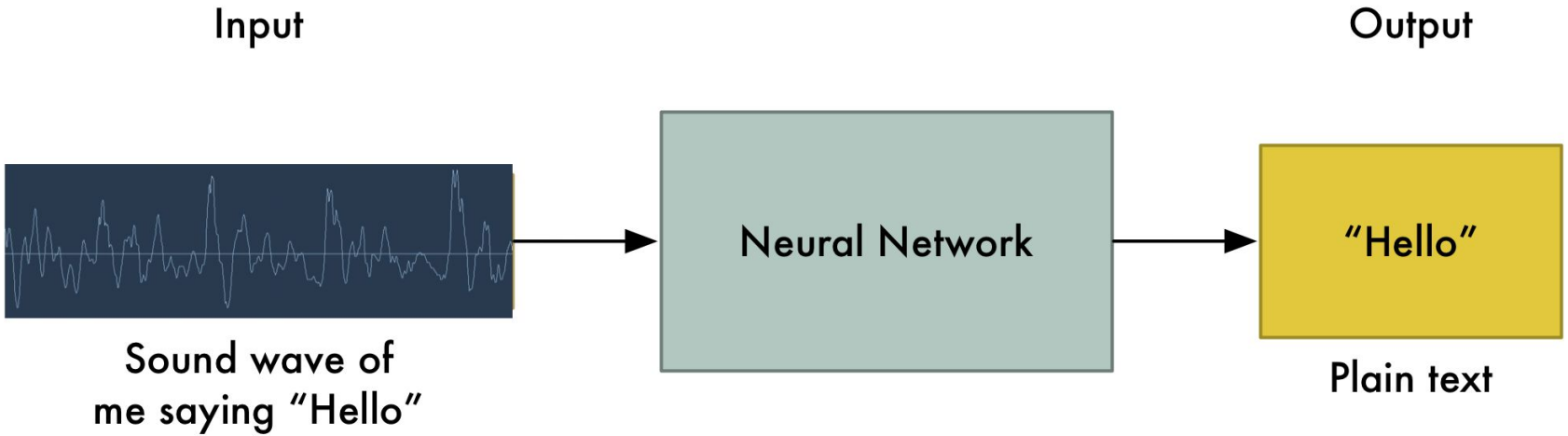
[Vinyals et al., CVPR 2015]

Translating Languages



[Sutskever et al.,NIPS 2014]

Speech Recognition



Neural Networks and Job Market

This guy didn't know about neural networks (a.k.a deep learning)

This guy learned about neural networks (a.k.a deep learning)

slido

Give some more applications of Neural Networks?

ⓘ Start presenting to display the poll results on this slide.

Outline (today)

- Module 1: History
- Module 2: Biological Neuron
- Module 3: M-P Neuron
- Module 4: Perceptron

Module 1 (History)

History of Neural Networks

Breakthrows	Year	People associated
Use of term "Neuron"	1891	H. W. G. von Waldeyer-Hartz
MP Neuron (Computation model)	1943	McCullah and Pitts
Perceptron Model	1957-58	Frank Rosenbellt

Perceptron may eventually may able to learn, make decisions and translate languages.

- Frank Rosenblatt (1957-1958)

History of Neural Networks

Breakthrows	Year	People associated
Multi Layer Perceptron (early deep learning model)	1965-68	Alexey Ivakhnenko
Limits of Perceptron	1969	Minsky and Papert
Gradient Descent	1986	Cauchy

History of Neural Networks

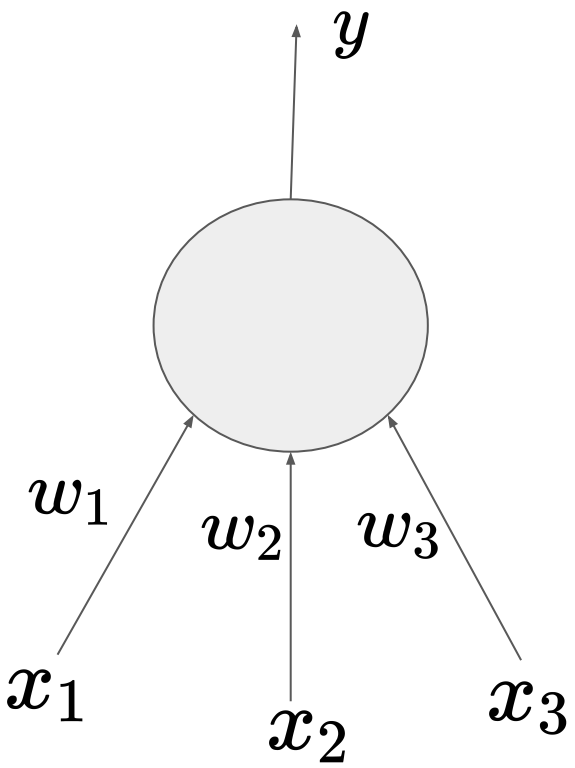
Breakthrows	Year	People associated
Universal Approximation Theorem	1989	George Cybenko
Unsupervised Pre training	2006	Hinton & Salakhutdinov
Handwriting, speech, GPUs	2009-2010	Graves, ...

History of Neural Networks

Breakthrows	Year	People associated
Visual Recognition (Alexnet): 16% error in imagenet	2012	Alex Krizhevsky, Ilya Sutskever and Hinton
Visual Recognition (VGG net)	2014	Simonyan and Zisserman
Visual Recognition (Resnet): 3.6% error in imagnet, better than human!	2016	Kaiming He et al.

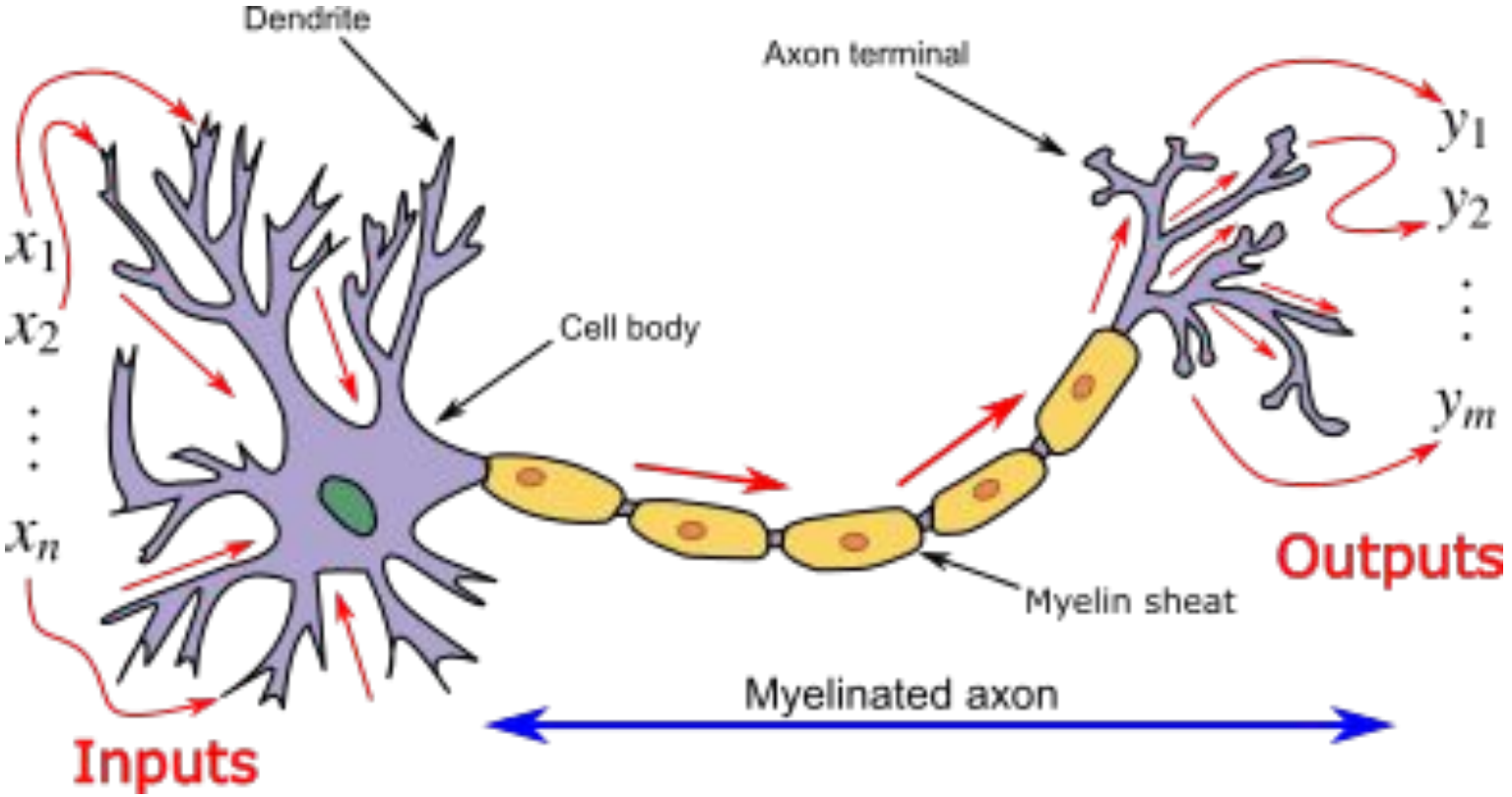
Module 2: Biological Neurons

Artificial Neurons

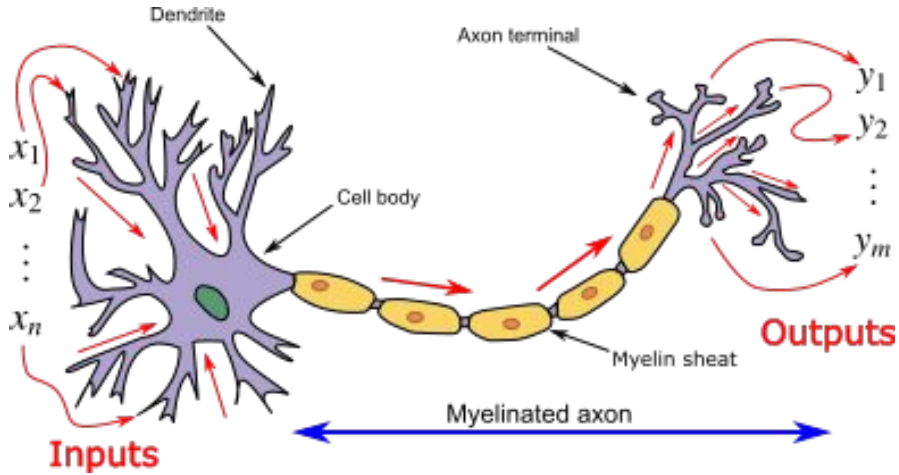


Why we call them neurons?

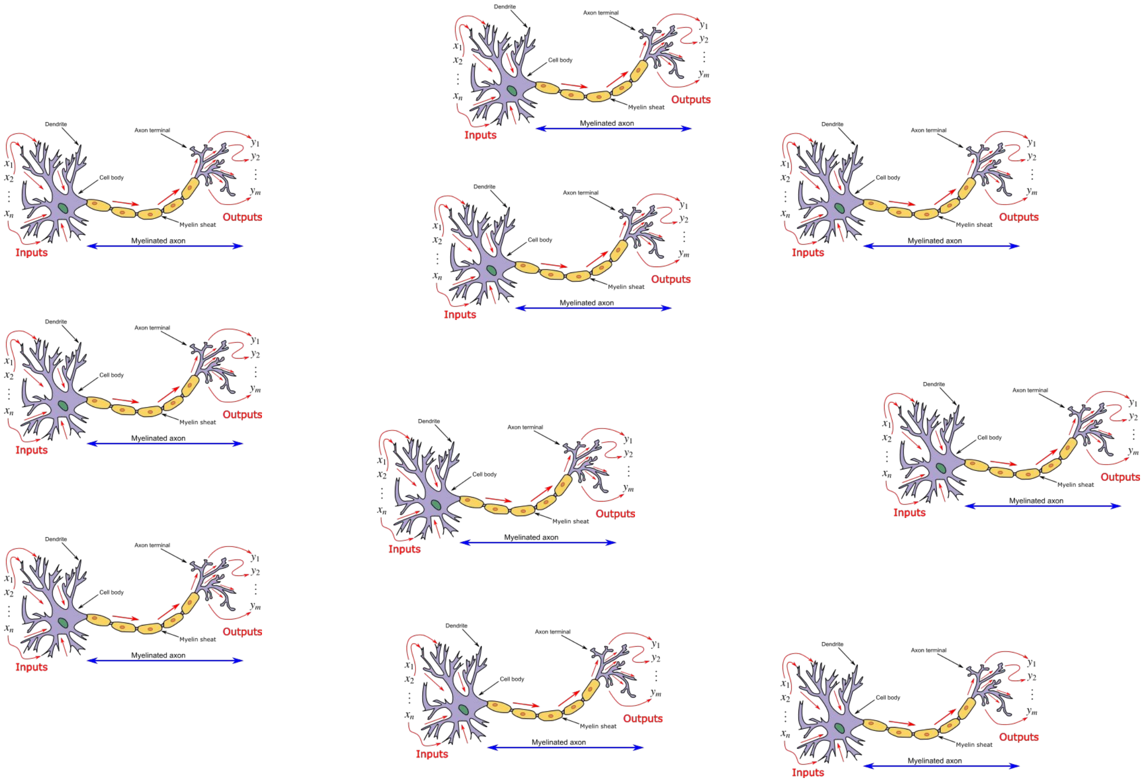
Biological Neurons



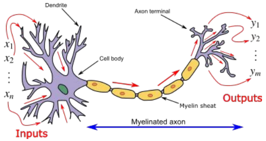
How biological neuron works?

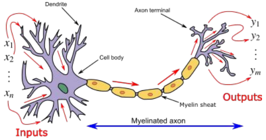


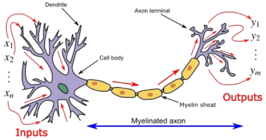
Massively Parallel Neurons

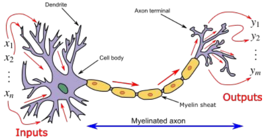


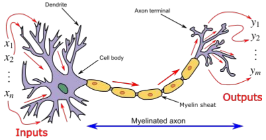
Massively Parallel Neurons

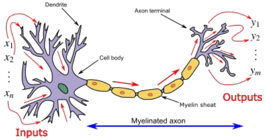


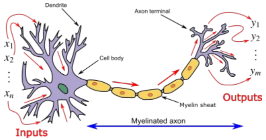


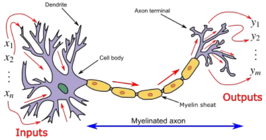


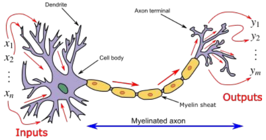


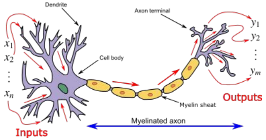




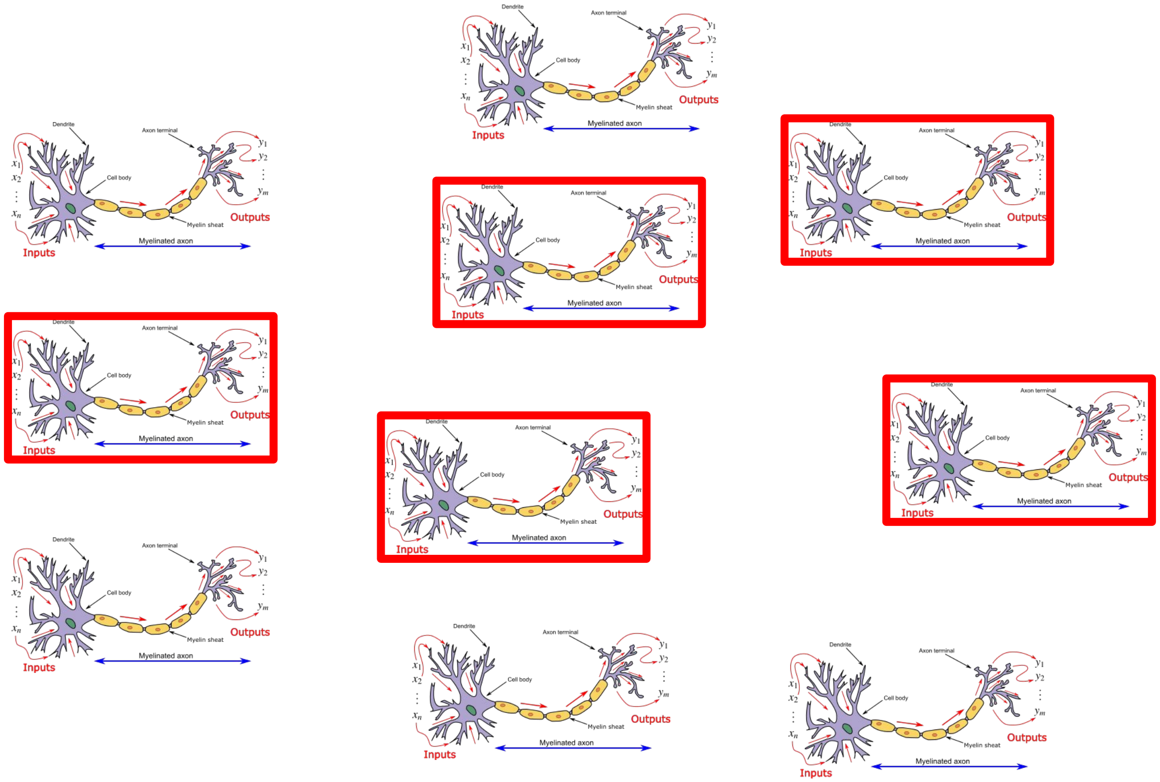




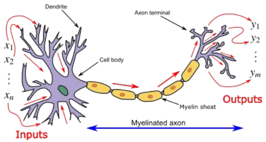


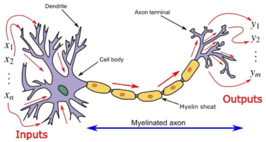


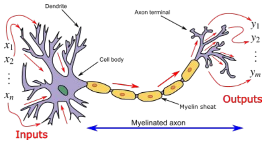
Massively Parallel Neurons

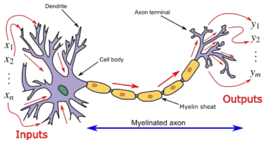


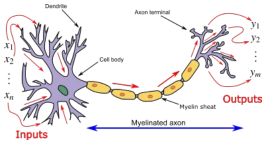
Massively Parallel Neurons

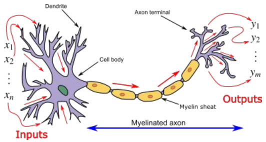


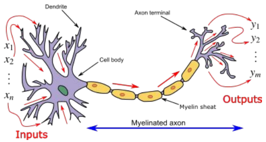


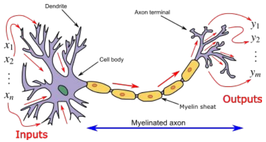


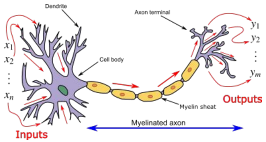


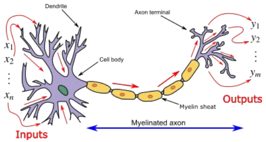




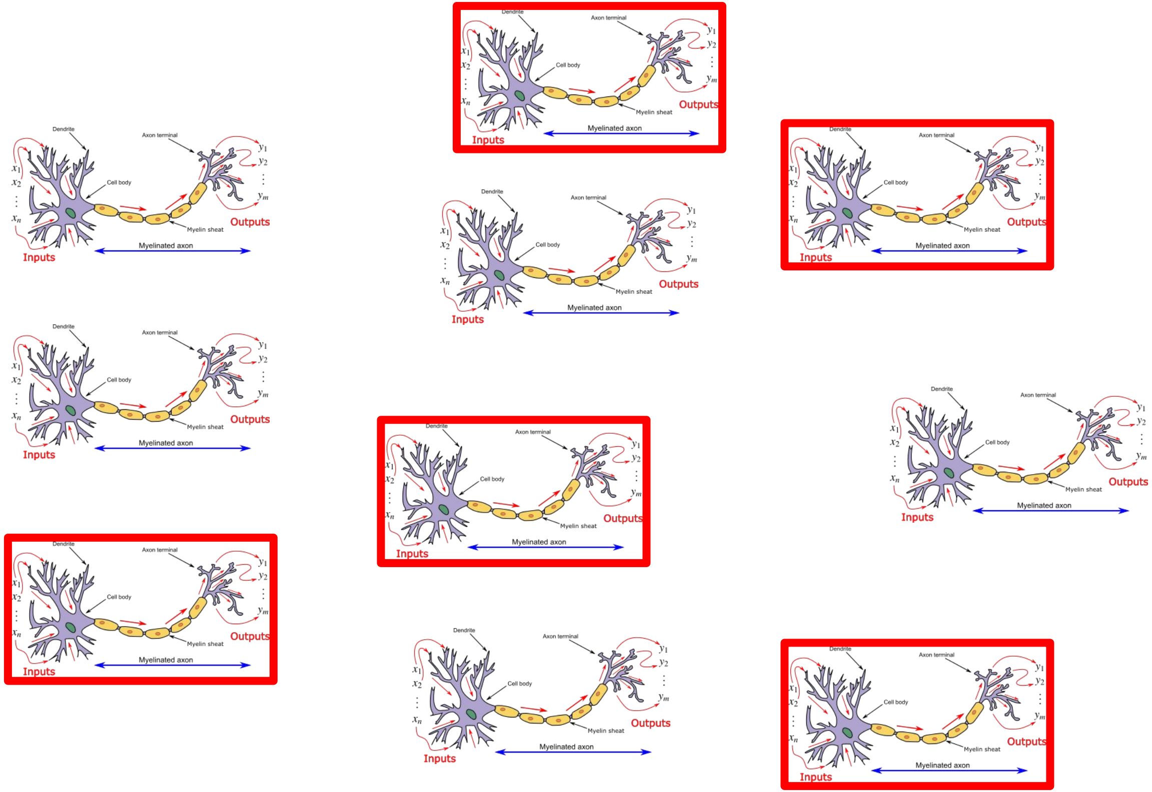






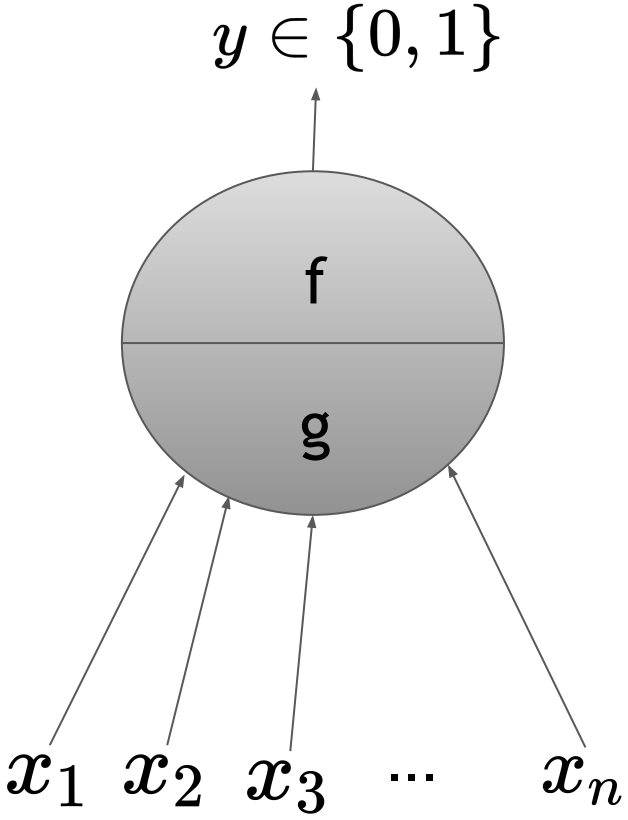


Massively Parallel Neurons



Module 3: McCulloch-Pitts Neuron

Highly simplified Computation Model



- Boolean input and Boolean output
- g aggregates input and f takes decision

$$g(\mathbf{x}) = \sum_i^n x_i$$

$$y = f(g(x)) = 1 \; if \; g(x) >= \theta$$

$$y = f(g(x)) = 0 \; if \; g(x) < \theta$$

Example

Estimating crowd in a famous temple

x_1 : Weather (Good/Bad)

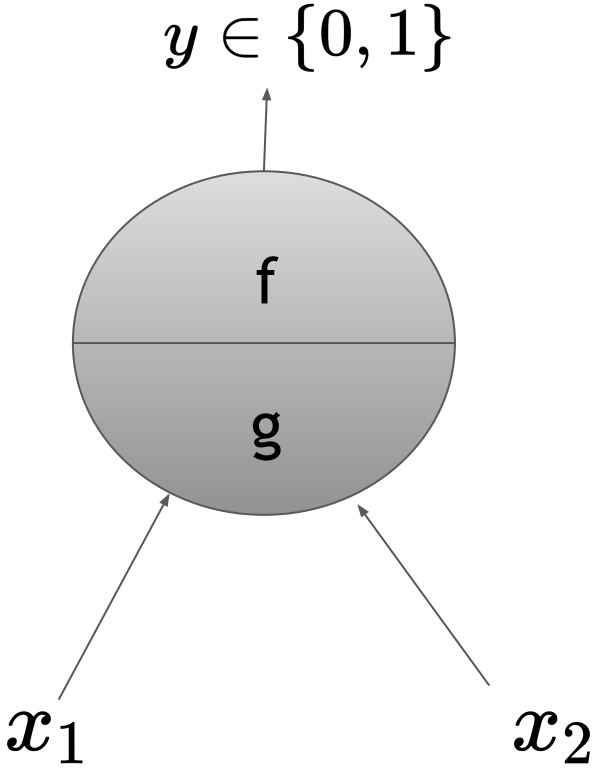
x_2 : Festival (Yes/No)

x_3 : Vacation (Yes/No)

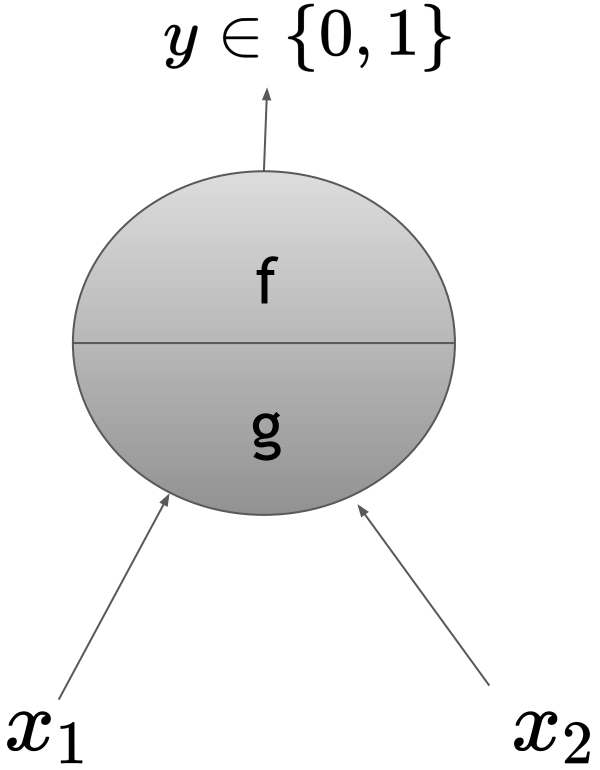
Implementing Boolean functions using MP neuron

Mapping $\sum_{i=1}^{n} x_i$ to a Boolean function

AND



AND

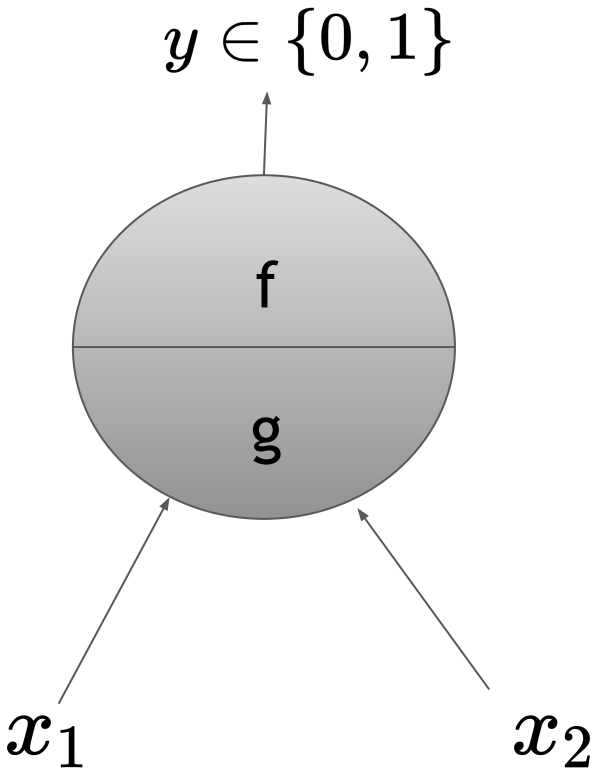


$$g(x) = x_1 + x_2$$

$$y = f(g(x)) = 1 \; if \; g(x) >= \theta$$

$$y = f(g(x)) = 0 \; if \; g(x) < \theta$$

AND



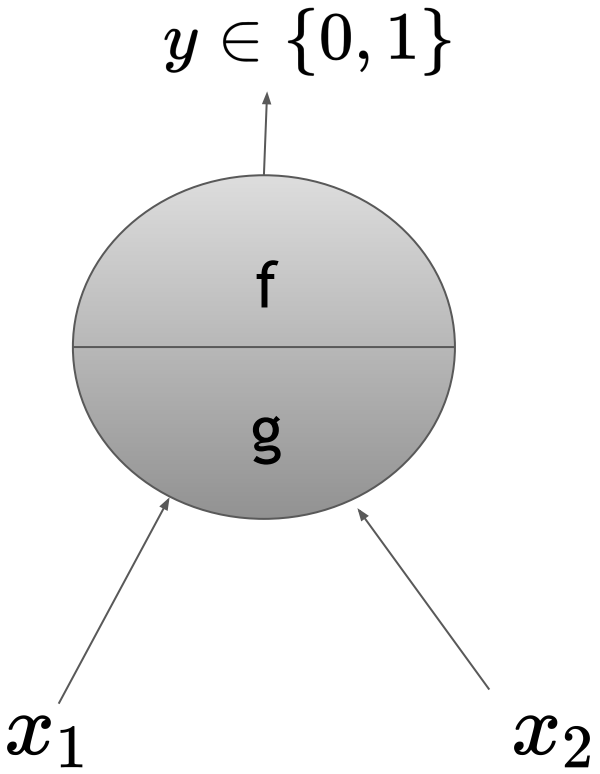
$$g(x) = x_1 + x_2$$

$$y = f(g(x)) = 1 \; if \; g(x) >= \theta$$

$$y = f(g(x)) = 0 \; if \; g(x) < \theta$$

x_1	x_2
0	0
0	1
1	0
1	1

AND



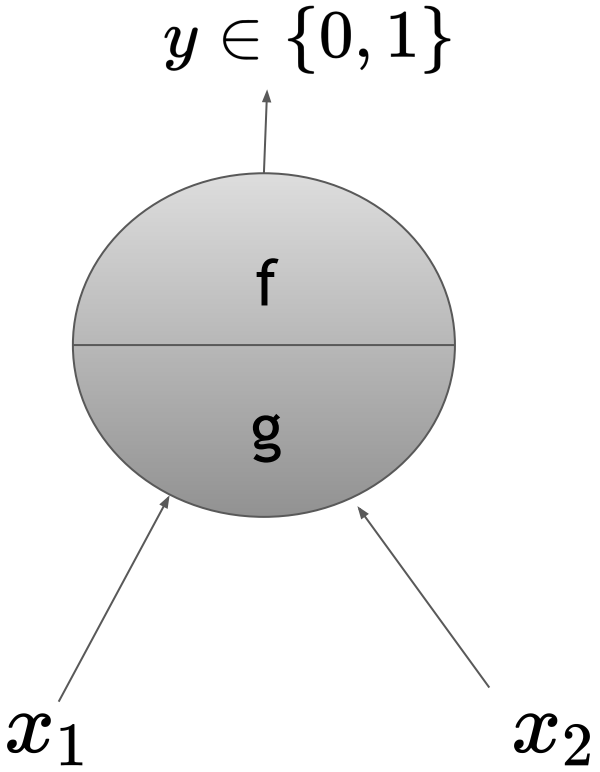
$$g(x) = x_1 + x_2$$

$$y = f(g(x)) = 1 \; if \; g(x) >= \theta$$

$$y = f(g(x)) = 0 \; if \; g(x) < \theta$$

x_1	x_2	g(x)
0	0	0
0	1	1
1	0	1
1	1	2

AND



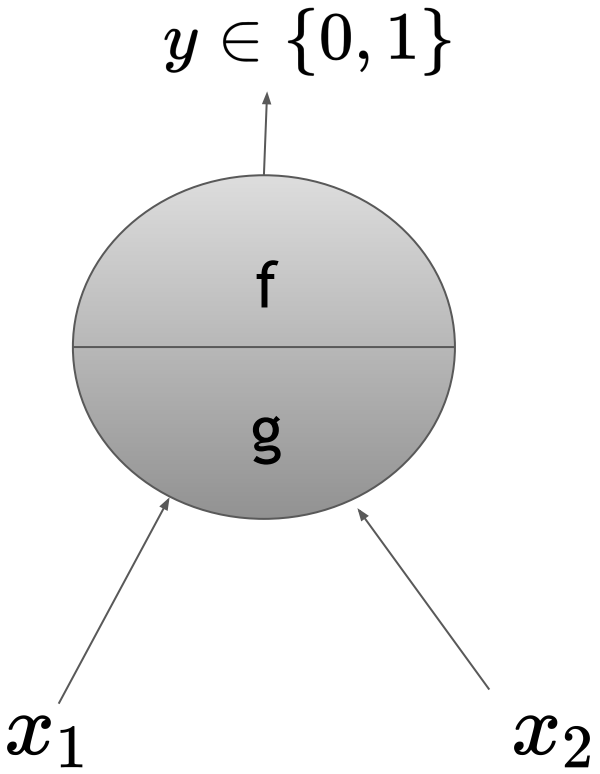
$$g(x) = x_1 + x_2$$

$$y = f(g(x)) = 1 \; if \; g(x) >= \theta$$

$$y = f(g(x)) = 0 \; if \; g(x) < \theta$$

x_1	x_2	g(x)	y= AND (x_1, x_2)
0	0	0	0
0	1	1	0
1	0	1	0
1	1	2	1

AND



$$g(x) = x_1 + x_2$$

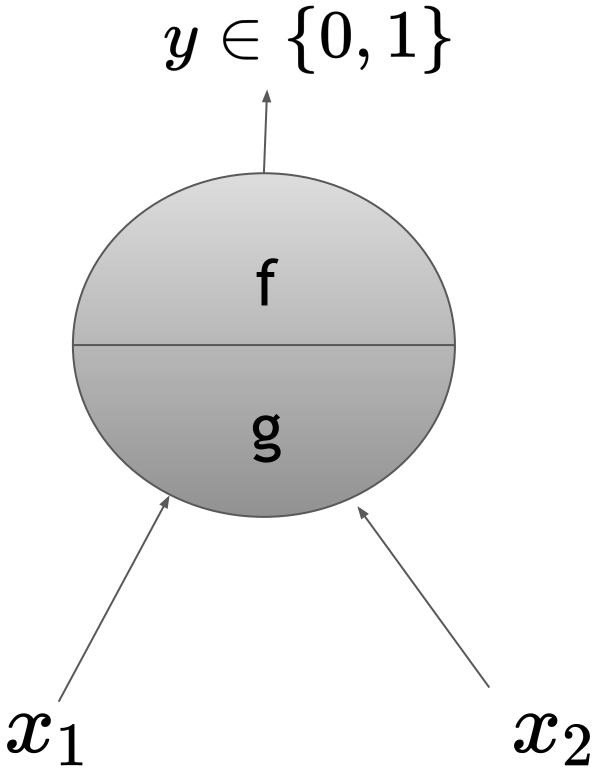
$$y = f(g(x)) = 1 \; if \; g(x) >= \theta$$

$$y = f(g(x)) = 0 \; if \; g(x) < \theta$$

x_1	x_2	g(x)	y= AND (x_1, x_2)
0	0	0	0
0	1	1	0
1	0	1	0
1	1	2	1

Q: When $y = 1$, then $g(x) >=$ to what?

AND



$$g(x) = x_1 + x_2$$

$$y = f(g(x)) = 1 \; if \; g(x) >= \theta$$

$$y = f(g(x)) = 0 \; if \; g(x) < \theta$$

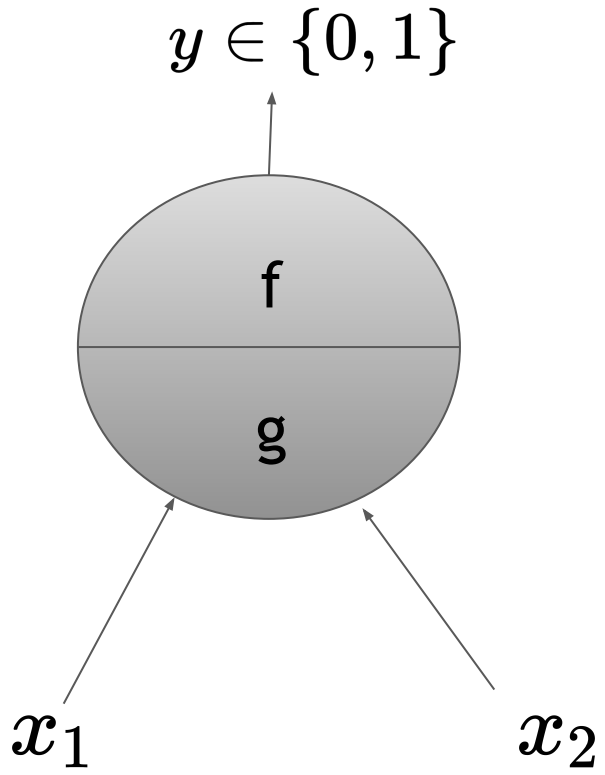
x_1	x_2	g(x)	y= AND (x_1, x_2)
0	0	0	0
0	1	1	0
1	0	1	0
1	1	2	1

slido

$$g(x) >=$$

ⓘ Start presenting to display the poll results on this slide.

AND



$y = 1$ When $x_1 + x_2 >= 2$

$y = 0$ When $x_1 + x_2 < 2$

$$g(x) = x_1 + x_2$$

$$y = f(g(x)) = 1 \; if \; g(x) >= \theta$$

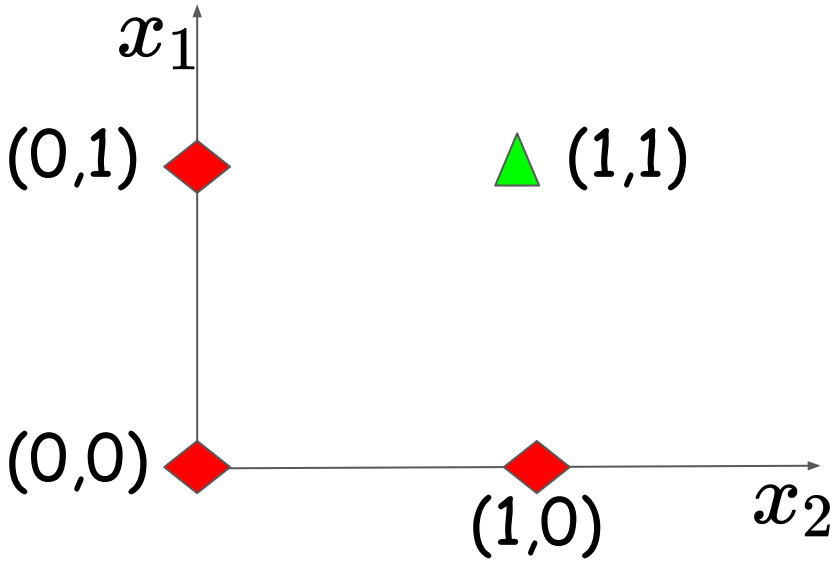
$$y = f(g(x)) = 0 \; if \; g(x) < \theta$$

x_1	x_2	g(x)	y= AND (x_1, x_2)
0	0	0	0
0	1	1	0
1	0	1	0
1	1	2	1

AND

$y = 1$ When $x_1 + x_2 >= 2$

$y = 0$ When $x_1 + x_2 < 2$



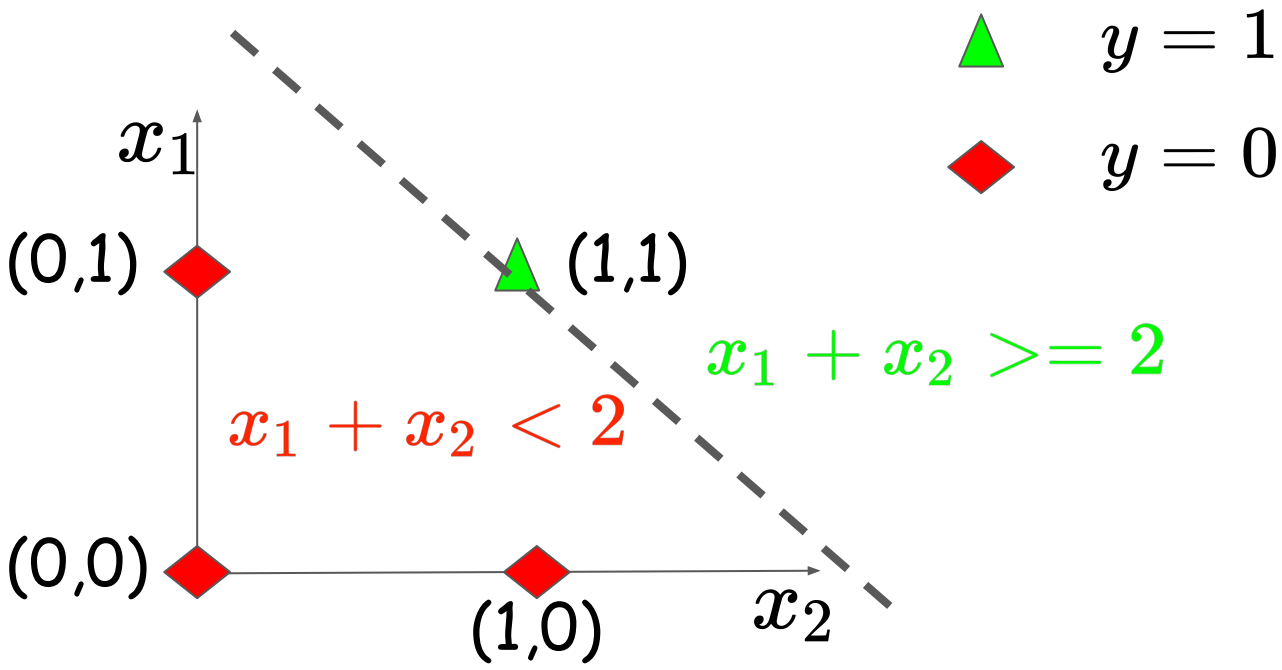
$y = 1$

$y = 0$

AND

$y = 1$ When $x_1 + x_2 >= 2$

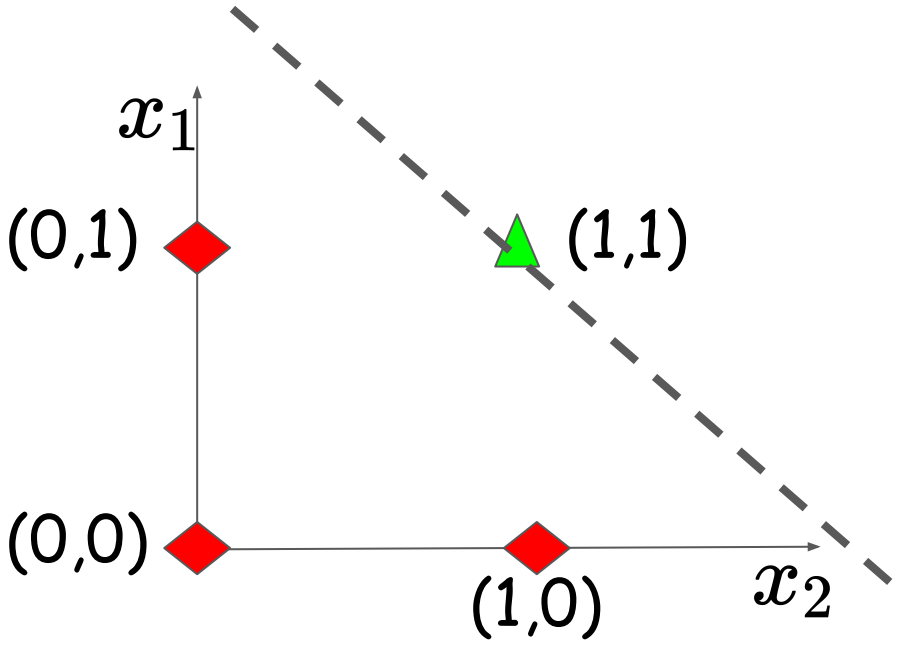
$y = 0$ When $x_1 + x_2 < 2$



AND

$y = 1$ When $x_1 + x_2 >= 2$

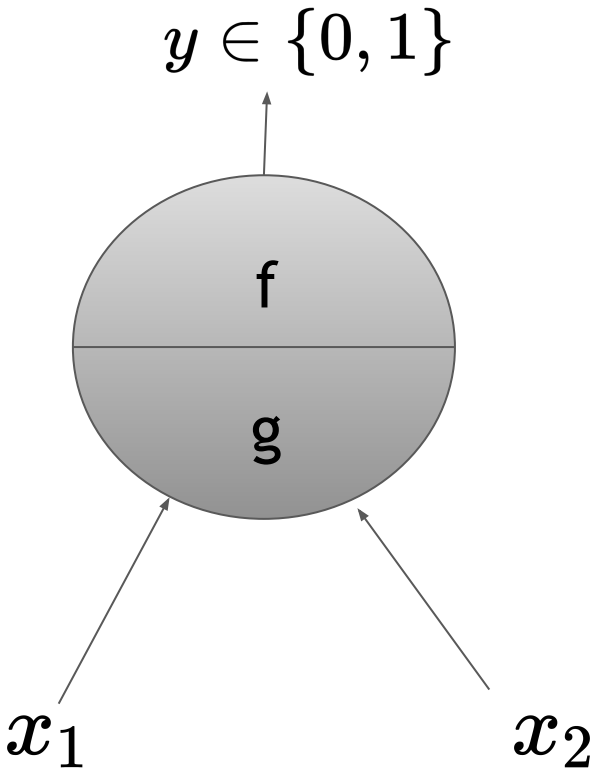
$y = 0$ When $x_1 + x_2 < 2$



$y = 1$

$y = 0$

OR



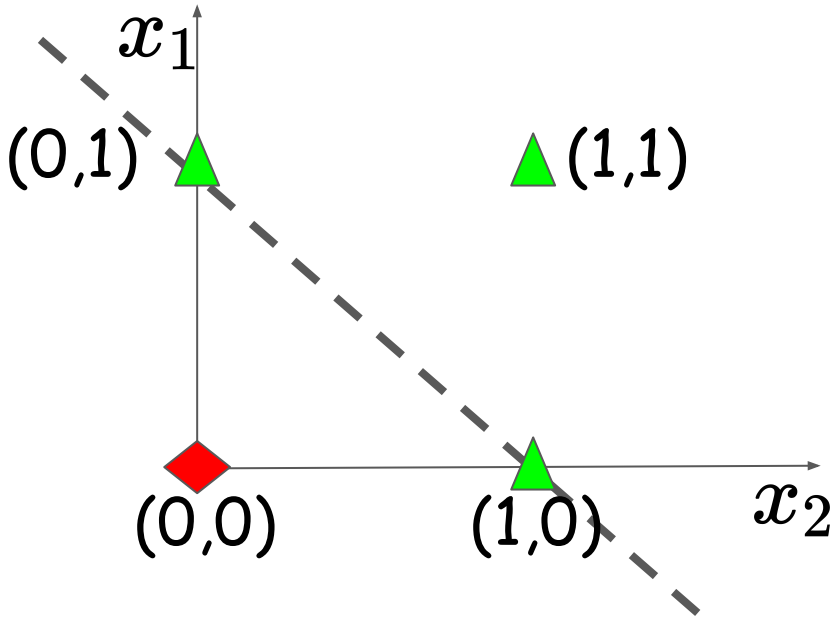
$$g(x) = x_1 + x_2$$

$$y = f(g(x)) = 1 \; if \; g(x) >= \theta$$

$$y = f(g(x)) = 0 \; if \; g(x) < \theta$$

x_1	x_2	g(x)	OR (x_1 , x_2)
0	0	0	0
0	1	1	1
1	0	1	1
1	1	2	1

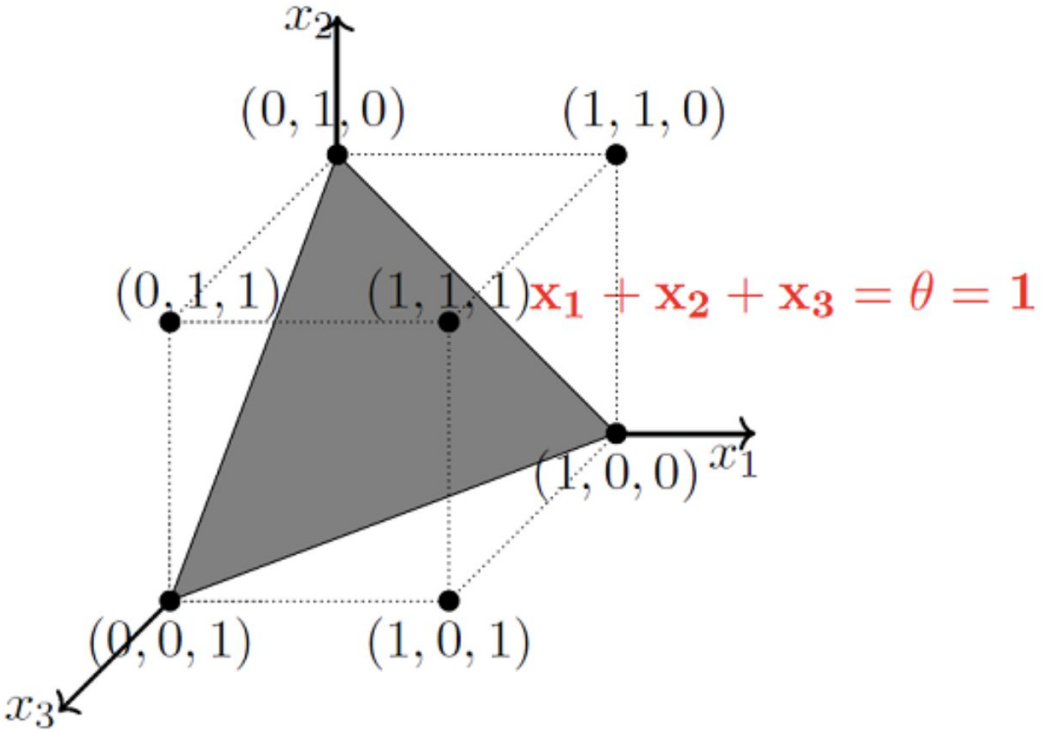
OR



$y = 1$

$y = 0$

More than 2-dimensions



Can any Boolean function be represented using MP neuron?

slido

What is the domain and range of a Boolean Function?

ⓘ Start presenting to display the poll results on this slide.

slido

Is this a Boolean function: 5x+6y where x and y are either 0 or 1.

ⓘ Start presenting to display the poll results on this slide.

Arbitrary Boolean function

x_1	x_2	x_3	$g(x)$	y
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	2	1
1	0	0	1	0
1	0	1	2	1
1	1	0	2	1
1	1	1	3	1

$$g(x) = x_1 + x_2 + x_3$$

$$y = \overline{x_1} x_2 x_3 + x_1 \overline{x_2} x_3 + x_1 x_2 \overline{x_3} + x_1 x_2 x_3$$

Arbitrary Boolean function

x_1	x_2	x_3	$g(x)$	y
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	2	1
1	0	0	1	0
1	0	1	2	1
1	1	0	2	1
1	1	1	3	1

$$g(x) = x_1 + x_2 + x_3$$

$$y = \overline{x_1} x_2 x_3 + x_1 \overline{x_2} x_3 + x_1 x_2 \overline{x_3} + x_1 x_2 x_3$$

Arbitrary Boolean function

x_1	x_2	x_3	$g(x)$	y
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	2	1
1	0	0	1	0
1	0	1	2	1
1	1	0	2	1
1	1	1	3	1

$$g(x) = x_1 + x_2 + x_3$$

$$y = \overline{x_1} x_2 x_3 + x_1 \overline{x_2} x_3 + x_1 x_2 \overline{x_3} + x_1 x_2 x_3$$

What will be θ?

Arbitrary Boolean function

x_1	x_2	x_3	$g(x)$	y
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	2	1
1	0	0	1	0
1	0	1	2	1
1	1	0	2	1
1	1	1	3	1

$$g(x) = x_1 + x_2 + x_3$$

$$y = \overline{x_1} x_2 x_3 + x_1 \overline{x_2} x_3 + x_1 x_2 \overline{x_3} + x_1 x_2 x_3$$

What will be decision boundary?

Arbitrary Boolean function

x_1	x_2	x_3	$g(x)$	y
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	2	1
1	0	0	1	0
1	0	1	2	1
1	1	0	2	1
1	1	1	3	1

$$g(x) = x_1 + x_2 + x_3$$

$$y = \overline{x_1} x_2 x_3 + x_1 \overline{x_2} x_3 + x_1 x_2 \overline{x_3} + x_1 x_2 x_3$$

What will be decision boundary?

$$x_1 + x_2 + x_3 >= 2$$

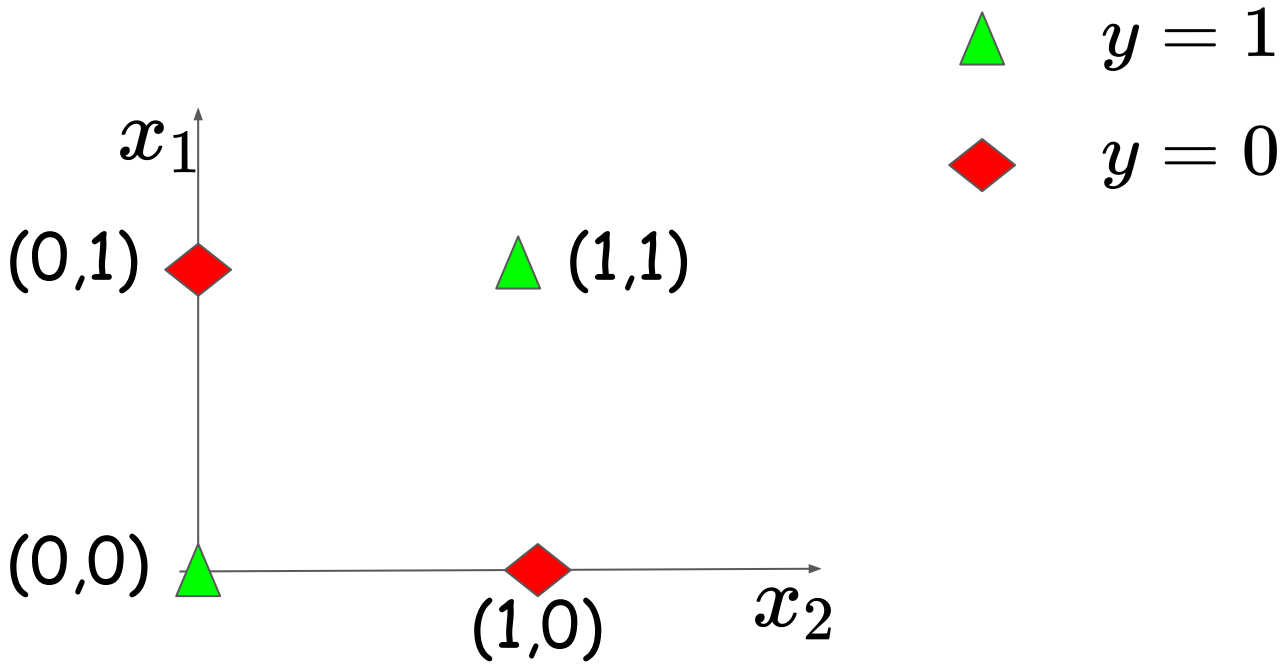
Arbitrary Boolean function

x_1	x_2	x_3	$g(x)$	y
0	0	0	0	1
0	0	1	1	0
0	1	0	1	1
0	1	1	2	0
1	0	0	1	1
1	0	1	2	0
1	1	0	2	1
1	1	1	3	1

$$g(x) = x_1 + x_2 + x_3$$

$$y = x_1 . x_2 + \overline{x_3}$$

X-OR



slido

Can you find a single line which separates green triangles to red diamonds?

ⓘ Start presenting to display the poll results on this slide.

Module 4: Perceptron

Three Qs

Q1. Does it work only for Boolean inputs?

Q2. Do we manually decide threshold?

Q3. Are all inputs equal?

1 Does it work only for Boolean inputs?

What if:

Example 1: Loan Approval by a bank using ML

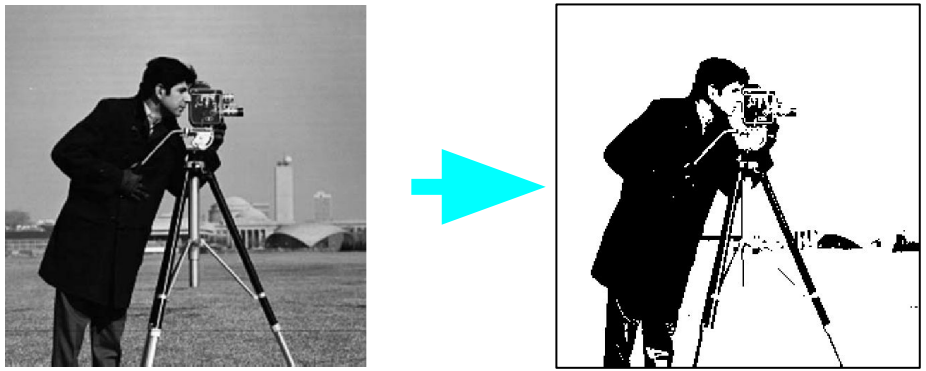
Inputs: Job status, total debt, account balance, average monthly savings, gender, age, residential status, ….

1 Does it work only for Boolean inputs?

What if:

Example 2: Thresholding a gray image

Inputs: Pixel values (0-255)



2. Do we manually decide threshold?

NO

What if we have an image of size 512 X 512 as input. The truth table will have $2^{512\times512}$ enteris!

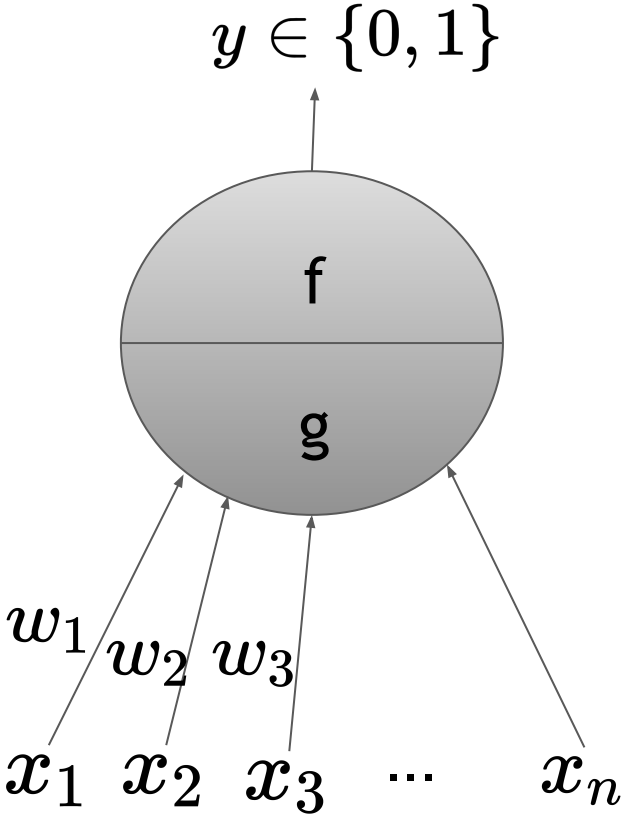
3. Are all input equal?

NO

Example 1: Loan Approval by a bank using ML

Inputs: Job status, total debt, account balance, average monthly savings, gender, age, residential status, ….

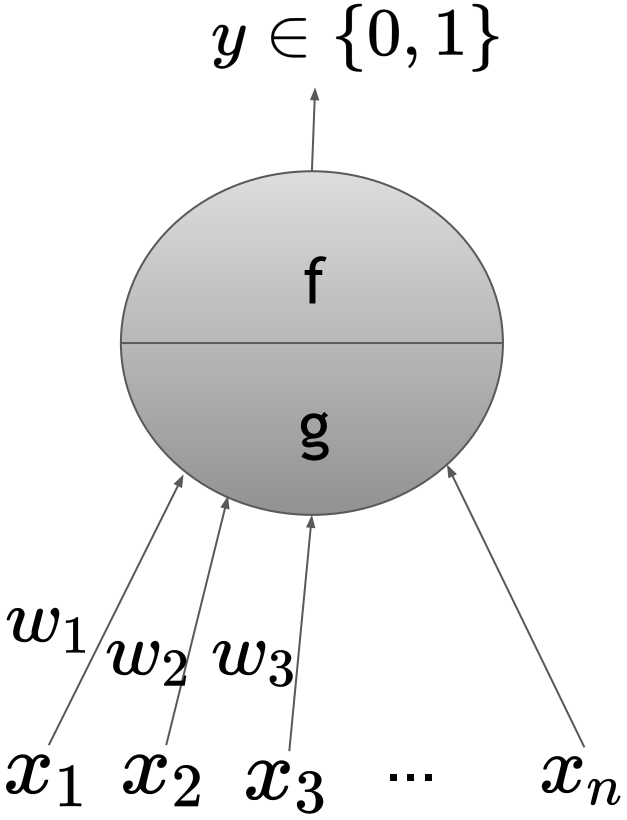
Perceptron



$$y = 1 \;\; if \;\; \sum_1^n w_i x_i >= \theta$$

$$y = 0 \;\; if \;\; \sum_1^n w_i x_i < \theta$$

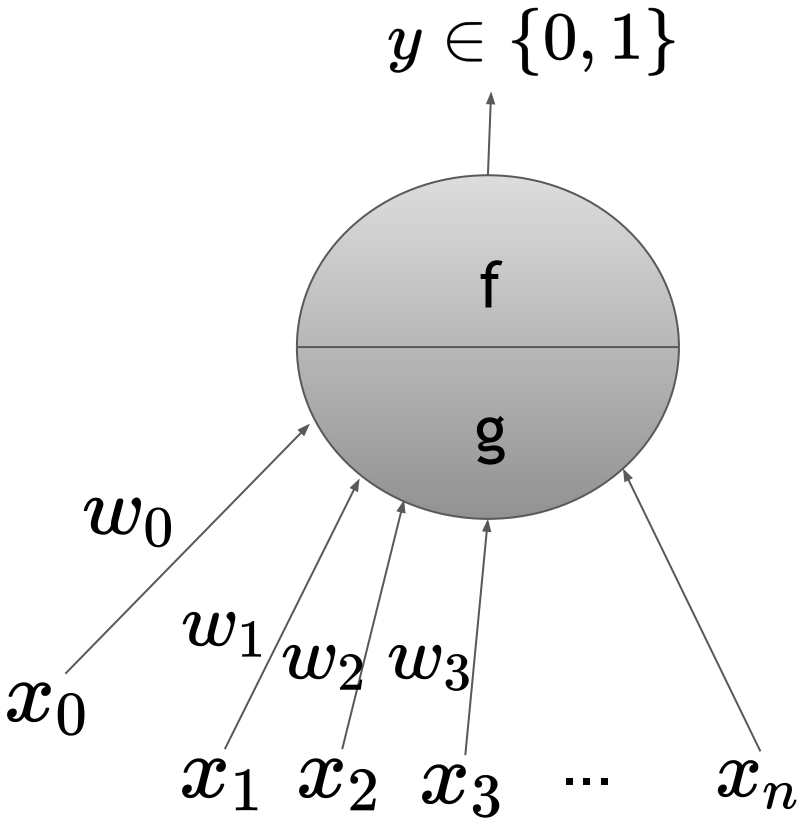
Perceptron



$$y = 1 \;\; if \;\; \sum_1^n w_i x_i - \theta >= 0$$

$$y = 0 \;\; if \;\; \sum_1^n w_i x_i - \theta < 0$$

Perceptron



$$y = 1 \;\; if \;\; \sum_1^n w_i x_i - \theta >= 0$$

$$y = 0 \;\; if \;\; \sum_1^n w_i x_i - \theta < 0$$

$$w_0 = -\theta \;\; and \;\; x_0 = 1$$

$$y = 1 \; if \; \sum_0^n w_i x_i >= 0$$

$$y = 0 \; if \; \sum_0^n w_i x_i < 0$$

Example:

x_1	x_2	y = x_1 AND x_2	Conditions
0	0	0	$w_0 + w_1 x_1 + w_2 x_2 < 0$
0	1	0	$w_0 + w_1 x_1 + w_2 x_2 < 0$
1	0	0	$w_0 + w_1 x_1 + w_2 x_2 < 0$
1	1	1	$w_0 + w_1 x_1 + w_2 x_2 >= 0$

Example:

x_1	x_2	y = x_1 AND x_2	Conditions
0	0	0	$w_0 < 0$
0	1	0	$w_0 + w_2 < 0$
1	0	0	$w_0 + w_1 < 0$
1	1	1	$w_0 + w_1 + w_2 >= 0$

Let us see it geometrically in paper and pen!