Einleitung: Entwicklungswege zur KI

Moritz Kirste, Markus Schürholz

Wir nennen uns selbst Homo sapiens – der weise Mensch. Erste Versuche, diese Weisheit zu beschreiben, zu verstehen, abzubilden und in Gesetzmäßigkeiten zu verwandeln, reichen bis in die Antike zurück und haben eine lange Tradition in der Philosophie, Mathematik, Psychologie, Neurowissenschaft und Informatik. Vielfach wurde versucht, den Begriff der Intelligenz – also die kognitive Leistungsfähigkeit des Menschen – besser zu verstehen und zu definieren. Als KI bezeichnet man traditionell ein Teilgebiet der Informatik, das sich mit der Automatisierung von intelligentem Verhalten befasst. Eine genaue Begriffsbestimmung ist jedoch kaum möglich, da auch alle direkt verwandten Wissenschaften wie Psychologie, Biologie, Kognitionswissenschaft, Neurowissenschaft an einer genauen Definition von Intelligenz scheitern.

Die Versuche, Intelligenz zu beschreiben und nachzubilden, lassen sich grob in vier Ansätze unterteilen, die sich mit menschlichem Denken, menschlichem Handeln, rationalem Denken und rationalem Handeln befassen (Russell et al. 2010). So gehört beispielsweise der berühmte Turing-Test (TURING 1950) in den Bereich menschliches Handeln, da bei diesem eine KI menschliches Handeln perfekt reproduziert, während moderne Programme zur Bilderkennung und damit verbundenen Entscheidungen eher im Bereich des rationalen Handelns verortet werden können. Neben den definitorischen Schwierigkeiten befasst sich ein Teil dieser philosophischen Debatte zur KI mit den Unterschieden und Konsequenzen zwischen erstens einer schwachen oder eingeschränkten KI (weak or narrow AI), welche spezielle Probleme intelligent lösen kann, zweitens einer starken oder generellen KI (strong/general AI), welche allgemeine Probleme ebenso gut wie Menschen lösen kann und drittens einer künstlichen Superintelligenz, welche die menschlichen Fähigkeiten weit übertrifft (Kurzweil 2001, Bostrom 2014).

Trotz dieser Vielzahl von Ansätzen und Definitionen lässt sich jedoch ein zentraler Aspekt benennen, den alle als KI bezeichnete Systeme aufweisen: Es ist der Versuch, ein System zu entwickeln, das eigenständig komplexe Probleme bearbeiten kann. Es gibt viele Möglichkeiten, das sehr heterogene Forschungsgebiet der KI und seiner vielen Unterkategorien zu beschreiben. Manche Ansätze befassen sich mit den Problemen, die auf dem Weg zur Intelligenz von Computersystemen auftreten, andere mit den Lösungsansätzen für diese Probleme und wiederum andere mit den Vergleichen zur menschlichen Intelligenz. Um die vielen Teilgebiete soll es hier nicht im

V. Wittpahl (Hrsg.), *Künstliche Intelligenz*,
DOI 10.1007/978-3-662-58042-4_1,

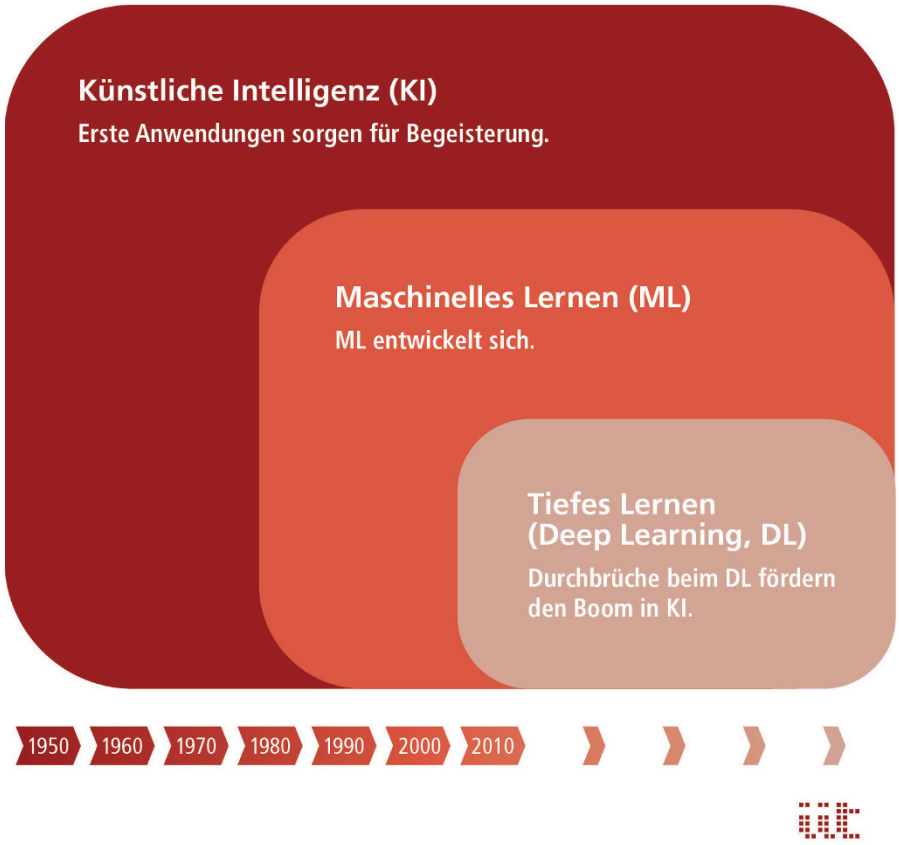

Abbildung A.1: Methoden der KI (eigene Darstellung in Anlehnung an Copeland 2016)

Einzelnen gehen.[2] Vielmehr sollen die wichtigsten Grundlagen der KI ohne den Anspruch auf Vollständigkeit erläutert werden (Abbildung A.1).

Die Anfänge

Am Anfang befasste sich die Entwicklung der KI häufig mit Spielen und mathematischen Repräsentationssystemen von Wissen und Entscheidungen, während seit dem Ende des 20. Jahrhunderts die Technik des maschinellen Lernens (Machine Learning,

[2] *Ausführliche Darstellungen zu Teilgebieten wie Verarbeitung natürlicher Sprache, Wissensrepräsentation, automatisches logisches Schließen, Planung und Wahrnehmung, Robotik und viele mehr finden sich in Russell et al. (2010) und Luger (2003).*

ML) und in jüngster Zeit das tiefe Lernen (Deep Learning, DL) große Erfolge verzeichnen konnten und letztlich das aktuell starke Interesse an KI verursachen.

Erste Ansätze der KI orientierten sich an klassischen Prinzipien der mathematischen Logik. In der Aussagenlogik können einfache logische Verknüpfungen wie UND, ODER, NICHT kombiniert und Aussagen mit einem Wahrheitsgehalt (WAHR, FALSCH) belegt werden, während in der Prädikatenlogik Argumente formuliert und auf ihren Wahrheitsgehalt überprüft werden können. Die ersten Systeme der KI waren logische Repräsentationssysteme, mit deren Hilfe sich einfache Schlussfolgerungen wie Aussage 1: „Die erste Konferenz zu KI fand 1956 am Dartmouth College statt", Aussage 2: „Claude Shannon hat an der ersten Konferenz zu KI teilgenommen" und Schlussfolgerung: „Claude Shannon war 1956 am Dartmouth College" nachvollziehen und beweisen lassen. KI-Systeme, die auf Logik basieren, werden natürlich für deutlich komplexere mathematische Beweise und Theoreme eingesetzt und werden mit Hilfe logischer Programmiersprachen wie PROLOG (Colmerauer und Roussel 1996) bis heute in modernen KI-Anwendungen wie WATSON von IBM genutzt (Lally und Fodor Paul 2011).

Ein beliebtes Anwendungsgebiet der KI war und ist das Gebiet der menschlichen Spiele (Samuel 1959). Dieser Ansatz ist naheliegend, denn die Fähigkeiten der KI lassen sich gut und vergleichbar daran messen, wie gut sie gegen den Menschen spielen oder diesen übertreffen. Der Vorteil dieser Spiele als Messlatte besteht in ihrem üblicherweise einfachen Regelsystem und einfach beschreibbaren Handlungsmöglichkeiten bei gleichzeitig, je nach Spiel, fast unbegrenzten Variationen. Schach beispielsweise hat sehr einfache Regeln, aber geschätzte 10^{120} Zugmöglichkeiten. Diese sehr große Zahl liegt außerhalb der menschlichen Vorstellungskraft und es ist bei einer derart hohen Anzahl zunächst unmöglich, dass ein Programm alle Möglichkeiten durchrechnet, um daraus die perfekte Spielstrategie zu entwickeln. Diese hohe Anzahl von Zugmöglichkeiten entsteht dadurch, dass jede Entscheidung, das heißt jeder mögliche Zug im Schach, wieder neue Entscheidungsalternativen und neue Züge, aber mit jeweils anderen Ausgangssituationen und immer so weiter hervorruft. Diese Entscheidungsvarianten können als Baum oder sogenannter Graph beschrieben werden, bei dem jedes Blatt beziehungsweise Knoten eine Möglichkeit – im Spiel ist das ein Spielzug – darstellt, aus der sich dann immer neue und andere bis ins Unendliche ergeben. So wie ein Baum wächst, so entfalten sich die möglichen Spielzüge in immer wieder neue Verzweigungen und Verästelungen bis ins quasi Unendliche aller möglichen Spielzüge. Einen solchen Baum nennt man Entscheidungsbaum (Decision Tree), und ganze Bereiche der Mathematik und Informatik beschäftigen sich mit der möglichst effizienten Suche in solchen verzweigten Graphen.

Eine sehr effektive Möglichkeit der Suche in Entscheidungsbäumen sind sogenannte Heuristiken. Eine Heuristik ist ein Verfahren, das innerhalb eines solchen zu durchsu-

chenden Graphen für jeden Punkt immer wieder die Sinnhaftigkeit einer weiteren vertieften Suche bestimmt und auf diese Weise verhindert, dass nach der besten Strategie lange – im schlimmsten Falle unendlich lange – gesucht wird. Beim Schach bedeutet dies, dass die möglichen Züge nach bestimmten Kriterien bewertet werden und die Möglichkeiten, die sich aus offensichtlich schlechten Zügen ergeben, nicht mehr weiter in Betracht kommen. Demnach führt die Heuristik dazu, dass ein Entscheidungsbaum ganz gezielt durchsucht wird, bis ein zufriedenstellendes Ergebnis herauskommt, das nicht unbedingt das bestmögliche Resultat sein muss. Entscheidungsbäume und die damit verbunden Heuristiken sind in der KI ein sehr effektives Verfahren für Problemstellungen, die durch ein klares und unveränderliches Regelsystem beschrieben werden können.

Auf die ersten Erfolge der KI im Bereich der Logik und Spiele folgten Versuche, die Verfahren auf allgemeinere Anwendungsfälle zu erweitern. In den 1970er Jahren entstanden Expertensysteme, die über Wenn-Dann-Beziehungen probieren, eine menschliche Wissensbasis in für Computer lesbare Informationen zu verwandeln. Mit den Möglichkeiten zu logischen Schlussfolgerungen und dem effektiven Suchen in diesen Wissensbasen mit Hilfe von Heuristiken konnten die Systeme zunächst einige Erfolge aufweisen und weckten in den 1980er Jahren große Erwartungen an die Möglichkeiten der KI. Ein wesentlicher Nachteil dieser Systeme ist jedoch der immense Aufwand bei der Erfassung menschlichen Wissens und der Umwandlung in die für das Expertensystem notwendige Wissensbasis. Anfang der 1990er Jahre wurden die großen Erwartungen an die KI enttäuscht: Viele Firmen, die zuvor für viel Geld Expertensysteme gekauft hatten, schafften diese wieder ab. Eine große Anzahl von Unternehmen, die solche Systeme angeboten hatten, verschwanden vom Markt. Diese Misserfolge führten gemeinsam mit einer signifikanten Reduktion von Forschungsgeldern im Bereich der KI ab Ende der 1970er Jahre zur ersten und zweiten Phase des sogenannten AI Winters (Crevier 1995).

Maschinelles Lernen

Trotz der genannten Rückschläge für die Forschung wurden in den 1980er Jahren die Grundlagen für den heute so zentralen Ansatz des ML gelegt. Die Grundidee ist einfach: Wie bringt man ein Computerprogramm, das eine bestimmte Aufgabe hat, dazu, aus Erfahrungen zu lernen und mit diesen Erfahrungen die Aufgabe in Zukunft besser zu erfüllen (Mitchell 2010)? Der Unterschied zu einem statischen Programm liegt darin, dass sich die Entscheidungsregeln über eine Rückkoppelung an das Erlernte anpassen (Abbildung A.2). ML unterteilt sich in die drei Hauptkategorien überwachtes Lernen (Supervised Machine Learning), unüberwachtes Lernen (Unsupervised Machine Learning) und verstärktes Lernen (Reinforcement Machine Learning), auf die im Folgenden näher eingegangen werden soll. Zusätzlich unterscheidet

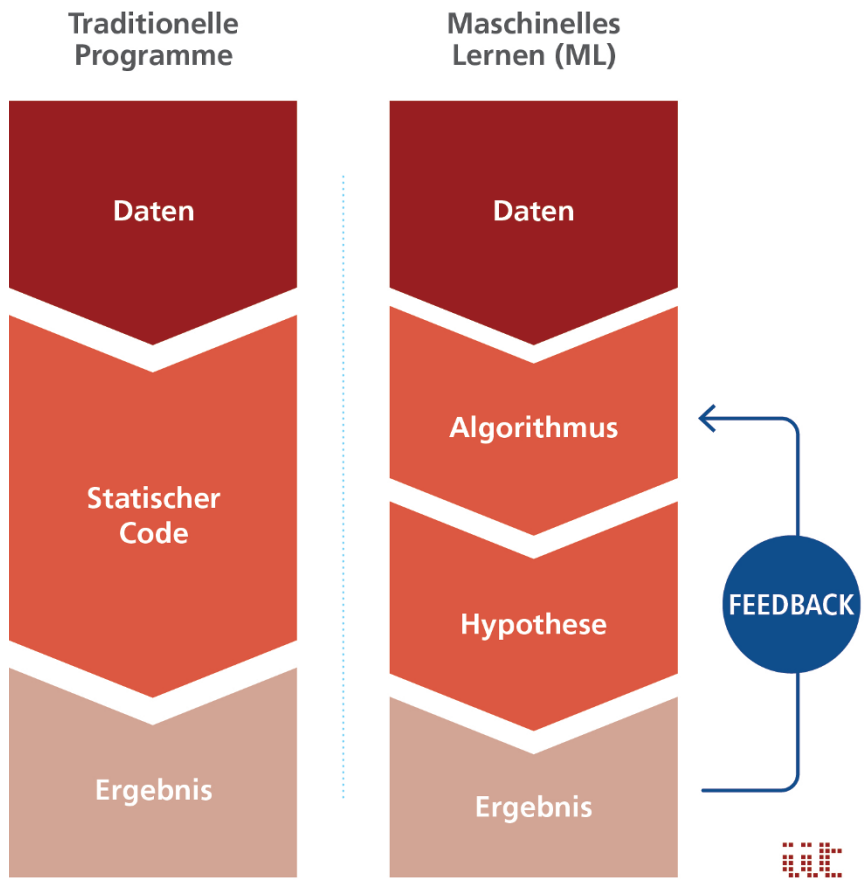

Abbildung A.2: Traditionelle Programme versus ML

man zwischen Offline- und Online-Lernsystemen. Bei dem ersten System findet das Lernen von Verhalten zunächst offline, also getrennt vom Anwendungsszenario, statt. Erst dann wird das Gelernte angewendet und nicht mehr verändert. Die Online-Lernsysteme hingegen lernen und verändern ihr Verhalten stets innerhalb des Anwendungsszenarios und passen sich beständig an.

Beim überwachten Lernen (Supervised Machine Learning) bekommt ein Computerprogramm bekannte Beispieldaten und wird auf eine gewünschte Interpretation und die damit verbundene Ausgabe trainiert. Das Ziel ist es, generelle Regeln zu finden, welche die bekannten Eingabedaten mit den gewünschten Ausgabedaten verbinden, und im Anschluss diese Regeln zu verwenden, um mit neuen Eingabedaten neue Ausgaben zu erstellen. In diesem Sinne hat das Computerprogramm etwas

gelernt, und mit diesem gelernten Wissen lassen sich dann Vorhersagen über künftige und bisher unbekannte Eingabe- und Ausgabedaten treffen. Es entsteht also eine Art eigenständiges Verhalten des Computerprogramms. Das einfachste Verfahren einer solchen Modellbildung ist das der Regression, welches sich an folgendem Beispiel erläutern lässt. Zwischen der Körpergröße und der Schuhgröße eines Menschen gibt es den einfachen linearen Zusammenhang: je größer der Mensch, desto größer auch der passende Schuh. Dieser Zusammenhang lässt sich als lineare Funktion darstellen, mit einer unabhängigen Eingangsvariable (Körpergröße) und einer abhängigen Ausgangsvariable (Schuhgröße). Durch das mathematische Verfahren der Regression werden nun die Parameter der Funktion ermittelt, und man erhält ein Modell, mit dem sich Schuhgrößen aus Körpergrößen vorhersagen lassen (siehe Abbildung A.3).

Ein zweites wichtiges Verfahren des überwachten Lernens ist das der Klassifikation. Dabei werden während des Lernprozesses jeweils mehrere Werte voneinander als Klassen unterschieden und bei der späteren Vorhersage einzelne Werte einer bestimmten Klasse zugeordnet. Beispielsweise könnte man mittels Klassifikation linke und rechte Füße unterscheiden, indem man alle Richtungen eines Fußes genau vermisst (Abbildung A.3). Oder man könnte ein einfaches Modell zur Kreditwürdigkeit erstellen, das auf den beiden Eingabewerten Einkommen und Ersparnisse beruht. Personen unterhalb einer bestimmten Einkommens- und Ersparnisgrenze wären demnach in der einen Klasse, nämlich der nicht kreditwürdigen, und oberhalb einer solchen Grenze in der anderen Klasse, der kreditwürdigen. Der Vorteil der Klassifikation besteht darin, dass immer aufgrund des Zusammenspiels mehrerer Werte beurteilt wird. Demzufolge würde eine Person mit zwar niedrigen Ersparnissen, dafür aber hohem Einkommen in der Klasse kreditwürdig eingeordnet werden.

Sowohl Regression als auch Klassifikation sind Vorhersagemodelle, die Aussagen über die Zukunft treffen können. Sie werden sehr effektiv beispielsweise im Bereich der Preisentwicklung, vorausschauenden Instandhaltung und Bilderkennung eingesetzt. Der Unterschied liegt in der Anwendung: Die Regression erlaubt Vorhersagen über stetige Werte, beispielsweise die Einkommensentwicklung einer Person, während bei der Klassifikation Klassen unterschieden werden, beispielsweise die Kreditwürdigkeit.

Unüberwachtes Lernen (Unsupervised Machine Learning) funktioniert ohne vorher bekannte Zuordnung und Kennzeichnung von Eingabedaten. Die möglichen Ergebnisse sind dabei gänzlich offen. Deshalb kann das Computerprogramm auch nicht trainiert werden, sondern muss vielmehr in den Daten Strukturen erkennen und diese in interpretierbare Informationen verwandeln. Ein anschauliches Verfahren des unüberwachten Lernens ist das Clustering, welches der zuvor beschriebenen Klassifikation ähnelt, mit dem Unterschied, dass beim Clustering die Klassifikationsklassen

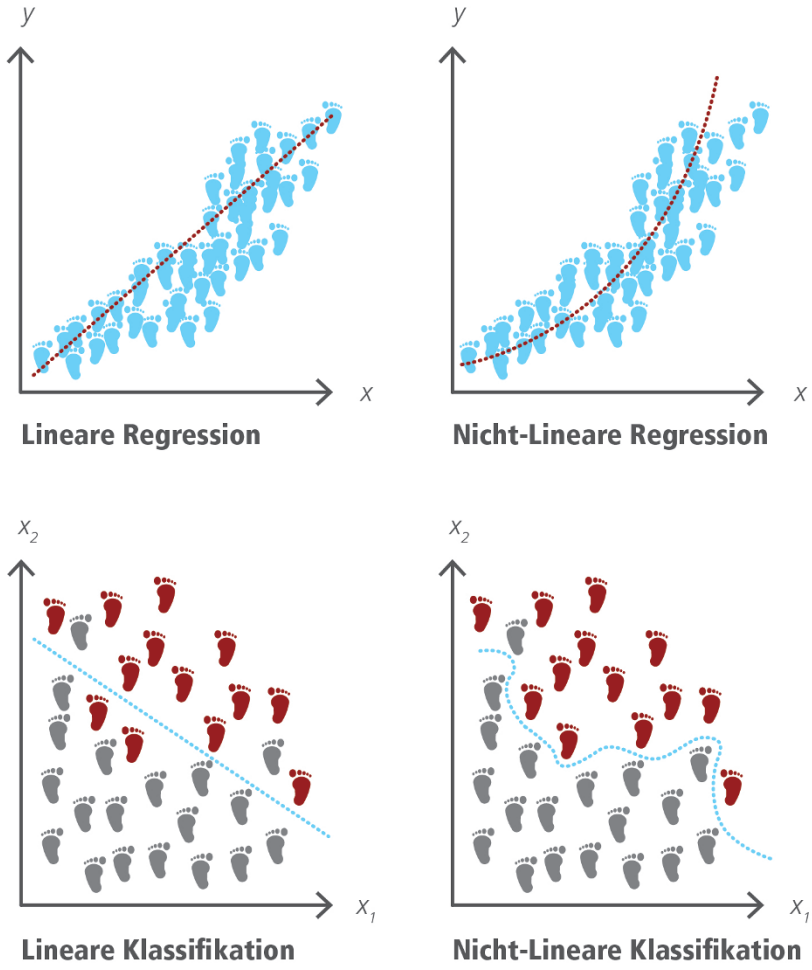

Abbildung A.3: Bei der linearen Regression (oben links) wird zwischen einer Eingangsvariable x (hier die Körpergröße) und einer Ausgangsvariable y (hier die Schuhgröße) ein linearer Zusammenhang hergestellt. Mit dem Modell lassen sich im Anschluss bisher noch unbekannte Werte vorhersagen. Dasselbe ist auch für einen komplizierteren nicht-linearen Zusammenhang möglich (oben rechts). Bei der Klassifikation (unten) werden die Eingangsvariablen für eine Unterteilung in verschiedene Klassen genutzt. In diesem Beispiel wird anhand von zwei Eingabewerten (x1 und x2) unterschieden, ob es sich um linke (grau) oder rechte (rot) Füße handelt. Auch bei der Klassifikation gibt es lineare (links) und nicht-lineare Verfahren (rechts).

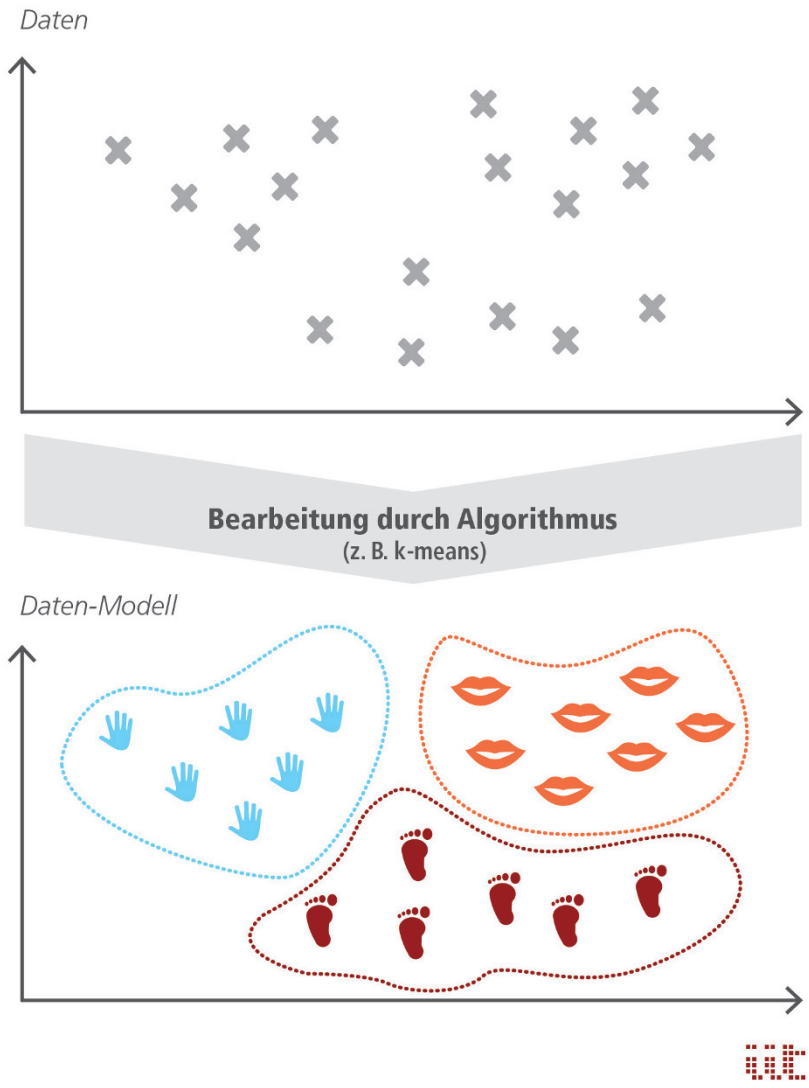

Abbildung A.4: Beim Clustering werden Eingabedaten durch Algorithmen (wie z. B. der bekannte k-means-Algorithmus) in Gruppen zusammengefasst. Alle Mitglieder dieser Gruppen haben ähnliche Merkmale – hier sind es Hände, Füße oder Münder. Auf diese Weise entsteht eine geordnete Struktur in den Daten und das zugehörige Modell kann für eine Interpretation genutzt werden.

erst dadurch entstehen, dass Ähnlichkeiten in den Daten erkannt und zu Gruppen zusammengefasst werden (Abbildung A.4). Weitere weniger anschauliche Verfahren sind Dimensionsreduktion und Hauptkomponentenanalyse sowie Dichteermittlung. Methoden des unüberwachten Lernens kommen in vielen alltäglichen Anwendungen zum Einsatz. So können Kaufverhalten und Nutzerverhalten im Onlinehandel vorhergesagt sowie Empfehlungssysteme beispielsweise für Filme erstellt werden (Netflix Prize o. J.).

Beim verstärkten Lernen (Reinforcement Machine Learning), der dritten Kategorie des ML, lernt ein Computerprogramm direkt aus den Erfahrungen. Hierzu interagiert es mit seiner Umgebung und erhält für richtige Ergebnisse eine Belohnung. Das Programm ist mit einem dressierten Tier zu vergleichen, indem es beispielsweise in einer Spielsituation dafür belohnt wird, wenn es das Spiel gewinnt. Das Ziel ist nun, dass das Programm sich die Konsequenzen seiner Handlung merkt und mit diesem Wissen versucht, seine Belohnung zu maximieren. Die Belohnung ist dementsprechend die Regelgröße, die in diesem Verfahren optimiert wird. Das zurzeit recht bekannte Beispiel für den Einsatz von verstärktem Lernen ist AlphaGo Zero, die Weiterentwicklung von AlphaGo.[3] AlphaGo Zero erlernte das Spiel Go mittels verstärktem Lernen ohne vorherige Kenntnis über das Spiel in nur drei Tagen so gut, dass es besser spielte als seine Vorgängerversion und weitaus besser als die weltbesten menschlichen Spieler (Silver et al. 2017). Verstärktes Lernen könnte sich in den nächsten Jahren als eine wichtige Technologie in der Automatisierung und insbesondere der Robotik erweisen (Kober et al. 2013). So erlernten etwa die Roboterarme der Firma Fanuc mittels verstärkten Lernens binnen weniger Stunden, ihnen bislang unbekannte Objekte sicher zu greifen und zu bewegen (Knight 2016).

Tiefes Lernen

Im Laufe der Zeit wurden unterschiedliche Ansätze, Methoden und (Software-)Technologien unter dem Namen KI entwickelt. Sie werden weiterhin erforscht und adaptiert. Der aktuelle KI-Boom beruht im Wesentlichen auf dem tiefen Lernen mit künstlichen neuronalen Netzen (KNN). So nennt man das Lernen mit Algorithmen, die Netzstrukturen von Nervenzellen nachbilden. „Tief" bedeutet in diesem Zusammenhang unabhängig von der genauen Netzstruktur, dass diese einige bis viele Schichten tief ist. Wie auch im Begriff KI schwingt im alltäglichen Wortgebrauch ein gewisser Hauch von „tiefem Verständnis" abstrakter Zusammenhänge mit. Obwohl sich das tiefe Lernen in Grundzügen an der Funktionsweise biologischer neuronaler Netze

[3] *AlphaGo ist das Programm der Firma Google Deep Mind, das die weltbesten Go-Spieler im März 2016 mühelos schlagen konnte.*

orientiert und viele Medien verkürzt nur von neuronalen Netzen sprechen, gibt es deutliche Unterschiede zum biologischen Vorbild.

Die Neurowissenschaft hat mittlerweile ein gutes Verständnis dafür entwickelt, wie ein einzelnes biologisches Neuron, z. B. eine Gehirnzelle, Information weiterverarbeitet. Dabei geben vorgeschaltete Neuronen elektrische Impulse über chemische Potenziale an ihren Synapsen an ein Neuron weiter. Das Neuron erhält im Zeitverlauf zahlreiche solcher Impulse und lädt sich dabei auf, bis ein Schwellenpotenzial erreicht ist. Dann feuert das Neuron einen eigenen Impuls über sein Axon, das einem großen Datenkabel entspricht, an dessen Ende der Impuls über die eigenen Synapsen des Neurons wieder an nachgeschaltete Zellen weitergegeben wird. Dieser Prozess findet kontinuierlich in allen Neuronen statt, die in ganz unterschiedlichen Netzwerkstruk-

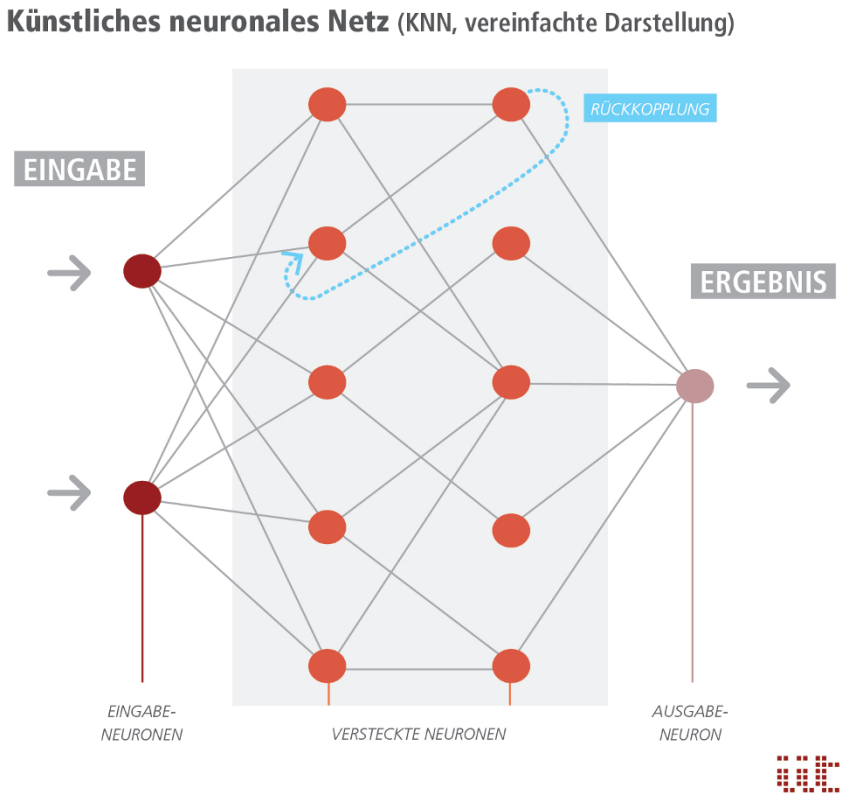

Abbildung A.5: In einem KNN werden Eingabewerte in Schichten versteckter Neuronen (hier beispielhaft zwei Schichten) verarbeitet. Wenn Rückkopplungen (hier der hellblau gepunktete Pfeil) eingesetzt werden, spricht man von einem rekurrenten Netz. Das Ergebnis der Berechnung sind die Ausgabewerte der Ausgabeneuronen (hier nur eins)

turen verschaltet sein können. Eine wesentliche Eigenschaft biologischer Neuronen ist dabei die Verschaltungsstärke oder Gewichtung, mit der ein Neuron seinen elektrischen Impuls jeweils individuell an zahlreiche andere Neurone überträgt. Diese Verschaltungsstärke bzw. ihre Änderung ist neben der Netzwerkstruktur der Neuronen eine wesentliche Eigenschaft für die Verarbeitung von Informationen in biologischen neuronalen Netzwerken.

Einzelne Neuronen können seit 1952 mit Hilfe des Hodgkin-Huxley-Modells simuliert werden (Hodgkin und Huxley 1952), wobei heute sowohl vereinfachte als auch komplexere Simulationsmodelle in Gebrauch sind. Die Simulation ganzer Netzwerke kann in Hinblick auf den Rechenaufwand sehr aufwendig sein. Aktuell werden insbesondere im Human Brain Project[4] große Netzwerke von Neuronen simuliert, perspektivisch sogar in der Größenordnung der Anzahl der biologischen Neuronen im menschlichen Gehirn.

Liest man über neuronale Netze im Bereich KI, so sind damit KNN gemeint, die nicht auf eine genaue Abbildung der biologischen Verhältnisse abzielen, sondern vielmehr nur abstrakt von der Modellierung biologischer neuronaler Netze motiviert sind. Sie setzen primär die Konzepte der Verschaltungsstärke bzw. Gewichtung und des Schwellenwerts informatorisch um. Solche KNN erfüllen ihren Zweck aber in aktuellen Anwendungen. Der KI-Boom speist sich vor allem daraus, dass die Konzepte neuronaler Netze auf bestimmter Hardware stark parallelisiert und effizient ausgeführt werden können (siehe Beitrag 1 „Hardware für KI“).

Die grundlegende Funktionsweise eines neuronalen Netzes ist in Abbildung A.5 dargestellt. Es erhält Eingabewerte, führt darauf Berechnungen durch und ermittelt schließlich die Ausgabewerte. Wie in der Abbildung dargestellt, fließen Informationen auf der linken Seite hinein, durchlaufen das Netz und fließen auf der rechten Seite verarbeitet hinaus. Dabei können in einem komplexeren Netz die Eingabewerte links beispielsweise die Farbwerte der Pixel eines Bildes sein und der Ausgabewert rechts eine Aussage, ob auf diesem Bild ein Hund erkennbar ist. In diesem Fall können die Ausgabewerte ein einfaches Klassifikationsergebnis, also beispielsweise eine 1 (wahr – Hund erkannt) oder 0 (falsch – kein Hund erkannt) sein. Die Ausgabewerte können aber auch eine beliebig komplexere Bedeutung haben. Bei jedem Verarbeitungsschritt werden die Werte aus der jeweils vorhergehenden Ebene weitergeleitet an die einzelnen Neuronen der nächsten Ebene. In einem Neuron der Folgeebene kommen also Werte mehrerer Neuronen an. Wie auch im biologischen Vorbild ist die

[4] *Das Human Brain Project ist ein seit 2013 von der Europäischen Kommission gefördertes Forschungsprojekt, an dem über zehn Jahre hinweg mehr als 100 Institutionen beteiligt sind. Die Gesamtkosten betragen mehr als eine Milliarde Euro.*

Gewichtung der Werte ein wesentliches Element des Netzes. Alle eingehenden Werte werden im Neuron in Hinblick auf ihre Gewichtung und den Schwellwert des jeweiligen Neurons zu einer Ausgabe verarbeitet, die es dann wiederum an mehrere Neuronen der Folgeschicht weitergibt. Dieser Prozess wiederholt sich bis zur letzten Ebene. Zwischen der ersten Schicht, den Eingabe-Neuronen, und der letzten Schicht, den Ausgabe-Neuronen, liegen die sogenannten versteckten Neuronen (Hidden Neurons). Aufgrund der Richtung des Informationsflusses nennt man ein solches Netz Feedforward-Netz. Möglich sind selbstverständlich auch komplexere Netzwerkstrukturen, in denen die Informationen gleichzeitig nach vorne und teilweise auch nach hinten fließen. Beispielsweise könnten die verarbeiteten Informationen einer Neuronenschicht nicht nur an die nächste Schicht weiterfließen, sondern auch an die vorhergehende Schicht zurückgekoppelt werden. Solche Netze bezeichnet man als rekurrente Netze. Die Rückkopplung kann eine Art von „Informationserinnerung" im Netz darstellen und je nach Anwendungsfall sinnvoll werden.

Ein leeres Netz muss zunächst trainiert werden, um seine gewünschte Funktion zu erfüllen. Die Gewichtungen an allen Stellen des Netzes müssen so justiert werden, dass das gewünschte Ergebnis erzielt wird. Beispielsweise müsste ein Netz erst lernen, ob auf Bildern ein Hund abgebildet ist oder nicht. Dieses Anlernen (Training) des Netzes ist dabei viel aufwendiger und rechenintensiver als die spätere Nutzung des Netzes zur Erkennung von Mustern (Inference). Eine Methode zum Anlernen ist die „Backpropagation", die zu den überwachten Lernmethoden gehört. Dabei fließen Eingabewerte in das Netz ein und das Netz errechnet Ausgabewerte. Anschließend wird verglichen, wie weit diese errechneten Ausgabewerte von den Ausgabewerten, die sich eigentlich richtigerweise aus den Eingabewerten ergeben müssten, abweichen. Diese Abweichung bzw. dieser Fehler muss so weit wie möglich gesenkt werden. Dazu werden die Gewichtungen innerhalb des Netzes angepasst. Dann durchlaufen die Eingabewerte wieder das Netz und produzieren neue Ausgabewerte, die wiederum einen gewissen Fehler haben. Dieser Vorgang wird wiederholt, bis der Fehler der Ausgabe ausreichend gering ausfällt. Dazu müssen zu allen Eingabewerten die richtigen Ausgabewerte bekannt sein. Beispielsweise könnte das Netz auf 10.000 Bilder trainiert werden, wobei sich auf vielen Bildern Hunde befinden und auf dem Rest nicht. Danach kann es idealerweise auf neuen unbekannten Bildern erkennen, ob ein Hund abgebildet ist oder nicht. Dabei wird es allerdings manchmal, hoffentlich möglichst selten, falsch entscheiden.

Wenn ein KNN wie oben beschrieben trainiert wird, dann handelt es sich um überwachtes Lernen. KNN können aber ebenfalls für unüberwachtes und für verstärktes Lernen eingesetzt werden.

Für das Beispiel der Erkennung von Hundebildern sind die skizzierten Arten bzw. Funktionen von neuronalen Netzen allerdings noch nicht ausreichend gut. Vielmehr

würde man dafür aktuell „faltende" neuronale Netze (Convolutional Neural Networks, kurz CNN) heranziehen. Faltungen sind mathematische Funktionen, die in der Software zahlreicher Hochtechnologien genutzt werden. In einem CNN kommen in verschiedenen Schichten Faltungen zum Einsatz, die Bildinformationen bzw. Merkmale abstrahieren. In Bildern mit möglichen Hunden sitzen oder laufen die Tiere natürlich nicht immer an der gleichen Stelle. Was einen Hund ausmacht, ist nicht seine Position im Bild, sondern es sind vielmehr Eigenschaften wie das abgebildete flauschige Fell, das im Bild bestimmte weiche Kanten zur Umgebung produziert, bestimmte Muster aus Augen, Schnauze und Ohren oder vier Beine mit hellen Pfoten an den Enden, die in bestimmten Positionen zueinander stehen. Diese Eigenschaften sind manchmal konkreter und manchmal abstrakter, sie finden sich aber nie in den reinen Rohdaten der Pixel eines Bildes. Deshalb funktioniert ein CNN so, dass es Teile des Bildes als Ganzes auswertet und so beispielsweise ein abstraktes Merkmal wie die flauschige, auf dem Bild leicht verschwommene Abgrenzung des Hundes von seiner Umgebung weiterverarbeitet. Das Bild wird also in den Schichten des Netzes abstrahiert und die abstrakteren Merkmale führen am Ausgang des Netzes zu der Entscheidung, ob ein Hund auf dem Bild zu sehen ist oder nicht.[5]

Eine weitere Methode im Bereich der neuronalen Netze sind sogenannte Generative Adversarial Networks (GAN) (Goodfellow et al. 2014). In gewisser Hinsicht kämpfen bei dieser Methode zwei Netzwerke gegeneinander. Dem eigentlich eingesetzten Netz, das lernen soll, wird ein Gegnernetz gegenübergestellt, das die Eingabewerte des lernenden Netzes erzeugt. Das Gegnernetz ist dabei aber so verschaltet, dass es lernt, Eingabewerte zu produzieren, die für das lernende Netz möglichst schlechte Ergebnisse mit einem hohen Grad an Fehlern liefern. Das Gegnernetz konfrontiert das lernende Netz also immer und immer wieder mit seinen Schwächen und führt es an seine Grenzen. Das Ergebnis dieser Auseinandersetzung ist, dass das lernende Netz exzellent wird und selbst mit schwierigen Eingabewerten zurechtkommt.

Unter KI, ML und DL versteht man heute eine ganze Reihe von Ansätzen, Probleme mit Hilfe von autonom agierenden und in diesem Sinne intelligenten Computerprogrammen zu lösen. In den folgenden Kapiteln soll auf bestimmte Teilbereiche der Technologie genauer eingegangen werden. Kapitel 1 beschreibt, welche Rechenhardware nötig ist, um neuronale Netze überhaupt effizient ausführen zu können. Kapitel 2 zeigt mit einer Übersicht zu Normung und Standardisierung auf, wie KI-Werkzeuge aktuell gehandhabt werden. In Kapitel 3 wird dargestellt, wie der Mensch mit komplexen KI-Systemen interagieren kann und könnte. Kapitel 4 befasst sich mit Ansätzen und Methoden im Anwendungsgebiet IT-Sicherheit.

[5] *Jeder, der schon einmal Google Fotos verwendet hat, kennt die Güte der Mustererkennung in Bildern (Computer Vision).*

Literatur

Bostrom, Nick (2014): Superintelligence. Paths, dangers, strategies. 1. ed. Oxford: Oxford University Press.

Colmerauer, Alain; Roussel, Philippe (1996): The birth of Prolog. In: Thomas J. Bergin (Hrsg.): History of programming languages II. [Second ACM SIGPLAN History of Programming Languages Conference (HOPL-II), April 20 - 23, 1993, Cambridge, Massachusetts]. New York, NY, Reading, Mass.: ACM Press; Addison-Wesley, S. 331–367.

Copeland, Michael (2016): What's the Difference Between Artificial Intelligence, Machine Learning, and Deep Learning? Online verfügbar unter https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/, zuletzt geprüft am 26.06.2018.

Crevier, Daniel (1995): AI. The tumultuous history of the search for artificial intelligence. [2. pr.]. New York, NY: Basic Books.

Diff Authors: Autonomous Weapons: an Open Letter from AI & Robotics Researchers. Online verfügbar unter https://futureoflife.org/open-letter-autonomous-weapons, zuletzt geprüft am 23.02.2018..

Goodfellow, I. J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S. et al. (2014): Generative Adversarial Networks. In: ArXiv e-prints.

Hodgkin, A.L.; Huxley, A. F. (1952): A quantitative description of membrane current and its application to conduction and excitation in nerve. In: The Journal of physiology 117 (4), S. 500–544.

Knight, Will (2016): This Factory Robot Learns a New Job Overnight. Online verfügbar unter https://www.technologyreview.com/s/601045/this-factory-robot-learns-a-new-job-overnight/, zuletzt geprüft am 23.02.2018.

Kober, Jens; Bagnell, J. Andrew; Peters, Jan (2013): Reinforcement learning in robotics. A survey. In: The International Journal of Robotics Research 32 (11), S. 1238–1274. DOI: 10.1177/0278364913495721.

Kurzweil, Ray (2001): The Law of Accelerating Returns. Online verfügbar unter http://www.kurzweilai.net/the-law-of-accelerating-returns, zuletzt geprüft am 23.02.2018.

Lally, Adam; Fodor Paul (2011): Natural Language Processing With Prolog in the IBM Watson System. The Association for Logic Programming. Online verfügbar unter https://www.cs.nmsu.edu/ALP/2011/03/natural-language-processing-with-prolog-in-the-ibm-watson-system/, zuletzt geprüft am 23.02.2018.

Luger, George F. (2003): Künstliche Intelligenz. Strategien zur Lösung komplexer Probleme. 4. Aufl., [Nachdr.]. München: Pearson Studium (Pearson Studien Informatik).

Mitchell, Tom M. (2010): Machine learning. International ed., [Reprint.]. New York, NY: McGraw-Hill (McGraw-Hill series in computer science).

Netflix Prize (o. J.): Online verfügbar unter https://www.netflixprize.com/community/topic_1537.html, zuletzt geprüft am 23.02.2018.

Russell, Stuart J.; Norvig, Peter; Davis, Ernest (2010): Artificial intelligence. A modern approach. 3. ed. Upper Saddle River NJ u.a.: Pearson Education (Prentice Hall series in artificial intelligence).

Samuel, A. L. (1959): Some Studies in Machine Learning Using the Game of Checkers. In: IBM J. Res. & Dev. 3 (3), S. 210–229. DOI: 10.1147/rd.33.0210.

Silver, David; Schrittwieser, Julian; Simonyan, Karen; Antonoglou, Ioannis; Huang, Aja; Guez, Arthur et al. (2017): Mastering the game of Go without human knowledge. In: Nature 550 (7676), S. 354–359. DOI: 10.1038/nature24270.

Turing, A. M. (1950): I.—COMPUTING MACHINERY AND INTELLIGENCE. In: Mind LIX (236), S. 433–460. DOI: 10.1093/mind/LIX.236.433.