

Maschinelles Lernen in Risikomodellen – Charakteristika und aufsichtliche Schwerpunkte

Konsultationspapier

Inhaltsverzeichnis

I. Maschinelles Lernen auf dem Vormarsch	3
II. Charakteristika von ML	5
1. Dimensionen und Beispiele	5
2. Methodik	7
2.1 Komplexität und Dimension des Hypothesenraums	7
2.2 Komplexität des Trainings	8
2.3 Adaptivität	8
3. Datengrundlage	9
4. Verwendung des Outputs	9
4.1 Bedeutung im Modell	9
4.2 Anwendungsbereich	9
4.3 Grad der Automatisierung	10
5. Auslagerung und IT-Infrastruktur	10
III. Aufsichtlicher Ansatz	11
1. Aufsichtspraxis hat Bestand	11
2. Methoden laden zur Datengläubigkeit ein	13
3. Erklärbarkeit rückt in den Fokus	14
4. Adaptivität: Modelländerungen sind schwerer zu erkennen	15
IV. Ausblick	18
V. Konsultation	18

I. Maschinelles Lernen auf dem Vormarsch

Die Diskussion um die Anwendung von künstlicher Intelligenz und maschinellem Lernen (kurz: ML-Methoden) gewinnt in letzter Zeit gerade im Bereich der Finanzdienstleistungen an Fahrt. Die Verfügbarkeit großer Datenmengen („Big Data“) zusammen mit gesteigerter Rechenleistung bilden die Grundlage dafür. Über standardisierte Prinzipien und Abläufe¹ wird die Entwicklung von ML-Methoden ebenso einfach wie zugänglich.

Der Einsatz von ML-Methoden kann dazu beitragen, Risiken genauer zu quantifizieren und die Qualität von Prozessen zu erhöhen und somit das Risikomanagement für Finanzunternehmen zu stärken.

Sowohl BaFin als auch Bundesbank haben sich bereits in mehreren Publikationen dazu geäußert: Am 16. Juli 2018 hat die BaFin ihren Bericht „Big Data trifft auf künstliche Intelligenz – Herausforderungen und Implikationen für Aufsicht und Regulierung von Finanzdienstleistungen“² (kurz: BDAI-Bericht) öffentlich zur Konsultation gestellt und am 28.02.2019 einen Überblick über die Ergebnisse zusammen mit einer ersten Bewertung veröffentlicht.³ Der BDAI-Bericht hat das Feld der Digitalisierung abgesteckt und Herausforderungen identifiziert. Im März 2020 hat die BaFin erläutert, warum sie einer generellen Genehmigungspflicht von Algorithmen kritisch gegenübersteht und wie sich ML in den risikoorientierten Aufsichtsansatz einfügt.⁴ Diese Überlegungen werden in den im Juni 2021 veröffentlichten „Prinzipien für den Einsatz von Algorithmen in Entscheidungsprozessen“ weiter konkretisiert.⁵ Die Bundesbank hat im November 2020 in ihrem Diskussionspapier „The Use of Artificial Intelligence and Machine Learning in the Financial Sector“⁶ grundsätzliche Überlegungen zum Umgang mit ML-Methoden im Kontext der prudentiellen Aufsicht veröffentlicht; hierin u. a. Aussagen zur Aufsichtsintensität von ML hergeleitet und die Bedeutung der Erklärbarkeit von ML-Methoden eingeordnet.

Auch international beschäftigen sich Aufseher und Regulatoren mit ML-Methoden. Zu nennen sind die Publikationen der De Nederlandsche Bank, die Prinzipien für den Einsatz von ML entwickelt haben⁷ sowie der ACPR, die die Erklärbarkeit von ML in den Mittelpunkt

¹ Dazu zählen DevOps, MLOps sowie Software-Bibliotheken, die gängige ML-Methoden beinhalten.

² BaFin, 2018, „Big Data trifft auf künstliche Intelligenz“, online abrufbar:

https://www.bafin.de/SharedDocs/Downloads/DE/dl_bdai_studie.pdf?__blob=publicationFile&v=3

³ BaFin, 2019, „Big Data trifft auf künstliche Intelligenz – Ergebnisse der Konsultation zum BaFin-Bericht“, online abrufbar:

https://www.bafin.de/SharedDocs/Veroeffentlichungen/DE/BaFinPerspektiven/2019_01/bp_19-1_Beitrags_SR3.html

⁴ BaFin, 2020, „Generelle Billigung von Algorithmen durch die Aufsicht? Nein, aber es gibt Ausnahmen“, online abrufbar:

https://www.bafin.de/SharedDocs/Veroeffentlichungen/DE/Fachartikel/2020/fa_bj_2003_Algorithmen.html;jsessionid=A8AACD58B7FCC16B44597DC152D1F7D4.1_cid503

⁵ BaFin, 2021, „Big Data und künstliche Intelligenz: Prinzipien für den Einsatz von Algorithmen in Entscheidungsprozessen“, online abrufbar:

https://www.bafin.de/SharedDocs/Downloads/DE/Aufsichtsrecht/dl_Prinzipienpapier_BDAI.html

⁶ Bundesbank, 2020, „Policy Discussion Paper, The Use of Artificial Intelligence and Machine Learning in the Financial Sector“, online abrufbar:

<https://www.bundesbank.de/resource/blob/598256/d7d26167bceb18ee7c0c296902e42162/mL/2020-11-policy-dp-aiml-data.pdf>

⁷ DeNederlandscheBank, 2019, „General principles for the use of Artificial Intelligence in the financial sector“, online abrufbar: <https://www.dnb.nl/media/jkbp2jc/general-principles-for-the-use-of-artificial-intelligence-in-the-financial-sector.pdf>

gestellt haben.⁸ Auch die EBA⁹ und die Bank of England¹⁰ haben ihre Sichtweisen auf datengetriebene Analysen und ML veröffentlicht. Die Hong Kong Monetary Authority legt mit Ihrer Veröffentlichung einen praxisnahen Fokus auf die Hürden der Implementierung von ML.¹¹

Dieses Konsultationspapier baut auf bisherigen nationalen und internationalen Veröffentlichungen auf und verknüpft die prudentiellen Risiken von ML mit der bestehenden Aufsichtspraxis. Dieser Ansatz vereinfacht nicht nur die Problemstellung, sondern folgt auch dem Prozess, den Finanzunternehmen bei der Einführung von ML-Methoden durchlaufen.

Der Fokus dieses Papiers liegt auf der Solvenzaufsicht, und hier insbesondere auf der Anwendung von ML-Methoden an Stellen, die aufsichtlich besonders relevant sind. Dazu zählen einerseits – als Ausnahme von dem Prinzip, dass Algorithmen keiner Genehmigungspflicht unterliegen – ML-Methoden, die Gegenstand aufsichtlicher Prüfungen und Erlaubnisverfahren sind, also in internen Modellen zur Berechnung der regulatorischen Eigenmittelanforderungen (Säule 1). Andererseits sind ML-Methoden relevant, wenn sie im Risikomanagement in Säule 2 zum Einsatz kommen.

Aspekte des Verbraucherschutzes sowie ethische Fragestellungen von ML spielen in diesem Zusammenhang eine vergleichsweise geringe Rolle und werden daher im Folgenden nicht behandelt.¹²

Es gibt keine einheitliche Definition von maschinellem Lernen, was zum einen an der Vielzahl unterschiedlicher Ansätze liegt, zum anderen daran, dass keine klare Trennlinie zu traditionellen Techniken existiert.¹³ Allerdings sind bestimmte Charakteristika bei ML-Methoden häufig besonders stark ausgeprägt, wodurch sie sich von traditionellen Techniken abheben. Ziel dieses Papiers soll es daher sein, solche charakteristischen Eigenschaften von ML-Methoden zu identifizieren, die für die Aufsicht relevante Konsequenzen haben, und Ideen zu entwickeln, wie entsprechende Weiterentwicklungen der aufsichtlichen Praxis aussehen könnten, um auf die verbundenen Risiken zu reagieren.

Dabei stellt sich auch die Frage, ob nicht nur die Aufsichtspraxis, sondern auch die regulatorischen Grundlagen selbst einer Überarbeitung bedürfen, und ob womöglich ein grundsätzlich neuer Aufsichtsansatz für ML-Methoden geschaffen werden muss. Ohne an dieser Stelle bereits zu viel vorwegzunehmen: Da die bestehenden regulatorischen Grundlagen technologieneutral formuliert sind, lassen sie sich weitgehend auf die Charakteristika von ML-Methoden übertragen, und nur an wenigen Stellen könnte eine Anpassung der regulatorischen Grundlagen notwendig werden.

⁸ Autorité de contrôle prudentiel et de résolution (ACPR), 2020, „Governance of Artificial Intelligence in Finance“, online abrufbar: <https://acpr.banque-france.fr/en/governance-artificial-intelligence-finance>

⁹ EBA, 2020, „Report on Big Data and Advanced Analytics“, online abrufbar: <https://www.eba.europa.eu/eba-report-identifies-key-challenges-roll-out-big-data-and-advanced-analytics>

¹⁰ Bank of England, 2019, „Machine learning in UK financial services“, online abrufbar:

<https://www.bankofengland.co.uk/report/2019/machine-learning-in-uk-financial-services>

¹¹ Hong Kong Monetary Authority (HKMA), 2019, „Reshaping Banking with Artificial Intelligence“, online abrufbar: https://www.hkma.gov.hk/media/eng/doc/key-functions/financial-infrastructure/Whitepaper_on_AI.pdf

¹² Im prudentiellen Mandat der Aufsicht werden diese Themen im Rahmen der operationellen Risiken geprüft.

¹³ Im Anhang sind einige bekannte Definitionen von Regulatoren und Entwicklern zusammengestellt, deren Vielfalt die Schwierigkeit einer präzisen Abgrenzung illustriert.

Dieses Papier spricht im Folgenden in erster Linie von „Banken“, die dargestellten Charakteristika und aufsichtlichen Konsequenzen lassen sich aber prinzipiell auf Versicherungen und andere Unternehmen übertragen, die mit der Entwicklung und Implementierung von ML-Methoden im Bereich der regulierten Finanzdienstleistungen befasst sind. Um einen Diskussionsprozess mit der Branche zu starten, ist das vorliegende Papier als Konsultationspapier ausgestaltet: Es enthält eine Anzahl von Fragen, deren Beantwortung zur Weiterentwicklung der Aufsichtspraxis beitragen soll.

Das Papier ist wie folgt gegliedert: Kapitel II identifiziert Charakteristika von ML-Methoden, die für die Ausgestaltung der Aufsichtspraxis von Bedeutung sein könnten. Auf dieser Basis diskutiert Kapitel III mögliche Änderungen der Aufsichtspraxis. Kapitel IV fasst die wichtigsten Ergebnisse zusammen. Kapitel V erläutert die Vorgehensweise der Konsultationsphase.

II. Charakteristika von ML

Dieses Papier entwickelt keine allgemeingültige Abgrenzung von ML-Methoden. Vielmehr soll für die Beurteilung von Modellen unter Säule 1 und 2 eine für aufsichtliche Zwecke hinreichende Abgrenzung entwickelt werden. Je nachdem, ob und welche ML-Charakteristika bei einer konkreten zu prüfenden Methodik vorliegen und wie stark diese ausgeprägt sind, gehen Aufsichtspraxis, Prüfungstechnik und -intensität darauf ein.¹⁴

1. Dimensionen und Beispiele

Die Charakteristika lassen sich anhand der drei Dimensionen des **AI/ML-Szenarios** gruppieren, dessen Konzept im Diskussionspapier der Bundesbank ausführlicher erläutert wird.¹⁵

Zusammen beschreiben (1) die **Methodik und Datengrundlage** die Komplexität und damit das mit dem ML-Verfahren verbundene Modellrisiko. Die (2) **Nutzung des Outputs** beinhaltet, welchen Stellenwert das Verfahren innerhalb des Risikomanagements einnimmt. Die Prüfungsintensität richtet sich dabei nicht nach der Unterscheidung zwischen (3) Eigenentwicklung und **Auslagerung** oder der zugrundeliegenden **IT-Infrastruktur**.

Tabelle 1: Charakteristika von ML-Methoden

AI/ML-Szenario	Charakteristika
1 Methodik und Datengrundlage	Komplexität und Dimension des Hypothesenraums Komplexität des Trainings Adaptivität Datenquellen

¹⁴ Diese Herangehensweise erinnert an den „Ententest“, siehe <https://de.wikipedia.org/wiki/Ententest>

¹⁵ Bundesbank, 2020, „The Use of Artificial Intelligence and Machine Learning in the Financial Sector“, online abrufbar: <https://www.bundesbank.de/resource/blob/598256/d7d26167bceb18ee7c0c296902e42162/mL/2020-11-policy-dp-aiml-data.pdf>

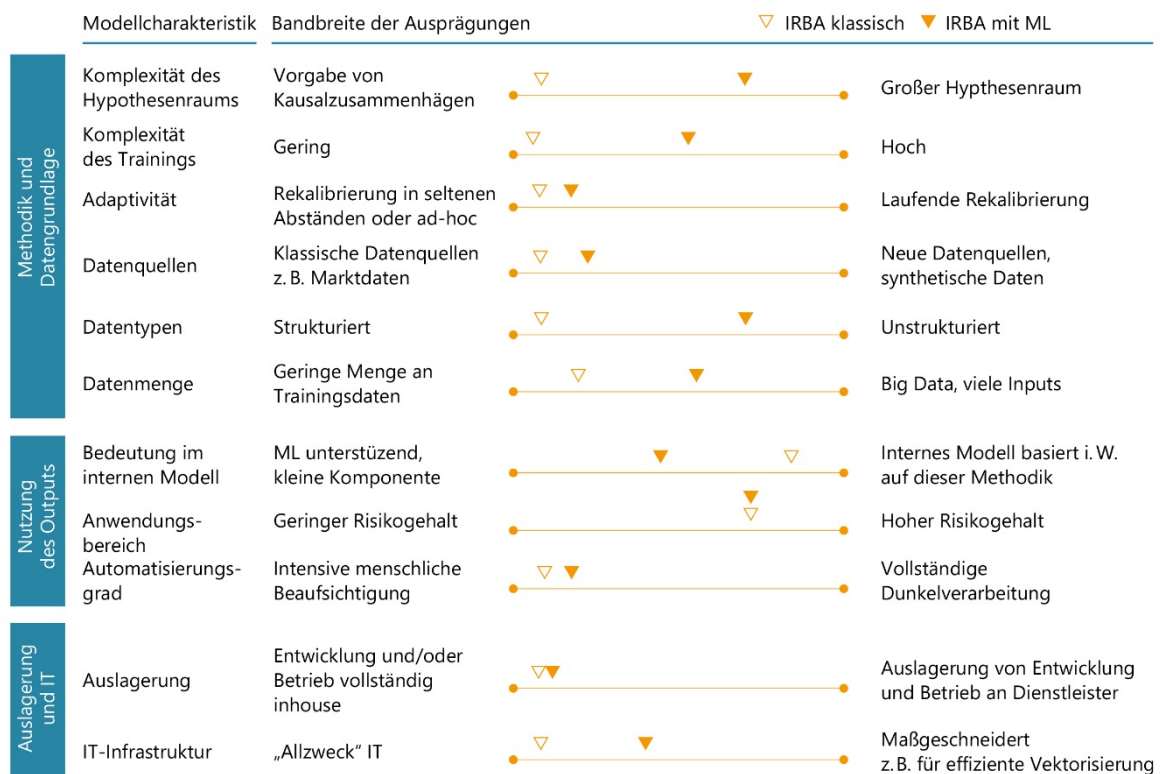
	Datentypen Datenmenge
2 Nutzung des Outputs	Bedeutung im Modell Anwendungsbereich Grad der Automatisierung
3 Auslagerung und IT	Auslagerung IT-Infrastruktur

ML-Methoden und klassische Ansätze bilden ein Kontinuum. Die folgende Abbildung ordnet beispielhaft zwei fiktive Anwendungsfälle aus dem Bereich der Ratingverfahren in die Charakteristika ein.

Anwendungsfall 1: Ein klassisches IRBA-Ratingverfahren schätzt mittels logistischer Regression die Ausfallwahrscheinlichkeit von Kreditnehmern. In das Modell geht eine kleine Anzahl von Variablen ein. Es handelt sich um strukturierte Daten.

Anwendungsfall 2: Für ein IRBA-Ratingverfahren wird ein mehrschichtiges neuronales Netz zur Schätzung der Ausfallwahrscheinlichkeit genutzt. Das Modell wird selten neu trainiert. Die Datengrundlage umfasst viele Variablen, sowohl strukturierte wie auch unstrukturierte Daten. Die Ergebnisse des neuronalen Netzes fließen zusammen mit den Ergebnissen klassischer Modelle in die Schätzung der Ausfallwahrscheinlichkeit ein.

Beim neuronalen Netz sind die dargestellten Charakteristika typischerweise stärker ausgeprägt. Andere ML-Methoden bzw. Anwendungsbereiche zeigen wiederum eigene Profile.



2. Methodik

2.1 Komplexität und Dimension des Hypothesenraums

Modelle in Säule 1 und 2 stellen einen vom Modellierer angenommenen kausalen Zusammenhang zwischen Input (wie etwa Markt- und Portfoliodaten) und Output (wie etwa Preise von Finanzinstrumenten oder Risikomaßzahlen) her. Dieser Zusammenhang wird auch Hypothese genannt. Das namensgebende Grundmerkmal des maschinellen Lernens besteht darin, dass der verwendete Algorithmus nur wenige Annahmen über die Struktur des zu beschreibenden Problems enthält, die z. B. in Form fester Regeln vorgegeben sind,¹⁶ statt dessen werden auch Parameter, die diese Problemstruktur beschreiben, anhand von Daten erlernt. Die zur Anpassung verfügbaren Hypothesen sind also deutlich vielfältiger, und der Hypothesenraum wird durch analytische Techniken nicht vollständig beherrscht, zumal seine Auswahl auch nicht primär durch Kausalitäten motiviert ist.

Der Modellierer gibt wie bei vielen traditionelleren Verfahren auch lediglich den zu durchsuchenden Raum aller von dem Verfahren darstellbaren Hypothesen vor (Hypothesenraum), und das Lernverfahren wählt eine konkrete Hypothese (in Form eines Satzes von Modellparametern) durch Optimierung aus. Die Hypothesen können hier aber beliebig komplex ausgestaltet sein. Letztendlich bestimmt der Modellierer durch die Wahl der

¹⁶ Ausgenommen davon sind sog. Expertensysteme, die ihrerseits die Gewichte von „Wenn-Dann-Regeln“ optimieren.

ML-Methode (z. B. Neuronales Netz, Random Forest, k-nearest neighbors) sowie deren Spezifikation (Modelldesign, Hyperparameter) die möglichen Hypothesen.

Einfache Hypothesenräume treten ebenso bei linearen oder logistischen Regressionen auf, die auch in traditionellen internen Modellen allgegenwärtig sind.

Eine deutlich komplexere Modellstruktur ist durch ein mehrschichtiges neuronales Netz (Deep Neural Network, DNN) gegeben. Hier wird die Hypothese durch die mehrstufige Verkettung nichtlinearer und linearer Abbildungen kodiert, der Hypothesenraum hat eine erheblich höhere Dimension. Die von einem DNN erlernte funktionale Beziehung zwischen Input und Output kann in der Regel nicht mehr durch eine einfache Beschreibung mittels mathematischer Formeln nachvollzogen werden, was als „Blackbox“-Eigenschaft bezeichnet wird.

Aus der eingeschränkten Nachvollziehbarkeit des Modellverhaltens ergeben sich Konsequenzen für die Modellentwicklung, die Validierung und die Rückschlüsse auf die Bedeutung der Datengrundlage. Es entstehen zusätzliche Herausforderungen für die Erklärbarkeit der Modellergebnisse, um ein ausreichendes Verständnis des Wirkungszusammenhangs sicherzustellen und die Anwendbarkeit der ML-Methode intern und extern zu rechtfertigen.

2.2 Komplexität des Trainings

Die Festlegung einer konkreten Hypothese geschieht in einem Lernverfahren, das als „Kalibrierung“ oder „Training“ bezeichnet wird.

Charakteristisch für das Training neuerer ML-Methoden sind die hohe Anzahl an Rechenoperationen durch die Vielzahl an Funktionsargumenten und Modellparametern, die komplexe Abfolge verschachtelter Rechenvorschriften oder iterative Verfahren. Dies führt zu Herausforderungen in der technischen Umsetzung, etwa bei der Verfügbarkeit erforderlicher Hardware-Ressourcen oder bei der numerischen Stabilität der Rechenverfahren.

Die Optimierung von nichtlinearen Funktionen in hochdimensionalen Räumen besitzt in der Regel keine eindeutige Lösung. Es hängt vom Trainingsansatz und möglicherweise vom Zufall ab, welches Optimum als konkrete Hypothese ausgewählt wird. Das kann Auswirkungen auf die Stabilität des ML-Verfahrens haben, wenn bei sukzessiven Trainings jeweils verschiedene Optima ausgewählt werden.

Mit steigender Komplexität der Trainingsverfahren rücken diese stärker in den aufsichtlichen Fokus.

2.3 Adaptivität

Manche ML-Methoden sind darauf ausgelegt, in hoher Frequenz oder sogar quasi kontinuierlich an neue Daten angepasst zu werden. Hierdurch verschwimmt die Abgrenzung zwischen Modellentwicklung und Modellbetrieb, sowie zwischen Modellpflege und

aufsichtlich zu beurteilender Modelländerung. Diese Abgrenzung ist insbesondere unter Säule 1 relevant, wenn Modelländerungen einer aufsichtlichen Zulassung bedürfen. Des Weiteren stellt sich die Frage der Validierbarkeit solcher Modelle und der Reproduzierbarkeit von Modellergebnissen. Auch die Gewährleistung einer fortlaufenden adäquaten Datenqualität spielt eine besondere Rolle. Mit zunehmender Adaptivität von Verfahren gewinnt aus aufsichtlicher Betrachtung eine klare Unterscheidung von Modellpflege und Modelländerung an Bedeutung.

3. Datengrundlage

Die zunehmende Verwendung großer Datenmengen kennzeichnet die aktuelle allgemeine Entwicklungen in den Banken, ist aber auch ein Charakteristikum von komplexen oder neuen ML-Methoden im engeren Sinne.

ML baut häufig auf einer größeren Anzahl von **Datenquellen** und einer Verknüpfung unterschiedlicher Datenquellen auf. Auch synthetische, also künstlich erzeugte Daten werden verwendet. Als **Datentyp** kommen auch unstrukturierte Daten zum Einsatz. ML-Methoden können oftmals mit einer Vielzahl von Input-Parametern im Modell umgehen, so dass sich eine große **Datenmenge** ergibt.

Die Leistungsfähigkeit von ML-Methoden wird nicht zuletzt durch die für das Training verfügbaren Datenmengen, ihrer Aussagekraft und ihrer Qualität bestimmt. Die Datengrundlage rückt damit noch stärker in den aufsichtlichen Fokus.

4. Verwendung des Outputs

4.1 Bedeutung im Modell

Die ML-Methode kann verschiedene Rollen innerhalb eines gesamten Modells einnehmen. Sie kann als unterstützende Komponente für ein Modell, bspw. zur Datenaufbereitung, oder als Teilkomponente in ein Modell eingehen, bspw. als Modul eines Ratingverfahrens. Mitunter stellt die ML-Methode auch den zentralen Bestandteil eines Modells dar. Oder sie wird außerhalb des internen Modells eingesetzt, zum Beispiel als Validierungsinstrument (challenger tool) oder als Proxy des „echten“ Modells innerhalb spezieller Anwendungsbereiche.

Mit steigender Bedeutung der ML-Methode im und neben dem Modell steigt die Intensität der aufsichtlichen Auseinandersetzung.

4.2 Anwendungsbereich

Der Anwendungsbereich beschreibt, welche Ergebnisse die ML-Methode liefert und mit welcher Relevanz diese in die Geschäftsprozesse des Unternehmens einfließen. Beispiele für den Anwendungsbereich können interne Modelle, Risikofrüherkennungsverfahren oder

Ratingverfahren sein. Je entscheidender sich der Anwendungsbereich auf die Risikosituation auswirkt, desto höhere Anforderungen stellt die Aufsicht an die ML-Methode.

4.3 Grad der Automatisierung

Der Grad der Automatisierung lässt sich in algorithmendeterminierte und algorithmenbasierte Prozesse unterteilen; entsprechend unterteilt sich auch das einhergehende operationelle Risiko. Algorithmendeterminierte Prozesse bezeichnen eine weitgehende Automatisierung in der weiteren Verwertung der Ergebnisse und damit höhere Risiken bei unzureichend überwachtem Einsatz von ML-Methoden. Algorithmenbasierte Prozesse hingegen setzen stärker auf menschlich kontrollierten Prozessen auf (was wiederum eigene Risiken mit sich bringt).

5. Auslagerung und IT-Infrastruktur

Auslagerungen an spezialisierte Dienstleister und der Einsatz einer spezifischen IT-Infrastruktur sind ebenfalls charakteristisch für ML-Methoden. FinTechs und BigTechs bieten modulare Systeme zur Erstellung von ML-Methoden an und stellen eine auf hohe Rechenleistung zugeschnittene IT-Infrastruktur zur Verfügung. Herausforderungen können entstehen, wenn ML-Methoden in sogenannte „Legacy IT“ integriert wird.

Die aufsichtliche Herangehensweise zur Beurteilung solcher Dienstleistungsverhältnisse ändert sich durch den Einsatz von ML-Methoden allerdings nicht; die einschlägigen Vorschriften wie etwa die „Bankaufsichtlichen Anforderungen an die IT“ (BAIT) decken auch diese Anwendungsfälle ab. Aus diesem Grund wird in Kapitel III nicht speziell auf diese Aspekte eingegangen.

Fragen zu Kapitel II:

- a) Halten Sie den Ansatz für zielführend, auf eine strikte Definition von ML-Methoden zu verzichten, und stattdessen die Aufsichts- und Prüfungspraxis anwendungsorientiert an den einzelnen Charakteristika der eingesetzten Methoden auszurichten?
- b) Welche weiteren Charakteristika von ML-Methoden können aus Ihrer Sicht für die Aufsichtspraxis oder auch die interne Modelle-Governance von Bedeutung sein?
- c) Welche Charakteristika gehören aus Ihrer Sicht nicht in diese Übersicht?
- d) In welchen relevanten Anwendungsbereichen kommen ML-Methoden bei Ihnen zum Einsatz bzw. wo planen Sie deren Umsetzung?

III. Aufsichtlicher Ansatz

Bereits in der Einleitung wurde angedeutet, dass ML-Methoden keine grundsätzlich neue Aufsichtspraxis erfordern. Für die in Kapitel II beschriebenen Charakteristika von ML-Methoden werden in diesem Abschnitt aufsichtliche Schlussfolgerungen gezogen. Hierbei wird auch dargestellt, inwieweit Anpassungen der Aufsichtspraxis an bestimmten Stellen erforderlich sind. Der aufsichtliche Ansatz wird nach dem Proportionalitätsprinzip und anhand der Charakteristika aus Kapitel II bestimmt. Dieses Kapitel gliedert sich nach den identifizierten Risiken von ML und ordnet diese in den Zyklus der Modellentwicklung und -pflege ein.

1. Aufsichtspraxis hat Bestand

In Säule 1 liegen umfangreiche Regeln zur Überprüfung und Genehmigung interner Modelle vor, die technologie-neutral formuliert sind und somit auch die Risiken von ML-Methoden

adressieren. In Säule 2 liefern prinzipienorientierte Anforderungen an das Risikomanagement und die IT eine solide Grundlage.^{17,18}

Die Aufsicht verfolgt systematisch einen auf bankgeschäftliche Prozesse ausgerichteten Prüfungsansatz, welcher allgemeine, übergreifende Prüfungsgebiete für die jeweils zu prüfende Risikoart (z. B. Kredit- oder Marktrisiko) festlegt und stetig an die aktuellen Gegebenheiten angepasst wird.

Die Aufsichtspraxis für ML-Methoden lässt sich daher aus dem bestehenden Rahmenwerk ableiten. Zugleich wird anhand einer Abweichungsanalyse aktuell erhoben – unterstützt auch durch diese Konsultation – in welchen Bereichen der aufsichtliche Prüfungsansatz dezidierter Konkretisierungen bedarf, um den Besonderheiten des Einsatzes von ML-Methoden gerecht zu werden.

Dabei wird neben den mathematisch-/methodischen Aspekten auch der prozessualen Einbettung der ML-Methoden Rechnung getragen, welche für einen kontrollierten und damit erfolgreichen und effizienten Einsatz gleichermaßen bedeutend sind.

Der aufsichtliche Fokus bei ML-Methoden liegt auf neuen oder deutlicher ausgeprägten Risiken. Diese zeigen sich bei der Datengrundlage, bei der Validierung (von der Modellentwicklung über Testverfahren bis zum Betrieb), bei Modelländerungen und bei der Steuerung.

Auf einer abstrakteren Ebene stellen neue Gesetzesvorhaben wie der Entwurf der KI-Verordnung durch die Europäische Kommission Aspekte des Verbraucherschutzes in Zusammenhang mit ML in den Vordergrund. Die Aufsicht steht vor der Aufgabe, neue Anforderungen widerspruchsfrei in die bestehenden Anforderungen von Säule 1- und Säule 2-Modellen einzufügen.

¹⁷ Die Rechtsgrundlagen bilden im Wesentlichen die CRR und die MaRisk für die Säule 1- bzw. Säule 2-Prüfungen, flankiert von Standards der EBA und des SSM. Für Säule 1-Modelle sind hier die Delegierten Verordnungen Nr. 529/2014 und 942/2015 anzuwenden. Die EBA/GL/2020/06 (Guidelines on loan origination and monitoring) nennt Anforderungen für die automatisierte Kreditwürdigkeitsprüfung und Kreditvergabe, die sich mit dem deutschen Aufsichtsansatz decken.

¹⁸ BaFin, 2021, „Big Data und künstliche Intelligenz: Prinzipien für den Einsatz von Algorithmen in Entscheidungsprozessen“, online abrufbar: https://www.bafin.de/SharedDocs/Downloads/DE/Aufsichtsrecht/dl_Prinzipienpapier_BDAI.html

Fragen zu Kapitel III.1:

- e) Liegen aus Ihrer Sicht bereits aufsichtliche Anforderungen in Regelwerken vor, die für den Einsatz von ML-Methoden hinderlich erscheinen? Ergeben sich aus Ihrer Sicht Widersprüche zwischen prudentiellen Regelungen für Säule 1- und Säule 2-Modelle und dem Entwurf der KI-Verordnung? Bitte geben Sie Referenzen auf die entsprechenden Regelwerke an und erläutern Sie die Herausforderungen.
- f) Inwiefern sind die Anforderungen der EBA/GL/2020/06 mit Bezug auf die automatisierte Kreditwürdigkeitsprüfung und Kreditvergabe auch für andere ML-Methoden in Säule 2 (MaRisk) aus Ihrer Sicht passend und sollten übernommen werden?
- g) Sehen Sie weitere Punkte, bei denen aus Ihrer Sicht eine Anpassung der bisherigen Aufsichtspraxis erforderlich ist, um ML-Verfahren und die damit verbundenen Risiken angemessen zu würdigen?
- h) Gehen mit ML-Methoden spezifische Risiken für die IT-Implementierung und das Auslagerungsmanagement einher? Sind z. B. sog. „Adversarial Attacks“ im Finanzbereich denkbar und sollten ML-Methoden besonders dagegen geschützt werden?

2. Methoden laden zur Datengläubigkeit ein

Die Datenqualität ist bereits ein Schwerpunkt aufsichtlichen Handelns. Die Charakteristika von ML-Methoden machen allerdings deutlich, dass die Datengrundlage hier insbesondere als Ausgangspunkt und Erfolgsfaktor anzusehen ist. Unstrukturierte Daten können mittlerweile durch bzw. für ML-Methoden erschlossen werden. Ferner erlauben ML-Methoden Berechnungen unter Einbeziehung zahlreicher Einflussgrößen. Es wird dem Modellierer einfach gemacht, ML zügig auf eine große Datengrundlage anzuwenden.

ML-Methoden lernen, was sie in den vorgegebenen Daten finden und replizieren dort vorhandene Muster. Die „Blackbox“-Eigenschaft kann dazu führen, dass Probleme durch scheinbar gute Performance verdeckt werden. Beispielsweise kommt es vor, dass Modelle Korrelationen zwischen Inputdaten erlernen, die keinen realen Zusammenhang darstellen, sondern auf zufälligen Eigenschaften des Lerndatensatzes beruhen (Model Overfitting).

ML-Methoden können große Datenmengen verwerten, deren Qualität fortlaufend sichergestellt werden muss. Dies betrifft neben der Modellentwicklung und -validierung auch die Modellanwendung.

Die Aufsicht erwartet, dass Banken zusätzliche Anstrengungen unternehmen, um die Qualität der Datengrundlage sicherzustellen. Dazu gehört insbesondere, dass die Trainingsdaten frei von systematischen Verzerrungen hinsichtlich der vom Modell zu erlernenden Wirkungszusammenhänge sind.

Fragen zu Kapitel III.2:

- i) Welche Herausforderungen sehen Sie bei der Auswahl der Daten und bei der Sicherstellung der Datenqualität von ML-Methoden?
- j) Welche Aspekte der Datenqualität werden durch die Anwendung von ML-Methoden aus Ihrer Sicht erleichtert?

3. Erklärbarkeit rückt in den Fokus

Je komplexer und höherdimensional der vom Modell abbildbare Hypothesenraum ist, desto schwieriger wird es, den funktionalen Zusammenhang zwischen Input und Output (d. h. die im Training konkretisierte Hypothese) verbal oder durch mathematische Formeln zu beschreiben, und desto weniger sind die Berechnungen durch Modellierer, Anwender, Validierer und Aufseher im Detail nachvollziehbar. Dies führt zu einer erschwerten Nachvollziehbarkeit der Modellierung und ggf. auch zu einer erschwerten Überprüfung der Validität der Modellergebnisse. Zusätzlich kann die Akzeptanz bei den Anwendern leiden.

Diese „Blackbox“-Eigenschaft kann als Preis für leistungsfähigere Modelle angesehen werden. Er kann durchaus gerechtfertigt sein, z. B. durch eine höhere Vorhersagegüte. Dem steht ein potenziell höheres Modellrisiko entgegen, abhängig davon, welche Bedeutung das Modell im Kontext der bankgeschäftlichen Prozesse hat.

Modellierer müssen begründen, warum Einbußen in der Nachvollziehbarkeit des Modells durch die erzielten Vorteile gerechtfertigt sind. Inwieweit eine „Blackbox“ aufsichtlich akzeptiert werden kann, hängt auch von der Behandlung des betreffenden Modells im Risikomanagement der Bank ab.

Dass ML-Methoden mitunter Zusammenhänge in den Daten lernen, die jeder realen Grundlage entbehren und keine generellen Schlüsse erlauben, kann durch die „Blackbox“ verborgen bleiben.

Anstelle der Nachvollziehbarkeit im Detail gewinnen daher die Erklärbarkeit und die Plausibilisierung des Modellverhaltens im Ganzen an Bedeutung.

Der Begriff der „Erklärbarkeit“ ist vielschichtig, da Modellierer, Validierer, Aufseher und Nutzer über einen unterschiedlichen fachlichen Hintergrund verfügen und unterschiedliche Informationsbedarfe haben. Um diesen adressatenspezifischen Anforderungen Rechnung zu tragen, wurden Techniken der „Explainable AI“ (XAI) entwickelt.

XAI-Techniken sind aus Sicht der Aufsicht vielversprechend, um die „Blackbox“-Eigenschaft abzumildern. XAI-Techniken stellen allerdings selbst Modelle mit Annahmen und Schwächen dar und befinden sich vielfach noch in der Erprobungsphase. Deshalb ergibt sich die Herausforderung, diese Techniken in einen strukturierten Prozess einzubinden. So sollte bspw. festgelegt werden, wann welche Techniken Verwendung finden, ob über globale oder lokale Ansätze gearbeitet werden soll, welche Auswahlkriterien und Stichprobengrößen lokale Erklärungen erfordern und welche Adressaten angesprochen werden sollten.

Banken müssen Validierungshandlungen durchführen, die auf die Modellstruktur zugeschnitten sind und die Modellrisiken ausreichend abdecken. Bei Verfahren maschinellen Lernens besteht bspw. durch die hohe Anzahl von Parametern eine besondere Gefahr des Overfittings. Zusätzlich zu den üblichen Ansätzen, wie etwa „out-of-sample“-Validierungen und Backtesting, können XAI-Ansätze dazu beitragen, Overfitting aufzudecken. Zusätzlich können auch synthetische oder Stress-/Extremszenarien sowie Tests gegen traditionelle Verfahren die Plausibilisierung und Erklärbarkeit der ML-Methode unterstützen.

Fragen zu Kapitel III.3:

- k) Welchen Einfluss hat die „Blackbox“-Eigenschaft Ihres Erachtens auf die Validierung der Verfahren?
- l) Welche Bedeutung messen Sie dem Trade-Off zwischen Performance und Erklärbarkeit bei?
- m) Bieten XAI-Techniken aus Ihrer Sicht (immer) einen Ausweg aus der „Blackbox“? Welche Verfahren haben sich als vielversprechend herausgestellt und bei welchen ML-Methoden?
- n) Wie sollte aus Ihrer Sicht eine der Methode nachgelagerte XAI in die Validierung einbezogen werden?

4. Adaptivität: Modelländerungen sind schwerer zu erkennen

Institute und Unternehmen sind dazu verpflichtet, der Aufsicht Änderungen von Säule 1-Modellen zu melden und diese ggf. erst nach Genehmigung in Betrieb zu nehmen. Die Grenze zwischen regelmäßiger Modellpflege und Modelländerung ist fließend und führt immer wieder zu Diskussionen mit der Aufsicht, zumal der Begriff der „Modelländerung“ auch vom jeweiligen aufsichtlichen Kontext abhängig ist.

Die Flexibilität und die teils hochfrequente Adaptivität, d. h. eine häufige Anpassung bspw. aufgrund neuer Daten, von ML-Verfahren erschweren eine klare Unterscheidung zwischen Anpassungen und Änderungen, die für die Aufsicht allerdings unverzichtbar ist. Die Notwendigkeit einer hochfrequenten Adaptivität sollte grundsätzlich gut begründet werden.

Dieses Dokument kann keine ML-spezifische Definition des Begriffs der Modelländerung geben, jedoch wird anhand von Beispielen illustriert, wie die aufsichtliche Einordnung erfolgen kann.

- Prozessuale Kriterien wie die erstmalige Verwendung einer ML-Methode für eine vorgegebene Aufgabe oder Veränderungen der Abläufe im Zusammenhang mit der Verwendung der Methodik können wie bei traditionellen Methoden zur Identifikation und Einordnung von Modelländerungen herangezogen werden.
- Ob das Re-Training, d. h. Training nach einem Update der Datengrundlage, einer ML-Methode als Änderung anzusehen ist, lässt sich durch einen Vergleich mit traditionellen Ansätzen nicht allgemeingültig beantworten:

Z. B. wird beim Kontrahentenrisiko die Rekalibrierung von Wahrscheinlichkeitsverteilungen der Risikofaktoren grundsätzlich als Modellpflege und nicht als Änderung betrachtet. Der von diesen Modellen darstellbare Hypothesenraum ist allerdings deutlich kleiner als bei modernen ML-Methoden, der funktionale Zusammenhang zwischen In- und Output wird durch die Rekalibrierung eines solchen Modells nicht fundamental geändert.

Das Re-Training einer ML-Methode kann hingegen auch strukturelle Anpassungen der abgebildeten Funktion wie etwa die Anzahl der Schichten eines DNN beinhalten. Auch wenn der vom Modell darzustellende funktionale Zusammenhang zwischen Input und Output sich möglicherweise nur geringfügig geändert hat (etwa wegen Marktbewegungen), kann die realisierte „Hypothese“ sich von der vorherigen so deutlich unterscheiden, dass sie aufsichtlich als Modelländerung anzusehen wäre.

Unabhängig von der Klassifikation können sich ML-Methoden durch ihre höhere Flexibilität und Adaptivität nach sukzessiven Trainings oder Anpassungen bereits nach kurzer Zeit weit vom Ursprungsmodell entfernt haben, in dem Sinne, dass Inputs mit gänzlich anderem Gewicht in die Outputs eingehen. Daher müssen dieser Anpassungsfähigkeit – mehr als bei traditionellen Techniken – Kontrollen entgegengestellt werden, die verhindern, dass die Modellstruktur sich grundlegend ändert, wodurch Erklärbarkeit und Validierung hinfällig sein können, ohne dass eine erneute aufsichtliche Bewertung stattfindet.

- Eine Erweiterung der verwendeten Risikofaktoren würde weiterhin als Modelländerung gewertet werden. Die Abgrenzung kann aber bei komplexen Datensätzen, insbesondere bei unstrukturierten Daten, schwierig sein.

Diese Beispiele zeigen auch die nötige Verschiebung des aufsichtlichen Fokus und der Herangehensweise bei Modellanpassungen:

- **Modellgenehmigung:** Bereits im Rahmen der Modellgenehmigung müssen die möglichen Techniken der Modellpflege (inkl. laufender interner Modellvalidierung und Modelländerung) untersucht werden. Die von der Bank zu erstellende Model Change Policy trägt bei der Einstufung von Anpassungen den Besonderheiten der verwendeten ML-Methode Rechnung. Möglicherweise können auch hier XAI-Techniken zur Anwendung kommen.
- **Kommunikation:** Die Aufsicht muss auch Aktivitäten der regelmäßigen Modellpflege verfolgen können, um rechtzeitig zu erkennen, ob durch eine Vielzahl schleichender Modelländerungen das Ursprungsmodell signifikant verändert wird.
- **Interne Validierung:** Stärker als bei traditionellen Techniken kommt der internen Validierung bei ML-Methoden die Aufgabe zu, kontinuierlich (und nicht nur im Jahresrhythmus) die Angemessenheit der eingesetzten Methoden zu überwachen und sicherzustellen. Die Validierung der Modelle muss deutlich machen, welche Änderung an der Modellstruktur welchen Effekt auf die Ergebnisse des Modells haben wird. Dadurch werden erhöhte Anforderungen an die Validierer gestellt. Nur so ist sichergestellt, dass ungewollte Effekte aus Modelländerungen frühzeitig identifiziert und vermieden werden. Bei hochfrequentem Retraining stellt sich die die Frage, welche Modellversion zu validieren ist (kritisch z. B. bei laufendem Retraining).

Viele ML-Methoden fallen in den nicht genehmigungspflichtigen Aufsichtsbereich der Säule 2. Damit ergibt sich eine größere Flexibilität für Retrainings und Änderungen des Modells, wobei auch hier bereits bestehende Anforderungen z. B. aus der MaRisk gültig bleiben. Aus Sicht der Aufsicht ist es dennoch entscheidend, trotz dieser Flexibilität den Trainingszyklus an dem Anwendungsfall anzupassen und entsprechend zu begründen, um ein Gleichgewicht zwischen Aktualität hinsichtlich der Daten und Erklär- sowie Validierbarkeit zu schaffen.

Fragen zu Kapitel III.4:

- o) Welche Fragen ergeben sich aus Ihrer Sicht zur Aufsichtspraxis in Bezug auf Modellanpassungen bei ML-Methoden?
- p) Sehen Sie für bestimmte ML-Methoden die Notwendigkeit sehr häufiger Retrainings?
- q) Werden ML-Methoden Ihres Erachtens eine Anpassung der Modell-Governance notwendig machen? Wie arbeiten klassische Modellierungseinheiten, Validierer und neue „Data Science“-Einheiten zusammen?

IV. Ausblick

Mit Blick auf die aufsichtliche Schwerpunktsetzung soll ergänzend zu den allgemeinen BDAI-Prinzipien Klarheit bei der Entwicklung und Anwendung von ML-Methoden im Kontext aufsichtlich relevanter Modelle in Säule 1 und Säule 2 geschaffen werden. Unternehmen wird damit ein Umfeld geschaffen, das Investitionen in ML ermöglicht und sie in die Lage versetzt, die Risiken der Methoden so früh wie möglich zu adressieren. Gleichzeitig muss der Impuls zum Einsatz von ML von den Unternehmen kommen. Die Aufsicht fordert nicht den Einsatz von ML-Methoden, solange „klassische“ Techniken zur Erfüllung der regulatorischen Anforderungen als geeignet angesehen werden.

Der nächste Schritt wird sein, über diese Konsultation in den Dialog mit Unternehmen einzutreten. Die Aufsicht sieht es außerdem als notwendig an, internationale Ansätze möglichst zu harmonisieren und sektorübergreifend gleichlautende Anforderungen an den Einsatz von ML-Methoden zu stellen. Die Digital Finance Strategy der EU Kommission¹⁹ leistet einen Beitrag zur europäischen Vereinheitlichung.

V. Konsultation

Wir laden Sie ein, BaFin und Bundesbank Ihre Stellungnahme zu diesem Konsultationspapier und die Antworten auf die gestellten Fragen schriftlich per E-Mail an Konsultation-11-21@bafin.de und ai-b3@bundesbank.de bis zum 30.09.2021 zukommen zu lassen.

Es ist nicht beabsichtigt, die eingereichten Stellungnahmen zu veröffentlichen. Stattdessen werden die Positionen zusammengefasst und anonymisiert veröffentlicht.

¹⁹ EU Kommission, 2020, „Strategie für ein digitales Finanzwesen in der EU“, online abrufbar: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0591>

Anhang

Häufig gestellte Fragen

Zählen auch Verfahren unter ML, die angenommene Wirkungszusammenhänge explizit abbilden?

Nicht zu ML im engeren Sinn – Lernen im statistischen Kontext bedeutet vielmehr, auf Basis von Beobachtungen der Vergangenheit Wirkungszusammenhänge und Gesetzmäßigkeiten zu erkennen, um in der Zukunft bessere Vorhersagen zu treffen.²⁰ Beim maschinellen Lernen steht ein bestimmtes Lernproblem im Mittelpunkt: Aus einer Sammlung von Daten (Input-Output) soll unter Ausnutzung von statistischen Schlüssen eine Funktion erlernt werden, die für neue Inputs den Output vorhersagt.²¹ Maschinelles Lernen wird genutzt, weil Modellierer den fachlichen Zusammenhang zwischen Input und Output weder selbst überblicken, noch Wenn-Dann-Bedingungen dafür finden können.²²

Fallen auch „klassische“, seit langem eingesetzte und der Aufsicht bekannte Modelle (bspw. logistische Regressionen) unter die Charakteristika von ML?

Auch in bestehenden internen Modellen werden selbstlernende Verfahren eingesetzt, die in unterschiedlichem Ausmaß die oben beschriebenen Charakteristika aufweisen. Neben einfachen Regressionsansätzen für individuelle Risikofaktoren (z. B. Short rate Modelle) finden sich z. B. in den internen Modellen von Versicherungsgruppen auch komplexe Ansätze zur Nachbildung von Eigenmittelveränderungen unter kombinierten Risikofaktorauslenkungen (Markt-, Kredit-, Zeichnungs- und operationelle Risiken), deren Spezifikation einen vergleichbar hohen Rechenaufwand erzeugt wie das Training eines neuronalen Netzes. Ihre Spezifikation kann auch „Trainingszyklen“ umfassen, etwa die iterative Hinzu- oder Herausnahme von Polynombausteinen anhand eines Informationskriteriums in Kombination anschließendem Retraining der erweiterten/gekürzten Funktion per Regression.

Der Hypothesenraum ist in den bekannten Modellen in der Regel stärker beschränkt als bei den oben genannten ML-Techniken, weshalb sich einige Folgeprobleme wie die „Blackbox“-Eigenschaft nicht oder deutlich weniger einstellen. Hinsichtlich der Validierung können jedoch vergleichbare Einschränkungen bei der schrittweisen Nachvollziehbarkeit der Berechnungen auftreten, so dass man auch hier auf die Plausibilisierung von Ergebnissen z. B. über „Out-of-Sample“-Vergleiche angewiesen ist.

²⁰ Vgl. Russel, S., Norvig, P., Artificial Intelligence: A Modern Approach, Dritte Auflage, 2016, S. 693

²¹ Der Fokus liegt hier zunächst auf überwachtem Lernen. Beim unüberwachten Lernens besteht die Lernaufgabe aus der Mustererkennung, indem Inputdaten (ohne Output/„Label“) systematisiert werden.

²² Vgl. Russel, S., Norvig, P., Artificial Intelligence: A Modern Approach, Dritte Auflage, 2016, S. 693ff; Bank of England, Machine learning in UK financial services, 2019

Steht die Aufsicht unüberwachtem Lernen kritischer gegenüber als überwachtem Lernen?

Nein. Der Anwendungsfall und die Ausprägung von ML stehen in direktem Zusammenhang. So löst jede Ausprägungsform unterschiedliche Probleme und eignet sich deshalb für unterschiedliche Anwendungsfälle. Eine Abwägung der Ausprägungen im Sinne ihres Risikogehaltes ist nicht möglich. Jeder Anwendungsfall ist entsprechend der Charakteristika individuell einzuordnen und zu prüfen. Grundsätzlich wendet die Banken- und Finanzaufsicht in Deutschland bei allen Ausprägungen von ML den gleichen Bewertungsmaßstab an. Gleichwohl überwiegen aktuell Anwendungsfälle mit überwachten Lernverfahren.

Tabelle 2: Ausprägungen von ML

Ausprägungen

Überwachtes Lernen: Algorithmus wird mittels Inputdaten auf ein bestimmtes Resultat, auch „Label“ genannt, trainiert. Das trainierte Modell kann dann auf neue Daten angewendet werden.

Unüberwachtes Lernen: Finden verborgener Muster oder Strukturen (Clustering, Assoziation, Dimensionsreduktion).

Bestärkendes Lernen: Lernaufgabe, bei der die eigenen Lösungsstrategien permanent wieder als Inputparameter aufgenommen und evaluiert werden.

Ausgewählte Beispiele von Definitionen des Begriffs „Machine Learning“

Tabelle 3: Ausgewählte Definitionen von ML

Wissenschaft	Regulatoren	IT-Unternehmen
<ul style="list-style-type: none"> ▪ „Set of methods that can automatically detect patterns in data, then use the uncovered patterns to predict future data, or to perform other kinds of decision making under uncertainty.“ Murphy, S. 1 ▪ „Machine Learning is the science (and art) of programming computers so they can learn from data.“ Geron, S. 2 ▪ „Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed.“ Samuel in Geron, S. 2 ▪ „A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E“ Mitchell in Geron, S.2 ▪ „Maschinelles Lernen bezweckt die Generierung von »Wissen« aus »Erfahrung«, indem Lernalgorithmen aus Beispielen ein komplexes Modell entwickeln. Das Modell, und damit die automatisch erworbene Wissensrepräsentation, kann anschließend auf 	<ul style="list-style-type: none"> ▪ „The standard on IT governance ISO/IEC 38505-1:2017 defines ML as a ‘process using algorithms rather than procedural coding that enables learning from existing data in order to predict future outcomes.’“ EBA, S. 14 ▪ „ML is a methodology whereby computer programmes fit a model or recognise patterns from data, without being explicitly programmed and with limited or no human intervention. This contrasts with so-called ‘rules-based algorithms’ where the human programmer explicitly decides what decisions are being taken under which states of the world.“ PRA, S. 6 ▪ „Machine learning may be defined as a method of designing a sequence of actions to solve a problem, known as algorithms, which optimise automatically through experience and with limited or no human intervention.“ FSB, S. 4 	<ul style="list-style-type: none"> ▪ „ML, a subset of AI, focuses on the ability of machines to receive data and learn for themselves without being programmed with rules. ML differs from traditional programming by allowing you to teach your program with examples rather than a list of instructions. [It] enables you to “train” an algorithm so that it can learn on its own, and then adjust and improve as it learns more about the information it is processing.“ Google ▪ „ML is a form of AI that enables a system to learn from data rather than through explicit programming. [...] ML enables models to train on data sets before being deployed. Some ML models are online and continuous [leading] to an improvement in the types of associations made between data elements. Due to their complexity and size, these patterns and associations could have easily been overlooked by human observation.“ IBM ▪ „ML is the process of using mathematical models of data to help a computer learn without direct instruction. It’s considered a subset of AI. ML uses algorithms to identify patterns within data, and those patterns are then used to create a data model that can make predictions. With increased data and experience, the results of machine learning are more accurate – much like how humans improve with more practice.“ Microsoft

neue, potenziell
unbekannte Daten
derselben Art angewendet
werden.“ Fraunhofer IAIS,
S. 8

Quellen der ML-Definitionen

1. A. Geron, 2019, „Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow“, 2. Aufl. O'Reilly.
2. T. Hastie, R. Tibshirani und J. Friedman (2009), „The Elements of Statistical Learning: Prediction, Inference and Data Mining“, 2. Aufl. Springer-Verlag.
3. G. James, D. Witten, T. Hastie und R. Tibshirani (2013), „An Introduction to Statistical Learning“, Springer-Verlag
4. K.P. Murphy, 2012, „Machine Learning: A Probabilistic Perspective“, MIT Press.
5. EBA, 2020, „Report on Big Data and Advanced Analytics“, online abrufbar: https://eba.europa.eu/sites/default/documents/files/document_library/Final%20Report%20on%20Big%20Data%20and%20Advanced%20Analytics.pdf
6. Bank of England, 2019, „Machine learning in UK financial services“.
7. Bundesbank, 2020, „Policy Discussion Paper, The Use of Artificial Intelligence and Machine Learning in the Financial Sector“, online abrufbar: <https://www.bundesbank.de/resource/blob/598256/d7d26167bceb18ee7c0c296902e42162/mL/2020-11-policy-dp-aiml-data.pdf>
8. Financial Stability Board (FSB), 2017, „Artificial intelligence and machine learning in financial services, Market developments and financial stability implications“, online abrufbar <https://www.fsb.org/wp-content/uploads/P011117.pdf>
9. Fraunhofer IAIS, 2018, „Maschinelles Lernen. Eine Analyse zu Kompetenzen, Forschung und Anwendung“, online abrufbar: [https://www.bigdata-ai.fraunhofer.de/content/dam/bigdata/de/documents/Publikationen/Fraunhofer Studie ML 201809.pdf](https://www.bigdata-ai.fraunhofer.de/content/dam/bigdata/de/documents/Publikationen/Fraunhofer_Studie_ML_201809.pdf)
10. Google, „Using AI for a social good“, online abrufbar: <https://ai.google/education/social-good-guide/?category=introduction>
11. IBM, 2020, „Data science and machine learning“, online abrufbar: <https://www.ibm.com/hk-en/analytics/machine-learning#:~:text=Machine%20learning%20enables%20models%20to,associations%20made%20between%20data%20elements>

12. Microsoft Azure, „What is Machine Learning?“, online abrufbar:
<https://azure.microsoft.com/en-us/overview/what-is-machine-learning-platform/>

Impressum

Herausgeber

Bundesanstalt für Finanzdienstleistungsaufsicht (BaFin)
Gruppe Kommunikation
Graurheindorfer Straße 108, 53117 Bonn
Marie-Curie-Straße 24 – 28, 60439 Frankfurt am Main
www.bafin.de

Redaktion und Layout

BaFin, Interne Kommunikation und Internet

Redaktion: Rebecca Frener
Tel.: +49 0 228 41 08 22 13
Kathrin Jung
Tel.: +49 0 228 41 08 16 28

Layout: Christina Eschweiler,
Tel.: +49 0 228 41 08 38 71

E-Mail: journal@bafin.de

Designkonzept

werksfarbe.com | concept + design
Humboldtstraße 18, 60318 Frankfurt
www.werksfarbe.com

Bezug

Das BaFinJournal* erscheint jeweils zur Monatsmitte auf der Internetseite der BaFin. Mit dem Abonnement des Newsletters der BaFin werden Sie über das Erscheinen einer neuen Ausgabe per E-Mail informiert.

Den BaFin-Newsletter finden Sie unter: www.bafin.de/Newsletter.

Disclaimer

Bitte beachten Sie, dass alle Angaben sorgfältig zusammengestellt worden sind, jedoch eine Haftung der BaFin für die Vollständigkeit und Richtigkeit der Angaben ausgeschlossen ist.

Ausschließlich zum Zweck der besseren Lesbarkeit wird im BaFin Journal auf die geschlechtsspezifische Schreibweise verzichtet. Alle personenbezogenen Bezeichnungen sind somit geschlechtsneutral zu verstehen.

* Der nichtamtliche Teil des BaFin Journals unterliegt dem Urheberrecht. Nachdruck und Verbreitung sind nur mit schriftlicher Zustimmung der BaFin – auch per E-Mail – gestattet.