Kurz erklärt: Deep Learning

Maschinen-schlau

Ramon Wartala

Soll ein IT-System komplizierte Sachverhalte erkennen, müssen leistungsfähige Hardware, moderne Algorithmen und neuronale Netze zusammenwirken.

Spätestens seit im März 2016 AlphaGo gegen den weltbesten Go-Profispieler Lee Sedol gewann, ist der Begriff Deep Learning in den Fokus der Öffentlichkeit gelangt. Dieses von der Google-Tochter DeepMind entwickelte System schaffte, woran viele Forscher lange Zeit nicht geglaubt haben: das Strategiespiel Go auf Großmeisterniveau zu beherrschen. Doch auch Nicht-Go-Spieler kommen im Alltag vermehrt mit Deep Learning in Berührung:

– Google Photos speichert nicht nur Fotos in der Cloud, sondern erkennt mit Deep-Learning-Algorithmen Orte und Personen und sortiert die Bildersammlungen danach. Dabei findet das System auch die „besten" Fotos einer Reise und erstellt daraus automatisch Vorschläge für ein Album. Darüber hinaus erkennt Google Photos mehr als 255 000 Sehenswürdigkeiten.

– Facebooks DeepText „versteht" Facebook-Posts, Kommentare und Nachrichten in über 20 Sprachen. Dabei erkennt DeepText die „Stimmung" eines Textes genauso wie Ereignisse, Personen oder Orte. DeepText kommt auch bei Facebooks Message-Bots zur Anwendung, um die Absicht einer Anfrage durch den Nutzer zu erkennen.

– Microsofts Skype-Translator übersetzt gesprochene Sprache simultan in 7 Sprachen und kann Textnachrichten in 50 Sprachen übersetzen.

Die Grundlagen zum Deep Learning legten künstliche neuronale Netze, deren Entstehung bis in die 40er-Jahre des vorigen Jahrhunderts zurückreicht. Zwei Entwicklungen führten ab 2005 zur Renaissance von Anwendungen mit künstlichen neuronalen Netzen: günstige und leistungsfähigere Hardware und immer größere Datenmengen.

Auf der Hardwareseite wuchs nicht nur die Prozessorleistung in jeder Generation, vor allem die Hardware moderner Grafikkarten machte mit jeder neuen Version einen großen Schritt nach vorn: Angetrieben durch die Spieleindustrie und deren Bestrebungen, standardisierte Bibliotheken für das schnelle Berechnen von 3D-Effekten in Echtzeit zu schaffen (wie DirectX oder OpenGL), führten etwa NVIDIA, AMD und Intel programmierbare Grafikprozessoren (GPU) ein. Diese erlauben das parallele Verarbeiten von Algorithmen aus der linearen Algebra, die in allen Formen künstlicher neuronaler Netze zur Anwendung kommen.

Programmierschnittstellen wie CUDA und OpenCL gestatten Entwicklern auch abseits von reiner Grafikausgabe, Algorithmen fürs schnelle Verarbeiten großer künstlicher neuronaler Netze auf GPUs zu implementieren. Moderne High-End-Grafikkarten wie NVIDIAs Titan X verfügen über 3500 Cores für eine massive Parallelisierung beim Rechnen. Im Verbund mit Cluster- und Cloud-Computing lässt sich Deep Learning zudem auf sehr große Datenbestände anwenden, um Muster zu erkennen und daraus Modelle zu erzeugen.

Zu Beginn der Forschung mit künstlichen neuronalen Netzen war die Anzahl der künstlichen Neuronen der Eingabe-, der Ausgabe- und der verdeckten Schichten stark begrenzt. Somit konnten nur einfache Sachverhalte (wie das Erkennen einfacher optischer Muster) gelernt werden. Von Deep Learning spricht man erst, wenn mehrschichtige neuronale Netze zum Einsatz kommen. Dabei werden die erkannten und gelernten Merkmale von Schicht zu Schicht weitergereicht, bis auch komplexe Merkmale erkannt werden.

Um etwa ein Gesicht zu erkennen, versuchen die ersten Schichten, Pixel zu erkennen, die zu einer Hautfläche gehören könnten, und separieren diese von den nicht dazugehörigen. Eine weitere Schicht sucht darauf aufbauend nach Formen, die ein Gesicht ausmachen, wie Augen, Nase, Ohren. Die nächste Schicht erkennt die identifizierten Gesichtsteile in ihrer Anordnung, um daraus abzuleiten, dass es sich wirklich um die Abbildung eines Gesichts handelt.

Von der Einzel- zur Mehrschichtenarchitektur

Für derartige Bilderkennungsaufgaben, aber auch für die Analyse natürlicher Sprache haben sich sogenannte Convolutional Neural Networks (CNN, etwa „faltende neurale Netze") als besonders geeignet erwiesen. Darüber hinaus existieren weitere, je nach Anwendungsfall und Aufgabe spezialisierte Deep-Learning-Netze.

Man unterscheidet zwei Arten: künstliche neuronale Netze für das überwachte und für das unüberwachte Lernen. Anwendungen mit dem Ziel, anhand von Beispielen zu lernen, fallen in die erste Kategorie. Um eigenständig Muster in großen Datenbeständen zu erkennen und diese zu klassifizieren, werden Netze der zweiten Kategorie verwendet.

Mit der Verbreitung der neuen Deep-Learning-Ansätze im maschinellen Lernen sind eine Reihe von Frameworks entstanden, welche die Implementierung von und das Experimentieren mit unterschiedlichen Netztypen und -topologien vereinfachen. Zu den bekannteren gehören Caffe, das an der Universität von Berkeley entwickelt wurde, Theano (Universität von Montreal) und Torch, das als Open-Source-Projekt von Google, Twitter, Facebook und weiteren Firmen getragen wird. In letzter Zeit erfreut sich das von Google zur Verfügung gestellte TensorFlow wachsender Beliebtheit. (tiw)

Ramon Wartala

hilft Firmen beim Einsatz von Big-Data-Technik im Hadoop-Umfeld.

Alle Links: www.ix.de/ix1611116