


# Themen

	Seite	
Maschinelles Lernen	306	1
Trainingsdaten, Parameter, Testdaten	308	2
Künstliches Neuron (Perzeptron)	310	3
Delta-Lernregel	312	4
Mehrlagiges Perzeptron	315	5
Large language models (LLM)	317	6
 Perzeptron in Python programmieren	325	7

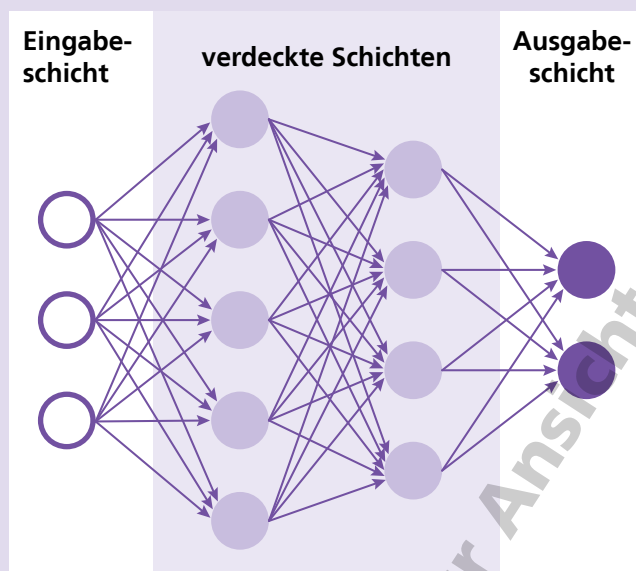
# Mehrlagiges Perzeptron

Mit dem 1957 vorgestellten Perzeptron und der Delta-Lernregel konnten Klassifizierungsaufgaben nur dann gelöst werden, wenn ihre Eingabedaten linear separierbar waren.

Erst 1985 gelang der Nachweis, dass mit einem mehrlagigen Perzeptron (englisch Multi-Layer Perceptron) auch nichtlineare Probleme lösbar sind. Das dabei verwendete Lernverfahren „Backpropagation of Error“ (frei übersetzt „Fehlerrückführung“) ist eine Erweiterung der für das Perzeptron entwickelten Delta-Lernregel.

Ein mehrlagiges Perzeptron ist ein künstliches neuronales Netz, das aus mindestens drei Schichten von Neuronen besteht:

- einer Eingabeschicht,
- einer oder mehreren verdeckten Schichten,
- einer Ausgabeschicht.



Jedes Neuron einer Schicht ist mit allen Neuronen der jeweils nächsten Schicht verbunden. Es gibt keine Verbindungen innerhalb einer oder zur vorherigen Schicht. Alle Verbindungen in einem mehrlagigen Perzeptron haben nur eine Richtung.

Die Neuronen der Eingabeschicht nehmen die Eingabedaten auf und leiten sie unverändert an alle Neuronen der ersten verdeckten Schicht weiter.

In den verdeckten Schichten findet die Verarbeitung der Daten statt. Je komplexer die Aufgabenstellung, desto mehr Schichten und Neuronen pro Schicht werden benötigt.

Die Neuronen der Ausgabeschicht erhalten ihre Eingaben von allen Neuronen der letzten verdeckten Schicht. Die Ausgabewerte der Neuronen in der Ausgabeschicht stellen die Vorhersage des mehrlagigen Perzeptrons für eine Eingabe dar.

Das Training erfolgt wie beim einfachen Perzeptron anhand klassifizierter Trainingsdaten. Der dazugehörige Lernalgorithmus gehört folglich zum Bereich des überwachten Lernens.

Vor dem Training werden die Gewichte aller Verbindungen mit kleinen Zufallszahlen belegt. Die mit diesen Gewichten und den eingegebenen Trainingsdaten erhaltenen Ausgabewerte werden mit den erwarteten Werten verglichen und der Fehler wird berechnet.

Bei der nun folgenden Fehlerrückführung wird der Fehler rückwärts von der Ausgabeschicht zur Eingabeschicht durch das Netzwerk weitergegeben, um alle Gewichte anzupassen. Der Trainingsprozess, bestehend aus Ausgabewert ermitteln, Fehler berechnen, Fehler rückführen und Gewichte anpassen, wird für alle Datensätze der Trainingsdaten wiederholt. Einen solchen Durchlauf nennt man auch Epoche. Über mehrere Epochen hinweg passen sich die Gewichte nach und nach an, bis alle Trainingsdatensätze richtig klassifiziert werden.

Neben dem mehrlagigen Perzeptron gibt es auch andere Arten von künstlichen neuronalen Netzen mit teilweise deutlich komplexerer Architektur. In rekurrenten, also rückgekoppelten neuronalen Netzen gibt es beispielsweise auch Verbindungen von Neuronen einer Schicht zu Neuronen einer vorangegangenen Schicht. Dadurch sind diese Netze für die Handschrift- oder Spracherkennung und Übersetzungsaufgaben geeignet, also für Problemstellungen, bei denen die Reihenfolge der Abarbeitung wichtig ist.

# Mehrlagiges Perzeptron

## Aufgabe 1

Beschreibe den Aufbau eines mehrlagigen Perzeptrons?

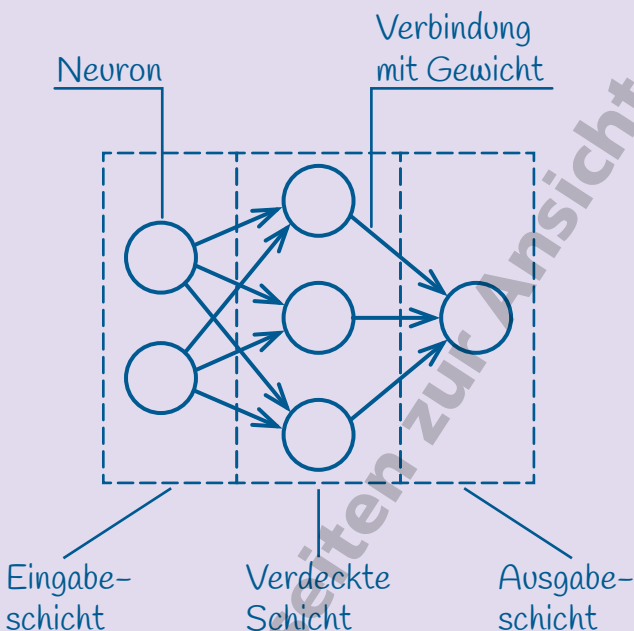
Ein mehrlagiges Perzeptron ist ein künstliches neuronales Netz, das aus mindestens drei Schichten besteht: einer Eingabeschicht, einer oder mehreren verborgenen Schichten und einer Ausgabeschicht.

Jedes Neuron einer Schicht ist mit allen Neuronen der jeweils nächsten Schicht verbunden. Es gibt keine Verbindungen innerhalb einer oder zur vorherigen Schicht.

Alle Verbindungen in einem mehrlagigen Perzeptron haben nur eine Richtung.

## Aufgabe 2

Skizziere ein mehrlagiges Perzeptron und beschrifte dessen Komponenten.



## Aufgabe 3

Wie funktioniert das Training eines mehrlagigen Perzeptrons?

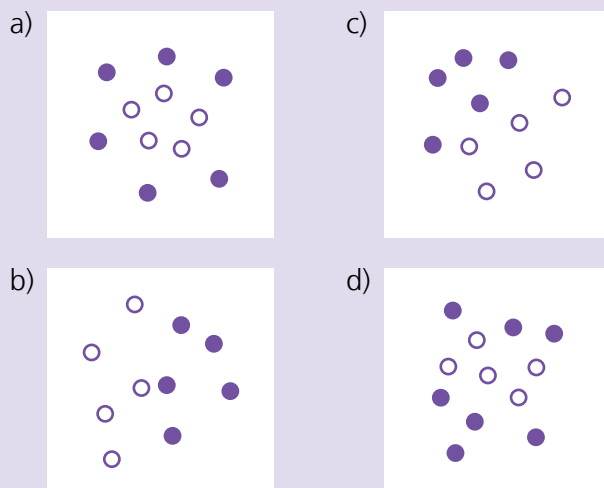
Vor dem Training werden die Gewichte aller Verbindungen mit Zufallszahlen belegt. Mit diesen Gewichten und den eingegebenen Trainingsdaten werden Ausgabewerte ermittelt. Die Ausgabewerte werden mit den erwarteten Werten verglichen und es wird ein Fehler berechnet.

Der Fehler wird von der Ausgabeschicht rückwärts bis zur Eingabeschicht durch das Netzwerk weitergegeben. Dabei werden alle Gewichte angepasst.

Der Trainingsprozess wird für alle Datensätze der Trainingsdaten wiederholt, bis alle Trainingsdatensätze richtig klassifiziert werden.

## Aufgabe 4

Entscheide, ob die folgenden Eingabedaten als Trainingsdaten für ein einfaches Perzeptron oder für ein mehrlagiges Perzeptron geeignet sind.



- a) mehrlagiges Perzeptron
- b) einfaches Perzeptron
- c) einfaches Perzeptron
- d) mehrlagiges Perzeptron