DECOMPOSING NEURAL NETWORKS

An applicant's guide to artificial learning 01.11.2022

LEUPHANA UNIVERSITÄT LÜNEBURG

ightarrow JENNIFER MATTHIESEN & TINO PAULSEN | WINTERSEMESTER 2022 |

RECAP INTRODUCTION

- **MACHINE LEARNING** gives computers the ability to learn without being explicitly programmed
- $-\!-\!$ ML requires <code>SAMPLES/DATA</code> and <code>FEATURES</code>
- NEURAL NETWORKS can find a non-linear decision boundary
- - $-\!-$ Class labels given
 - Like e.g. "The Teachable Machine"

 $-\!\!-\!\!$ No class labels are given



INTRO NN

- A biological neuron
- The artificial imitation: Perceptron
- ---- Non-linear- decisions
- Cognitive psychology:
 - Geons
 - Bigram detectors
- Feature detection in images



NEURAL NETWORK WHY NEURAL COMPUTING?

— To understand how the brain actually works — The brain is big and complicated, and it gets damaged when you poke it.

- —To solve practical problems by using novel learning algorithms inspired by the brain
 - Learning algorithms can be very useful even if they are not how the brain actually works.
- To investigate, how the imitation is different from the biological brain.
 - It is similar, better or worse?

























ч

Kretzschma









$$x \in X \longrightarrow f(x) \hat{y}$$





 $x \in X \longrightarrow f(x) \hat{y}$

 \hat{y}

f(x)

 $x\in X$ ————

 \hat{y}

 x_1

f(x)

 \hat{y}

 x_1

 x_2

 x_3

f(x)











 \bigotimes



 \bigotimes





 \bigotimes







 \bigotimes


























Ex: We want to decide to go outside or not.





Ex: We want to decide to go outside or not.





Ex: We want to decide to go **outside or not**.





x_1	$*\omega_1$	+	x_2	*	ω_2	
		L				













HOW IT LOOKS LIKE...





Ex: We want to decide to go **outside or not**.



.



$$\sigma\!\left(\sum_{i=1}^n x_i st \omega_i
ight)$$

































Ex: We want to decide to go **outside or not**. 2-2- outside • Not outside 0.6 WEATHER 1.5 $\sum_{i=1}^{n}$ σ 0 FRIENDS -2.9 $\sigma\left(-2.9 + \sum_{i=1}^{n} x_i * \omega_i\right)$ $\sigma\left(\begin{array}{c}0.7\end{array}\right) = \frac{1}{(1 + e^{-0.7})} = 0.668$ 0.5 0.5 1.5 0 WEATHER











TIME FOR DESMOS








$$\sigmaigg(-2.9+\sum_{i=1}^n x_i*\omega_iigg)$$

$$\sigmaigg(-2.9+\sum_{i=1}^n x_i*\omega_iigg)=0.5$$

$$\sigmaigg(-2.9+\sum_{i=1}^n x_i*oldsymbol{\omega}_iigg)=0.5$$
 $\sigmaigg(-2.9+(x_1*6+x_2*1igg)=0.5$

$\sigmaigg(-2.9+\sum_{i=1}^n x_i*\omega_iigg)=0.5$

$$\sigma\left(\begin{array}{c} -2.9 + (x_1 * 6 + x_2 * 1) \\ 1 \\ \frac{1}{1 + e^{-(-2.9 + (x_1 * 6 + x_2 * 1))}} = 0.5 \end{array}\right)$$

NEURAL NETWORK

A PERCEPTRON

0.5 0 1.5

1

 σ

 σ

NEURAL NETWORK A PERCEPTRON

$$\sigma \left(-2.9 + \sum_{i=1}^{n} x_i * \omega_i
ight) = 0.5$$

 $\left(-2.9 + (x_1 * 6 + x_2 * 1)
ight) = 0.5$
 $\frac{1}{1 + e^{-(-2.9 + (6x_1 + x_2))}} = 0.5$

$$egin{aligned} &\sigmaigg(-2.9+\sum_{i=1}^n x_i*\omega_iigg)=0.5\ &\sigmaigg(-2.9+(x_1*6+x_2*1igg))=0.5\ &rac{1}{1+e^{(2.9-(6x_1+x_2))}}&=0.5 \end{aligned}$$

$\sigma \left(-2.9 + (x_1 * 6 + x_2 * 1) \right) = 0.5$

$$\frac{1}{1+e^{(2.9-(6x_1+x_2))}} =$$

NEURAL NETWORK

0.5

1.5

0.5-

0

$\sigma \left(-2.9 + (x_1 * 6 + x_2 * 1) \right) = 0.5$

$$\frac{\chi}{1+e^{(2.9-(6x_1+x_2))}} = \frac{\chi}{2}$$

 $\sigmaigg(-2.9+\sum_{i=1}^n x_i*\omega_iigg)=0.5$

NEURAL NETWORK

A PERCEPTRON

NEURAL NETWORK A PERCEPTRON

$$\sigma\left(-2.9 + \sum_{i=1}^{n} x_i * \omega_i\right) = 0.5$$

$$\sigma\left(-2.9 + (x_1 * 6 + x_2 * 1)\right) = 0.5$$

$$\frac{\chi}{1 + e^{(2.9 - (6x_1 + x_2))}} = \frac{\chi}{2}$$

$$1 + e^{(2.9 - (6x_1 + x_2))} = 2$$

NEURAL NETWORK

 $\frac{\chi}{1+e^{(2.9-(6x_1+x_2))}} = \frac{\chi}{2}$

 $1 + e^{(2.9 - 6x_1 - x_2)} = 2$

A PERCEPTRON

87

NEURAL NETWORK A PERCEPTRON

Ex: We want to decide to go outside or not.

$$\sigma \left(-2.9 + \sum_{i=1}^{n} x_i * \omega_i \right) = 0.5$$

$$\sigma \left(-2.9 + (x_1 * 6 + x_2 * 1) \right) = 0.5$$

$$\frac{\chi}{1 + e^{(2.9 - (6x_1 + x_2))}} = \frac{\chi}{2}$$

$$1 + e^{(2.9 - 6x_1 - x_2)} = 2 \mid -1$$

NEURAL NETWORK A PERCEPTRON

Ex: We want to decide to go **outside or not**.

$$\sigma\left(-2.9 + \sum_{i=1}^{n} x_i * \omega_i\right) = 0.5$$

$$\sigma\left(-2.9 + (x_1 * 6 + x_2 * 1)\right) = 0.5$$

$$\underbrace{\mathcal{X}}_{1+e^{(2.9-(6x_1+x_2))}} = \underbrace{\mathcal{X}}_{2}$$

$$1 + e^{(2.9-6x_1-x_2)} = 2 \mid -1$$

$$e^{(2.9-6x_1-x_2)} = 1$$

Ex: We want to decide to go **outside or not**.

$$\sigma \left(-2.9 + \sum_{i=1}^{n} x_i * \omega_i \right) = 0.5$$

$$\sigma \left(-2.9 + (x_1 * 6 + x_2 * 1) \right) = 0.5$$

$$\frac{\chi}{1 + e^{(2.9 - (6x_1 + x_2))}} = \frac{\chi}{2}$$

$$1 + e^{(2.9 - 6x_1 - x_2)} = 2 \mid -1$$

$$e^{(2.9 - 6x_1 - x_2)} = 1 \mid \log(1)$$

1

1

 σ

 σ

90

NEURAL NETWORK A PERCEPTRON

Ex: We wa t to doo!do t .

We want to decide to go outside or not.

$$\sigma\left(-2.9 + \sum_{i=1}^{n} x_i * \omega_i\right) = 0.5$$

$$(-2.9 + (x_1 * 6 + x_2 * 1)) = 0.5$$

$$\frac{\chi}{1 + e^{(2.9 - (6x_1 + x_2))}} = \frac{\chi}{2}$$

$$1 + e^{(2.9 - 6x_1 - x_2)} = 2 | -1$$

$$e^{(2.9 - 6x_1 - x_2)} = 1 | \log(0)$$

$$2.9 - 6x_1 - x_2 = 0$$

A PERCEPTRON Ex: We want to decide to go outside or not.

NEURAL NETWORK

$$\sigma\left(-2.9 + \sum_{i=1}^{n} x_i * \omega_i\right) = 0.5$$

$$\sigma\left(-2.9 + (x_1 * 6 + x_2 * 1)\right) = 0.5$$

$$\frac{\chi}{1 + e^{(2.9 - (6x_1 + x_2))}} = \frac{\chi}{2}$$

$$1 + e^{(2.9 - 6x_1 - x_2)} = 2 \mid -1$$

$$e^{(2.9 - 6x_1 - x_2)} = 1 \mid \log(0)$$

$$2.9 - 6x_1 - x_2 = 0$$

$$x_2 = 2.9 - 6x_1$$

Ex: We want to decide to go **outside or not**.

NEURAL NETWORK

A PERCEPTRON

BUT WHAT ABOUT **NON-LINEAR** DECISIONS?

0L

 \bigotimes

DMPII | DECOM. NEURAL NETWORKS

 \bigotimes

DMPII | DECOM. NEURAL NETWORKS

95

DMPII | DECOM. NEURAL NETWORKS

DMPII | DECOM. NEURAL NETWORKS









TIME FOR DESMOS



https://www.desmos.com/calculator/1czmp32hjs





 \bigotimes



















TIME FOR DESMOS



https://www.desmos.com/calculator/4dtlyi0bjg?



INTRO COGNITIVE PSYCHOLOGY

125

 \bigotimes

- GEONS (geometry icons) are simple shapes
- From the network theory known as recognition by components (RPC model) (Hummel & Biederman 1992)
- According to Biederman (1987, 1880), we need (at most) 36 different GEONS to describe every object.



- GEONS (geometry icons) are simple shapes
- From the network theory known as recognition by components (RPC model) (Hummel & Biederman 1992)
- According to Biederman (1987, 1880), we need (at most) 36 different GEONS to describe every object.



- GEONS (geometry icons) are simple shapes
- From the network theory known as recognition by components (RPC model) (Hummel & Biederman 1992)
- According to Biederman (1987, 1880), we need (at most) 36 different GEONS to describe every object.



- GEONS (geometry icons) are simple shapes
- From the network theory known as recognition by components (RPC model) (Hummel & Biederman 1992)
- According to Biederman (1987, 1880), we need (at most) 36 different GEONS to describe every object.



- GEONS (geometry icons) are simple shapes
- From the network theory known as recognition by components (RPC model) (Hummel & Biederman 1992)
- According to Biederman (1987, 1880), we need (at most) 36 different GEONS to describe every object.



Let's do a little experiment:
 Try to read the following letter combination:

Let's do a little experiment:

Try to read the following letter combination:

"PIRT"

Let's do a little experiment:
 Try to read the following letter combination:

— Let's do a little experiment:

Try to read the following letter combination:

"ITPR"

Let's do a little experiment:
 Try to read the following letter combination:

Let's do a little experiment:
 Try to read the following letter combination:

Let's do a little experiment:

Try to read the following letter combination:

"HICE"

Let's do a little experiment:
 Try to read the following letter combination:

Let's do a little experiment:

Try to read the following letter combination:

"HCEI"

Let's do a little experiment:
 Try to read the following letter combination:

- So why are "PIRT" or "HICE" easier to detect then "ITRP" and "HCEI"?
- We do not just detect words, but rather bigrams
- The bigram detector detects letter
 combinations and will be triggered by lowerlevel detectors

- So why are "PIRT" or "HICE" easier to detect then "ITRP" and "HCEI"?
- We do not just detect words, but rather bigrams
- The bigram detector detects letter combinations and will be triggered by lowerlevel detectors



- So why are "PIRT" or "HICE" easier to detect then "ITRP" and "HCEI"?
- We do not just detect words, but rather bigrams
- The bigram detector detects letter combinations and will be triggered by lowerlevel detectors
- You have seen the letter pairs "HI" ("HIT", "HIGHT") and "CE" ("FACE", "MICE") before
 - Benefits from priming



FEATURE DETECTION IN IMAGES

 $\circ \bigtriangleup$
FEATURE DETECTION IN IMAGES



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

 \bigotimes

until ~13:15



SUMMARY OF TODAY

- **A NEURON** has a cell body, dendrites as input preceptors and Axons for the output
- A PERCEPTION is the computer science model of a neuron. It can "learn" a linear decision
- A MULTI-LAYER PERCEPTRON can learn, non-linear decisions, but is sensitive to the weights.
- -GEONS:
 - —Are the "building blocks" for more complex shapes in cognition theory

- FEATURE NET IN THE BRAIN:

— Bigram detectors and priming