

Artificial Neural Networks

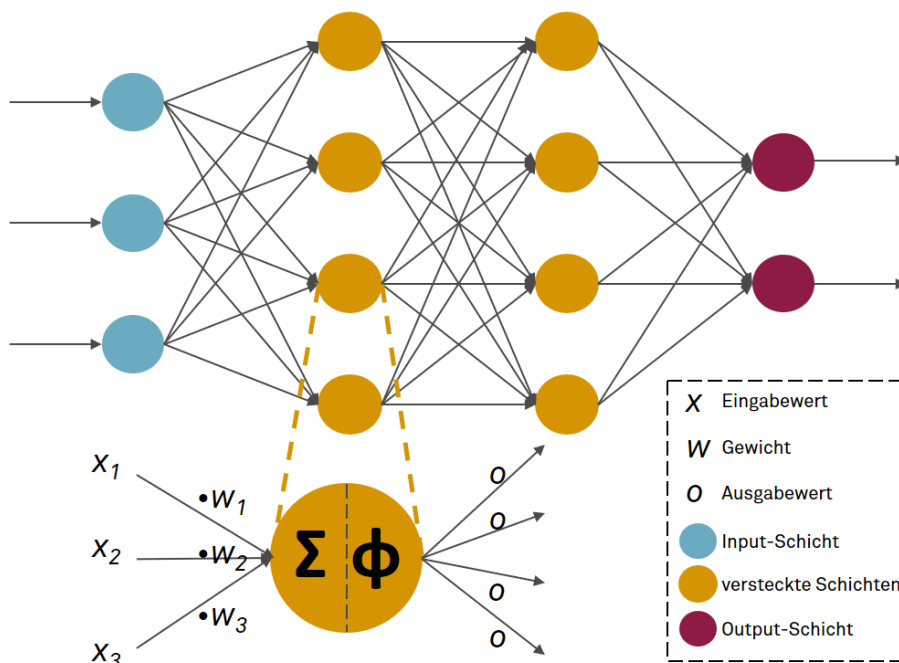
Künstliche Intelligenz nach dem Vorbild des Gehirns

Künstliche neuronale Netze sind das architekturelle Paradigma hinter Deep Learning und vielen anderen Facetten des maschinellen Lernens. Obwohl bereits in den 1950er-Jahren erfunden, wurde ihr Einsatz erst in der letzten Dekade technisch praktikabel. Die Ergebnisse sind zunehmend von beeindruckenderer Qualität.

Definition

Die Inspiration für künstliche neuronale Netze liegt in den Strukturen des biologischen Gehirns. Ein Netz besteht aus einzelnen Neuronen, die jeweils Eingabewerte von anderen Neuronen im Netz annehmen, sie mathematisch verarbeiten und den resultierenden Ausgabewert wiederum an andere Neuronen übergeben. Typischerweise verfügt ein neuronales Netz über Input- und Output-Neuronen, an denen Daten von außen an das Netz herangetragen beziehungsweise die Ergebnisse abgelesen werden. Das Netz als Ganzes lässt sich trainieren, um komplexe KI-Probleme zu lösen.

Innerhalb eines Neurons werden die Eingabewerte mit einer Reihe entsprechender Gewichte multipliziert. Die Summe der resultierenden Produkte bildet dann die Eingabe zu einer Aktivierungsfunktion, welche die Ausgabe des Neurons erzeugt. Es sind diese Gewichte, die während der Trainingsphase des Netzes allmählich angepasst werden, damit das Netz seine Aufgabe erlernt.



1. Einsatzgebiete

- Bilderkennung und -klassifikation
- Verarbeitung natürlicher Sprache
- Analyse molekularer Strukturen
- Empfehlungen und Vorhersagen jeglicher Art

2. Funktionale Bausteine

- Klassifizierung (supervised)
- Clustering (unsupervised)
- Merkmalerkennung
- Regularisierung

Artificial Neural Networks

3. Begünstigende Entwicklungen

- gestiegene Rechenleistung
- Verfügbarkeit großer Datenmengen für Trainingszwecke
- Entwicklung spezialisierter Hardware (GPUs)
- Online-Informationsaustausch unter Experten

4. Topologien

- Multilayer-Perzeptron
- Restricted Boltzmann Netzwerk
- Convolutional Neural Network
- Long-Short-Term-Memory-Netzwerk

Die gängigste Netztopologie ist das Multilayer-Perzeptron, das aus einer Input-Schicht, mehreren versteckten Schichten und einer Output-Schicht besteht: Alle Neuronen in jeder Schicht sind mit allen Neuronen der nachfolgenden Schicht verbunden. Zu den weiteren relevanten Netzarchitekturen gehören das Convolutional Neural Network, das seine Input-Daten

scant, um enthaltene Objekte unabhängig von ihrer Position zu erkennen. Das Restricted Boltzmann Network hat keine Output-Schicht, sondern berechnet die Wahrscheinlichkeit, dass neue Input-Daten seinen erlernten Mustern entsprechen. Das Long-Short-Term-Memory-Netzwerk (kurz LSTM) ermöglicht es hingegen, dass Werte von einer Schicht gespeichert und von einer späteren Schicht ausgelesen werden. Ebenfalls erwähnenswert ist die Neural Architecture Search, ein evolutionäres Verfahren, um die optimale Netztopologie für ein gegebenes Problem auszuloten.

Bei den meisten Netztopologien ist die Anzahl der versteckten Schichten ein wichtiger Konfigurationsparameter. Lange Zeit wurde mit nur wenigen Schichten gearbeitet, was rein mathematisch-theoretisch hätte ausreichen müssen, um beliebig komplexe Probleme abzudecken. Die Erfahrung der letzten Jahre hat aber gezeigt, dass ein Netz mit mehr Schichten erfolgreicher trainiert werden kann und weit überlegene Ergebnisse liefert. Hinter dem Begriff *Deep Learning* verbirgt sich eine dementsprechende Schichtentiefe.

Referenzszenario

Der Einsatz künstlicher neuronaler Netze lohnt sich erst, wenn einfachere Alternativen nicht mehr infrage kommen. Es müssen ausreichende Rechenleistung und Trainingsdaten zur Verfügung stehen und es muss akzeptabel sein, dass die Funktionsweise eines trainierten Netzes weitgehend unverstanden bleiben wird.

Potenzial

Da künstliche neuronale Netze komplett neue Ergebnisfelder ermöglichen haben und, so die Annahme, noch weitere ermöglichen werden, ist ihr starker Einfluss auf den allgemeinen technologischen Fortschritt und damit auf die Geschäftswelt unumstritten.

Reifegrad

Am breiten IT-Markt werden die Funktionsweisen und praktischen Anwendungsmöglichkeiten künstlicher neuronaler Netze momentan ausgiebig ausgelotet. Deshalb erscheinen sie als typisches Hype-Thema. Gleichzeitig wurde das Thema längst über mehrere Jahrzehnte hinweg vor allem in der akademischen Forschung beleuchtet, was ihm wiederum eine gewisse systemische, verlässliche Reife verleiht.

Marktübersicht

Da sich KI-Algorithmen leicht nachimplementieren lassen, unterscheiden sich die führenden Frameworks wie Google Tensorflow, Microsoft Cognitive Toolkit, Theano und Matlab Neural



Buzzword Factor (Ent./Customer)

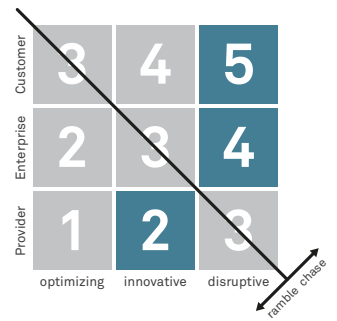
1 low	2 medium	3 high
----------	-------------	-----------

Entry Barrier (Provider)

1 low	2 medium	3 high
----------	-------------	-----------

Benefit Level (Provider)

1 low	2 medium	3 high
----------	-------------	-----------



Network Toolbox vor allem nichtfunktional. Microsoft, Google und Amazon bieten zudem Deep-Learning-as-a-Service an.

Neben Software-Anbietern haben auch Anbieter spezieller Hardware-Komponenten den Markt bereits befüllt, allen voran Nvidia. Denn Grafikkartenkomponenten, wie sie normalerweise für 3D-Spiele zum Einsatz kommen, eignen sich perfekt zur Rechenbeschleunigung künstlicher neuronaler Netze.

Derweil arbeitet die Khronos Group an einem standardisierten Beschreibungs- und Austauschformat für neuronale Netze namens NNEF¹, das den Ex- und Import vereinfacht.

Alternativen

Gleich mehrere alternative Ansätze stehen zur Auswahl. Insbesondere dann, wenn qualifizierte Daten für ein verlässliches Training des künstlichen neuronalen Netzes fehlen, bietet sich die herkömmliche Programmierung an. Dazu gesellen sich Entscheidungsbäume und Regelsysteme. Aber auch klassische, statistische Algorithmen der künstlichen Intelligenz bieten in vielen Fällen adäquate Lösungsoptionen.

Pro	Contra
Qualität der Ergebnisse	keine Debug-Möglichkeit bei Fehlverhalten
Abbildung beliebig komplexer Problemräume	hoher Bedarf nach Rechenressourcen während des Trainings
keine Einschränkungen hinsichtlich der Struktur der Input-Daten	teilweise überzogene Kundenerwartungen

¹ Neural Network Exchange Format, www.khronos.org/nnef