

ABSCHNITT 3

ARTIFICIAL NEURAL NETWORKS

EIN KERNKONZEPT MASCHINELLEN LERNENS

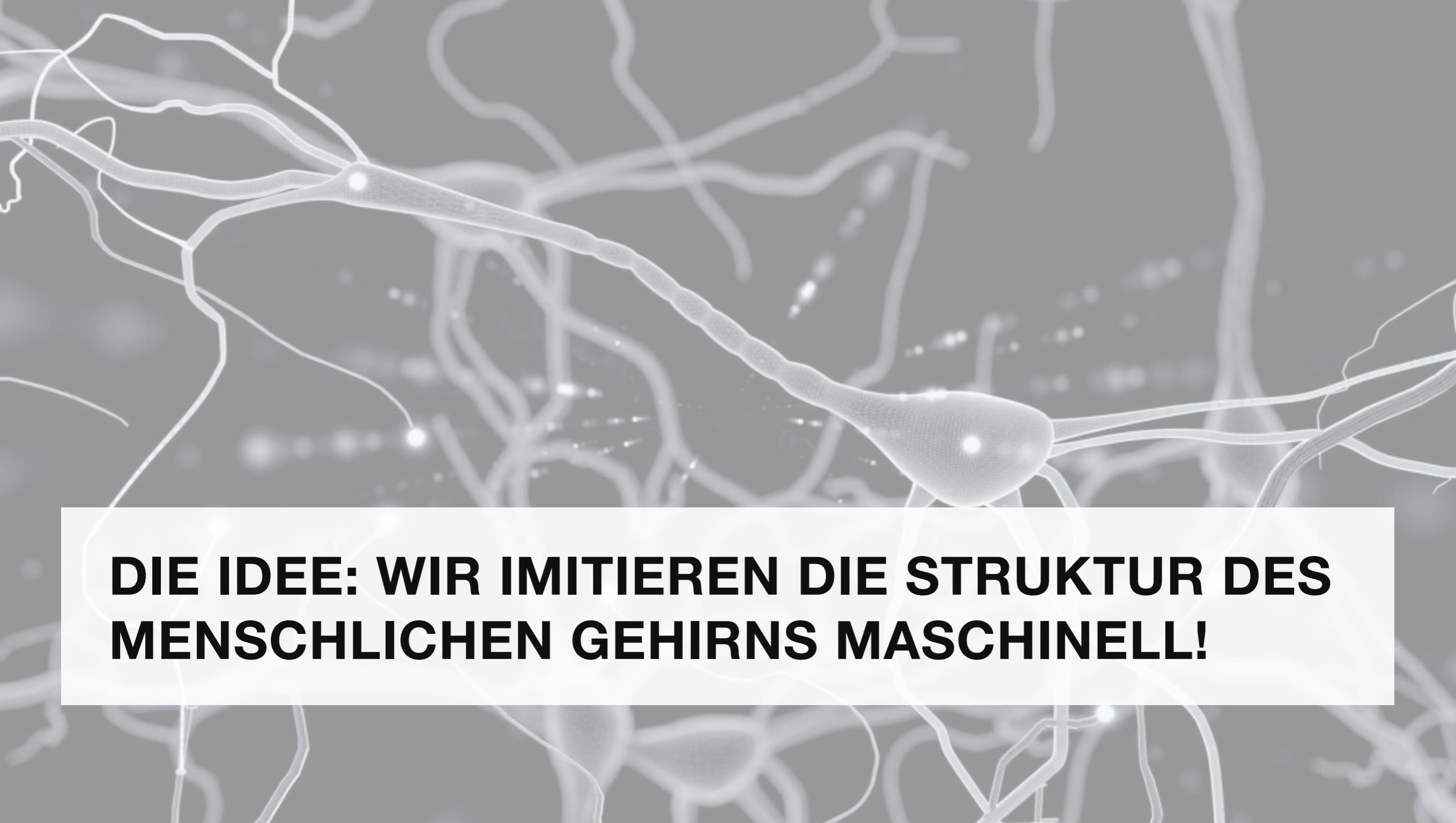


ARTIFICIAL INTELLIGENCE IS NOT
COMING FOR YOU.



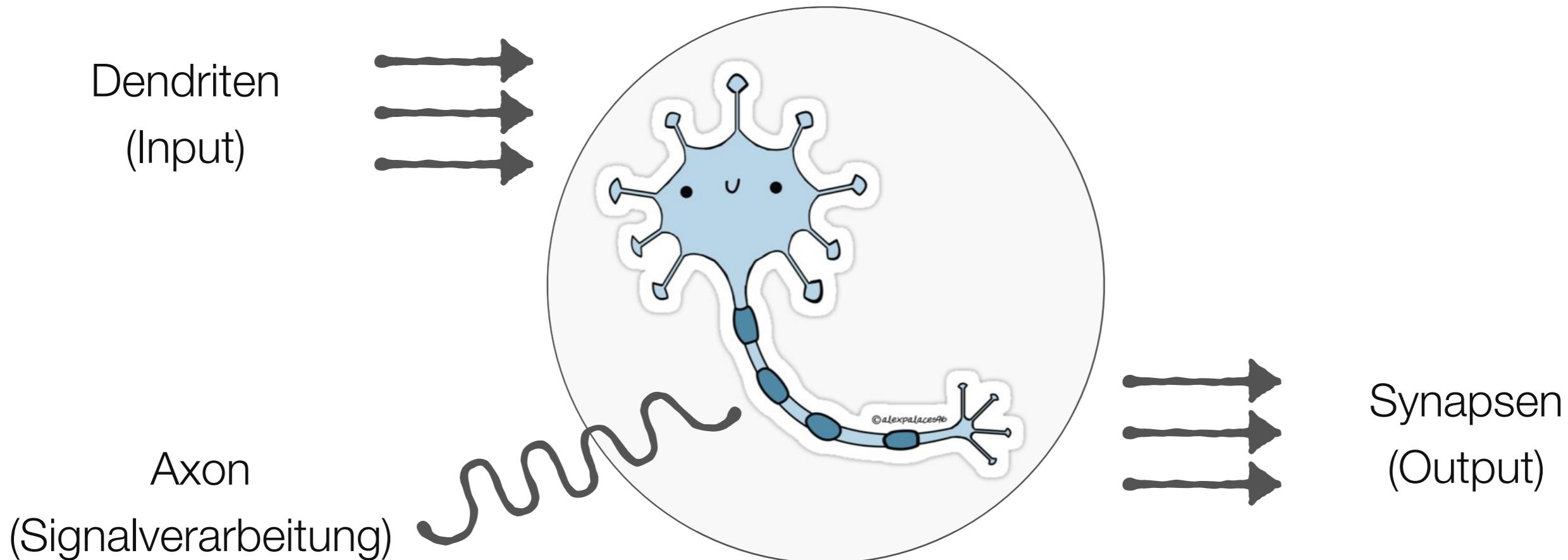
A close-up photograph of a hand holding a wooden spoon, stirring a vibrant stir-fry in a dark-colored wok. The stir-fry contains various vegetables like green beans, red and yellow bell peppers, broccoli, and snap peas. The background is blurred, focusing on the cooking action.

**ARTIFICIAL INTELLIGENCE IS NOT
COMING FOR YOU.**

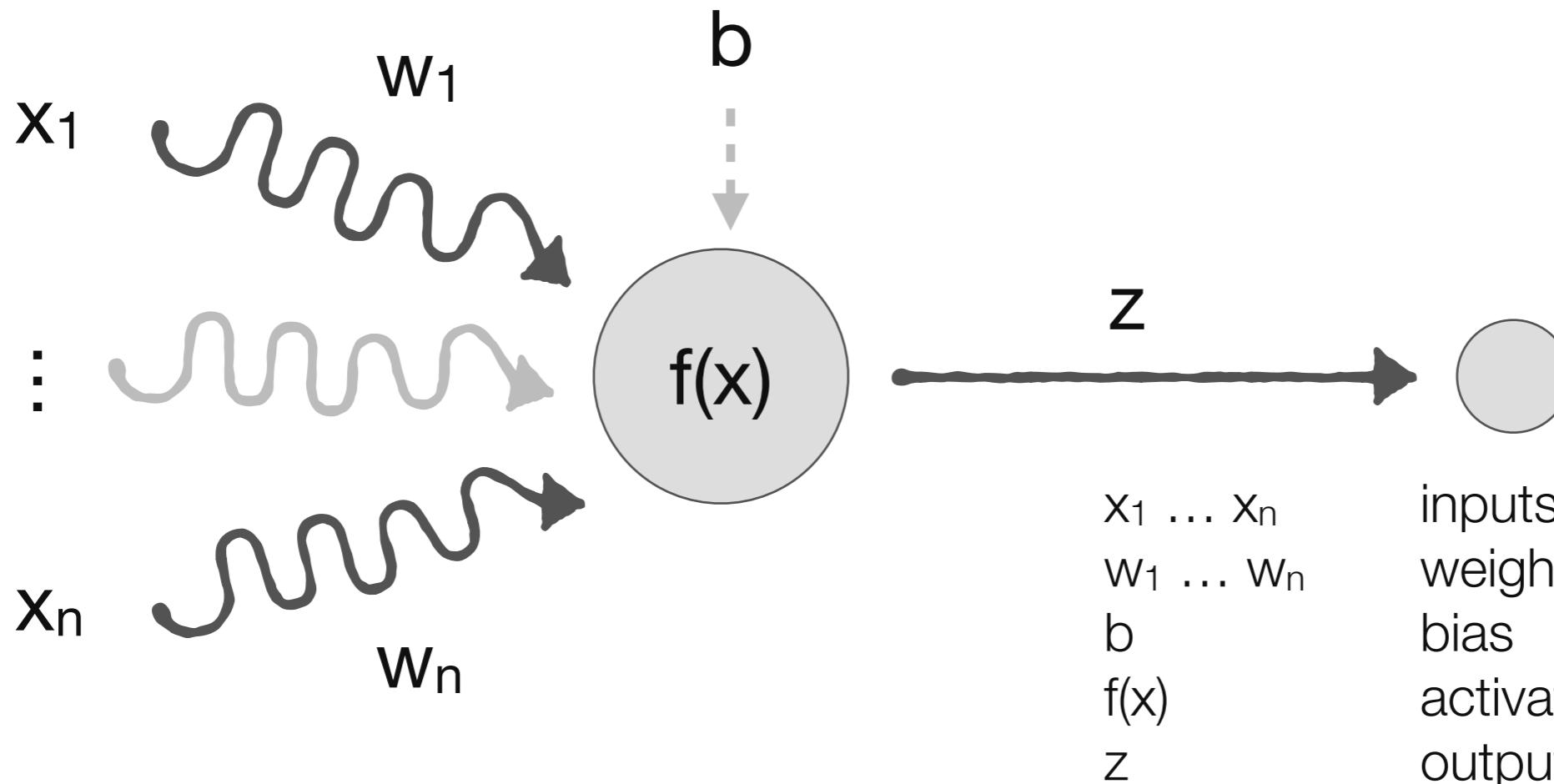


**DIE IDEE: WIR IMITIEREN DIE STRUKTUR DES
MENSCHLICHEN GEHIRNS MASCHINELL!**

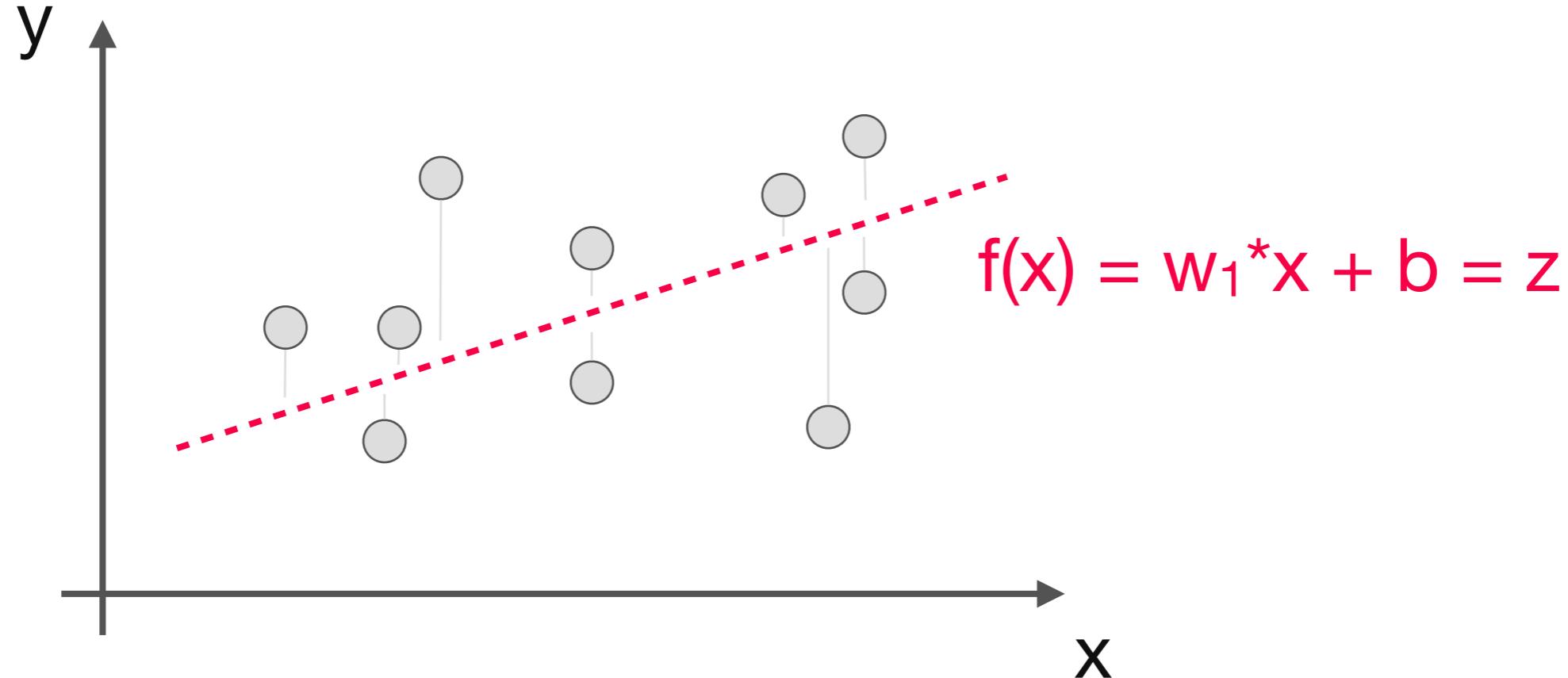
DIE EINFACHSTE FORM EINES NEURONALEN NETZES: INSPIRIERT DURCH DAS NEURON.



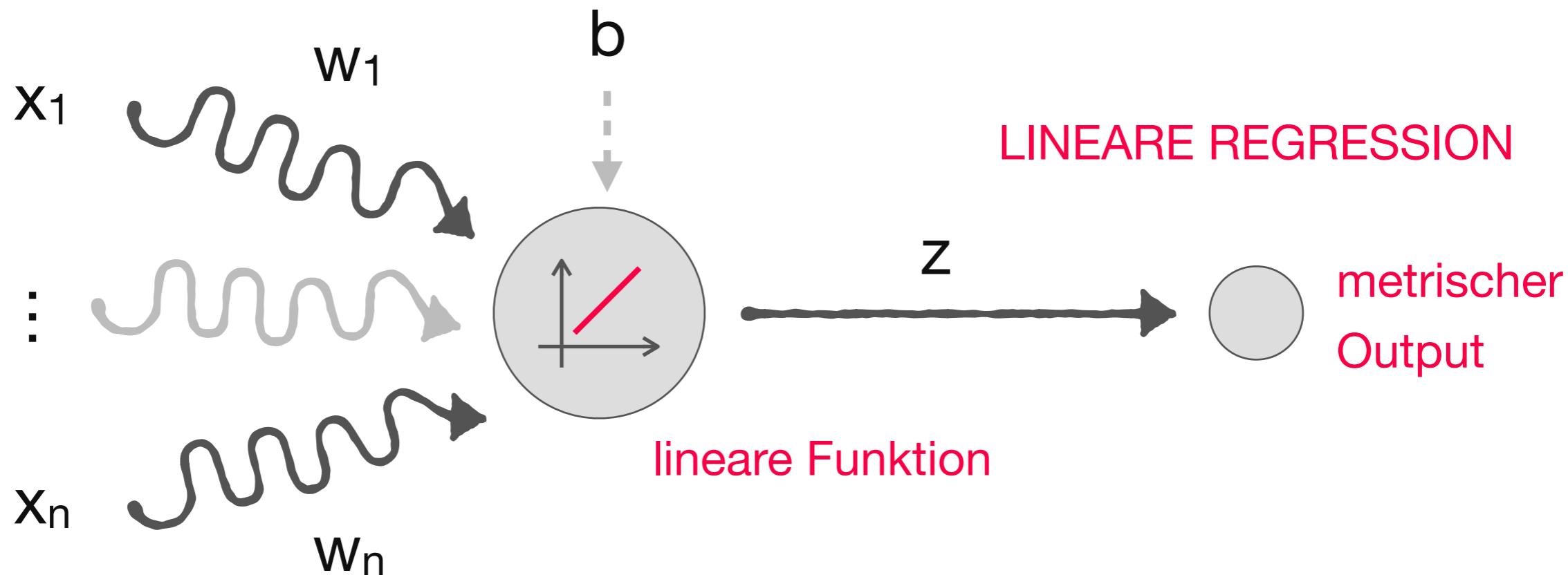
DIE EINFACHSTE FORM EINES NEURONALEN NETZES: DAS PERCEPTRON.



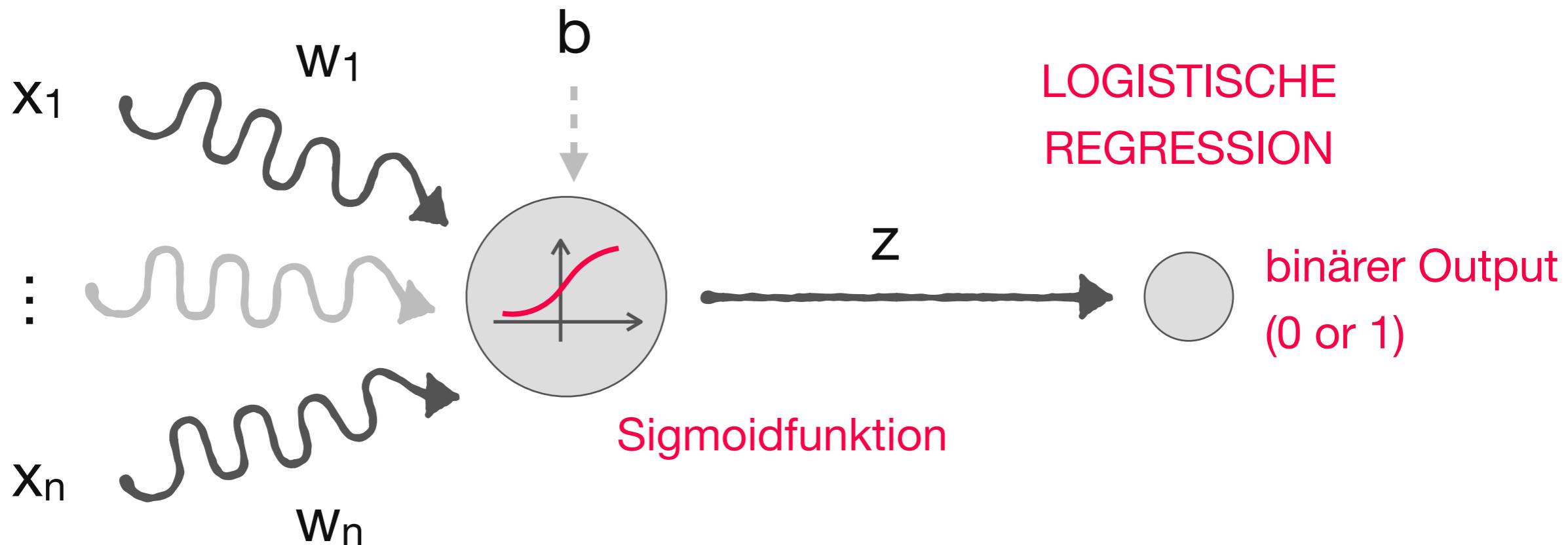
DAS PERCEPTRON - EIN GESCHWISTER DER LINEAREN REGRESSION?



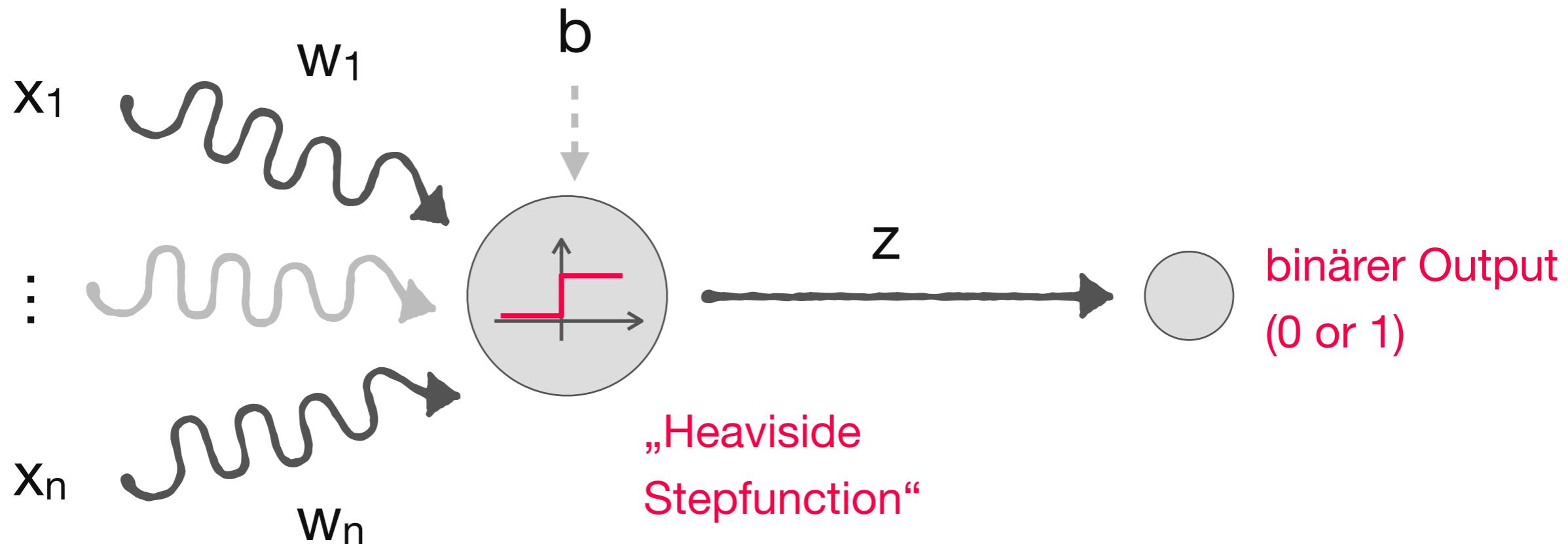
DAS PERCEPTRON - EIN GESCHWISTER DER LINEAREN REGRESSION? JA UND NEIN.



DAS PERCEPTRON - EIN GESCHWISTER DER LINEAREN REGRESSION? JA UND NEIN.



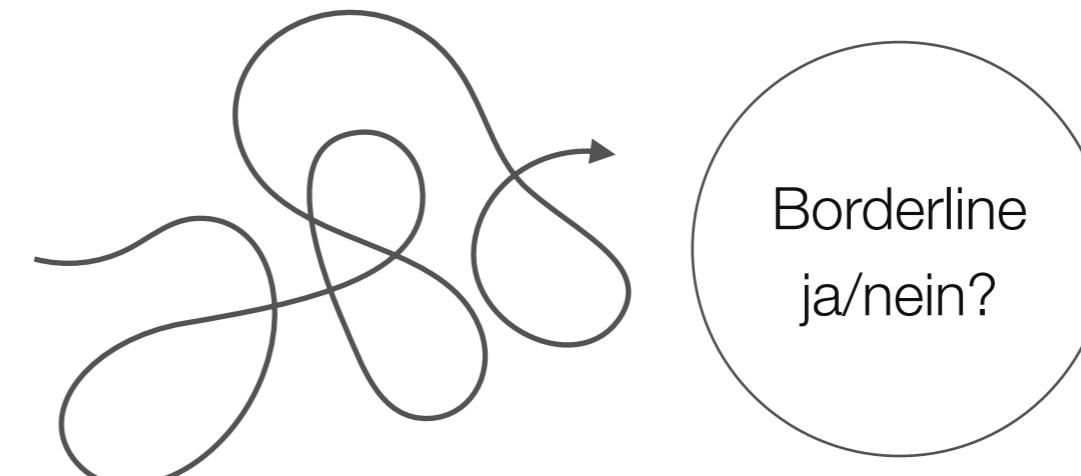
DAS PERCEPTRON - EIN GESCHWISTER DER LINEAREN REGRESSION? JA UND NEIN.



NICHT ALLE PROBLEME KÖNNEN DURCH EIN REGRESSIONSMODELL GELÖST WERDEN.

Welche Personen entwickeln angesichts welcher Risiko- und Schutzfaktoren eine Borderline Störung?

Kindesmissbrauch
emotionale Fähigkeit A, B, C, ...
genetischer Faktor A, B, C, ...
sozialer Faktor A, B, C, ...
...

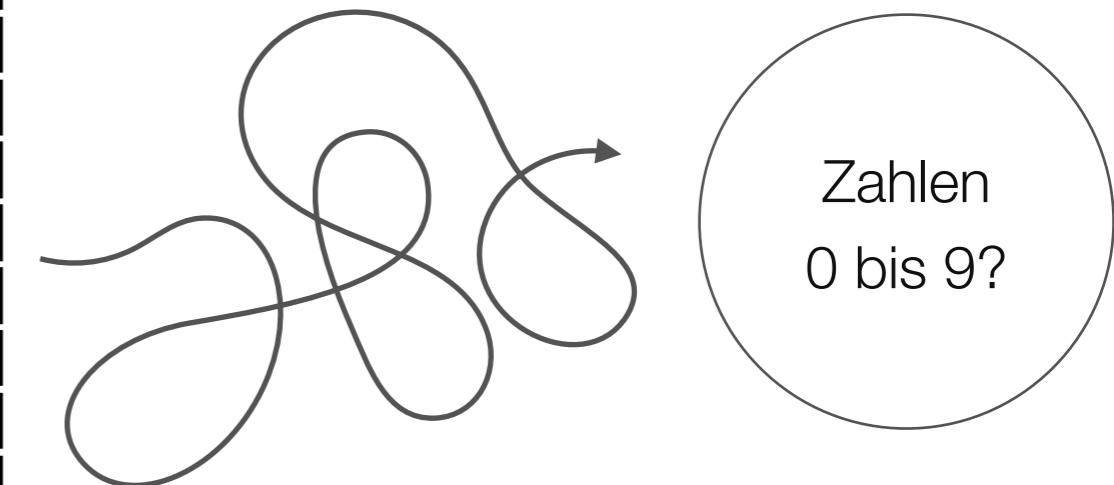


NICHT ALLE PROBLEME KÖNNEN DURCH EIN REGRESSIONSMODELL GELÖST WERDEN.

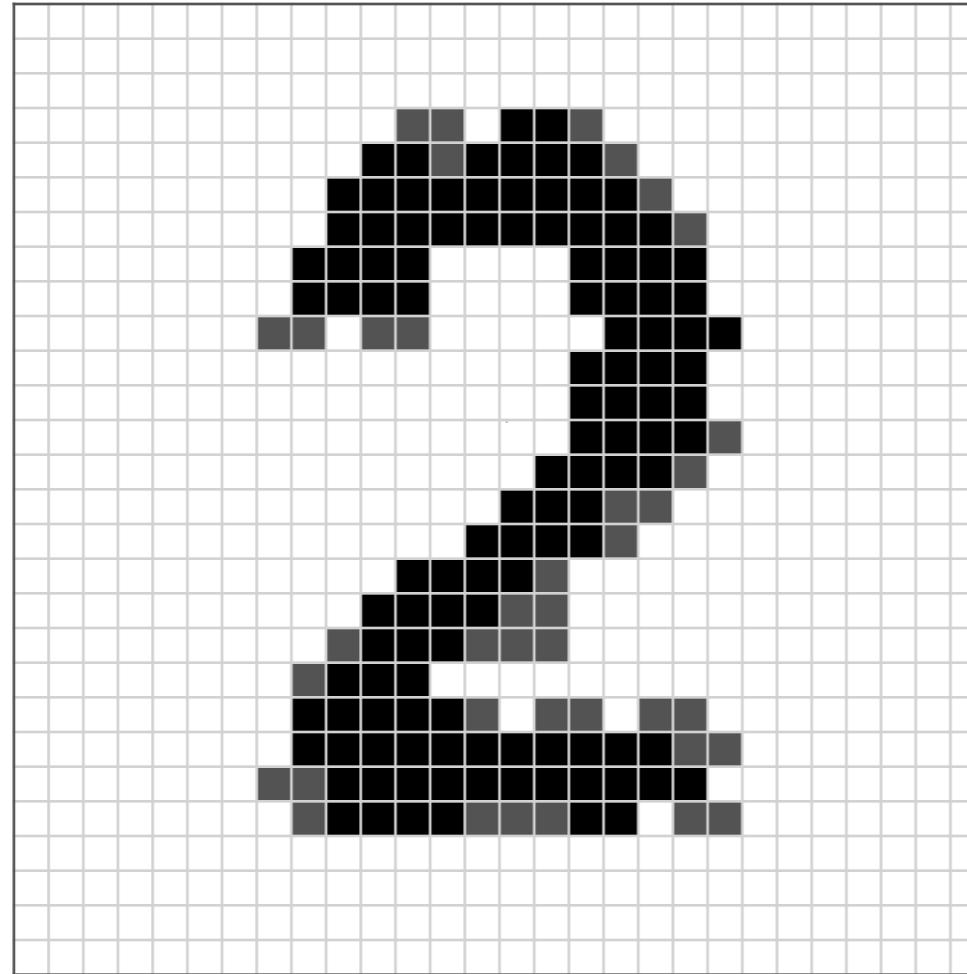
Ein klassisches Machine Learning Beispiel: Das Erkennen handgeschriebener Zahlen.

MNIST Datenbank
60.000 Bilder mit labels
28x28 Pixel
mit Graustufen Werten

0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9



EIN MNIST BILD-BEISPIEL.

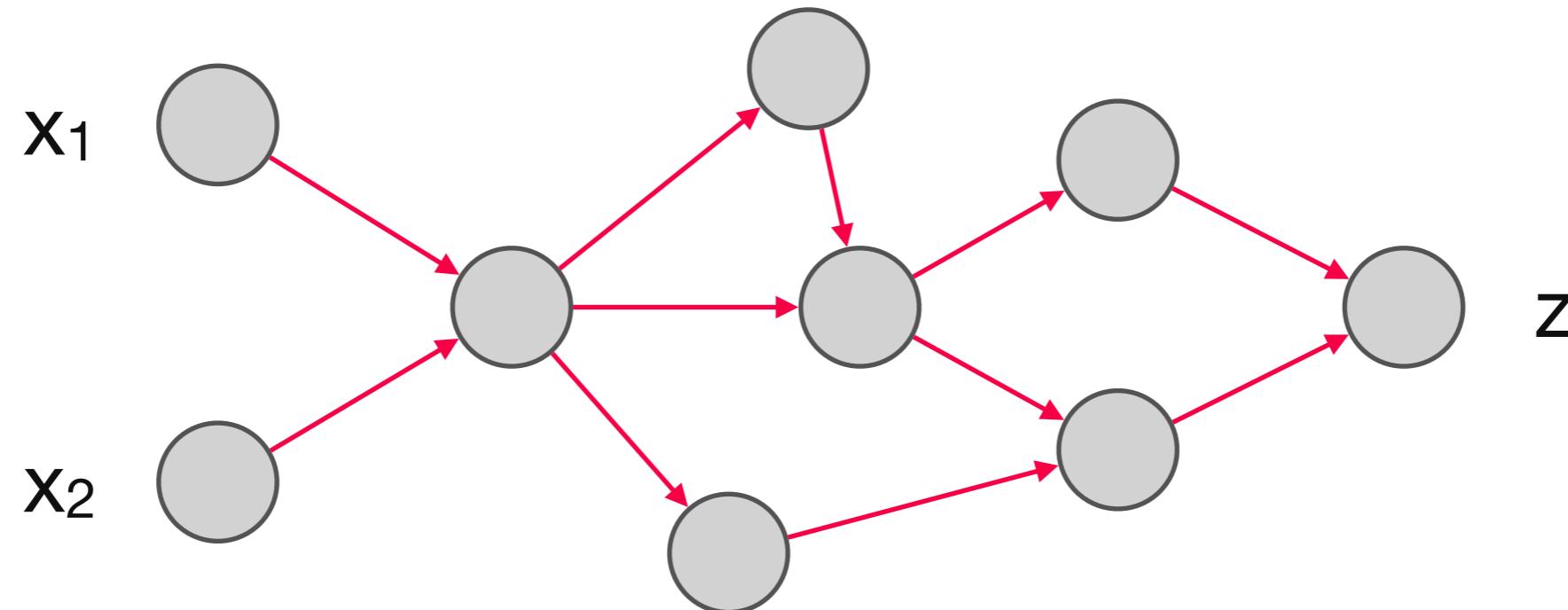


28×28
= 784 features

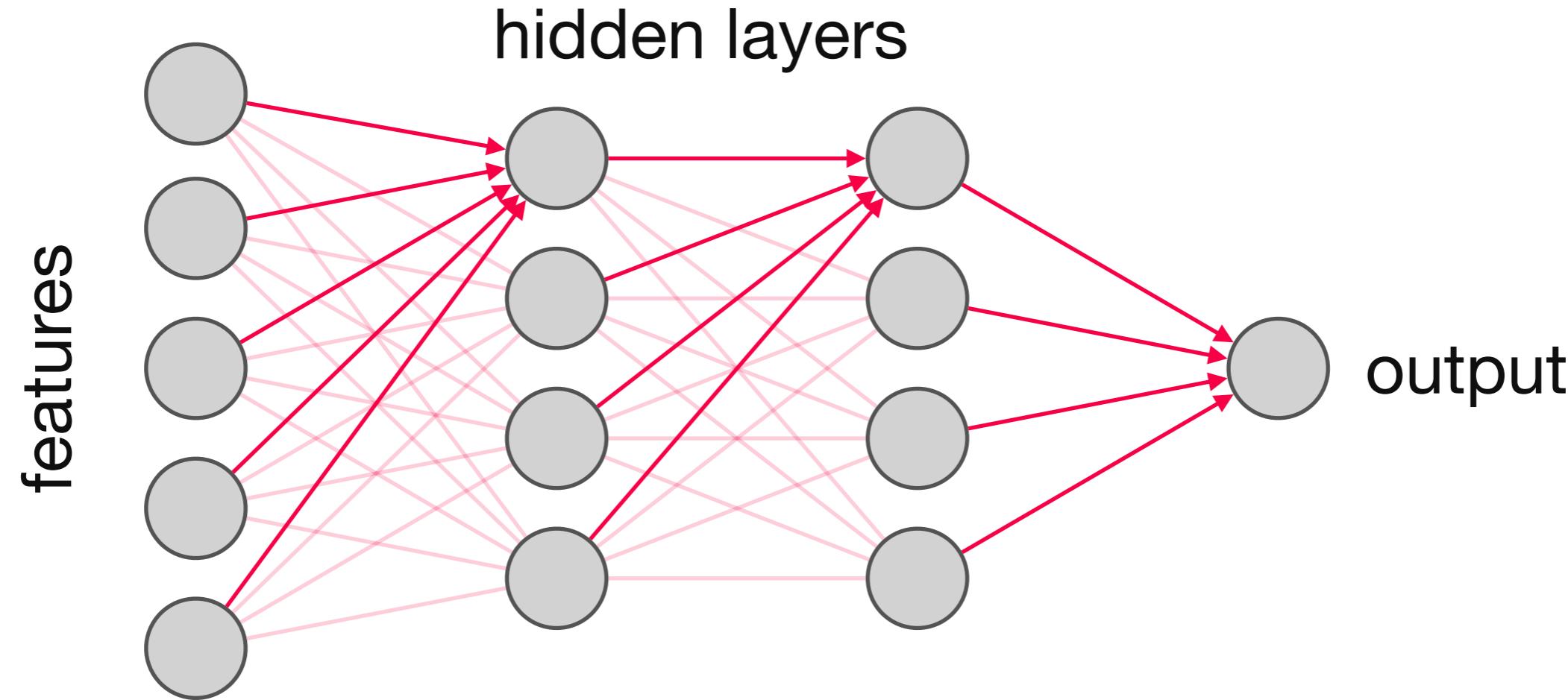
EINE LÖSUNG KOMPLEXER PROBLEME: STACK IT UP!

AUS DEM PERCEPTRON EIN FEEDFORWARD NEURAL NETWORK BAUEN.

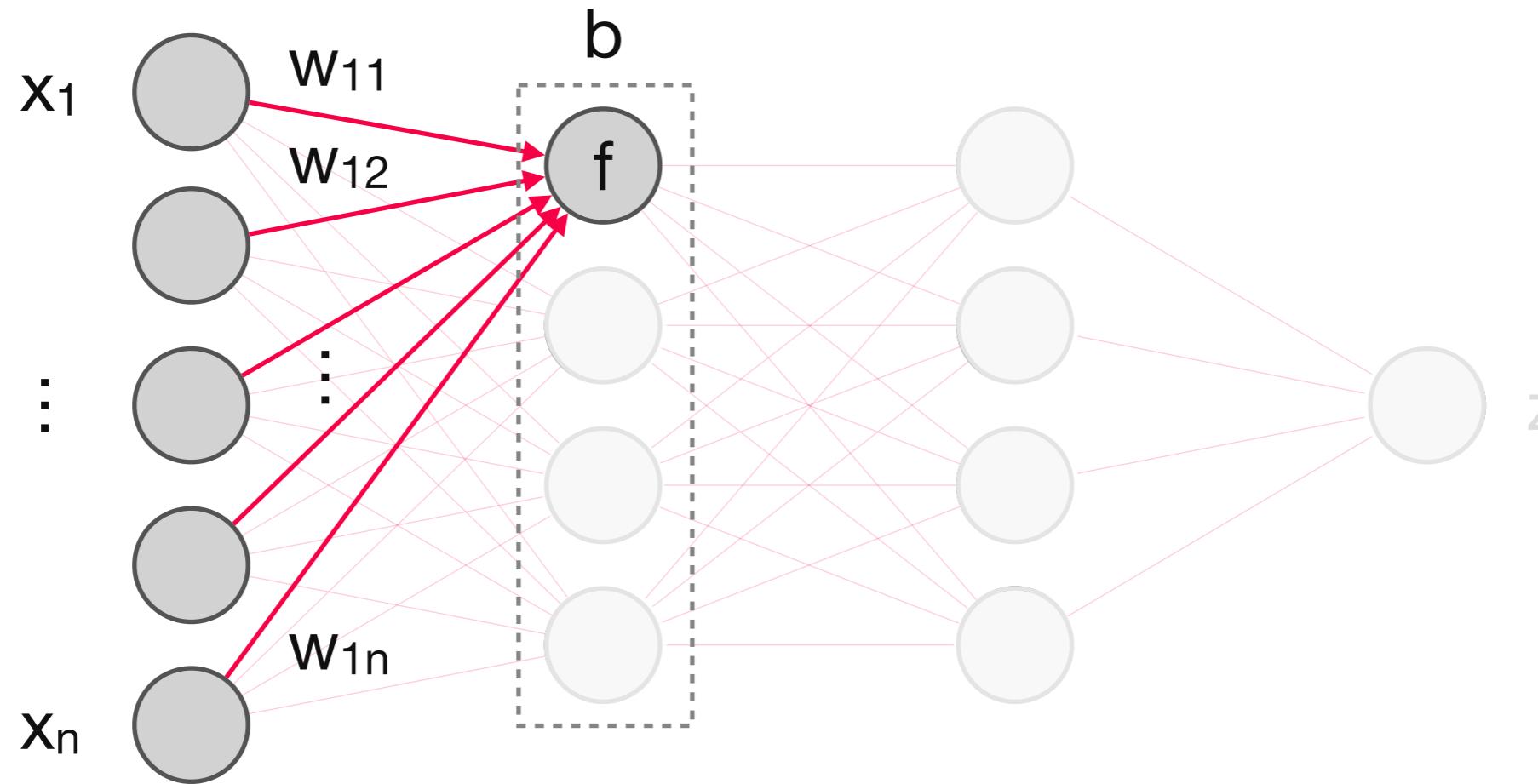
Auch Multilayer Neural Perceptron genannt.



EIN NN IST AUFGEBAUT AUS INPUT LAYER, OUTPUT LAYER & HIDDEN LAYERS.



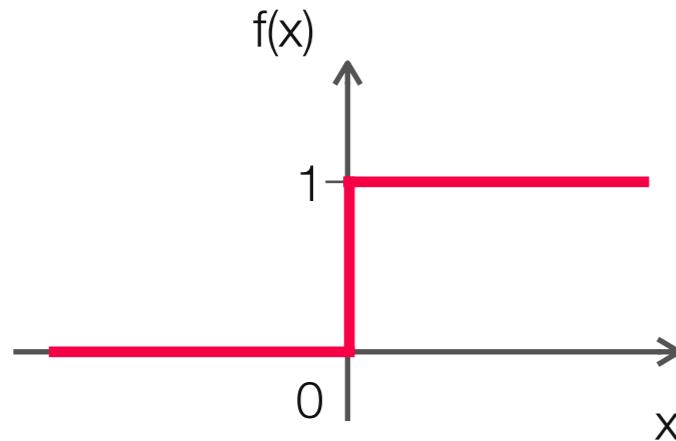
DIE GEWICHTE SIND DIE STÄRKE DES ZUSAMMENHANGS ZWISCHEN NEURONEN.



DIE GÄNGIGSTEN AKTIVIERUNGSFUNKTIONEN.

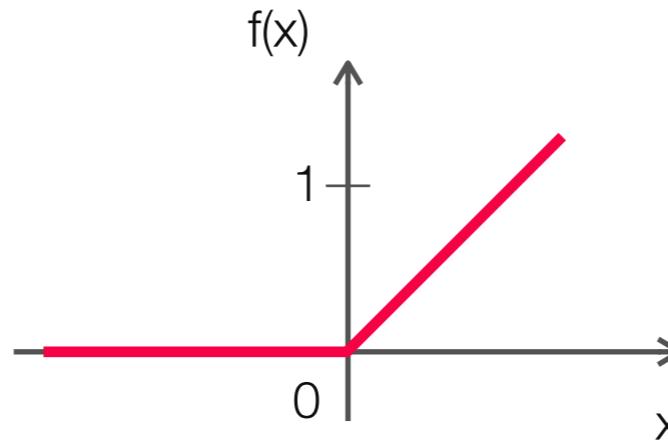
Heaviside Step Function

$$x \mapsto \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$



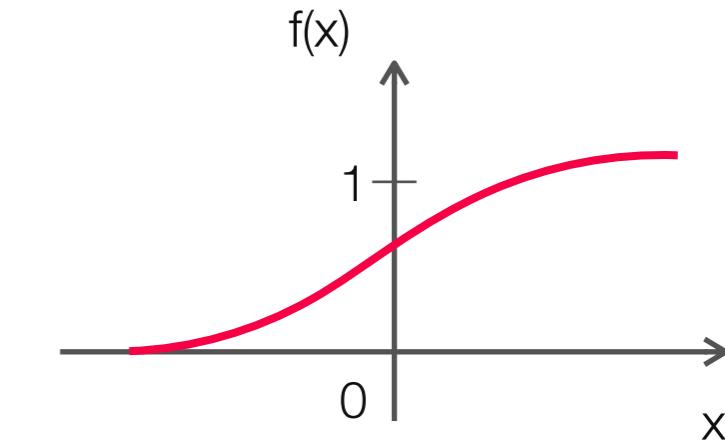
RELU Funktion

$$x \mapsto \begin{cases} 0 & x < 0 \\ x & x \geq 0 \end{cases}$$



Sigmoidfunktion

$$\text{sig}(t) = \frac{1}{1 + e^{-t}}$$



EINIGE NN-MERKMALE SIND VORGEGEBEN - VIELE MÜSSEN BESTIMMT WERDEN!

Wie viele Neuronen hat das input layer?

Wie viele Neuronen hat das output layer?

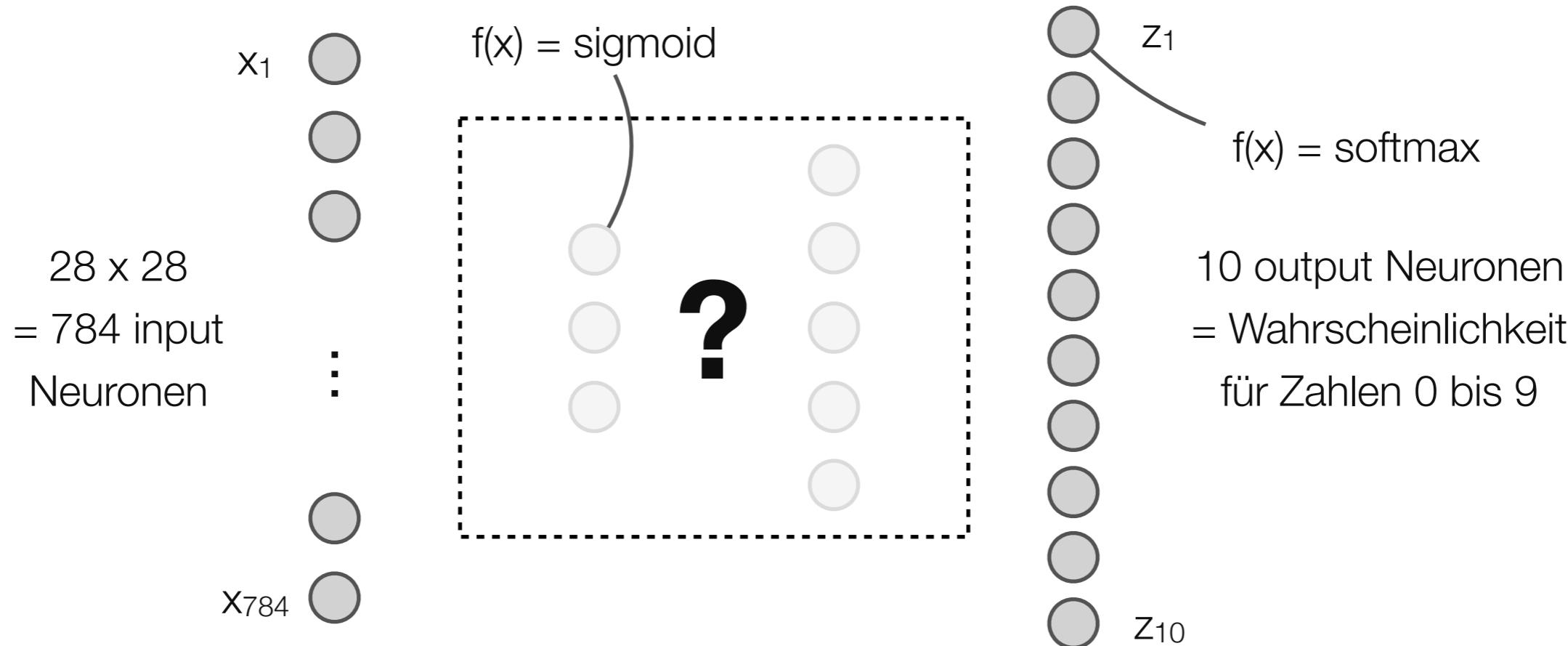
Wie viele hidden layer gibt es?

Wie viele Neuronen hat ein hidden layer?

Welche Aktivierungsfunktion haben die output Neuronen?

Welche Aktivierungsfunktionen haben Neuronen der hidden layer?

BEISPIEL: EIN NN ZUR ERKENNUNG HANDGESCHRIEBENER ZAHLEN.



WIE SIND ALL DIESE MERKMALE ZU INTERPRETIEREN?

Allein ein NN mit 784 input Neuronen, 10 output Neuronen und 2 hidden layer à 16 Neuronen hat rund 13.000 Parameter!



WIE SIND ALL DIESE MERKMALE ZU INTERPRETIEREN?

Allein ein NN mit 784 input Neuronen, 10 output Neuronen und 2 hidden layer à 16 Neuronen hat rund 13.000 Parameter!



ABER INTERPRETERBARKEIT IST UNSER A & O!

Nicht immer.

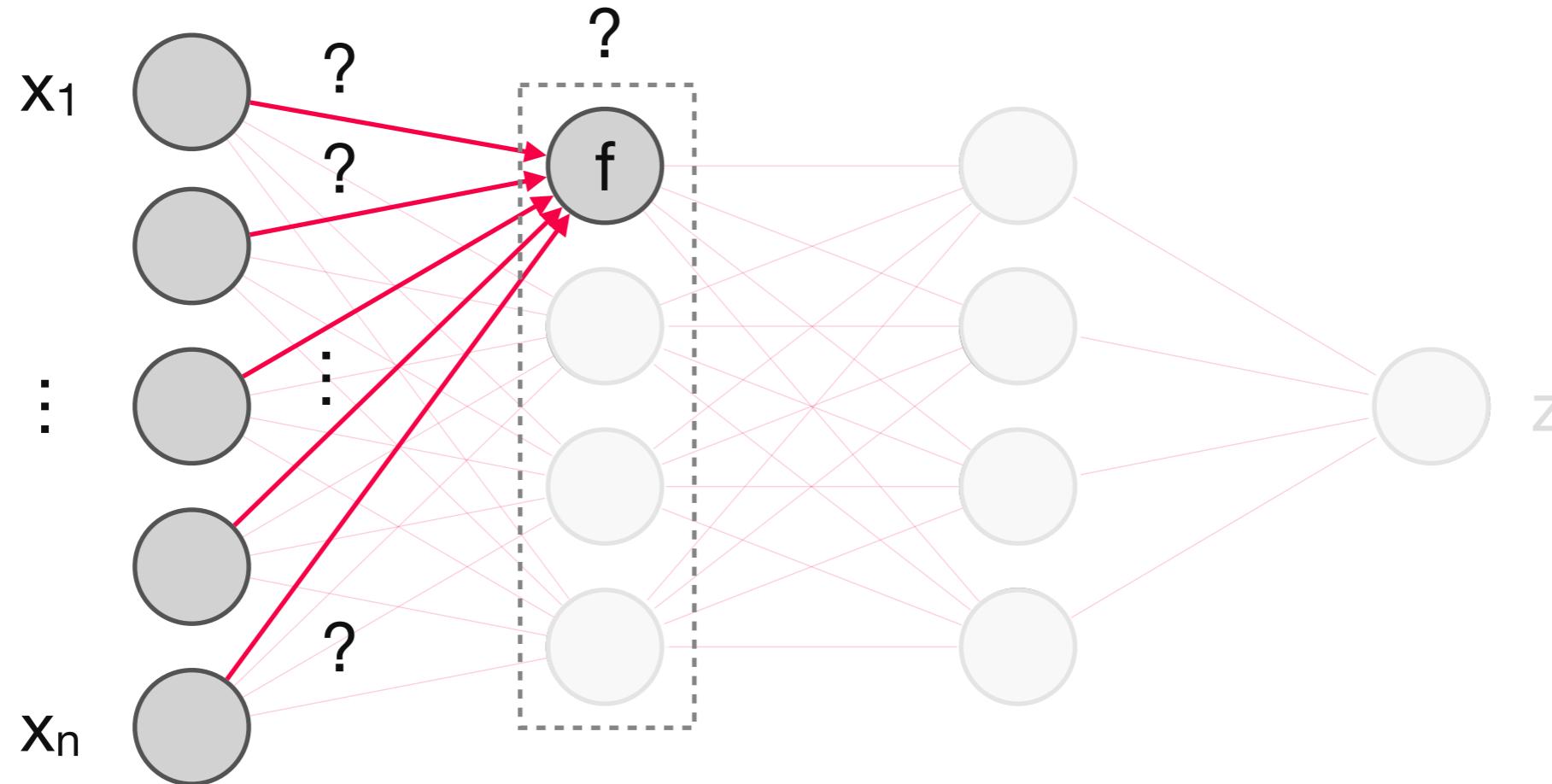
Measuring the predictability of life outcomes with a scientific mass collaboration

How predictable are life trajectories? We investigated this question with a scientific mass collaboration using the common task method; 160 teams built predictive models for six life outcomes using data from the Fragile Families and Child Wellbeing Study, a high-quality birth cohort study. Despite using a rich dataset and applying machine-learning methods optimized for prediction, the best predictions were not very accurate and were only slightly better than those from a simple benchmark model. Within each outcome, prediction error was strongly associated with the family being predicted and weakly associated with the technique used to generate the prediction. Overall, these results suggest practical limits to the predictability of life outcomes in some settings and illustrate the value of mass collaborations in the social sciences.

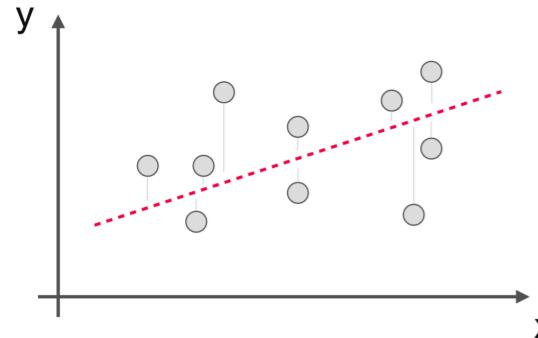
life course | prediction | machine learning | mass collaboration

WIE LERNT SO EIN NEURONALES NETZ DENN NUN EIGENTLICH?

SCHRITT EINS: DIE PARAMETER DES NN ZUFÄLLIG INITIALISIEREN!



SCHRITT ZWEI: DIE ANPASSUNGSGÜTE MESSEN.



$$\Delta (\text{grey circle} - \text{red circle})^2$$

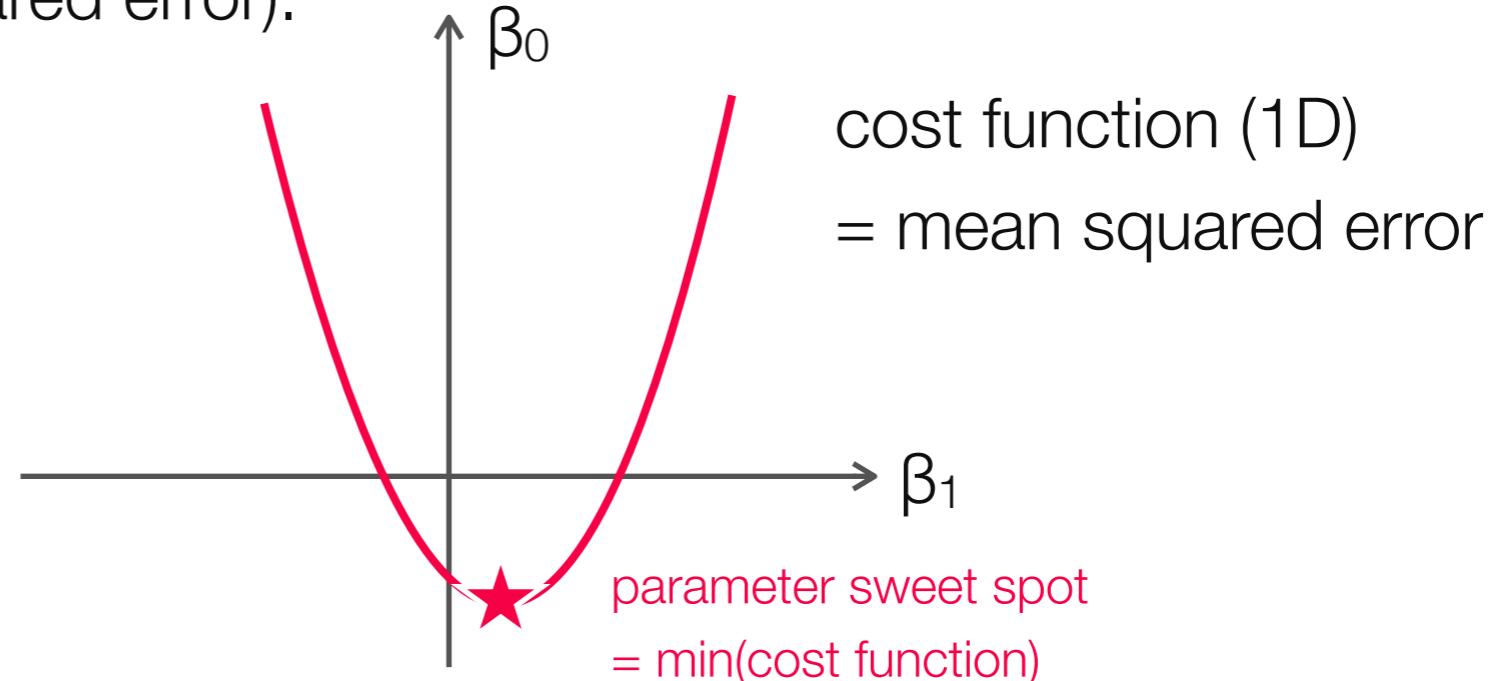
$$\Delta (\text{grey circle} - \text{red circle})^2$$

:

$$\Delta (\text{grey circle} - \text{red circle})^2$$

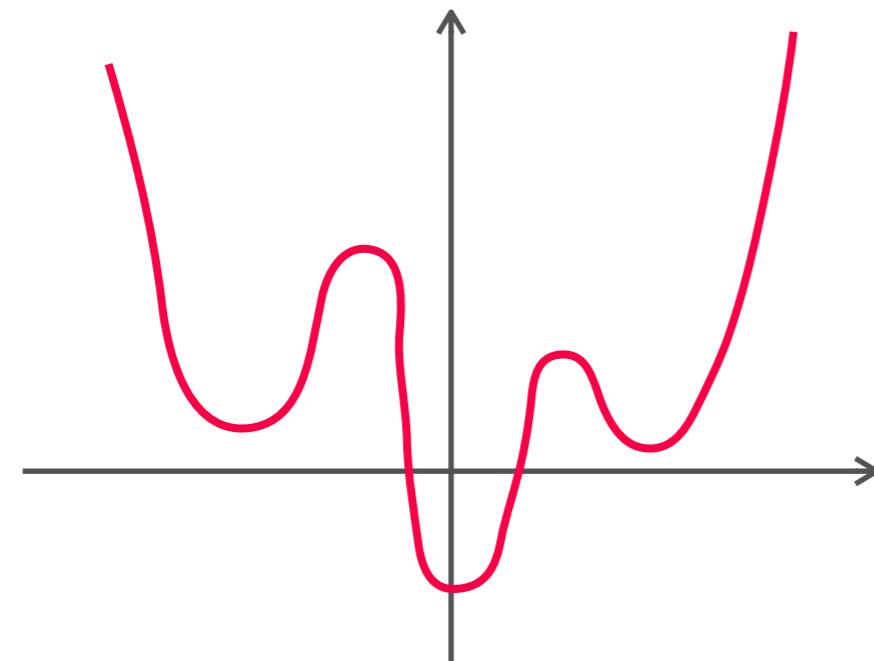
$$\underline{\Sigma / N}$$

Genau wie bei der linearen Regression messen wir den durchschnittlichen quadrierten Fehler (mean squared error).

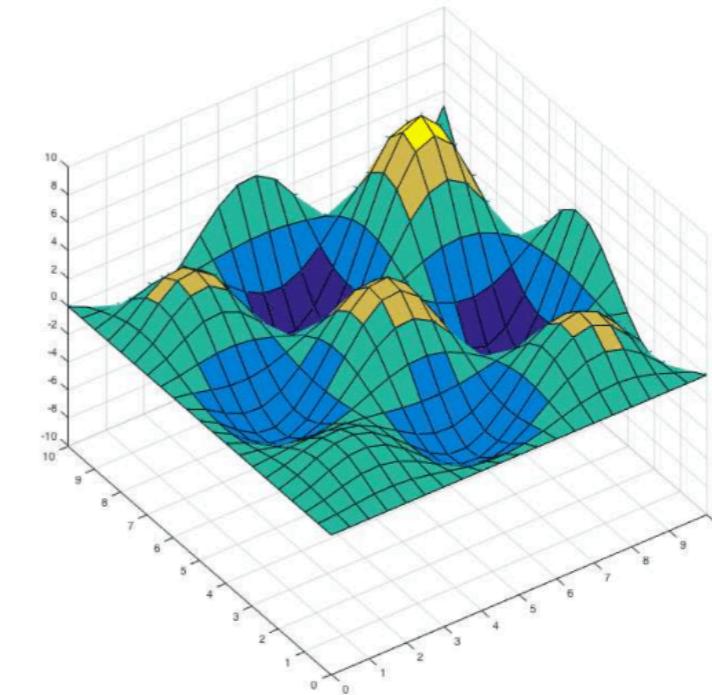


BEI NEURONALEN NETZEN IST DIE SACHE ETWAS KOMPLIZIERTER.

Die cost function ist nicht konvex.

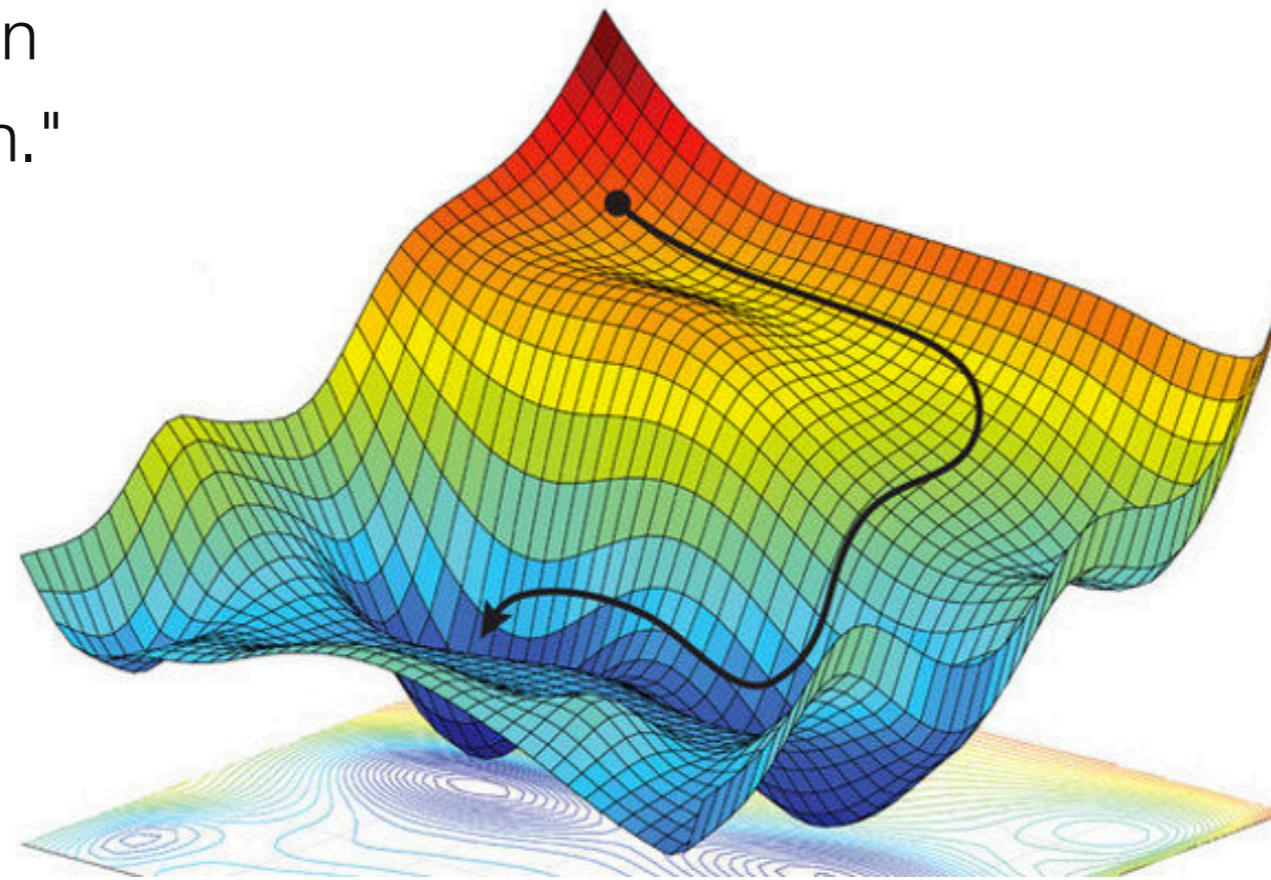


Die Dimensionalität ist hoch.



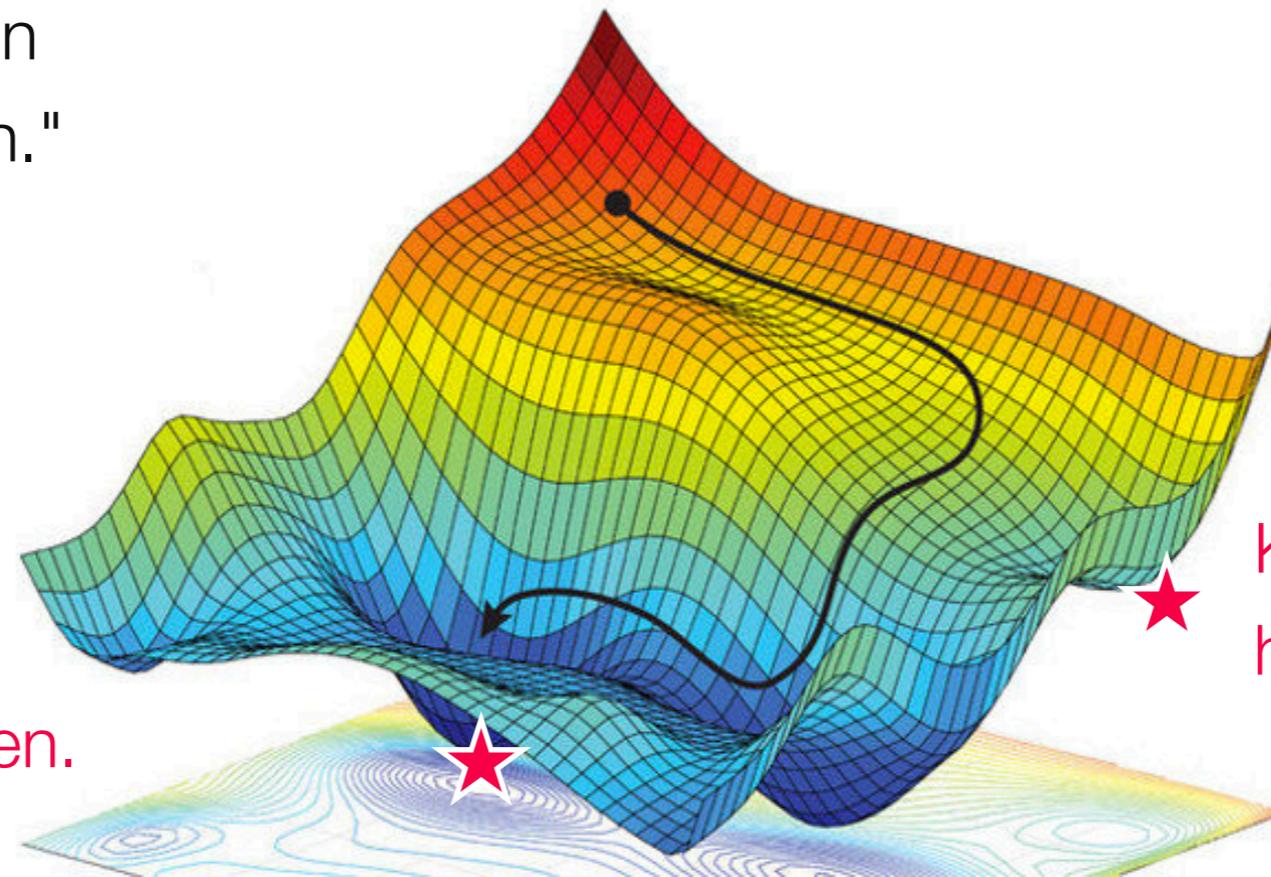
DIE LÖSUNG: DAS „GRADIENT DESCENT“ VERFAHREN.

„Die cost function herunterrutschen.“



DIE LÖSUNG: DAS „GRADIENT DESCENT“ VERFAHREN.

„Die cost function herunterrutschen.“



Könnte hier enden.

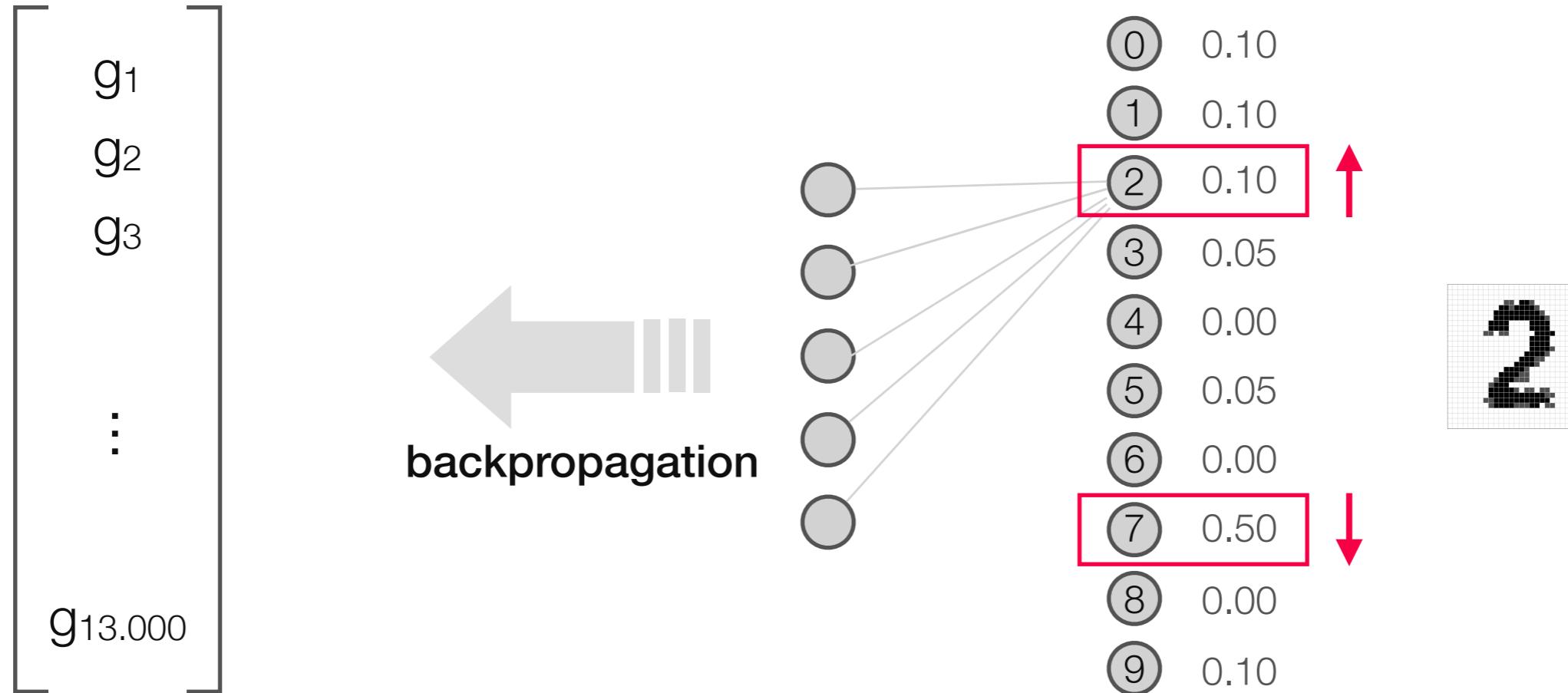
Könnte aber auch hier enden.

SCHRITT DREI: „BACKPROPAGATION“ ODER DIE PARAMETER UPDATEN.

CREDIT ASSIGNMENT PROBLEM

Das Problem zu bestimmen, wie die Leistung eines gesamten Systems (d.h. des NN) auf dessen einzelne Komponenten (d.h. dessen Parameter) zurückzuführen ist.

ZUM GLÜCK HABEN WIR DEN GRADIENTEN BERECHNET.



**AND THAT'S THAT:
NEURAL NETWORKS!**

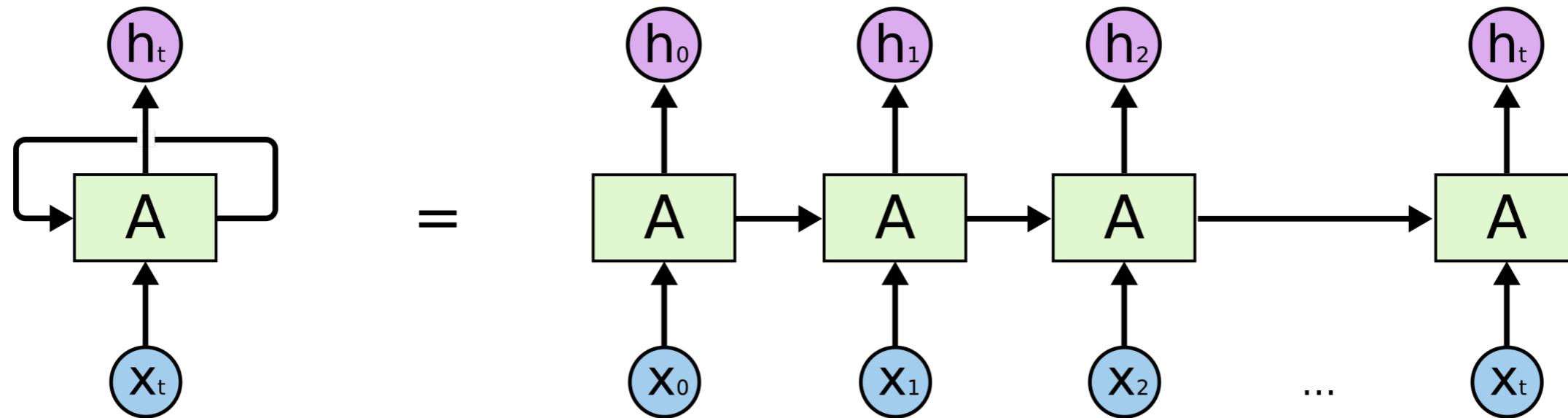
**NOCH EINIGE ERWEITERTE
BEMERKUNGEN...**

DEN ALGORITHMUS NOCH EFFIZIENTER MACHEN: „STOCHASTIC GRADIENT DESCENT“

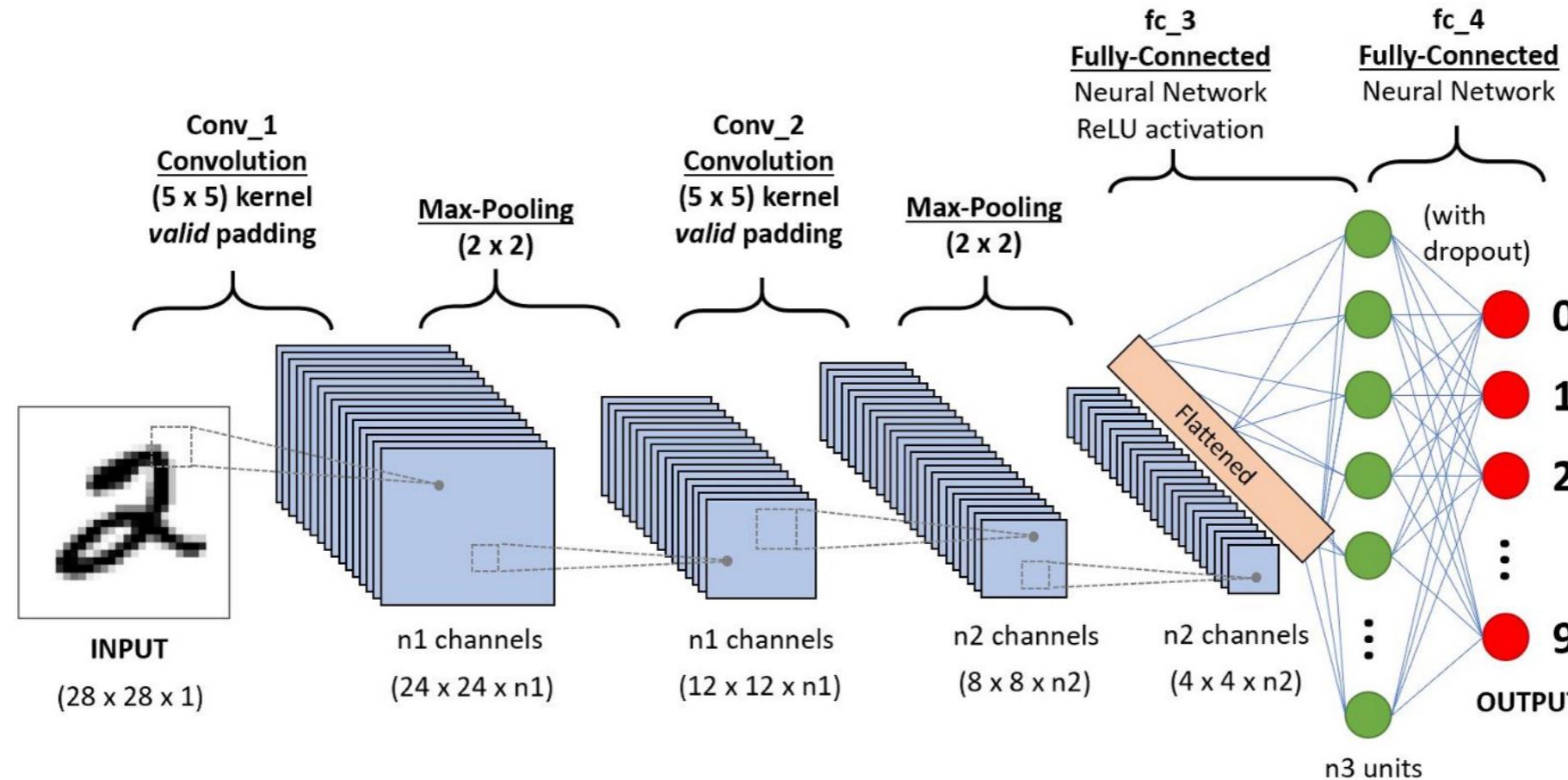
Bei großen Trainingsdatensätzen ist es ziemlich ineffizient der Gradienten für jeden einzelnen Datenpunkt zu berechnen und danach das Backpropagation Verfahren anzuwenden.

Daher wird der Datensatz in sogenannte *mini batches*, also kleinere Datenpakete unterteilt. Gradient Descent und Backpropagation werden dann pro mini batch durchgeführt und insgesamt über alle batches iteriert.

FORTGESCHRITTENE NN ARCHITEKTUREN: RECURRENT NEURAL NETWORKS (RNN)



FORTGESCHRITTENE NN ARCHITEKTUREN: CONVOLUTIONAL NEURAL NETWORKS (CNN)



GIB DEM KIND EINEN NAMEN: NICHT ALLE NEURONALEN NETZE HEISSEN AUCH SO.

CNNs

LeNet

AlexNet

VGGNet

GoogleNet

ResNet

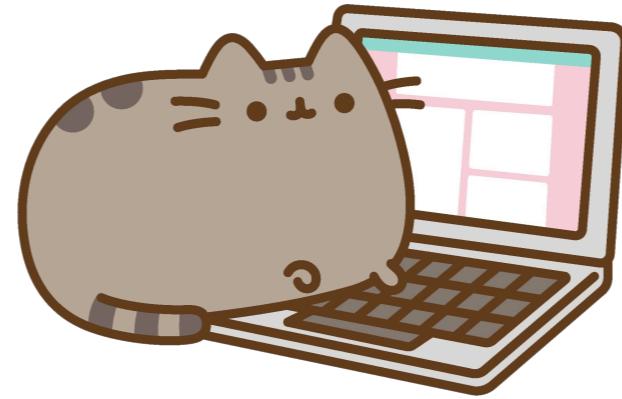
ZFNet

OFFENSICHTLICH IST DAS ALLES ZIEMLICH HARDWARE-INTENSIV...

KANN MEIN COMPUTER DAS ÜBERHAUPT?

Ja. Computer sind und werden immer effizienter.

Trotzdem kommt es auf die Spezifikationen an....



THANK YOU!

NOCH TIEFER EINTAUCHEN: WERTVOLLE QUELLEN

Eine Einführungsserie zu neuronalen Netzen auf YouTube mit wundervollen Illustrationen von 3Blue1Brown:

https://www.youtube.com/playlist?list=PLZHQQObOWTQDNU6R1_67000Dx_ZCJB-3pi

Eine interaktive Visualisierung neuronaler Netze von Tensorflow:

<https://playground.tensorflow.org/#activation=tanh&batchSize=10&dataset=circle®Dataset=reg-plane&learningRate=0.03®ularizationRate=0&noise=0&networkShape=4,2&seed=0.52441&showTestData=false&discretize=false&percTrainData=50&x=true&y=true&xTimesY=false&xSquared=false&ySquared=false&cosX=false&sinX=false&cosY=false&sinY=false&collectStats=false&problem=classification&initZero=false&hideText=false>