Fundamentals of Artificial Neural Networks

Agenda

Intro

History

Neural Networks

Perceptrons

Outlook

Outlook

Multiple sessions (3) on Artificial Neural Networks (ANNs)

- Today: basic ANNs: perceptrons
- Next: learning and more complex ANNs
- Next: applied ANNs

Intro

Computer Science (MSc)

Teaching Mathematics (MSc)

Windesheim

Cognitive Psychology (PhD)

Xsens

Windesheim + UT + Saxion

History

1943-1956 Mathematical foundations

1952-1969 Discovery

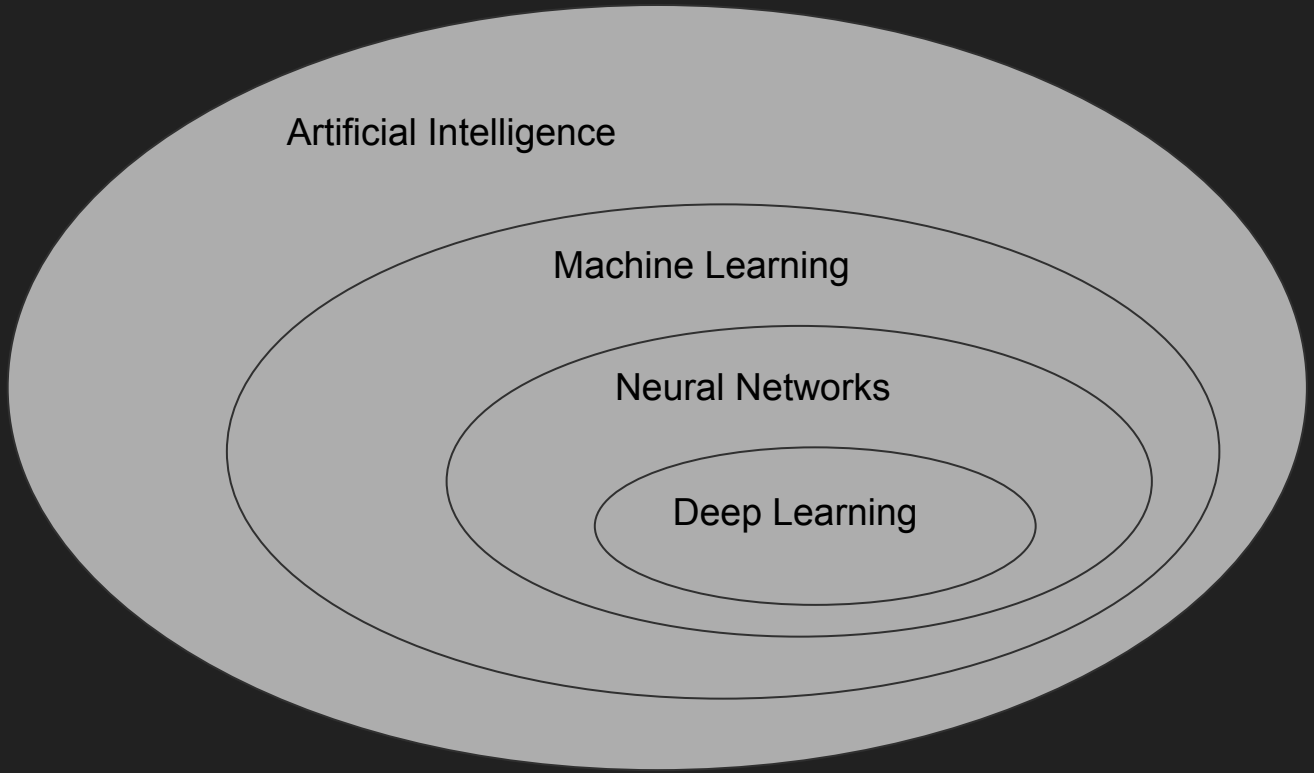
1966-1974 Reality check

1969-1979 Knowledge based systems

1987-2000 Return of Neural Networks

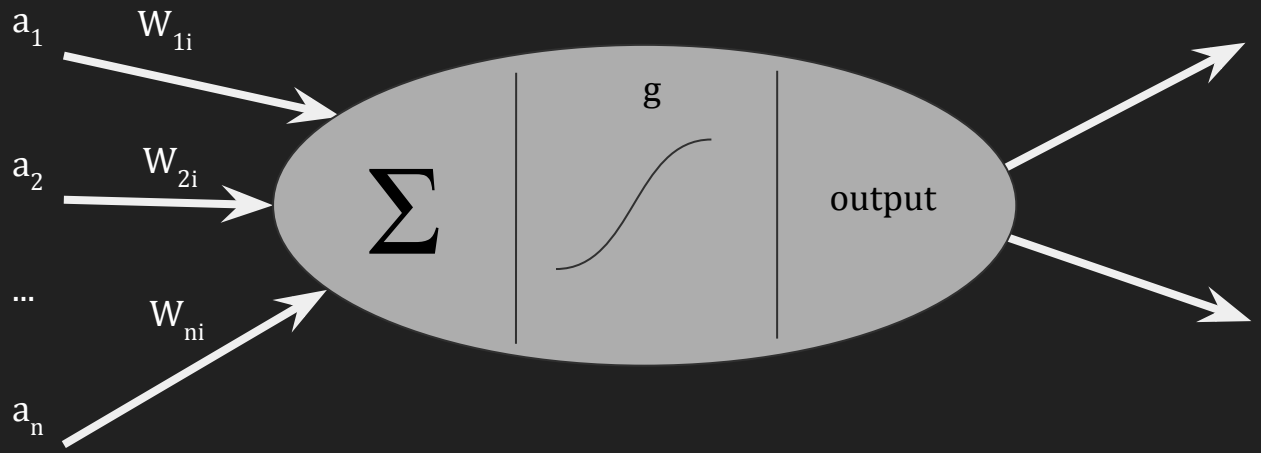
2015-now Deep Learning Hype

The bigger picture

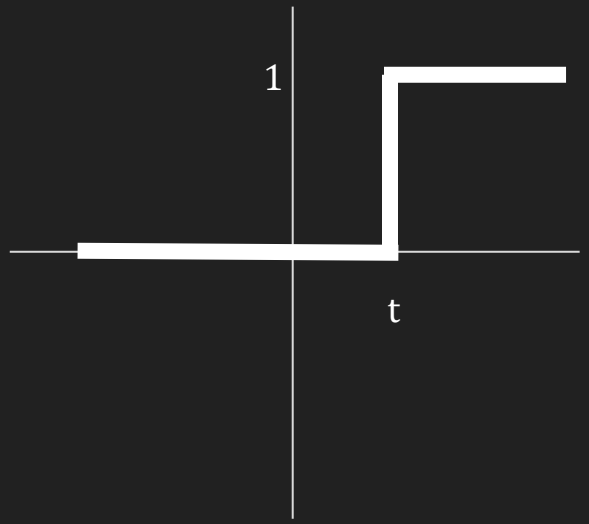


ANN concepts

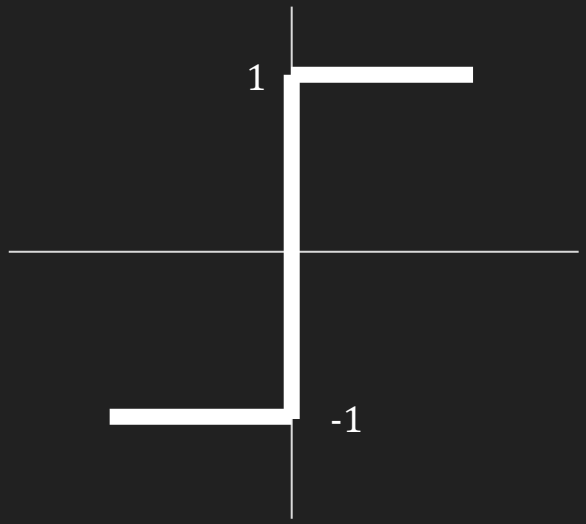
- Node
- Link
- Input
- Weight
- Activation function
- Output
- Layer
- “Knowledge”



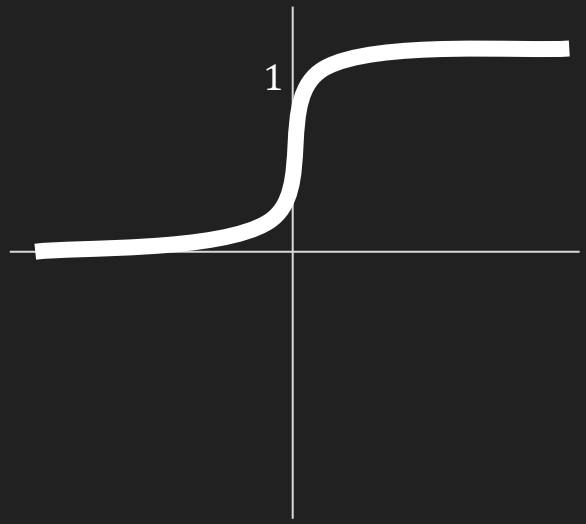
Common Activation Functions



Step function

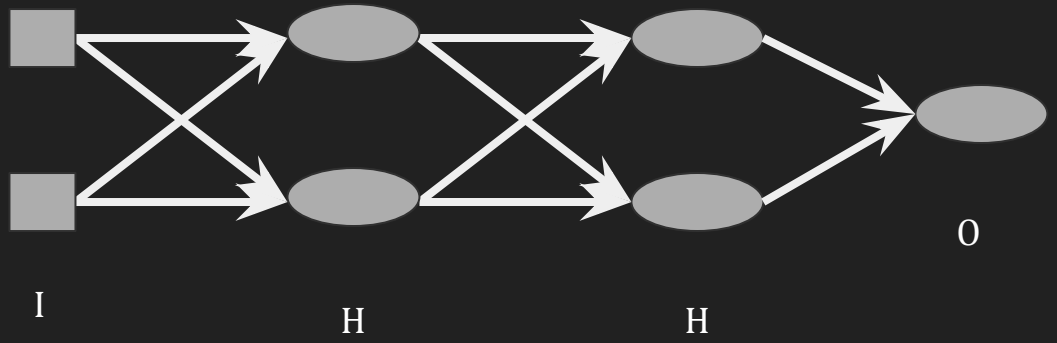


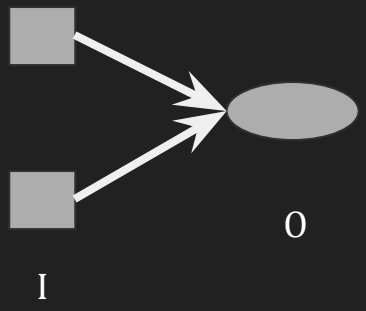
Sign function

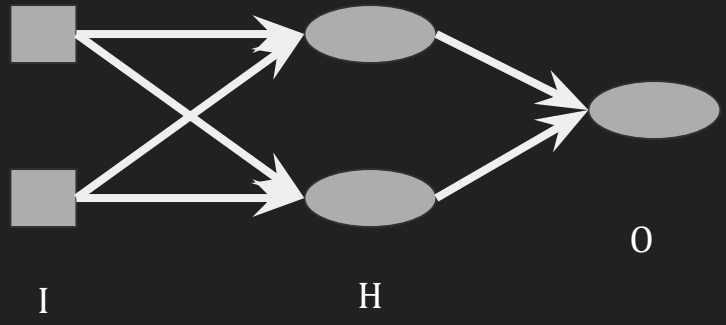


Sigmoid function

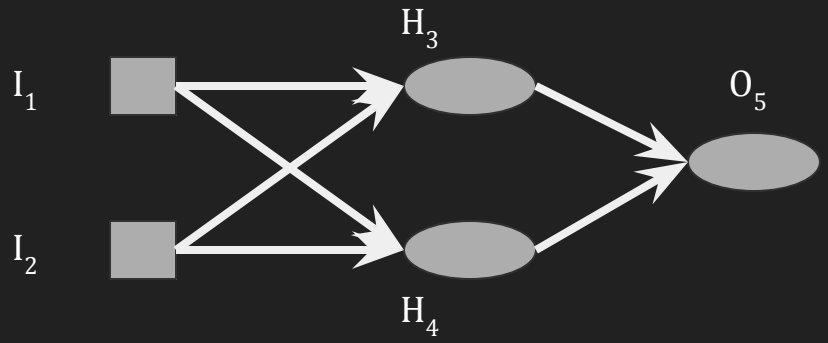
Layers







Feed Forward (different from brains!!!)



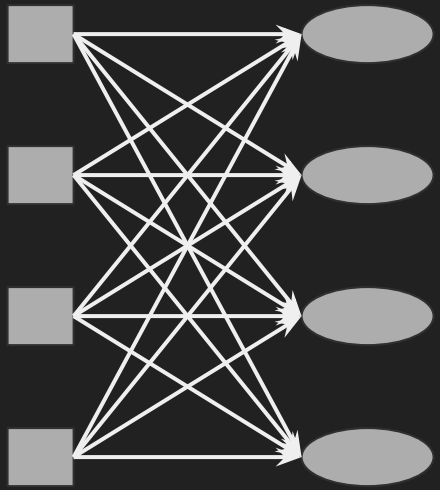
Basically, it is a function

$O_5 = f(I_1, I_2)$

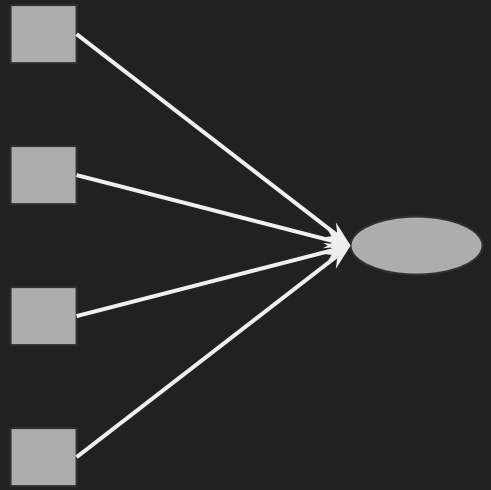
Now, what is learning?

Perceptrons

No hidden layers



I W O



I W O

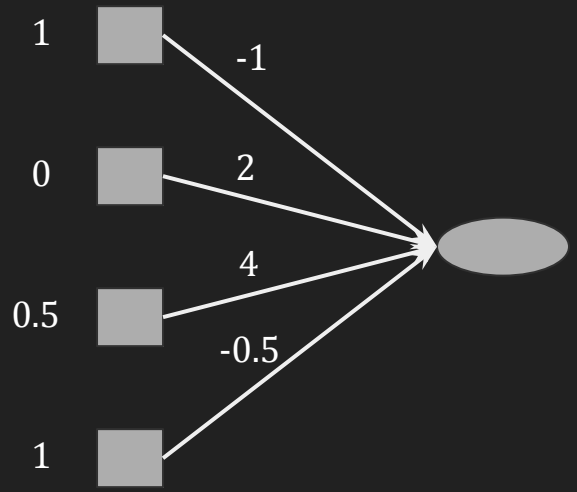
$$O = \text{Step}_0(\sum W_j I_j) = \text{Step}_0(\mathbf{W} \cdot \mathbf{I})$$

$$I_1 * W_1 + I_2 * W_2 + \ldots$$

What is the output?

Activation

$Step_0$

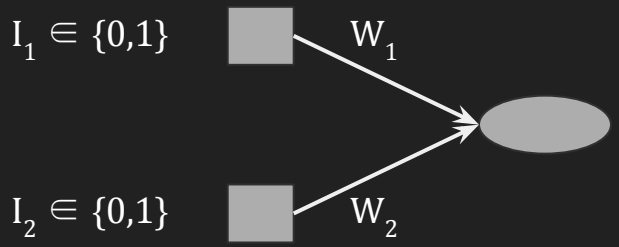


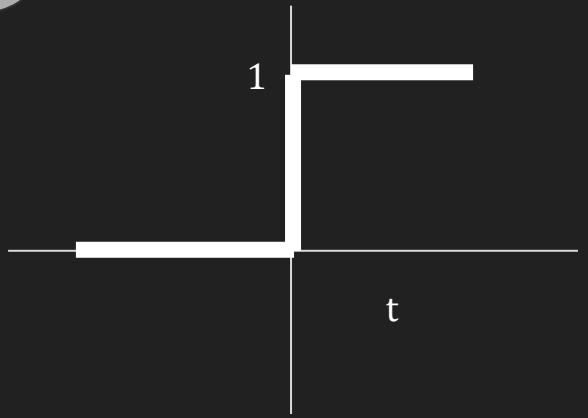
And for

$Step_1$

Represent boolean functions

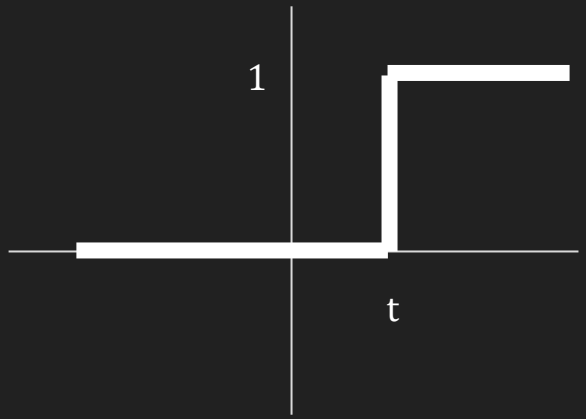
I_1	I_2	O
0	0	
0	1	
1	0	
1	1	



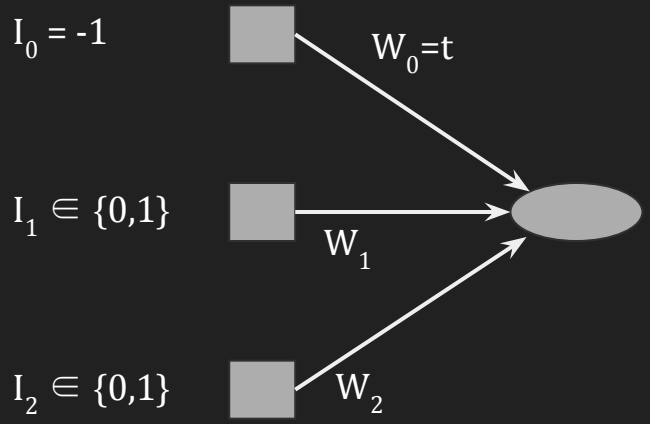


Step function

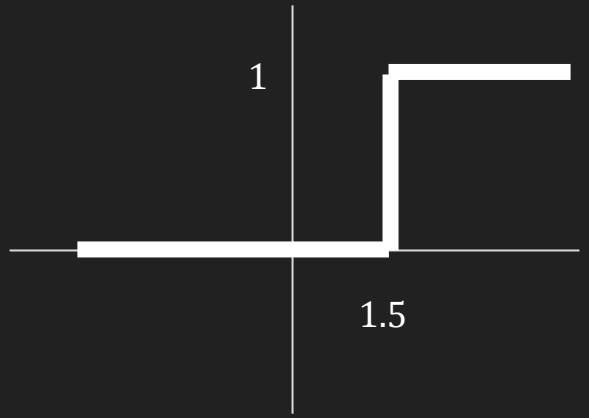
Represent boolean functions



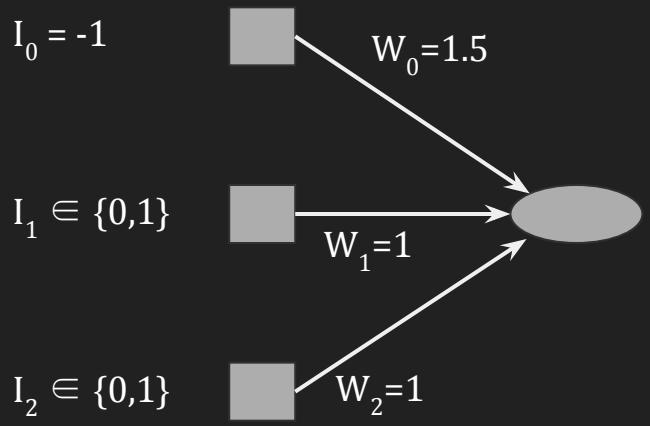
Step function



Represent boolean functions

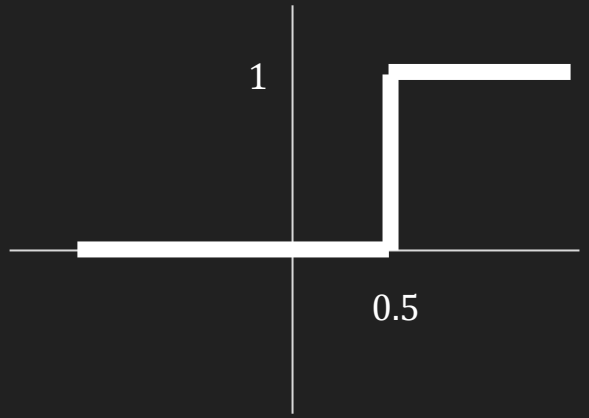


Step function

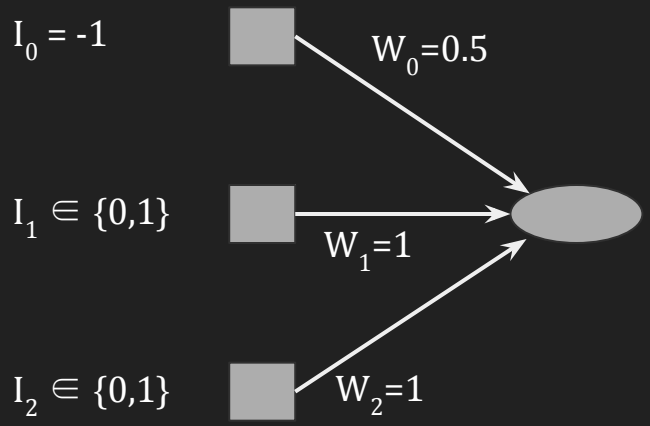


I_1	I_2	O
0	0	
0	1	
1	0	
1	1	

Represent boolean functions

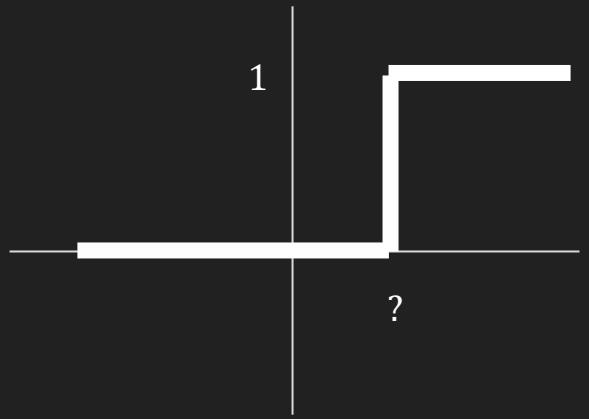


Step function

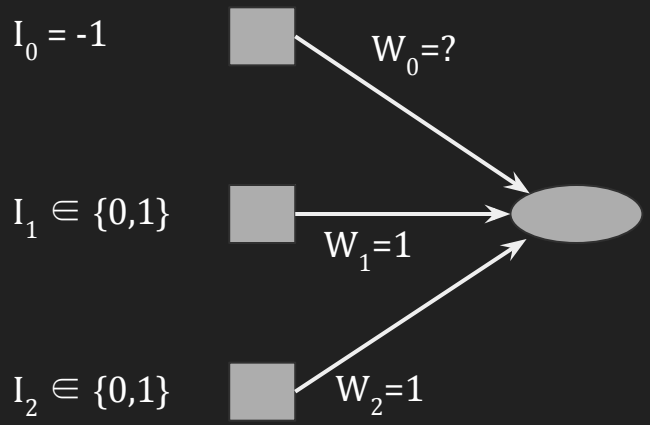


I_1	I_2	O
0	0	
0	1	
1	0	
1	1	

Represent boolean functions



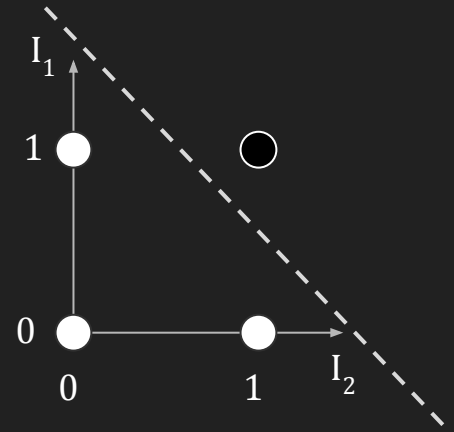
Step function

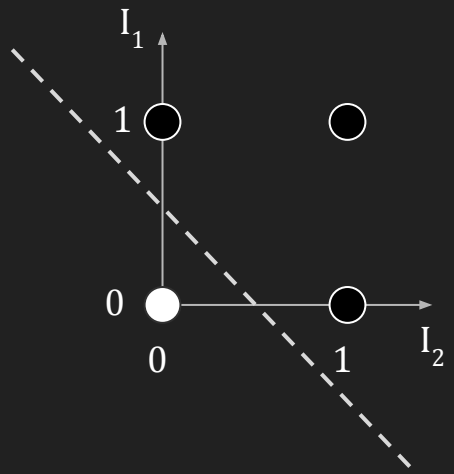


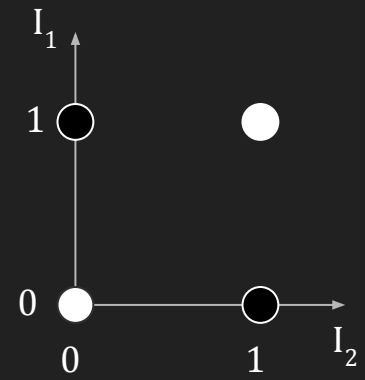
XOR

I_1	I_2	0
0	0	0
0	1	1
1	0	1
1	1	0

Linear Separable (two inputs / dimensions)







Why more layers?

Single layers networks can represent some functions, but not all

Why must the activation function be “non-linear” ?

Outlook

Learning in a perceptron

More complex networks

Learning in complex networks

Resources

Book: Artificial Intelligence A Modern Approach

Auteur: Stuart Russell