Erster Teil
Einführung

§ 1 Einleitung

„Kreative Maschinen" – schon diese zwei Worte muten widersprüchlich an, es könnte gar ein Oxymoron vermutet werden. Der Gegensatz scheint nicht größer sein zu können: Maschinen arbeiten stur vorgegebene und bis ins kleinste Detail definierte Aufgaben ab, während „Kreative" – Maler, Schriftsteller, Bildhauer, Beschäftigte der „Kreativbranche" und andere – danach streben, immer Neues und Unerwartetes und Ungewöhnliches zu schaffen. Dennoch mehren sich die Stimmen, die von „kreativer künstlicher Intelligenz", „kreativen Algorithmen" und „Roboterjournalismus" berichten.[1] Einige halten dagegen, Kreativität und Maschinen passen nicht zusammen, schließen sich gegenseitig aus, oder Kreativität sei den Menschen (und vielleicht noch den Tieren) vorbehalten.[2]

Mit den Methoden des Machine Learnings – also Ansätzen, Maschinen lernen zu lassen, sei es aus Beispielen oder im Rahmen des Versuch-und-Irrtum-Prinzips – scheint sich ein Weg zu öffnen, Maschinen von der Anforderung vollständiger Programmierung im Voraus zu befreien und ihnen die Möglichkeit zu geben, sich veränderten Bedingungen anzupassen. Diese angestrebte Abkehr von der vollkommen vorbestimmten Abarbeitung menschlicher Anweisungen könnte Maschinen womöglich auch die Fähigkeit der Kreativität näherbringen, sei es im Rahmen des kreativen Problemlösens oder unter Umständen auch im Kontext der Erzeugung von Ergebnissen, die, wenn sie von einem Menschen geschaffen worden wären, als Werke im Sinne des Urheberrechts eingeordnet werden könnten. Ob die Resultate als Kunst bezeichnet werden können, wenn sie von Maschinen produziert

1 Vgl. z. B. folgende Medienberichte: *Álvarez*, Wenn Computer Kunst schaffen (Tagesspiegel online vom 09.03.2020); *Epstein* et al., iScience Nr. 23 2020; *Vogel*, Kreativität und Codes (sueddeutsche.de vom 16.09.2020); *Kremp*, Roboterjournalismus auf MSN online (Spiegel.de vom 30.05.2020); sowie die Ausführungen zu Roboterjournalismus in *Habel*, Roboterjournalismus; eine Zusammenstellung aktueller kreativer (?) Systeme präsentiert *Miller*, The Artist in the Machine; Internetadressen befinden sich der Lesbarkeit halber im Literaturverzeichnis.

2 Vgl. z. B. *Blöchl*, „KI ist nicht kreativ" (sueddeutsche.de vom 27.11.2020); *Schwab*, 3 reasons why AI will never match human creativity (fastcompany.com vom 25.04.2019); *Miller*, Can machines be more creative than humans? (TheGuardian.com vom 04.03.2019).

werden, bleibt zu klären und ist stark von der gesellschaftlichen Interpretation und Bewertung des Kunstbegriffes abhängig.

Im Kontext selbstlernender und produzierender Systeme ergeben sich zahlreiche urheberrechtliche Fragestellungen. Diese reichen von der grundsätzlichen Frage, ob Maschinen Urheberrechte zugestanden werden könnten, bis zu praktischen Fragen in Bezug auf Schutzmöglichkeiten für solche vermeintlich kreativen Systeme und deren Erzeugnisse.

A. *Problemstellungen*

Im Jahr 2018 befasste sich bereits der Wissenschaftliche Dienst der Bundesregierung mit einigen dieser Fragestellungen.[3] Ziel dieser Arbeit ist es, eine deutlich tiefergehende Betrachtung vorzunehmen und dabei besonderes Augenmerk auf die technische Umsetzung der Konzepte kreativer Maschinen in Form von Machine Learning-Algorithmen zu legen. Notwendigerweise wird dabei auch auf das Konzept der künstlichen Intelligenz (KI) einzugehen sein, die praktischen Ausführungen werden sich jedoch stets auf maschinelles Lernen als Anwendungsfall der KI beziehen.

Dabei stehen die folgenden Problemstellungen im Mittelpunkt:

- Werden im Rahmen des Trainings von Machine Learning-Modellen urheberrechtlich relevante Handlungen vorgenommen, und wenn ja, gibt es eine Schranke zugunsten des Trainings von Modellen maschinellen Lernens mit urheberrechtlich geschützten Werken (in etwa analog oder gar unter Anwendbarkeit der Text- und Data-Mining-Schranke)?
- Wie können – und sollten – Machine Learning-Modelle urheberrechtlich geschützt werden?
- Machine Learning-Modelle werden zunehmend auch zur Erzeugung möglicherweise urheberrechtlich geschützter Werke eingesetzt (Bilder, literarische Werke, Musik,…). Die verwendeten Algorithmen werden immer komplexer, wodurch der Einfluss des menschlichen Inputgebers auf den Output immer geringer wird. Ist das, was im Rahmen des Einsatzes solcher generativer Modelle als Erzeugnis dieses Prozesses entsteht, noch dem Menschen zuzurechnen? Ist das Erzeugnis noch eine persönliche geistige Schöpfung, wenn der Schaffensprozess durch eine „Black Box" geht, die der Mensch nicht vollständig versteht? Oder müssen solche

3 *Deutscher Bundestag*, Künstliche Intelligenz und Machine Learning. Eine urheberrechtliche Betrachtung.

Erzeugnisse als urheberrechtslos, sprich gemeinfrei, verstanden werden? Welche Möglichkeiten gibt es, die Erzeugnisse dennoch zu schützen?
– Welche Zusammenhänge bestehen zwischen Kreativität und (Urheber-) Recht? Was ist Kreativität, und worin besteht der Zusammenhang mit dem Konzept der Intelligenz? Können Maschinen kreativ sein? Wie wirkt sich das auf die Urheberrechte aus?

Im Rahmen dieser Arbeit erfolgt also eine umfassende urheberrechtliche Betrachtung des Machine Learning-Prozesses samt entstehender Erzeugnisse, die aufgrund der starken Orientierung an der tatsächlichen Umsetzung von Machine Learning in der Praxis zur Unterstützung der Beantwortung urheberrechtlicher Fragen herangezogen werden kann. Diese Fragen können etwa ausgehen von einem Benutzer eines auf einer Internetseite bereitgestellten Machine Learning-Systems zur Erzeugung von Bildern oder Musik, der wissen möchte, ob er die Urheberrechte an dem damit erzeugten Ergebnis hat; oder auch von einer Softwareentwicklerin, die sich darüber vergewissern möchte, wie das aufwendig konzipierte und implementierte Machine Learning-Modell urheberrechtlich geschützt ist. Um diesen Praxisbezug herzustellen und die relevanten Technologien zu identifizieren, wurden im Rahmen der Abfassung dieser Arbeit immer wieder Gespräche mit Praktikern aus den entsprechenden Fachgebieten geführt.

B. Gang der Arbeit

Diese Arbeit wird sich dem Thema entsprechend der oben genannten vier Problemstellungen annehmen.

Zunächst erfolgt im ersten Teil eine knappe, übersichtsvermittelnde Einführung in die Themen der künstlichen Intelligenz und des maschinellen Lernens, allerdings ohne dabei tief auf die technischen Grundlagen einzugehen. Stattdessen sind diese jeweils an der Stelle behandelt, an der sie für das Verständnis vorausgesetzt werden.

Der zweite Teil befasst sich mit der Frage, inwiefern der Machine Learning-Prozess urheberrechtlich relevante Handlungen umfasst, und ob diese von Schrankenbestimmungen erfasst werden. Dabei wird sowohl die aktuelle Fassung des Urheberrechtsgesetzes berücksichtigt wie auch die zu erwartenden Änderungen durch die Umsetzung der DSM-Richtlinie 2021.

Der dritte Teil analysiert die urheberrechtliche Situation in der Entwicklung von Machine Learning-Systemen. Dabei erfolgt die Untersuchung nicht

abstrakt, sondern stets anhand für die Praxis relevanter Implementierungen von Machine Learning-Systemen. Besonderes Augenmerk liegt hier auf der Identifikation von Schutzgegenständen und der Bewertung im Rahmen des Computerprogramm- und Datenbankschutzes.

Daran logisch anschließend befasst sich der vierte Teil mit von Machine Learning-Systemen generierten Erzeugnissen und deren urheberrechtlicher Schutzfähigkeit. Insbesondere wird es darum gehen, den Zurechnungszusammenhang zwischen Urheber und Werk im Kontext der Anforderung einer menschlich-gestalterischen Tätigkeit zu untersuchen. Zudem wird ein Schema vorgeschlagen, das die Zusammensetzung gängiger Machine Learning-Modelle berücksichtigt und das dabei hilft, bei einer Vielzahl an dem Entstehungsprozess eines Machine Learning-Erzeugnisses Beteiligter einen Urheber zu ermitteln.

Der fünfte Teil wirft einen Blick über den juristischen Tellerrand hinaus auf Intelligenz und Kreativität und ihre – untrennbare? – Verbindung mit dem Menschen. Im Zentrum des fünften Teils steht die Kreativität, insbesondere in der Form der „Computational Creativity". Zudem wird zu diskutieren sein, ob de lege ferenda ein Schutz hypothetisch autonom geschaffener Werke erforderlich ist oder erforderlich werden könnte.

C. Grenzen der Arbeit

Auch wenn diese Arbeit den Anspruch stellt, insbesondere in Bezug auf die Untersuchung des Schutzes der Machine Learning-Modelle einen starken Praxisbezug aufzuweisen, können doch nicht alle verfügbaren Technologien Berücksichtigung finden. Es mussten daher einige Technologien herausgegriffen werden, die als weit verbreitet und daher sehr praxisrelevant eingeordnet werden.

Zudem unterliegen diese Technologien rapidem technischen Wandel. Die diskutierten Frameworks werden regelmäßig aktualisiert, teilweise geht damit auch eine Änderung der technischen Abläufe und Funktionen einher. Diese Arbeit kann daher, wenn technische Details aus den Frameworks und Bibliotheken bzw. Anwendungsschnittstellen diskutiert werden, nur den während der Erstellung der Arbeit aktuellen Stand wiedergeben. Anhand dessen wird ein Bild gezeichnet, das auch auf zukünftige Entwicklungen anwendbar ist.

Des Weiteren beschränkt sich diese Arbeit auf die Bewertung nach dem deutschen Urheberrecht. Andere Schutzmöglichkeiten, etwa Patentschutz

oder der Schutz der Geschäftsgeheimnisse, für Modelle oder auch für Erzeugnisse, wurden bewusst ausgeklammert, weil diese nicht in der angemessenen Tiefe hätten berücksichtigt werden können und teilweise auch andernorts bereits ausführlich behandelt worden sind. Es wird jedoch an den entsprechenden Stellen auf weiterführende Literatur hingewiesen. Gleiches gilt für die Berücksichtigung anderer Rechtsordnungen.

§ 2 Grundlagen

A. Künstliche Intelligenz

Der Titel dieser Arbeit verwendet weder die Begriffe „künstliche Intelligenz“ noch „Machine Learning“, sondern spricht schlicht von „kreativen Maschinen“. Der Begriff der „Maschinen“ sollte keine Assoziation mit einer großen, dampfbetriebenen Industriemaschine hervorrufen, sondern vielmehr einen Gegensatz zum Begriff des Natürlich-Menschlichen darstellen. Gemeint sind insbesondere „Machine Learning“-Systeme, die in der Regel dem Oberbegriff der künstlichen Intelligenz zugeordnet werden.

Inzwischen dürfte sich der Begriff der „künstlichen Intelligenz“ (KI) – auch im Umfeld der Rechtswissenschaft – weit verbreitet haben, wenn auch die Vorstellungen dazu noch sehr subjektiv und möglicherweise romantisiert oder mit Vorurteilen belegt sind, die vielfach von gefährlichem Halbwissen oder Hörensagen und Science Fiction-Medien erzeugt und befeuert werden.

Was ist also gemeint, wenn von künstlicher Intelligenz die Rede ist? Gleich zu Beginn ist es wichtig, Fiktion von Realität zu trennen. Was ist reine Fantasie? Was funktioniert heute schon? Und was scheitert noch am Stand der Technik, wäre aber theoretisch realisierbar? Um für den weiteren Verlauf der Arbeit eine grundlegende Diskussionsbasis zu schaffen, wird dieses Kapitel KI – losgelöst von rechtlichen Fragestellungen – beleuchten. Dazu erfolgt zunächst ein Überblick über das KI-Thema (und was diese Arbeit darunter versteht). Daran schließen einige Erklärungen zur Funktionsweise zweier betrachteter Technologien an. Das Kapitel schließt mit Ausführungen zu praktischen Anwendungen der genannten Methoden, wobei der Fokus auf solchen Anwendungen liegt, die potenziell urheberrechtlich relevante Handlungen nahelegen.

I. Begriff der künstlichen Intelligenz

Der Begriff der „künstlichen Intelligenz“ gibt Anlass zu der Frage, ob er einen offensichtlichen Gegensatz zur „natürlichen Intelligenz“ darstellen soll. *Herberger* setzt sich mit den Begriffen des Künstlichen und der Intelligenz ausführlich auseinander, und stellt dabei zunächst fest, dass eine direkte

Übersetzung des aus dem Englischen stammenden Begriffs der „Artificial Intelligence" in „künstliche Intelligenz", hält man sich an die historische Bedeutung der Wörter, nicht so recht zu passen scheint.[4] In dieser Arbeit wird auf die Repetition andernorts durchgeführter etymologischer Herleitungen[5] verzichtet und folgender, vereinfachter, intuitiver Definitionsansatz als Grundlage gewählt:

Als „künstlich" versteht der Mensch alles, was nicht im Ursprung natürlich entstanden ist, also was nicht lebendig ist und nicht schlicht aus der Evolution dem natürlichen Lauf der Dinge folgt; „künstlich" ist also etwas, bei dessen Entstehung nachgeholfen wurde, genauer, was der Mensch vollständig – nicht im Wege der Fortpflanzung – erschaffen hat.[6]

Nach dieser Definition wäre „künstliche Intelligenz" eine vom Menschen gesteuerte oder entwickelte Intelligenz. Diese Annahme basiert allerdings darauf, dass der Begriff der „künstlichen Intelligenz" tatsächlich auf „Intelligenz" abstellt und nicht eigentlich auf etwas ganz anderes abzielt.[7] Wie auch später noch festzustellen sein wird, ist auch der Begriff der Intelligenz nicht leicht zu fassen. Was also meint künstliche Intelligenz?

1. Ursprung

Um diese Frage zu beantworten, erscheint es hilfreich, die Ursprünge der KI(-Forschung) zu erkunden. Die offiziellen Anfänge der Wissenschaft um künstliche Intelligenz liegen wohl im Jahr 1956 am Dartmouth College in Hanover, New Hampshire, USA.[8]

Die Idee, den Menschen – bzw. menschliche Gedankenvorgänge – künstlich nachzubauen, ist jedoch viel älter und naturgemäß eng mit der Entwicklung der Computer verbunden.[9] Von Maschinen, die dem Menschen lästige Arbeit abnehmen – sei es, weil hierzu großer Muskelkraftaufwand oder umfangreiche Denkleistung erforderlich ist – ist der Sprung zu solchen Automaten, die mitunter Sparringspartner im intellektuellen Duell sein könn-

4 *Herberger*, NJW 2018, 2825, 2826 f..
5 Ders., NJW 2018, 2825 ff..
6 Vgl. z. B. https://www.duden.de/rechtschreibung/kuenstlich (Stand: 22.02.2021).
7 Dies bezweifelt auch *Herberger*, NJW 2018, 2825 f..
8 *McCarthy* et al., Dartmouth Conference Proposal 1955.
9 Einen ausführlichen chronologischen Überblick über die Anfänge der KI samt Einordnung in die Computergeschichte gibt *McCorduck*, Machines Who Think, S. XXIII ff..

ten, naheliegend (und doch – aus der Perspektive der Anfangszeit – scheinbar unerreichbar).

So kommentierte *Ada Lovelace* bereits 1843 eine Version der „Analytical Engine" von *Charles Babbage*,[10] und entwickelte darin für damalige Zeiten sehr futuristische Gedanken:

> „Supposing, for instance, that the fundamental relations of pitched sounds in the science of harmony and of musical composition were susceptible of such expression and adaptations, the engine might compose elaborate and scientific pieces of music of any degree of complexity or extent."[11]

Hier wird allerdings deutlich, dass weniger Bestrebungen im Mittelpunkt stehen, die menschliche Intelligenz oder die Struktur des Gehirns künstlich nachzubauen, sondern vielmehr eine wissenschaftliche Basis und Berechenbarkeit – in diesem Falle der Musik – zu erforschen, um daraus möglicherweise Regeln aufzustellen, die dann wiederum programmatisch einen Computer „komponieren" lassen könnten. Auch wenn hier noch kein direkter Bezug zur Schaffung künstlicher Intelligenz bestand, so sind diese Gedanken doch Zeugnis dafür, dass der Mensch schon früh Vorstellungen davon hatte, wie klassischerweise dem menschlichen Verstand vorbehaltene Vorgänge maschinell abgebildet werden könnten.

1943 veröffentlichen *McCulloch* und *Pitts* ihre Ausarbeitung „A Logical Calculus of the Ideas Immanent in Nervous Activity"[12], in der sie sich mit der logischen Abbildung neuronaler (Gehirn-)Aktivitäten beschäftigen und umgekehrt feststellen, dass für jede logische Operation ein diese abbildendes Netz gefunden werden kann. Der Bezug zur späteren KI-Forschung ist offensichtlich, insbesondere bei Betrachtung künstlicher neuronaler Netze als Modelle des maschinellen Lernens.

---

10 Die Analytical Engine ist ein früher Entwurf einer mechanischen Rechenmaschine, *Babbages* Notizen dazu sind z. B. hier einsehbar: https://collection.sciencemuseumgroup.org.uk/documents/aa110000020 (Stand: 22.02.2021).

11 *Lovelace*, Scientific Memoirs 1843, 666, 694, Volltext abrufbar unter https://www.fourmilab.ch/babbage/sketch.html (Stand: 22.02.2021), übersetzt: „Angenommen, beispielsweise, dass die fundamentalen Beziehungen von Tönen in der Wissenschaft von Harmonie (Harmonielehre) und musikalischer Komposition fähig wären zu solchem Ausdruck und Anpassungen [wie etwa Nummern im Kontext der abstrakten Wissenschaft von Rechenoperationen], dann könnte die Maschine elaborierte und wissenschaftliche Musikstücke jeglichen Grades an Komplexität oder Ausmaß komponieren."

12 *McCulloch*/*Pitts*, The Bulletin of Mathematical Biophysics Band 5 Nr. 4 1943, 155 ff..

Kurz darauf verfasst *Alan Turing* „Intelligent Machinery“[13], ein Papier, das jedoch nie als solches veröffentlicht wurde. Auch *Turing* orientiert sich darin an der Struktur des menschlichen Gehirns und weist bereits darauf hin, dass auch für die Maschine ein Lernprozess unumgänglich sei,[14] eine Erkenntnis, die heute noch im Kontext des Machine Learnings zentral für die Umsetzung „künstlicher Intelligenz“ ist.

Auch *Isaac Asimov's* berühmte drei Gesetze der Robotik[15] stammen aus der Zeit vor den Anfängen der KI-Wissenschaft: Sein Buch „I, Robot“ erschien bereits 1950, und damit im selben Jahr, in dem auch *Turing* seinen *Turing Test* publiziert.[16]

Aus dem „Proposal for the Dartmouth Summer Research Project on Artificial Intelligence“ von *McCarthy, Minsky, Rochester* und *Shannon* aus dem Jahr 1955 geht hervor, dass die KI-Forschung sich mit Lernprozessen auseinandersetzen soll. Die vorgeschlagene Studie basiert (wie ja ähnlich auch schon von *Ada Lovelace* mehr als einhundert Jahre zuvor formuliert) auf der Annahme, dass jeder Lernprozess so exakt definiert bzw. beschrieben werden kann, dass eine Maschine ihn nachzuahmen in der Lage ist.[17]

Auch wenn das Projekt wohl aus Gründen mangelnden Commitments der beteiligten Wissenschaftler nie so stattfinden konnte wie geplant, hat es gleichwohl den Namen der Wissenschaft geprägt[18] und den sozialen Grundstock an personellen Verbindungen zum effizienten Wissensaustausch gelegt.[19] Es

13 *Turing*, Intelligent Machinery, 107 ff..

14 „It is pointed out that the potentialities of the human intelligence can only be realized if suitable education is provided. The investigation mainly centres round an analogous teaching process applied to machines.“ Ders., Intelligent Machinery, 107.

15 Die Gesetze lauten:
„1. Ein Roboter darf kein menschliches Wesen (wissentlich) verletzen oder durch Untätigkeit (wissentlich) zulassen, dass einem menschlichen Wesen Schaden zugefügt wird.
2. Ein Roboter muss den ihm von einem Menschen gegebenen Befehlen gehorchen – es sei denn, ein solcher Befehl würde mit Regel eins kollidieren.
3. Ein Roboter muss seine Existenz beschützen, solange dieser Schutz nicht mit Regel eins oder zwei kollidiert.“,*Asimov*, I, Robot, S. 44 f., Übersetzung aus dem Englischen: Ders., Meine Freunde, die Roboter: Erzählungen, S. 67.

16 *Turing*, Mind LIX Nr. 49 1950, 433 ff..

17 *McCarthy* et al., Dartmouth Conference Proposal 1955.

18 *McCorduck*, Machines Who Think, S. 114, andere gängige Begriffe für Forschung auf dem Gebiet – im Jahr 1956 – waren etwa „Automata Studies“, *Shannon/McCarthy*, Automata Studies. (AM-34), oder „Complex Information Processing“, *McCorduck*, Machines Who Think, S. 115.

19 Dies., Machines Who Think, S. 130.

klingt also bereits an, dass „Artificial Intelligence“ weniger die direkte Nachahmung der menschlichen Intelligenz meint,[20] sondern vielmehr erforscht werden soll, wie der – bis dato dem Lebewesen vorbehaltene – Lern- bzw. Problemlösungsprozess maschinell modelliert werden kann.

2. KI heute

Nicht nur die Technik, sondern auch das Forschungsgebiet der KI hat sich seit den 1950er Jahren weiterentwickelt. Es sind Teilgebiete entstanden: Computer Vision, Robotik, Sprachverarbeitung (Natural Language Processing – NLP), Entscheidungstheorien und insbesondere maschinelles Lernen.[21] Bis heute ist unklar, wie „künstliche Intelligenz“ tatsächlich zu definieren ist (was auch daran liegen mag, dass selbst die natürliche Intelligenz mit ihrer Definierbarkeit zu kämpfen hat).[22]

Möglicherweise liegt die Herausforderung, KI zu definieren, auch darin, dass KI sich regelmäßig selbst „erledigt“: Sobald ein maßgeblicher Erfolg erzielt wird, also eine Aufgabe oder ein Problem gelöst wird, das scheinbar nur mit KI zu lösen ist, wird diese Lösung schnell zum Standard, und die Messlatte für künstliche Intelligenz steigt.[23] Sobald der Mensch erkennt, wie das Problem gelöst wird, verschwindet der „magische Effekt“ der Intelligenz; „Das kann dann keine Intelligenz mehr sein“, „Das ist ja einfach nur ein Computerprogramm“, wird dann oft gesagt, weshalb viele Fortschritte, die der KI-Forschung zuzuschreiben sind, nicht die Anerkennung bekommen, die sie verdienen.[24] Wenn wir jedoch alles, was – sobald erreicht – als ultimative künstliche Intelligenz gelten soll, nach der tatsächlichen Implementierung als reines Werkzeug abtun, ist es kein Wunder, dass es unmöglich scheint, KI abschließend zu definieren. Und so ist es auch nachvollziehbar, dass sich die Definition, was künstliche Intelligenz ausmacht, ständig verändert und nur punktuell fassen lässt.

---

20 So sieht es allerdings wohl *Zech*, ZfPW 2019, 198, 199, der dies auch als das Verständnis der Dartmouth Conference auffasst.

21 *McCorduck*, Machines Who Think, S. 417.

22 Dies., Machines Who Think, S. 419; *Craglia* et al., Artificial intelligence: A European perspective, S. 19; vgl. dazu außerdem auch § 12 A..

23 So auch *McCorduck*, Machines Who Think, S. 423, *Kaplan*, Artificial intelligence: What everyone needs to know, S. 37.

24 Vgl. auch *McCorduck*, Machines Who Think, S. 423.

a) Definitionsansätze

Gleich vorweg: diese Arbeit erhebt weder den Anspruch, eine neue, uneingeschränkt gültige Definition für KI aufzustellen, noch will sie alle jemals erdachten Definitionen wiedergeben. Vielmehr erfolgt an dieser Stelle eine Destillation des bisher andernorts Gesagten (soweit dies unter Berücksichtigung der schieren Menge der Definitionsversuche möglich ist), um eine für das weitere Vorgehen gangbare Diskussionsgrundlage zu schaffen. Aufbauend auf den Grundlagen und Anfängen der KI-Wissenschaft, die soeben dargelegt wurden, wird im Folgenden ausgegangen von den „klassischen Definitionsansätzen". Sodann werden alternative Vorschläge kurz angerissen, um schlussendlich bei einer strukturellen Definition zu landen.

aa) „Klassische" KI-Definitionen

Definitionen aus der Anfangszeit der KI beschäftigen sich mit dem Kerngebiet der algorithmischen Umsetzung bzw. Nachahmung menschlicher Intelligenz. So wurde im Dartmouth-Konferenz-Proposal (1955) das „Künstliche Intelligenz-Problem" beschrieben als „die Aufgabe, eine Maschine dazu zu bringen, sich so zu verhalten, dass, würde ein Mensch sich so verhalten, diese Verhaltensweise als intelligent bezeichnet würde".[25]

Minsky stellt bereits im Vorwort zu „Semantic Information Processing" (1968) klar, dass Künstliche Intelligenz (bzw. Artificial Intelligence) die „Wissenschaft [ist], Maschinen dazu zu bringen, das zu tun, was Intelligenz erfordern würde, wenn es durch Menschen getan würde"[26] und liegt damit sehr nah an der Definition der Dartmouth-Konferenz: Auch hier wird für einen Vergleich bzw. den Maßstab des Erreichens künstlicher Intelligenz der Mensch herangezogen.

Russell/Norvig scheinen einen anderen Definitionsansatz zu verfolgen: Sie beschreiben „das Gebiet der KI" als den Versuch, „den Beweis für das Verständnis [von Intelligenz] zu führen, indem es intelligente, technische Systeme erschafft",[27] und schlüsseln sodann vier Bereiche auf, denen sie

25 *McCarthy* et al., Dartmouth Conference Proposal 1955; *Ertel*, Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung, S. 1.
26 *Minsky*, Semantic Information Processing, S. V.
27 *Russell/Norvig*, Künstliche Intelligenz, S. 22.

bestehende Definitionen zuweisen: Menschliches Denken, rationales Denken, menschliches Handeln und rationales Handeln.[28]

Kaplan (2016) hingegen definiert „die Essenz der KI" als die Fähigkeit, in einem angemessenen zeitlichen Rahmen passende generalisierende Schlüsse auf der Grundlage begrenzter Daten zu ziehen,[29] und führt dann weiter aus, dass Lernen ein Prozess des Generalisierens sei, bei dem frühere Erfahrungen in zukünftige Analysen einbezogen werden.[30] Auffällig ist: hier wird kein Vergleich mehr zu menschlichen Fähigkeiten oder dem Menschen insgesamt gezogen.

Kaplan stellt allerdings auch fest, dass dies nicht nur die „Essenz der KI" beschreibt, sondern gleichfalls die „Essenz der Intelligenz".[31] Er unterscheidet also gar nicht mehr zwischen der menschlichen und der künstlichen Intelligenz, sondern wählt eine abstrakte Definition, die es dann später technisch umzusetzen gilt, nur um dann wenig später zu konstatieren:

> „We may not be able to define AI just yet, but in the meantime I'm confident that most people feel, as U.S. Supreme Court justice Potter Stewart famously said of pornography, 'I know it when I see it'."[32]

Die *WIPO* (2020) definiert KI in ihrem „Revised Issues Paper" vom 21.05.2020 ähnlich wie Minsky, nämlich als „Disziplin der Computerwissenschaften, die darauf abzielt, Maschinen und Systeme zu entwickeln, die Aufgaben ausführen können, von denen angenommen wird, dass sie menschliche Intelligenz erfordern, ohne oder mit nur eingeschränkter menschlicher Intervention."[33]

Auch denkbar ist eine Umformung des englischen Akronyms AI in „Augmented Intelligence", also „verbesserte" oder „erweiterte" Intelligenz, wobei auch hier die Interpretationen in verschiedene Richtungen gehen: So könnte Augmented Intelligence als eine Beschreibung der Partnerschaft zwischen Mensch und Maschine verstanden werden,[34] oder andererseits als zweite Stu-

28 *Russell/Norvig*, Künstliche Intelligenz, S. 23.
29 *Kaplan*, Artificial intelligence: What everyone needs to know, S. 5.
30 Ders., Artificial intelligence: What everyone needs to know, S. 6.
31 Ders., Artificial intelligence: What everyone needs to know, S. 5.
32 Ders., Artificial intelligence: What everyone needs to know, S. 7, Nachweis zu dem Zitat des Richters dort in Fn. 10.
33 *WIPO*, Revised Issues Paper, S. 3: „Artificial Intelligence (AI) is a discipline of computer science that is aimed at developing machines and systems that can carry out tasks considered to require human intelligence, with limited or no human intervention."
34 *Hurwitz* et al., Augmented Intelligence: The Business Power of Human–Machine Collaboration, S. 2.

fe dreier Reifestadien von KI (Assisted Intelligence, Augmented Intelligence, Autonomous Intelligence)[35].

Sascha Lobo stellt schließlich 2019 fest, im Alltag ließe sich KI „synonym zu »lernende Mustererkennung« verwenden“[36]. Ein Bezug zum Begriff der Intelligenz wird hier gar nicht mehr hergestellt, die Definition klingt vielmehr wie eine (zu kurz greifende, aber in der aktuellen Praxis wohl sehr treffende) Reduktion von KI auf maschinelles Lernen.

Für die „klassische“ Definition des Begriffs der KI gestaltet sich die Einigung auf einen allgemeinen Begriff erkennbar schwierig.

bb) Alternative Definition

Die oben (§ 2 A.I.2.) genannten und heute zur KI gezählten Teilgebiete der KI betrachtend drängt sich der Gedanke auf, ob das, was erreicht werden soll, überhaupt noch unter den Begriff der Intelligenz zu fassen ist. Schließlich soll diese Wissenschaft scheinbar nicht nur das erforschen und nachzumodellieren versuchen, was im Inneren des menschlichen Gehirns vorgeht, sondern auch Robotik (also der nach außen sichtbare, mechanische Teil, der notwendig ist, um einen tatsächlichen Menschen zu simulieren), Computer Vision (also das Sehvermögen) und die Sprachverarbeitung (als Kommunikationsmöglichkeit) umfassen.

Eventuell wäre es also passender, das Gesamtgebiet der „künstlichen Intelligenz“ mit seinen Teilgebieten eher als „Lebewesensimulation“, „Artificial Life“ oder „Artificial Human“ zu bezeichnen, und die künstliche Intelligenz auf das zu beschränken, was tatsächlich im Kopf vorgehen soll. Schließlich funktioniert das Gehirn auch ohne die Möglichkeit, zu sehen, zu sprechen, zu hören, oder Körperteile zu bewegen (obgleich ein Umkehrschluss nicht funktionieren mag).

cc) Hier: Strukturelle Definition

Ein eher nüchterner, jedoch auch in dieser Arbeit vertretener Ansatz betrachtet KI als Oberbegriff für die oben genannten Teilgebiete (Robotik, Computer

35 *Mohanty/Vyas*, How to Compete in the Age of Artificial Intelligence: Implementing a Collaborative Human-Machine Strategy for Your Business, S. 13.

36 *Lobo*, Realitätsschock, S. 221.

Vision etc.) denen auch das maschinelle Lernen zuzuordnen ist.[37] Durch diese strukturelle Sicht auf künstliche Intelligenz nähert sich der Begriff der Realität: Es wird in zahlreichen Teildisziplinen versucht, menschliche Fähigkeiten nachzubilden. Und das, was im Rahmen dieser Arbeit behandelt wird – der „Denkvorgang" oder „Lernprozess", wenn man so will – begründet das Forschungsgebiet des maschinellen Lernens. Angesichts dessen, was auch schon die „Urmütter" und „Urväter" der KI erreichen wollten,[38] scheint dieser Begriff ohnehin passender. Diese Arbeit befasst sich mit dem, was in Code und Daten gefasst werden kann, also mit dem, was „im Kopf der Maschine vorgeht", weshalb vorwiegend eine Beschreibung als maschinelles Lernen erfolgt. Wenn in den folgenden Kapiteln dennoch einmal von KI die Rede ist, so ist darunter – hier – maschinelles Lernen zu verstehen.

b) Funktionsmäßige Kategorisierung

Sobald eine gangbare Definition gefunden wurde, erfolgt üblicherweise eine funktionale Unterteilung nach der Leistungsfähigkeit von KI-Systemen. Unterschieden wird nach schwacher und starker KI,[39] oft auch spezifische und generelle KI genannt.[40] Diese Zuordnung ist nicht ganz unumstritten. Die Begriffe „stark" und „schwach" werden im KI-Umfeld auch eingesetzt, um zwischen zwei Perspektiven auf KI zu unterscheiden: Die eine Seite ist überzeugt davon, dass KI-Maschinen bereits jetzt oder in der Zukunft einen „Geist" haben werden, während die andere auf dem Standpunkt steht, dass KI-Technologien Intelligenz weniger duplizieren als simulieren.[41] Zudem veranschaulicht diese Unterscheidung in sich auch die Hoffnung, vielleicht auch das Ziel der KI-Forschung, eine generelle, omnipotente KI zu schaffen,

37 So auch *EC HLEG AI*, A definition of AI, 7; *Herberger*, NJW 2018, 2825, 2827; *Linke*, GRUR Junge Wissenschaft 2019, S. 30 und S. 37; ebenfalls auf eine konkrete Definition verzichtend und auf Gebiete bzw. Elemente zurückgreifend *Kaulartz/Braegelmann*, Rechtshandbuch Artificial Intelligence, Kap. 1 Rn. 9.

38 Vgl. aa).

39 So auch *Kaulartz/Braegelmann*, Rechtshandbuch Artificial Intelligence, Kap. 1 Rn. 10; ebenso *WIPO*, Revised Issues Paper, S. 3 f..

40 Unterscheidung nach stark und schwach geht zurück auf *Searle*, The Behavioral and Brain Sciences 3 1980, 417; ebenfalls Ausführungen zur Unterscheidung machen u. a. *Lenzen*, Natürliche und künstliche Intelligenz, S. 15; *Datenethikkommission der BReg*, Gutachten der Datenethikkommission der Bundesregierung, S. 59; *Koncsik*, Quantum Mind, S. 17; *Specht-Riemenschneider*, FS Taeger, 711, 713.

41 *Kaplan*, Artificial intelligence: What everyone needs to know, S. 68.

und stellt dem die Realität gegenüber: Der letzteren Kategorie der „starken" KI dürften wohl bis dato noch keine Systeme angehören. Vielerorts wird eine Kategorisierung dieser Art vorgenommen, für die Zwecke dieser Arbeit ist diese Unterscheidung nicht relevant und wird daher nicht weiter erläutert.

II. Zusammenfassung

Um die zu Beginn des Kapitels aufgeworfene Frage zu beantworten: Nein, „künstliche Intelligenz" soll – zumindest in dieser Arbeit – keinen Gegensatz zum Begriff der „natürlichen Intelligenz" darstellen, sondern nur als Oberbegriff verschiedener Disziplinen verstanden werden, von denen besonders eine – das maschinelle Lernen – im Fokus der Betrachtung stehen wird.

B. Maschinelles Lernen

Ein Thema, dass schon zu Dartmouth-Zeiten[42] untersucht wurde, und nach wie vor große Aufmerksamkeit bekommt, ist das Modellieren von Lernprozessen, und damit verbunden der quasi selbständige Wissensaufbau durch Algorithmen. Dieses Feld wird heute unter dem Oberbegriff des maschinellen Lernens gefasst. Es gehört ebenso zu den Anfängen der KI, und seine Erforschung hat sogar bereits vor der Dartmouth-Konferenz begonnen: Maschinelles Lernen geht wohl zumindest zurück auf *Warren McCulloch* und *Walter Pitts* und deren Forschung in 1943, auch wenn das maschinelle Lernen bis in die 1990-er Jahre nicht vorrangig erforscht wurde,[43] was auch an der sehr begrenzten zur Verfügung stehenden Rechenkraft und Speicherkapazität gelegen haben könnte.[44] Die Haupttreiber dieses Forschungsfeldes waren die Steigerung dieser technischen Möglichkeiten, die zunehmende Digitalisierung bzw. digitale Erfassung von Daten, einfacherer Zugang zu denselben (v. a. unterstützt durch das Internet) sowie günstige hochauflösende digitale Sensoren.[45]

42 Vgl. aa).
43 *Kaplan*, Artificial intelligence: What everyone needs to know, S. 32.
44 Ders., Artificial intelligence: What everyone needs to know, S. 39; *Goodfellow* et al., Deep Learning Handbuch, S. 20.
45 *Kaplan*, Artificial intelligence: What everyone needs to know, S. 39.

I. Definitionen

Zunächst eine grundlegende Begrifflichkeitsklärung: „Maschine" in „maschinelles Lernen" meint nicht zwangsläufig eine „Maschine" im klassischen Sinne. Mit „Maschine" wird oft ein mechanisches Gerät assoziiert, das – mehr oder weniger automatisiert – etwa in der Produktion von Gegenständen eingesetzt wird, oder aber zunehmend auch Computer im Hardwaresinne. Im Kontext von „Maschinellem Lernen" bzw. „Machine Learning" zielt „Maschine" aber wohl eher in die Software-Richtung: So wurde der Begriff bereits in den 50er-Jahren im KI-Kontext vor allem als Synonym für „Programm" verwendet, mutmaßlich daraus folgend, dass *Turing* seine abstrakten Prozeduren bereits als „Maschine" bezeichnete.[46]

Es geht also um lernende Algorithmen, oder auch um Lernen *durch* Algorithmen (denn von einem *lernenden* Algorithmus könnte erwartet werden, dass er sich verändert – und das ist nicht immer der Fall).

1. Algorithmus

Algorithmen lassen sich definieren als „Verfahren zur schrittweisen Umformung von Zeichenreihen" oder als „Rechenvorgang nach einem bestimmten, sich wiederholenden Schema",[47] allerdings wird der Begriff häufig auch verwendet, wenn von Computerprogrammen die Rede ist. Jedoch sollten „Computerprogramm" und „Algorithmus" nicht synonym verwendet werden, denn ein Algorithmus (als allgemeine Regel zur Lösung eines Problems) kann durch verschiedenste Computerprogramme in eine für einen Computer verständliche Sprache umgesetzt werden. In dieser Arbeit werden Algorithmen in ihrer eigentlichen Form nicht tiefgehend thematisiert, stattdessen werden Computerprogramme behandelt (hier häufig kurz „Code" genannt).

2. Lernen

Der Machine Learning-Forscher *Goodfellow* definiert Lernen als das „Mittel zum Erlangen der Fähigkeit, die für das Durchführen der Aufgabe benötigt

46 *McCorduck*, Machines Who Think, S. 119.
47 https://www.duden.de/rechtschreibung/Algorithmus (Stand: 22.02.2021).

wird".[48] Lernen ist für ihn also ein Werkzeug – möglicherweise in Form eines Algorithmus – das eingesetzt werden kann, um eine Fähigkeit zu erwerben, die zur Zielerreichung erforderlich ist. Da es hier nur um das Lernen im Kontext von Machine Learning geht, werden Definitionen etwa aus dem Gebiet der Pädagogik hier nicht berücksichtigt.

3. Machine Learning

Machine Learning hingegen beschreibt *Goodfellow* als den Versuch, „Regeln zu finden, die vermutlich für die meisten Elemente der betrachteten Menge korrekt sind".[49] Hieraus lässt sich schon erkennen, dass Machine Learning vor allem Mustererkennung ist, präziser formuliert: Das Erlernen der Fähigkeit, Muster zu erkennen.

Der Herausforderung, dem Computer – bzw. dem Algorithmus – das Lernen beizubringen, näherte man sich von unterschiedlichen Seiten. In den letzten Jahren haben sich dabei einige Standards herausgebildet, die im Folgenden zunächst überblicksmäßig dargestellt und anschließend näher erklärt werden. Mit den dargestellten Ansätzen soll das System lernen, Antworten zu geben, oder konkreter: Muster zu erkennen, zum Beispiel, indem ein Objekt auf einem Bild erkannt („Was siehst du hier?", könnte die Frage lauten, auf die das Programm die Antwort geben soll) oder das Alter einer Person geschätzt wird (hier würde dann konkret gefragt: „Wie alt ist diese Person?"). Dabei wird grundsätzlich unterschieden zwischen Klassifikation und Regression, sowie überwachtem, unüberwachtem und verstärkendem Lernen.[50]

4. Modell

Für diese Arbeit ist der Begriff des *Modells* wesentlich, insbesondere erfolgt die Verwendung im Rahmen von „Machine Learning-Modellen" sehr häufig. Der Vielfalt an möglichen Bedeutungen dieses Begriffes halber erfolgt an dieser Stelle eine für diese Arbeit geltende Konkretisierung.

Die grundlegende Bedeutung des Begriffs ist nach *Brockhaus* zweigeteilt: Zum einen kann „Modell" einen Entwurf oder ein Muster meinen, zum

48 *Goodfellow* et al., Deep Learning Handbuch, S. 109.
49 Dies., Deep Learning Handbuch, S. 130.
50 Vgl. auch Dies., Deep Learning Handbuch, S. 108, 115.

anderen kann „Modell“ aber auch in der Bedeutung „Vorbild, Beispiel“ verwendet werden.[51] Die Wortherkunft liegt im italienischen *modello* bzw. im lateinischen *modulus* (Maß, Maßstab).[52]

Unterschieden wird also zwischen der Bedeutung, die als „Modell“ das Original versteht, also einen Gegenstand, der im Betrachtungsmittelpunkt steht, und der nachgebildet oder bearbeitet wird, und andererseits der Vorstufe eines zu schaffenden Ergebnisses. Beide Bedeutungen liegen nahe an dem, was der Modellbegriff im Machine Learning meint, aber treffen es nicht ganz. Hierzu ist eine Definition aus dem Machine Learning-Kontext hilfreicher: Im *Google Machine Learning Glossary* wird Modell definiert als eine „Repräsentation dessen, was ein Machine Learning-System aus Trainingsdaten gelernt hat.“[53]

Diese Repräsentation ist jedoch nicht als ein abstraktes, undurchschaubares Konstrukt zu verstehen, sondern vielmehr als ein konkretes Gebilde, das aus Zahlenwerten und Funktionsinformationen zusammengesetzt ist. Dabei gibt es Modellarten, die sich etabliert haben – wie etwa künstliche neuronale Netze oder Random Forests – die den Entwicklern zur Auswahl stehen, und die sie dann anpassen, trainieren und evaluieren. Das Endergebnis des Entwicklungsprozesses im Machine Learning, also das Gebilde, das eingesetzt wird, um Vorhersagen auf neuen Daten zu treffen, Daten zu sortieren oder Gemälde zu erzeugen, heißt (immer noch) Modell. Die Formen, die das ML-Modell annehmen kann, werden in § 6 D. beschrieben.

II. Untersuchte Technologien

Freilich wird hier nur auf einen kleinen Ausschnitt heute verfügbarer Technologien eingegangen, nämlich auf diejenigen, die im weiteren Verlauf der Arbeit zum Verständnis der Argumentation hilfreich sind: Entscheidungsbäume bzw. Random Forest-Modelle und künstliche neuronale Netze.[54]

---

51 Vgl. dazu den *Brockhaus*-Eintrag „Modell“ in https://brockhaus.de/ecs/enzy/article/modell-allgemein (Stand: 22.02.2021).

52 Vgl. *Brockhaus* „Modell“, https://brockhaus.de/ecs/enzy/article/modell-allgemein (Stand: 22.02.2021).

53 Vgl. *Google Machine Learning Glossary*, https://developers.google.com/machine-learning/glossary#model (Stand: 22.02.2021).

54 Die folgenden Ausführungen in Abschnitt § 2 B.II.1. und § 2 B.II.2. wurden teilweise bereits abgedruckt in *Käde/von Maltzan*, CR 2020, 66 ff..

1. Entscheidungsbäume bzw. Random Forest-Modelle

Schon der Name „Random Forest" verspricht zumindest irgendein Zufallselement bei der Ermittlung der Antwort auf die gestellte Frage. Diese Systeme sagen zum Beispiel die Wahrscheinlichkeit einer bestimmten Antwort auf eine datenbezogene Frage voraus, indem sie relevante Eigenschaften der Datensätze identifizieren und deren Beitrag zum Ergebnis berechnen. Random Forest-Modelle[55] basieren auf Entscheidungsbäumen.[56] In einer Baumstruktur mit je zwei Zweigen – Ja/Nein, Wahr/Falsch, 1/0 – durchläuft ein Datensatz mehrere Entscheidungsknoten von oben nach unten.[57]

Oft wird zur Veranschaulichung das Beispiel der Überlebenswahrscheinlichkeit von Passagieren auf der Titanic herangezogen. Je nach Geschlecht, Reiseklasse, Alter und Anzahl der Familienmitglieder könnte vorhergesagt werden, wie wahrscheinlich eine fiktive Person den Untergang der Titanic überlebt hätte. Ein Entscheidungsbaum-Machine Learning-Algorithmus[58] („Modell") erhält Datensätze mit Label – im Beispiel Passagiere mit Angabe über Überleben oder Sterben – und identifiziert aus dieser Vielzahl von Datensätzen anhand von Statistik die relevantesten Entscheidungskriterien (Features) und Schwellenwerte für die Entscheidungsknoten („das Modell wird trainiert"). Nach den Schwellenwerten entscheidet sich, ob der Datensatz zum rechten oder linken Kindknoten weiter wandert. Der Prozess wiederholt sich für jeden neuen Entscheidungsknoten. Der Baum baut sich also sukzessive in eine Richtung selbst auf. Die Features ergeben sich aus den zur Verfügung stehenden Daten (beispielsweise Alter, Geschlecht, Reiseklasse). Am Ende des Baumes – in Abbildung 2.1 als „A" und „B" gekennzeichnet – steht die Entscheidung, die dem Modell den Namen verleiht. Die „Entscheidung" entspricht den zuvor für die Trainingsdaten definierten Labels.

Bei dem Einsatz klassischer Entscheidungsbäume besteht jedoch die Gefahr der Überanpassung („Overfitting"[59]): Ein zu lange durchgeführtes oder nicht breit genug gestaltetes Training des Modells führt dazu, dass es für die gegebenen Datensätze optimale Ergebnisse liefert, aber für unbekannte Datensätze unbrauchbar ist – eventuell werden sogar irrelevante Merkmale

55 Zurückgehend auf *Breiman*, Machine Learning 45 Nr. 1 2001, 5.
56 *Sammut*/*Webb*, Encyclopedia of Machine Learning and Data Mining.
57 Vgl. zur Erklärung auch *Beierle*/*Kern-Isberner*, Wissensbasierte Systeme, S. 107.
58 Vgl. *Rokach*/*Maimon*, Data Mining with Decision Trees: Theory and Applications, S. 5.
59 *Goodfellow* et al., Deep Learning Handbuch, S. 123.

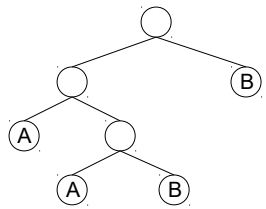

Abbildung 2.1: Entscheidungsbaum, Quelle: eigene Darstellung.

der Datensätze (z. B. die Haarfarbe der Passagiere) in die Überlebenschancen miteinbezogen.

Random Forest-Modelle sollen dem entgegenwirken, indem nicht nur ein einzelner Baum zum Einsatz kommt, sondern eine vom Entwickler bestimmbare Vielzahl (daher „Forest" – „Wald"),[60] vgl. Abbildung 2.2. Außerdem wird neben der „Bagging-Technik"[61] nur eine zufällig gewählte Teilmenge der Features in die Bäume einbezogen, wodurch eine Menge unterschiedlicher Bäume entsteht. Letztendlich durchläuft dadurch jeder Datensatz verschiedene Bäume und wird dementsprechend auf andere Features analysiert. Am Ende jedes Baumes weist das System jedem Datensatz ein Label zu (beispielsweise „überlebt" oder „nicht überlebt") – auch in Abbildung 2.2 als „A" (letzte linke Knoten) und „B" (letzte rechte Knoten) dargestellt. Sodann wird das Ergebnis jedes Baumes für einen Datensatz als Votum gewertet, und am Ende gegenübergestellt, welche Klasse für den Datensatz überwiegt, was letztendlich in der Klassifizierung resultiert.[62]

2. Künstliche neuronale Netze

Der Ursprung der „Faszination KI" lag (nachvollziehbarerweise) unter anderem in dem Wunsch, menschliche Intelligenz nachzubilden.[63] Da liegt es nahe, dass einige Ansätze der KI-Technologien sich zumindest teilweise an den grundlegenden Strukturen des menschlichen Gehirns orientieren.

60 *Breiman*, Machine Learning 45 Nr. 1 2001, 5, 6.

61 Bagging, kurz für „Bootstrap Aggregating", beschreibt eine Datensatzauswahlmethode, bei der aus der Gesamtmenge der Datensätze Teilmengen gebildet werden, die aber Duplikate enthalten dürfen, Ders., Machine Learning 45 Nr. 1 2001, 5, 5; *Goodfellow* et al., Deep Learning Handbuch, S. 285.

62 *Breiman*, Machine Learning 45 Nr. 1 2001, S. 7.

63 Vgl. § 2 A..

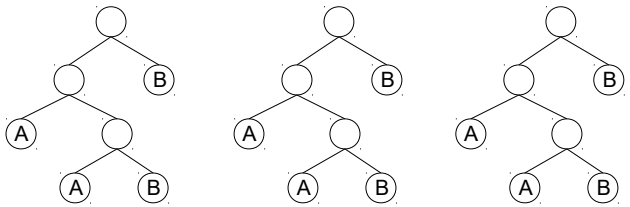

Abbildung 2.2: Random Forest, Quelle: eigene Darstellung.

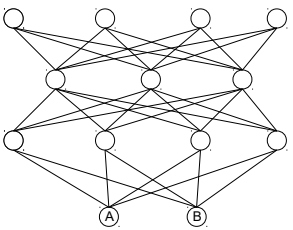

Abbildung 2.3: Künstliches neuronales Netz mit vier Schichten, Quelle: eigene Darstellung.

Wir meinen heute zu wissen, dass das Gehirn im wesentlichen – einfach gesagt – ein Netzwerk spezialisierter Neuronen (Nervenzellen) ist, die durch Synapsen miteinander verbunden sind.[64] Dabei kann jede Nervenzelle durch mehrere Synapsen mit anderen Neuronen kommunizieren.[65] Neuronen werden durch Energie aktiviert, und der aktivierende elektrische Impuls wird mittels chemischer Überträgersubstanzen weitergeleitet.[66] Verbindungen, die viel genutzt werden, werden verstärkt.[67]

Sogenannte KNNs, insbesondere in der Variante tiefer Neuronaler Netze (Deep Learning) sind zu unterschiedlichen Graden von der (vermuteten)

64 *Liebmann*/*Gertz*, Basiswissen Neuroanatomie, S. 1.
65 Dies., Basiswissen Neuroanatomie, S. 3; *Huggenberger* et al., Neuroanatomie des Menschen, S. 6.
66 *Liebmann*/*Gertz*, Basiswissen Neuroanatomie, S. 1.
67 *Rahmann*/*Rahmann*, Das Gedächtnis: Neurobiologische Grundlagen, S. 224.

Struktur des menschlichen Gehirns inspiriert.[68] Mit ihnen wird nicht etwa – wie der Name vermuten lässt – direkt versucht, Gehirnfunktionen nachzubauen, sondern vielmehr werden diese Modelle eingesetzt um Annäherungen an statistische Funktionen zu erreichen und dadurch statistische Generalisierungen zu erzielen[69] (auch dies ist für den mathematischen Laien als Mustererkennung vorstellbar).

In einer visuellen Darstellung ähneln die KNNs auf den ersten Blick einer Baumstruktur.[70] Der auffälligste Unterschied ist, dass ein Knoten (Neuron) mehrere Vorgänger haben kann. Größere Unterschiede sind erst auf den nächsten Blick – im Folgenden anhand des Titanic-Beispiels – zu erkennen.

Ein einfaches KNN untersucht einen Datensatz in der ersten Schicht („Input Layer") nach verschiedenen Kriterien (im hiesigen Beispiel etwa die Eigenschaften der Titanic-Passagiere). Die erste Schicht hat also so viele Input-Neuronen wie Eigenschaften untersucht werden sollen.[71] Die letzte Schicht („Output Layer") liefert das Ergebnis (Überlebenswahrscheinlichkeit – in Abbildung 2.3 markiert als „A" und „B"). Die hauptsächliche Arbeit wird im Rahmen des Deep Learnings in den dazwischenliegenden „Hidden Layers"[72] verrichtet. In diesen werden – je nach Art des Netzwerkes mithilfe unterschiedlichster mathematischer Funktionen – Feinjustierungen vorgenommen oder verschiedene Aspekte des Inputs betrachtet. In unserem Beispiel könnten etwa die Neuronen der Hidden Layers analysieren, ob bestimmte Kombinationen von Eigenschaften vorliegen (zum Beispiel Alter > 12 und oberste Reiseklasse in einem Neuron, Alter < 12 und Anzahl Familienmitglieder < 4 in einem weiteren, etc.) und dementsprechend

---

68 *Kaplan*, Artificial intelligence: What everyone needs to know, S. 28; *Goodfellow et al.* weisen darauf hin, dass nicht alle Deep Learning-Forscher sich von den Neurowissenschaften inspirieren lassen, für den eigentlichen Nachbau von Gehirnfunktionen wird auf das Gebiet der Computational Neuroscience verwiesen, *Goodfellow* et al., Deep Learning Handbuch, S. 18, von dieser Inspiration rührt auch die Bezeichnung als künstliches „neuronales" Netz her, Dies., Deep Learning Handbuch, S. 186.

69 Dies., Deep Learning Handbuch, S. 186.

70 Hier zur Verdeutlichung der Ähnlichkeit ebenfalls vertikal dargestellt – in der Regel erfolgt die Abbildung künstlicher neuronaler Netze in der Horizontalen.

71 *Nielsen*, Neural Networks and Deep Learning, Kap. 1, The Architecture of Neural Networks.

72 Der Name suggeriert, dass hier etwas im Versteckten passiert. Dies ist irreführend, denn „Hidden" weist lediglich darauf hin, dass keine Schnittstelle – z. B. zum Nutzer – besteht, bzw. dass es sich weder um die Input- noch um die Output-Schicht handelt, vgl. auch Ders., Neural Networks and Deep Learning, Kap. 1, The Architecture of Neural Networks.

Wahrscheinlichkeitswerte anpassen, welche dann wiederum gewichtet weitergeleitet werden (hier wieder die Parallele zu den energieweiterleitenden Synapsen im Gehirn). Bei künstlichen neuronalen Netzen wird die Struktur vor dem Training festgelegt, das Netz „wächst" also im Unterschied zu Entscheidungsbäumen nicht. Variabel sind nur die sogenannten Gewichte („Weights") und Verzerrungsparameter[73] („Biases"). Diese Werte werden im Rahmen des Trainingsvorgangs verändert und optimiert.

Dies ist eine stark vereinfachte Darstellung der möglichen Funktionsweise solcher Modelle. Für das Verständnis der Arbeit ist eine tiefergehende Einsicht in die verwendeten statistischen Funktionen etc. nicht erforderlich und wird dem Leser an dieser Stelle erspart.[74] Wichtig festzustellen ist jedoch, dass die Art und die „Architektur" des Modells einen wesentlichen Einfluss auf die Einsatztauglichkeit und Genauigkeit des Modells haben. Der Entwickler steht also stets vor der schwierigen Aufgabe, die richtige Anzahl Schichten, Neuronen je Schicht, Rückkopplungen, Aktivierungsfunktionen und dergleichen auszuwählen.[75]

3. Sonderfall: Generative Adversarial Networks

Während in den bisherigen Ausführungen Modelle im Vordergrund standen, die Daten analysieren und Werte ausgeben, erfolgt nun die Einführung einer sich davon doch wesentlich unterscheidenden Art: Sogenannte generative Modelle, bzw. insbesondere generative adversariale Netzwerke (GANs).[76] Diese Sonderform in Gestalt einer Kombination von KNNs zielt nicht darauf ab, einen Wert zu schätzen, sondern vielmehr darauf, Daten – in Form von Bildern, Texten, Musik etc. – zu erzeugen. So komplex die Ergebnisse erscheinen mögen, so simpel ist die Idee dahinter – hier erklärt anhand der Erzeugung von Bildern, auf denen Menschen zu erkennen sind:

---

73 *Goodfellow* et al., Deep Learning Handbuch, S. 121: „Diese Bezeichnung leitet sich aus der Sichtweise ab, dass ohne jegliche Eingabe die Ausgabe der Transformation zu [dem Achsenschnittpunkt] *b* verschoben ist. Der Begriff darf nicht mit einer statistischen Verzerrung [..] verwechselt werden.".

74 Stattdessen wird verwiesen auf Dies., Deep Learning Handbuch, *Kaplan*, Artificial intelligence: What everyone needs to know und *Nielsen*, Neural Networks and Deep Learning.

75 Ders., Neural Networks and Deep Learning, Kap. 1, The architecture of neural networks.

76 Eingeführt 2014 von *Goodfellow* et al., Generative Adversarial Nets.

Zwei künstliche neuronale Netze werden miteinander verbunden – ein generativer und ein adversarialer (gegnerischer) Part. Der adversariale Part ist ein klassifizierendes Modell, das schon vortrainiert wurde mit Bildern von Menschen, und das der im vorigen Abschnitt beschriebenen Funktionsweise entspricht. Der generierende Part wurde nicht vortrainiert, sondern bekommt vom Entwickler lediglich Vorgaben wie zum Beispiel die Größe des zu erzeugenden Bildes. Sodann erzeugt das generierende Modell zunächst quasi zufälliges „Rauschen", setzt also beliebige Pixel auf beliebige Werte. Das Ergebnis wird dann zum Input des gegnerischen Parts. Dieser hat nun die alleinige Aufgabe, zu unterscheiden, ob der Input aus den ihm bekannten, „echten" Bildern stammt oder durch das generierende Modell erzeugt wurde – muss also unterscheiden zwischen „Original" und „Fake".[77] Diese zwei Kategorien sind die Klassen, die dem gegnerischen Part vorgegeben sind und stellen die Trainingsaufgabe für den gegnerischen Part dar. Dem generierenden Part wird zurückgemeldet, ob der gegnerische Part das Fake als ein solches erkannt hat (bzw. mit welcher Wahrscheinlichkeit es für ein solches gehalten wurde), und startet dann den nächsten Versuch, indem einige Pixelwerte angepasst werden, bis die Wahrscheinlichkeit sinkt, dass es sich um ein Fake handelt. Das generierende Modell versucht also, das gegnerische Modell „auszutricksen", indem es seine Erzeugnisse so optimiert, dass das gegnerische Modell Fake und Original nicht mehr unterscheiden kann.[78] Gleichzeitig versucht das gegnerische Modell, seine Fähigkeit, zwischen Original und Fake zu differenzieren, zu optimieren. Auf diese Weise entstehen eindrucksvolle Ergebnisse: Einige Beispiele, die auf diese Weise entstanden sind, sind das 2018 versteigerte „Gemälde" des französischen Künstlerkollektivs *Obvious* „Edmond de Belamy"[79] sowie das Projekt „Next Rembrandt",[80] in dessen Rahmen der adversariale Part mit Rembrandt-Werken trainiert wurde und der generative Part die Aufgabe bekam, Werke zu generieren, die der adversariale Part für „einen echten Rembrandt" halten würde.

Da hier zumindest der Art der Ergebnisse der GANs nach potenziell „Werke" entstehen können, lohnt es sich, generative Modelle und insbesondere den Entstehungsprozess der Erzeugnisse genauer zu betrachten. Damit wird sich der vierte Teil dieser Arbeit befassen, weitere Beispiele werden in § 10 C. untersucht.

---

77 *Goodfellow* et al., Generative Adversarial Nets, S. 1.
78 Dies., Generative Adversarial Nets, S. 2.
79 Vgl. https://obvious-art.com/portfolio/edmond-de-belamy/ (Stand: 22.02.2021); vgl. zur Erklärung außerdem § 10 C.III..
80 Vgl. https://nextrembrandt.com (Stand: 22.02.2021).

III. Einige grundlegende Begriffe

Es folgt eine Erläuterung einiger grundlegender Begriffe, die teilweise schon zur Erklärung des Machine Learnings herangezogen werden mussten. Dort waren sie für das Verständnis der wesentlichen Konzepte nicht zwingend erforderlich, andererseits ist für das Verständnis der Begriffe eine Vorstellung von Machine Learning förderlich, weshalb diese Reihenfolge gewählt wurde. Es handelt sich um eine Auswahl von Begriffen aus dem Kontext des maschinellen Lernens, die im Rahmen dieser Arbeit immer wieder zur Verwendung kommen. Auf umfangreiche Ausführungen wird allerdings verzichtet, es erfolgt eine Beschränkung auf die für das Kontextverständnis hinreichenden Informationen.

1. Klassifikation, Klassifizierung

Die Begriffe Klassifikation und Klassifizierung werden in der deutschsprachigen Machine Learning-Literatur weitgehend synonym verwendet. Beides bezeichnet laut Duden einen Vorgang der Einordnung in eine Klasse.[81] Im Kontext des Machine Learnings sind klassifizierende Modelle einfach gesprochen Systeme, die einem gegebenen Eingangswert einen Ausgangswert aus einem vorgegebenen Katalog zuweisen.[82] Zur Verdeutlichung ein Beispiel: Dem System wird ein Bild zugeführt, auf dem ein Haus und ein Baum abgebildet sind. Wenn das System gelernt hat, (unter anderem) Häuser und Bäume zu erkennen, wird es die Klassen „Haus“ und „Baum“ ausgeben.

81 Vgl. Duden „Klassifikation“, https://www.duden.de/rechtschreibung/Klassifikation (Stand: 22.02.2021) sowie Duden „Klassifizierung“, der Eintrag verweist für die Wortbedeutung auf die Klassifikation und lässt vermuten, dass beide Worte synonym verstanden werden können, https://www.duden.de/rechtschreibung/Klassifizierung (Stand: 22.02.2021), wenngleich z. B. Wikipedia darauf hinweist, dass Klassifikation eigentlich als Ergebnis der Klassifizierung zu verstehen sei, https://de.wikipedia.org/wiki/Klassifizierung#Begriffsabgrenzung (Stand: 22.02.2021).

82 *Goodfellow* et al., Deep Learning Handbuch, S. 109 mit weiteren Beispielen.

2. Regression

Regression (bzw. eigentlich präziser *Regressionsanalyse*) hingegen beschreibt einen Vorgang der Annäherung an einen Wert.[83] Das System versucht, eine Frage, die die Angabe eines Wertes erwartet, durch Annäherung an im Training Gelerntes zu beantworten.[84] Beispiel: Dem System wird ein Bild zugeführt, auf dem ein Mensch abgebildet ist. Wenn das System gelernt hat, das Aussehen von Menschen und Alterswerte in Relation zu setzen, wird es ein geschätztes Alter ausgeben. Die Regressionsanalyse ist ein statistisches Verfahren, auf dessen tiefergehende, aber hier für den weiteren Verlauf unnötige Erklärung an dieser Stelle verzichtet wird.

Goodfellow et al. erläutern neben Klassifizierung und Regression noch einige weitere klassische ML-Aufgaben – darunter auch die Umwandlung unstrukturierter Daten in Text (Bsp: Text-to-Speech-Umwandlung), maschinelle Übersetzungen, Anomalieerkennungen, Sprachsynthese (erlaubt es etwa digitalen Assistenten zu „sprechen", ohne dass vorher ganze Wörter oder Phrasen vom Menschen eingesprochen wurden) u. v. m..[85]

3. Überwachtes Lernen (Supervised Learning)

Überwachtes Lernen wird die Lernmethode genannt, wenn die Daten, mit denen das System lernt, bereits mit Labels (Klassen) versehen wurden. Auf das obige Beispiel der Klassifikation übertragen: Wenn das System die Fähigkeit, das Haus und den Baum zu erkennen, dadurch gelernt hat, dass ihm zuvor eindeutig als Haus und Baum bezeichnete Bilder im Rahmen des Lernprozesses „vorgeführt" wurden, wird von überwachtem Lernen gesprochen. Analoges gilt für ein System, das mit Regression arbeitet.[86]

---

83 Hinweis: hier ist der statistische Wortgebrauch relevant, es handelt sich nicht um Regression im wirtschaftlichen Sinne, vgl. auch Duden „Regression", https://www.duden.de/rechtschreibung/Regression (Stand: 22.02.2021).

84 *Goodfellow* et al., Deep Learning Handbuch, S. 109.

85 Dies., Deep Learning Handbuch, S. 109 ff..

86 Vgl. auch *Ertel*, Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung, S. 313; *Goodfellow* et al., Deep Learning Handbuch, S. 115.

4. Unüberwachtes Lernen (Unsupervised Learning)

Unüberwachtes Lernen hingegen überlässt dem Algorithmus das Auffinden und Zuweisen von Gemeinsamkeiten, bzw. Klassen. Die Aufgabe könnte heißen „Finde Gemeinsamkeiten auf diesen Bildern!“.[87]

5. Lernen durch Verstärkung (Reinforcement Learning)

Reinforcement Learning beschreibt einen Prozess, der ohne Trainingsdaten auskommt. Hierbei könnte sich das System zum Beispiel durch Ausprobieren der Lösung nähern, wobei zuvor programmatisch „Belohnungen“ und „Bestrafungen“ für gute respektive schlechte oder aufwendige Lösungen vorgesehen sind. Das System versucht, während es den Weg zur Lösung sucht, die Belohnungen zu maximieren.[88] Der Input kann sich etwa aus der Aufzeichnung von Umgebungsdaten durch Sensoren ergeben.[89]

6. Parameter, Gewichte, Biases, Hyperparameter, Topologie, Schichten, Netzstruktur, Architektur

Die in der Überschrift genannten Begriffe lassen sich am besten gemeinsam erklären. Sie sind grundsätzlich zu teilen in zwei Gruppen:

– Parameter, Gewichte, Biases
– Hyperparameter, Topologie, Schichten, Netzstruktur, Architektur.

Die erste Gruppe beschreibt den Teil der ML-Modelle, der im Rahmen des Trainings automatisch angepasst und optimiert wird. Gewichte, bzw. engl. *Weights*, und *Biases* sind Begriffe, die insbesondere im Kontext von KNNs relevant sind. Dort beschreiben sie die veränderlichen Werte, die – einfach gesagt – den Anteil eines Inputs am Ziel-Output ausmachen. Für eine ausführlichere Beschreibung vgl. § 2 B.II.2.. Der allgemeinere (und in dieser

87 Ausführlich zu Unsupervised Learning: *Hinton*/*Sejnowski*, Unsupervised Learning; *Goodfellow* et al., Deep Learning Handbuch, S. 115.
88 Vgl. auch *Ertel*, Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung, S. 313.
89 Vgl. auch *Goodfellow* et al., Deep Learning Handbuch, S. 116 f. mit weiteren Nachweisen zu Reinforcement Learning.

Arbeit überwiegend verwendete) Begriff ist der der Parameter. Dieser ist auch für andere ML-Modelle zutreffend und erfasst etwa für *Random Forest*-Modelle, die mit *Python* erzeugt werden, die variablen Split-Schwellenwerte.

Die Begriffe der zweiten Gruppe haben gemeinsam, dass sie im Rahmen des Trainings nicht automatisiert angepasst und zudem nicht innerhalb des eigentlichen Algorithmus eingestellt werden.[90] Topologie, Netzstruktur und Architektur können synonym verwendet werden und beschreiben unter anderem, wie viele Schichten ein Modell hat und wie diese aufgebaut sind. Schichten (engl. Layer) sind – im Falle eines KNN – Neuronen, die parallel abgearbeitet werden (vgl. dazu Abbildung 2.3: das abgebildete KNN besteht aus 13 Neuronen, die auf 4 Schichten verteilt sind).

Zu den Hyperparametern gehören alle Angaben und Einstellungen, die für ein Modell vorgenommen werden – für ein KNN etwa die Anzahl der Schichten und der Neuronen je Schicht, für einen Random Forest u. a. die maximale Tiefe des Baumes. Veränderungen an den Hyperparametern sind stets Gestaltungsentscheidungen des Entwicklers, um die Modellperformanz zu optimieren. Weil sie durch die Einstellungen das Modell, würde man es visuell abbilden, im Aufbau verändern, werden die Hyperparameter auch als Netzstruktur (speziell für KNN), Architektur des Modells oder Topologie bezeichnet, obwohl diese Begriffe darüber hinwegtäuschen können, dass eine Sammlung von Hyperparametern mitunter auch zum Beispiel die von den Neuronen zu verwendenden Funktionen enthält. Diese Arbeit verwendet den Begriff der Hyperparameter. Falls einmal die Begriffe Struktur, Architektur oder Topologie genannt werden, so sind damit immer die Hyperparameter gemeint.

7. Trainingsprozess, Trainingsdaten, Trainingsergebnisse

Auch wenn der Trainingsprozess sich im Detail für verschiedene Modelle und Aufgaben unterscheiden kann,[91] erfolgt hier eine verallgemeinerte Beschreibung für klassifizierende künstliche neuronale Netze, um einen Überblick zu verschaffen. Damit die beschriebenen Modelle tatsächlich Ergebnisse liefern, werden diese mit Trainingsdaten auf ihre zugrundeliegende Aufgabe trainiert. *Trainingsdaten* sind zu analysierende Daten (Bilder, Texte, Zahlen, …), die – für den überwachten Lernprozess – bereits den korrekten Kategorien zugeord-

90 *Goodfellow* et al., Deep Learning Handbuch, S. 107.
91 Dies., Deep Learning Handbuch, S. 114.

net sind („gelabelt sind“).[92] Die Zuordnung kann beispielsweise geschehen, indem die die Daten enthaltenden Dateien entsprechend benannt werden oder indem die Dateien in den Labels entsprechenden Ordnern abgelegt werden.[93] Üblicherweise separiert der Entwickler von den Trainingsdaten noch einige *Test-* bzw. *Validierungsdaten*, um das fertig trainierte Modell mit Daten testen zu können, die demselben nicht bekannt sind.

Für das Training wird im überwachten Lernen versucht, einen Zusammenhang zwischen Input (Trainingsdaten) und Output (Label) zu finden.[94] Für künstliche neuronale Netze ist in der Regel vorgegeben, welche Eigenschaften der Trainingsdaten analysiert werden sollen. Dabei durchläuft jeder Trainingsdatenpunkt das Modell einzeln. In jedem Neuron wird ein unterschiedlicher Aspekt des Inputs untersucht (bei Bildern z. B. die einzelnen Pixel – die erste Schicht hätte dann so viele Inputneuronen wie das Bild Pixel hat, evtl. noch erweitert um die Farbwerte[95]). Um zu messen, wie gut das Modell schon gelernt hat, wird eine Zielfunktion definiert.[96] Wenn es sich dabei um eine Fehlerquotenfunktion handelt, gilt: je niedriger der Fehler (also die Abweichung des Ergebnisses von dem Ergebnis, das durch das Label vorgegeben ist[97]), desto besser findet die Formel (das Modell) zu dem Ergebnis, das dem Label entspricht.[98] Wenn der Fehler nahe Null ist, findet das Modell für alle Trainingsdaten den richtigen Output, und hat dementsprechend gute Gewichte und Biases gewählt.

Gut ist das Modell, wenn es auch für Daten, die nicht in den Trainingsdaten enthalten waren – sogenannte Testdaten – mit möglichst geringer Fehlerquote die richtigen Ausgaben erzeugt.[99] Dazu wird eine (von der Zielfunktion verschiedene) Funktion konzipiert, die die Gewichte der Neuronen so optimiert,

92 *Goodfellow* et al., Deep Learning Handbuch, S. 115; im Vorgang des unüberwachten Lernens sind keine Labels erforderlich, da aus den Daten Zusammenhänge erkannt werden sollen, die auch dem Datensammler noch unbekannt sind.

93 Vgl. z. B. die Ausführungen in https://www.tensorflow.org/tutorials/load_data/images (Stand: 22.02.2021).

94 Vgl. *Goodfellow* et al., Deep Learning Handbuch, S. 154 f..

95 Dies., Deep Learning Handbuch, S. 109.

96 Dies., Deep Learning Handbuch, S. 113 f..

97 Der geneigte Schuldrechtler denkt hier „Abweichung der Ist-Beschaffenheit von der Soll-Beschaffenheit“…

98 Es gibt verschiedene Möglichkeiten, die Performanz eines Modells zu messen, die auch von der Art des Modells abhängt – die Kostenfunktion für Klassifizierungsaufgaben (auch Mean Squared Error, Fehlerquote genannt) ist eine davon, vgl. auch *Nielsen*, Neural Networks and Deep Learning, Kap. 1, Learning with gradient descent und *Goodfellow* et al., Deep Learning Handbuch, S. 114, 119.

99 Dies., Deep Learning Handbuch, S. 114.

dass die Fehlerquote sinkt,[100] ohne dabei aber die Trainingsdaten auswendig zu lernen: Denn das Modell soll ja auch unbekannte Daten erfolgreich verarbeiten. Um das sicherzustellen, werden die Testdaten eingesetzt.

Trainingsergebnisse sind die errechneten bzw. optimierten Parameter. Die Unterscheidung zwischen Trainingsdaten und Trainingsergebnissen ist erkennbar wichtig, wird jedoch nicht an allen Stellen konsequent durchgehalten.[101] Diese Arbeit verwendet die Begriffe so, wie sie hier definiert wurden.

IV. Technische Einordnung

Häufig besteht Unklarheit darüber, was „KI" bzw. ein „ML-Modell" eigentlich ist, insbesondere welcher technische „Charakter" hier zugrunde liegt. Handelt es sich um eine Datei, um ein Programm, oder um etwas ganz anderes? Hierauf wird in § 6 vertieft eingegangen, weil es dort besonders relevant wird. An dieser Stelle sei nur so viel gesagt: Ein ML-Modell funktioniert nicht ohne Computerprogrammcode, das Herzstück (oder „Hirnstück"?) besteht jedoch in der Regel in einer Vielzahl von Werten, die in Dateiform – losgelöst vom Computerprogramm – vorliegen können. Und „das Herzstück" ist letztendlich, Werte eingesetzt und ausgeschrieben, eine enorm lange mathematische Formel mit sehr vielen Parametern. Ein ML-System lässt sich zerlegen in vier Kernkomponenten: einen Datensatz, der analysiert werden soll, ein Modell, das der Datensatz durchläuft, eine mathematische Kostenfunktion, die die Performanz misst, und einen Optimierungsalgorithmus, der das Modell verändert.[102] Diese Komponenten sind in der Implementierung (also jenseits der statistisch-mathematischen Theorie) verbunden durch Computerprogrammcode. Dabei ist es irrelevant, ob es sich um ein Random Forest-Modell oder ein künstliches neuronales Netz handelt.

100 *Goodfellow* et al., Deep Learning Handbuch, S. 119.

101 Vgl. z. B. *Kaulartz/Braegelmann*, Rechtshandbuch Artificial Intelligence, Kap. 7.1 Rn. 49 ff., dort ist die Rede von „Trainingsdaten", gemeint sind jedoch vermutlich „Trainingsergebnisse".

102 *Goodfellow* et al., Deep Learning Handbuch, S. 108.

1. Verhältnis des Machine Learning zu Text und Data Mining

Der Begriff des Text und Data Mining (TDM) taucht bereits heute sowohl in EU-Richtlinien (vgl. z. B. Art. 3 DSM-RL) als auch im deutschen Urheberrechtsgesetz auf (vgl. § 60d UrhG), und fällt sogleich auch immer wieder im Kontext des maschinellen Lernens. An dieser Stelle werden die beiden Begriffe für ein besseres Verständnis ins Verhältnis gesetzt.

Aus der BT-Drs. 18/12329 ergibt sich, dass der deutsche Gesetzgeber unter Text und Data Mining die Auswertung einer „Vielzahl von Texten, Daten, Bildern und sonstigen Materialien" versteht, „um so neue Erkenntnisse zu gewinnen."[103]

Grundsätzlich ist zu unterscheiden zwischen Text Mining und Data Mining: Ersteres analysiert unstrukturierte Textdaten, letzteres befasst sich mit strukturierten Daten, z. B. aus Datenbanken.[104] Dabei wird Machine Learning insbesondere für die Zwecke des Data Minings eingesetzt[105] Machine Learning fokussiert auf die Vorhersage bzw. Wahrscheinlichkeitsschätzung unbekannter Werte, während TDM sich in erster Linie mit bekannten Daten befasst und diese auf Muster, Strukturen und Zusammenhänge analysiert. Teilweise werden TDM und ML synonym verwendet,[106] oder nur das unüberwachte Lernen dem Data Mining gleichgesetzt[107]. Sinnvoll erscheint es, zumindest den Lernprozess des Machine Learnings, der mit bestehenden, sicheren Daten arbeitet, als (Text und) Data Mining-Verfahren zu betrachten. Dafür spricht, dass die Aufgabe – Muster in den Daten zu erkennen – die gleiche ist, auch wenn Machine Learning später noch einen Schritt weiter geht und daraus Vorhersagen für unbekannte Daten ableiten will oder gar ganz neue Daten erzeugt. Dies erleichtert zumindest auch die urheberrechtliche Bewertung des Lernprozesses, und ermöglicht es, die für das TDM bestehenden (und entstehenden) Regelungen auch auf den Lernprozess des ML anzuwenden. Falsch wäre es jedoch, TDM und ML grundsätzlich immer synonym zu verwenden, dafür sind die Begriffe nicht trennscharf genug definiert.

103 BT-Drs. 18/12329 S. 22.
104 *Klass*, Wirtschaftsinformatik & Management 11 Nr. 4 2019, S. 267.
105 Ders., Wirtschaftsinformatik & Management 11 Nr. 4 2019, S. 267.
106 Vgl. z. B. *Faul*, A Concise Introduction to Machine Learning., S. XVII.
107 *Bell*, Machine Learning, S. 4.

2. Verhältnis des Machine Learning zu Expertensystemen

Ein naheliegender Versuch, den menschlichen Verstand nachzubilden, liegt darin, einer Maschine möglichst viel Wissen zu vermitteln, und die Maschine aus der ihr zur Verfügung stehenden Wissensbasis Antworten geben zu lassen. Solche „Expertensysteme" bzw. auch allgemeiner „wissensbasierte Systeme"[108] genannten Ansätze werden seit 1974 entwickelt.[109] *Beierle* fasst bestehende Definitionen von Expertensystemen wie folgt zusammen:

> „Ein Expertensystem ist ein Computersystem (Hardware und Software), das in einem gegebenen Spezialisierungsbereich menschliche Experten in Bezug auf ihr Wissen und ihre Schlussfolgerungsfähigkeit nachbildet."[110]

Zusammengesetzt sind solche Systeme grundsätzlich aus einer Wissensbasis und einer wissensverarbeitenden Komponente,[111] hinzu kommen eine Wissenserwerbs- und eine Erklärungskomponente sowie eine Dialogkomponente.[112] Sowohl Verfahren maschineller Lernsysteme als auch allgemein Data Mining können grundsätzlich dazu eingesetzt werden, Expertensysteme zu implementieren,[113] so ist es etwa möglich, die das Expertenwissen repräsentierenden Regeln deduktiv aus Daten zu lernen. Es ist jedoch auch hier zu vermeiden, die Begriffe des maschinellen Lernens und der Expertensysteme, oder auch Expertensysteme und KI, synonym zu verwenden.

V. Zusammenfassung

Festzustellen ist, dass Modelle maschinellen Lernens keine gänzlich undurchsichtige mystischen Gebilde sind, sondern vielmehr Software bzw. Algorithmen, die auf mathematischen bzw. statistischen Funktionen basieren und eine große Menge an Daten benötigen, um im Einsatz Ergebnisse in Form von Wahrscheinlichkeitswerten liefern zu können.

---

108 *Puppe*, XPS-99, Vorwort; zur Unterscheidung vgl. *Beierle/Kern-Isberner*, Wissensbasierte Systeme, S. 11.
109 *McCorduck*, Machines Who Think, S. XXVIII identifiziert das Programm MYCIN von *T. Shortliffe* als erstes „Expertensystem".
110 *Beierle/Kern-Isberner*, Wissensbasierte Systeme, S. 12.
111 Dies., Wissensbasierte Systeme, S. 17.
112 Dies., Wissensbasierte Systeme, S. 18.
113 Vgl. für einige Beispiele Dies., Wissensbasierte Systeme, S. 99 ff.; *Looney*, Expert Systems With Applications 6 1993, 129 ff..