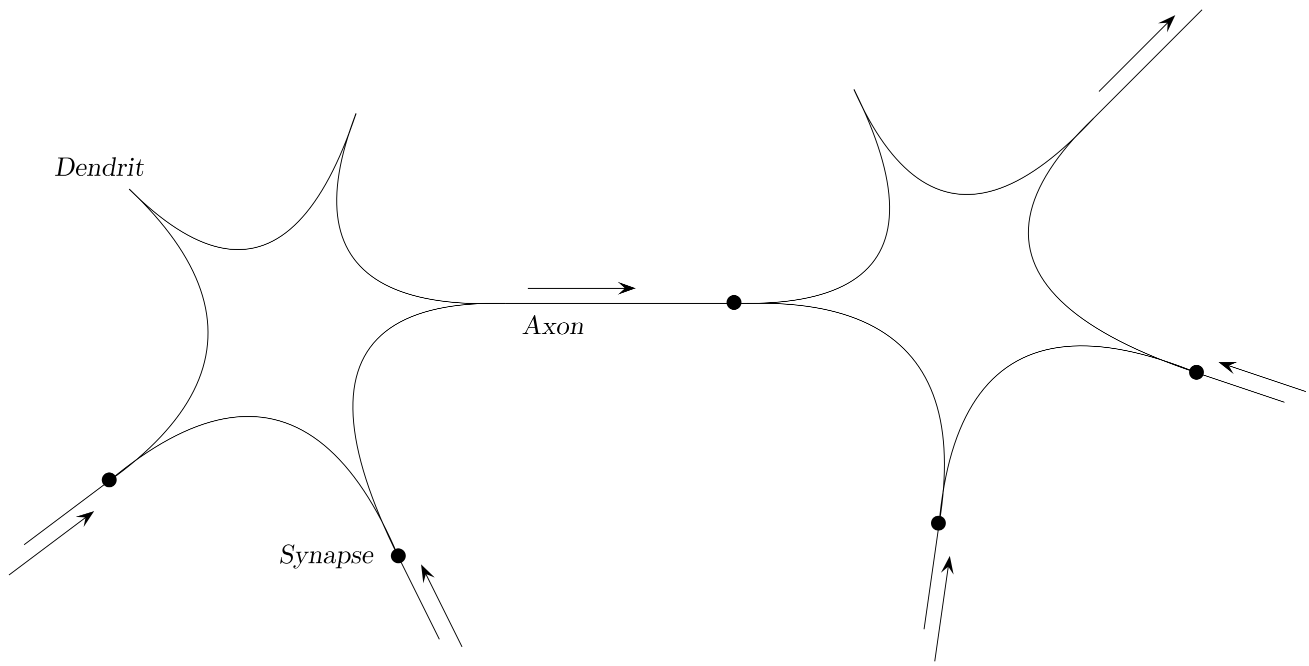
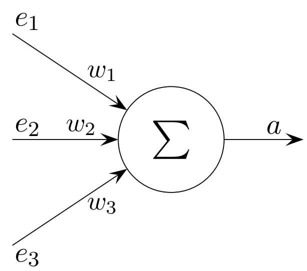
Neuronale Netze

Ein Neuron ist eine Nervenzelle und besteht aus einem Zellkörper, aus dem mehrere Fortsätze, die Dendriten, hervorgehen. Einer von ihnen, das Axon, überträgt die Ausgangssignale der Zelle über seine Verästelungen an andere Neuronen. Die Signale der benachbarten Neuronen liegen an den Synapsen an, die die Signale verstärken, abschwächen oder hemmen können. Das Neuron summiert alle gewichteten Eingangssignale auf und wird erst dann zu einem Ausgangssignal angeregt, wenn diese Summe einen bestimmten Schwellenwert erreicht. Das menschliche Gehirn besitzt etwa 10^{10} Neuronen und jedes von ihnen ungefähr 10000 synaptische Verbindungen zu anderen Neuronen.



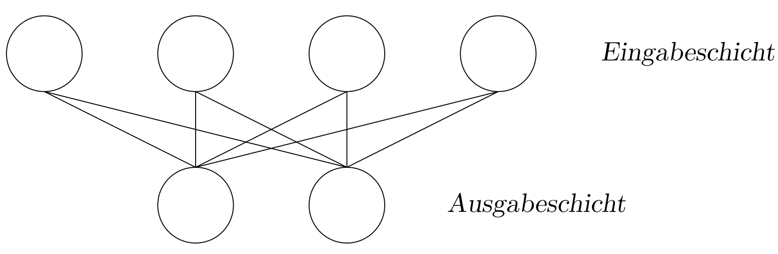
Das Neuronenmodell mit den Eingängen e_1, e_2, e_3, dem Schwellenwert d und der Ausgabe a orientiert sich am biologischen Vorbild.
Eine mögliche Ausgabefunktion wäre:



$$\begin{cases} 0 & \text{für} \quad e_1w_1 + e_2w_2 + e_3w_3 < d \\ 1 & \text{sonst} \end{cases}$$

w_1, w_2, w_3 sind die Gewichte des Neurons,
verstärkend für $w > 1$, abschwächend für $0 < w < 1$, hemmend für $w \leq 0$.

Durch Kopplung mehrerer Neuronen entstehen neuronale Netze.
Im einfachsten Fall liegen lediglich eine Eingabe- und Ausgabeschicht vor.



Ein Lernverfahren für zweischichtige Netze

1. Geben Sie ein neuronales Netz an, dass die logische Und-Funktion (Oder-Funktion) berechnet. Untersuchen Sie, ob auch für n $(n > 2)$ binäre Funktionsargumente ein Ausgabeneuron ausreicht.

Ein Netz, in dem die Neuronen in Schichten angeordnet sind, heißt Perzeptron (perception, engl. Wahrnehmung) und eignet sich zur Erkennung von Mustern. Die Gewichte und die Schwellenwerte der Neuronen sind im Allgemeinen nicht bekannt und müssen der jeweiligen Aufgabenstellung schrittweise in einer Lernphase angepasst werden, zu Beginn können sie zufällig gewählt werden. Anschließend werden dem Netz wiederholt die Eingaben vorgelegt, die es lernen soll zu unterscheiden, und es wird jeweils die tatsächliche Netz-Ausgabe mit der gewünschten verglichen. Für jedes Neuron wird die Differenz

$$\Delta = a_{\text{soll}} - a_{\text{ist}}$$

bestimmt. Um sie zu verkleinern, werden die Gewichte und der Schwellenwert gemäß

$$w_{\text{neu}} = w_{\text{alt}} + \alpha \cdot \Delta \cdot e \qquad \text{(Veränderung ist proportional zu } \Delta \text{ und zur Eingabe } e\text{)}$$

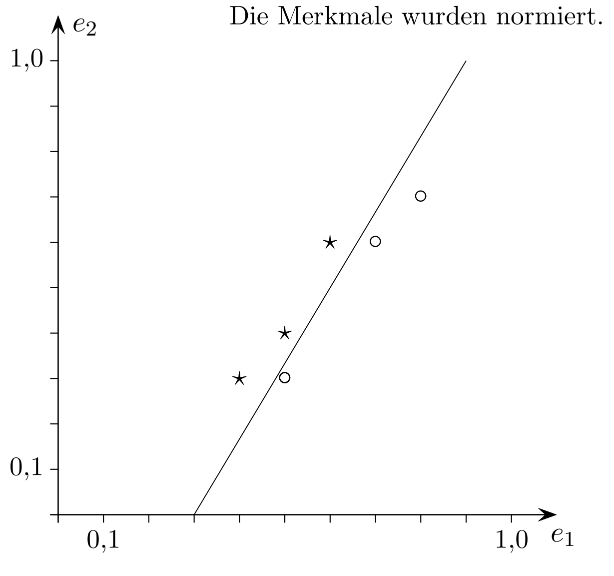
$$d_{\text{neu}} = d_{\text{alt}} + \alpha \cdot \Delta$$

für jedes e-w-Paar nach einem Vorschlag des Psychologen F. Rosenblatt (1958) geändert.

Damit die Veränderungen pro Lernschritt nicht zu stark in Richtung einer einzigen Eingabe erfolgen, sollte die Lernrate α nicht zu groß und dann allmählich kleiner gewählt werden.

Für mehrschichtige Netze wurde erst 1986 ein Lernalgorithmus entdeckt.

2. Die Personen ($\star$,$\circ$) zweier Gruppen sind durch die korrelierenden Merkmale e_1 und e_2 (z. B. Tabakkonsum und Lungenkrebs) gekennzeichnet. Es soll ein neuronales Netz entwickelt werden, mit dem die Gruppenzugehörigkeit erkannt werden kann.

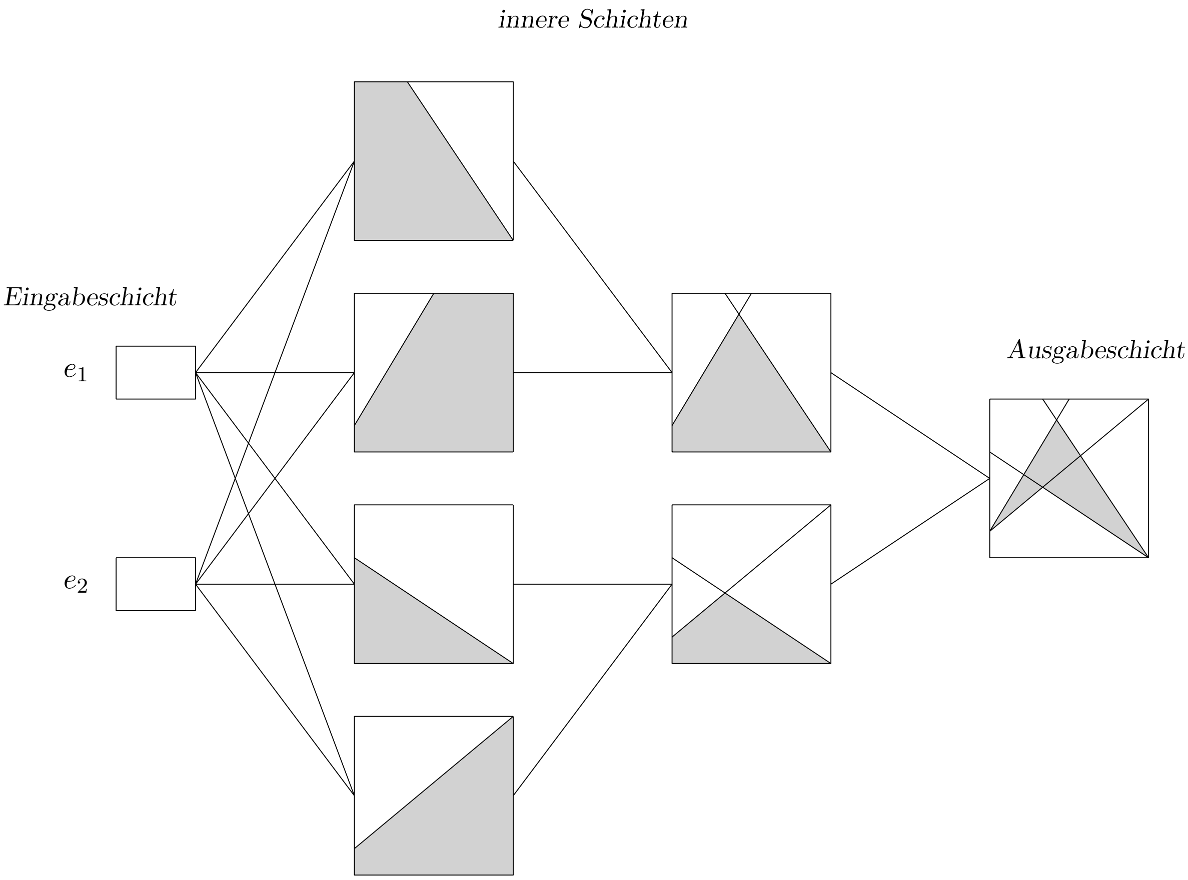


Wie kann diese Problemstellung verallgemeinert werden?

Fähigkeiten mehrschichtiger Netze

Bei einer zweidimensionalen Eingabe e_1, e_2 überprüft ein Neuron, ob $e_1w_1 + e_2w_2 < d$ oder $e_1w_1 + e_2w_2 \geq d$ ist. Für feste Gewichte w_1, w_2 stellt $e_1w_1 + e_2w_2 = d$ eine Geradengleichung dar.
Ein Neuron kann erkennen, ob eine Eingabe unterhalb oder oberhalb bzw. auf der Geraden liegt.
In der Lernphase lassen die Geraden den Fortschritt erkennen.
Bei einer dreidimensionalen Eingabe erhalten wir die Gleichung einer Ebene $e_1w_1 + e_2w_2 + e_3w_3 = d$, die die Punktmengen separiert. Auch höherdimensionale Eingaben sind denkbar.
Nun zeichnen sich Probleme ab, die mit einem zweischichtigen Netz nicht mehr lösbar sind, da die Punktmengen nicht durch eine Gerade getrennt werden können. Hierzu gehört die Berechnung der Exklusiv-oder-Funktion.

Erläutern Sie das gegebene neuronale Netz.

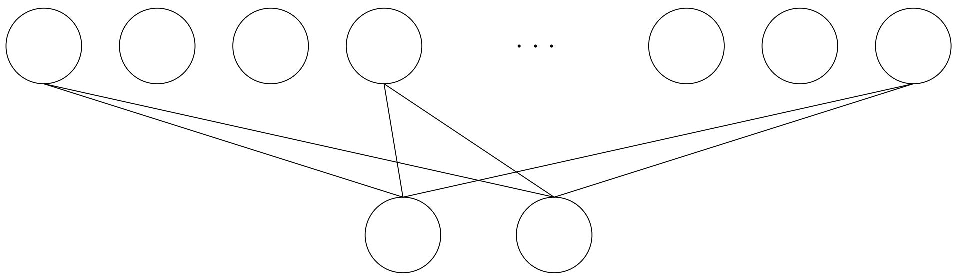


Das Modell von Kohonen

1982 stellte Teuvo Kohonen vom biologischen Vorbild geleitet ein neuronales Netz vor, das ähnliche Reize durch benachbarte Neuronen verarbeitet. Das gedanklich einfache Modell ermöglicht eine Vielzahl überraschender und bedeutungsvoller Anwendungen. Es soll an einem Problem erläutert werden.

> Gesucht ist ein möglichst kurzer Weg, der zwei Städte verbindet
> und auf dem mehrere gegebene Orte liegen.
> Alle benachbarten Orte sind vereinfachend geradlinig miteinander verbunden.

Wir ordnen n Neuronen in einer Reihe an. Ihre Ausgabe wird nicht benötigt. Jedes Neuron dieser Schicht ist mit zwei Neuronen für die Eingabe der Ortskoordinaten vernetzt.



Ziel ist nun, den Orten nachbarschaftserhaltend Neuronen zuzuordnen. Hierzu wird bei Eingabe der Koordinaten eines Ortes jeweils dasjenige Neuron (Erregungszentrum) bestimmt, dessen Gewichtsvektor dem Eingabevektor am ähnlichsten ist.

Ist einem Ort ein Neuron N zugeordnet, werden die Gewichte von N in Richtung der Eingabe verändert und die Gewichte der benachbarten Neuronen denen von N angeglichen, und zwar um so stärker je dichter sie an N liegen. Die Nachbarschaftsumgebung und die Anpassungsrate werden während der Lernphase verkleinert.
Die Anzahl n der Neuronen sollte deutlich (dreimal) größer sein als die Anzahl der Orte. Damit bei unserem Problem Anfang und Ende des Weges dem ersten und letzten Neuron fest zugeordnet werden, wählen wir als deren Gewichte die jeweiligen Ortskoordinaten und die übrigen Gewichte zufallsbedingt.

Welche Änderungen müssten vorgenommen werden, um das Rundreiseproblem zu lösen?

Um das Erregungszentrum zu ermitteln, kann auch das Neuron mit dem maximalen gewichteten Eingang $e_1w_1 + e_2w_2$ bestimmt werden. Dies führt aufgrund der Eigenschaften des Skalarprodukts unter gewissen Voraussetzungen hinsichtlich der Vektorlängen zu gleichwertigen Ergebnissen, wie Kohonen zeigte.
Sein Modell macht Ähnlichkeitsbeziehungen durch Distanzen sichtbar. Es ist auch geeignet, drei- und höherdimensionale Eingaben auf eine Fläche abzubilden. Dies entspricht einer Abstraktion.

Informatik

Mathematik