

ANITA GIERBL

MARCO SCHREYER

PETER LEIBFRIED

DAMIAN BORTH

KÜNSTLICHE INTELLIGENZ IN DER PRÜFUNGSPRAXIS

Eine Bestandsaufnahme aktueller Einsatzmöglichkeiten und Herausforderungen

Während künstliche Intelligenz die Arbeitsweise verschiedener Berufsgruppen zunehmend und nachhaltig verändert, steckt ein solcher Wandel im Bereich der Wirtschaftsprüfung derzeit in seinen Anfängen. Der nachfolgende Beitrag soll Einsatzmöglichkeiten und Herausforderungen des maschinellen Lernens (ML), eines Teilgebiets der künstlichen Intelligenz, im Kontext der Abschlussprüfung aufzeigen.

1. TECHNOLOGISCHER WANDEL UND PRÜFUNGSPRAXIS

Bedingt durch die technologischen Entwicklungen in den letzten Jahrzehnten steht der Wirtschaft eine Vielzahl neuer Möglichkeiten zur Verfügung, um Wertschöpfungsprozesse effizienter und effektiver zu gestalten. Technologische Innovationen wie Industrie 4.0, Internet of Things (IoT) oder Robotic Process Automation (RPA) spiegeln sich in einer Zunahme der innerhalb von IT-Systemen erfassten Datenvolumina wider. Zeitgleich schaffen datenintensive Systeme (z. B. Enterprise Resource Planning [ERP]) die Voraussetzungen, um Geschäfts- und Finanzprozesse komplett oder teilweise zu digitalisieren. Eine Option, von welcher viele Unternehmen im Rahmen ihrer Digitalisierungsbestrebungen derzeit Gebrauch machen [1]. Dieser Wandel beeinflusst zunehmend auch die Prüfungspraxis im Hinblick auf die Erfassung prüfungsrelevanter Informationen sowie die Aufbereitung von Jahresabschlüssen.

In einem solchen Umfeld ist die Wirtschaftsprüfung zunehmend herausgefordert, weiterhin eine ausreichende Sicherheit der Prüfung (ISA 200) zu gewährleisten. Um diese Herausforderung zu bewältigen, müssen Prüfgesellschaften sich zunehmend selbst «digitalisieren», d. h. sich in gleichem Tempo wie die zu prüfenden Unternehmen entwickeln [2]. Durch neuartige Prüfsoftware und neue Prüfroutinen vollzieht sich, subsumiert unter dem Begriff

Audit 4.0, eine zunehmende Digitalisierung der Abschlussprüfung [3].

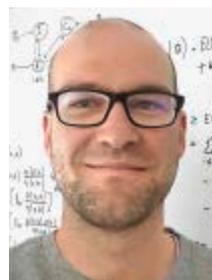
Im Kontext dieses Wandels forcieren Wirtschaftsprüfer zunehmend auch den Aufbau von Fähigkeiten im Bereich der künstlichen Intelligenz (KI bzw. Artificial Intelligence [AI]). Die Anwendung von KI verspricht, die für die Durchführung einer Prüfung notwendigen Ressourcen zu reduzieren (Effizienz), gleichzeitig ein höheres Mass an Prüfungssicherheit zu erreichen (Effektivität) und folglich auch die Kosten einer Prüfung zu senken [4].

2. KÜNSTLICHE INTELLIGENZ UND MASCHINELLES LERNEN

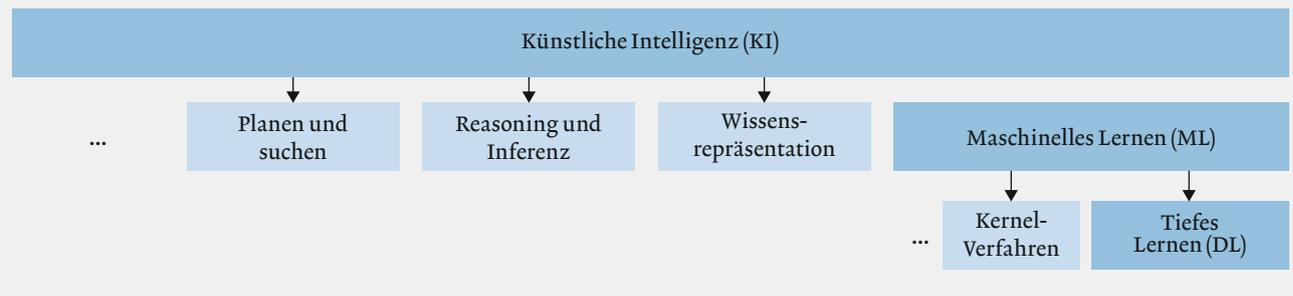
Die künstliche Intelligenz ermöglicht Unternehmen die Entwicklung neuer Applikationen, die sich grundlegend von traditionellen Softwarelösungen unterscheiden. Das Paradigma der Entwicklung von KI-Systemen unterscheidet sich von der klassischen Softwareentwicklung und lautet: «Lernen statt programmieren» [5]. Als Teildisziplin der künstlichen Intelligenz (vgl. *Abbildung 1*) beschreibt das maschinelle Lernen ein automatisiertes Verfahren der kontinuierlichen Modellbildung [6]. Der Aspekt des Lernens bezieht sich hierbei auf die Tatsache, dass sich die Modellqualität im Laufe des Lernprozesses kontinuierlich verbessern kann. Das Lernen selbst erfolgt über einen iterativen Trainingsprozess, in welchem einem Modell kontinuierlich Daten zu-



ANITA GIERBL,
M.A. (HSG),
DOKTORANDIN
UNIVERSITÄT ST. GALLEN,
FACHASSISTENZ
SWISS GAAP FER,
AUDIT PWC



MARCO SCHREYER, DIPL.
WIRT.-INF., WISSENSCHAFT-
LICHER MITARBEITER AM
LEHRSTUHL FÜR KÜNST-
LICHE INTELLIGENZ UND
MASCHINELLES LERNEN,
INSTITUTE OF COMPUTER
SCIENCE, UNIVERSITÄT
ST. GALLEN (HSG)

Abbildung 1: **BEISPIELHAFTE TAXONOMIE KÜNSTLICHE INTELLIGENZ** [8]

geführt werden, um neue und zugleich robuste Regelmäßigkeiten bzw. «Muster» in umfangreichen Datenmengen zu erlernen [7].

Ein solcher iterativer Trainingsprozess kann anhand unterschiedlicher Lernverfahren vollzogen werden, die im Allgemeinen in (1) überwachtes und (2) unüberwachtes Lernen klassifiziert werden:

Das «überwachte Lernen» (supervised learning) beschreibt ein Lernverfahren, in welchem dem Algorithmus im Rahmen des Trainingsprozesses das gewünschte Zielverhalten immer wieder aufgezeigt wird (z. B. dem Erkennen von Gegenständen in Bildern). Die Qualität des erlernten Modells kann hierbei jederzeit anhand entsprechender Kennzeichnungen (Labels) der Daten überwacht, d. h. mit dem gewünschten oder richtigen Lösungsverhalten verglichen werden. Solche Lernverfahren finden oftmals in Anwendungsbeispielen mit «prädiktivem» Charakter ihre Anwendung.

Das «unüberwachte Lernen» (unsupervised learning) beschreibt hingegen ein Lernverfahren, bei dem es dem angewandten Algorithmus selbst obliegt, mehrdimensionale Regelmäßigkeiten aus den Daten zu lernen. Im Vordergrund steht hierbei, die erlernten Regelmäßigkeiten von den strukturalosen Unregelmäßigkeiten («Rauschen») zu unterscheiden, z. B. die selbstständige Gruppierung von Gegenständen anhand ihrer Form und Grösse. Eine explizite Vorgabe des gewünschten Zielverhaltens durch einen menschlichen Experten erfolgt im Rahmen des Lernprozesses nicht. Unüberwachte Lernverfahren finden deshalb oftmals in Szenarien mit «detektivem» Charakter Anwendung.

Eine junge und überaus erfolgreiche Teildisziplin des maschinellen Lernens ist das «tiefe Lernen» (Deep Learning [DL]). Das tiefe Lernen bezeichnet den Einsatz selbstlernender, künstlicher neuronaler Netze. Hierbei handelt es sich

um ein biologisch inspiriertes Verfahren, dessen Funktionsweise dem Aufbau des menschlichen Gehirns nachempfunden ist [9]. Analog zu den Nervenzellen innerhalb des Kortex bestehen solche Netze aus einer Vielzahl von Schichten miteinander verknüpfter künstlicher Neuronen. Das Attribut «tief» bezieht sich in diesem Kontext auf die hohe Zahl von bis zu mehreren hundert Schichten künstlicher Neuronen. Die Vielzahl an Schichten und Neuronen ermöglicht es tiefen Lernverfahren, auf Grundlage einer Vielzahl von Trainingsbeispielen komplexe nichtlineare Muster zu erlernen, um anschliessend diese Zusammenhänge in den Daten automatisch wiederzuerkennen.

Die Anwendung tiefer Lernverfahren führt derzeit zu bahnbrechenden Fortschritten in Bereichen der Bild- und Textanalyse. Viele Ideen des tiefen Lernens sind jedoch nicht neu und lassen sich auf bereits in den 1950er-Jahren gemachte Entdeckungen zurückführen [10]. Die Renaissance künstlicher neuronaler Netze wird aktuell durch die Verfügbarkeit umfangreicher und auswertbarer Datenbestände, die Weiterentwicklung künstlicher neuronaler Netzarchitekturen und den einfachen Zugang zu leistungsstarken Prozessoren begünstigt.

3. EINSATZ MASCHINELLER LERNVERFAHREN IN DER PRÜFUNGSPRAXIS

Für die Prüfungspraxis kann konstatiert werden, dass sich die Anwendung maschineller Lernverfahren aktuell noch in einer Forschungs- und Entwicklungsphase befinden [11]. Die Potenziale der Anwendung sind jedoch in vielen Bereichen der ordentlichen Abschlussprüfung bereits heute deutlich erkennbar [12]. Im Allgemeinen kann der Prozess der Abschlussprüfung in die drei Hauptphasen Prüfungsplanung, Prüfungsdurchführung und Prüfungsabschluss sowie in jeweils assoziierte Teilphasen unterteilt wer-

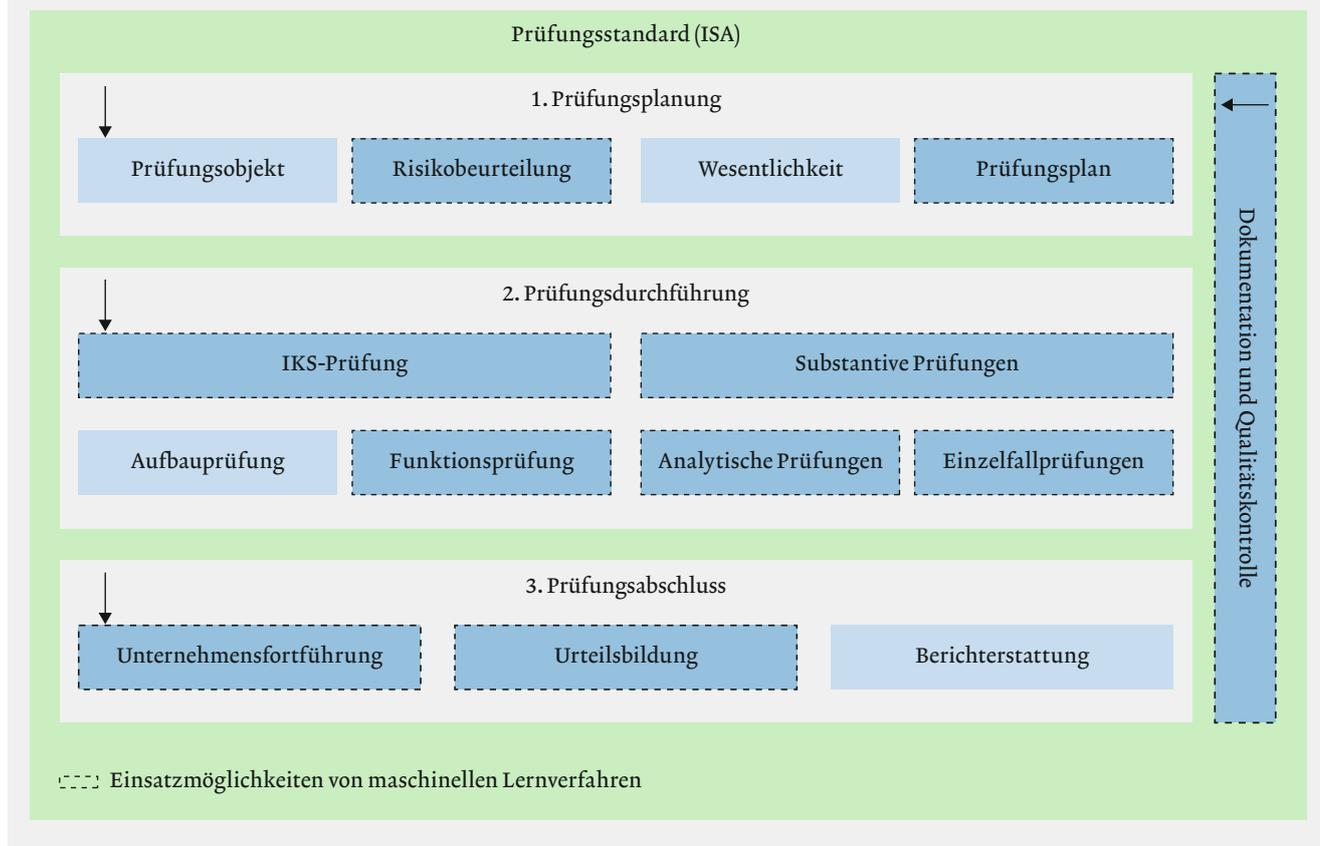


PETER LEIBFRIED,
PROF. DR., PRÄSIDENT
DER FACHKOMMISSION
SWISS GAAP FER,
INHABER DES KPMG-
LEHRSTUHL FÜR AUDIT
UND ACCOUNTING,
UNIVERSITÄT ST. GALLEN



DAMIAN BORTH,
PROF. DR., INHABER DES
LEHRSTUHL FÜR
KÜNSTLICHE INTELLIGENZ
UND MASCHINELLES
LERNEN, DIREKTOR AM
INSTITUTE OF COMPUTER
SCIENCE (ICS),
UNIVERSITÄT ST. GALLEN

Abbildung 2: **EINSATZMÖGLICHKEITEN VON MASCHINELLEN LERNVERFAHREN IM PRÜFUNGSPROZESS** [14]



den [13]. Begleitet wird der Prüfungsprozess durch eine fortlaufende Dokumentation und Qualitätskontrolle (vgl. *Abbildung 2*).

Die nachfolgenden Ausführungen bieten einen Überblick über erste ML-basierte Entwicklungen und deren Einsatz in der Prüfungspraxis und orientieren sich an den verschiedenen Phasen des Prüfprozesses.

3.1 Prüfungsplanung. Ziel der Prüfungsplanung ist es, ein Prüfprogramm zu erstellen, welches ein Prüfungsurteil mit hinreichender Sicherheit (Effektivität) und mit vertretbarem Aufwand (Wirtschaftlichkeit und Effizienz) bis zum vereinbarten Termin ermöglicht [15].

3.1.1 Risikobeurteilung. Maschinelle Lernverfahren eignen sich im Besonderen zur initialen Risikobeurteilung im Hinblick auf betrügerische Handlungen (Fraud) (ISA 240) und wesentliche Falschdarstellungen (Anomalien) (ISA 315) im Kontext des sogenannten Journal Entry Testing [16]. Dies überrascht nicht, wenn man bedenkt, dass es sich heutzutage bei Buchungssätzen um strukturierte elektronische Datensätze handelt, welche unmittelbar durch Algorithmen verarbeitet werden können. Sowohl überwachte als auch unüberwachte Lernverfahren können angewandt werden, um digitale Buchungssätze zu analysieren, z. B. um Regelmäßigkeiten in Buchungen zu erkennen und Buchungssätze semantisch zu gruppieren [17]. Hierdurch werden mehr-

dimensionale Buchungssystematiken erkannt, die für den menschlichen Prüfer nur schwer ersichtlich sind.

Die gewonnenen Erkenntnisse können anhand bekannter Darstellungsformen visuell aufbereitet werden [18]. Dem Prüfer wird es hierdurch erleichtert, Schlussfolgerungen für die Prüfung abzuleiten. Erste Lösungen in diesem Bereich werden heute bereits in der Prüfungspraxis eingesetzt, so z. B. Mindbridge Ai, PwC GL.ai und EY Helix GL. Die genannten Verfahren haben gemeinsam, dass sie Analyseergebnisse in Dashboards visualisieren und mittels Drill-down-Funktionen einen vertiefenden Einblick in die erkannten Buchungssystematiken und -abweichungen ermöglichen.

3.1.2 Prüfungsplan. Im Rahmen der Prüfungsplanung sind Wirtschaftsprüfer angehalten, Art, Umfang, Zeitpunkt und Personaleinsatz der Prüfung zu planen (ISA 300). Bedingt durch den Umstand, dass Prüfgesellschaften mehrere Mandate mit zeitlicher Überschneidung prüfen, resultiert ein erhöhter Planungs- und Abstimmungsaufwand. Dies gilt insbesondere für die Verfügbarkeit von fachlichen und personellen Ressourcen. Bereits in dieser frühen Phase der Prüfung ist es deshalb denkbar, maschinelle Lernverfahren einzusetzen, um Prüfgesellschaften bei der effektiven Ressourcenauslastung zu unterstützen.

Auch in diesem Bereich sind erste Praxislösungen ersichtlich, wie z. B. der durch Aspaara Algorithmic Solutions [19] entwickelte MatchingCore. Dabei wird anhand maschineller

Lernverfahren die Einsatzplanung der Mitarbeiter entsprechend den verschiedenen Prüfungsmandaten optimiert. Im Rahmen des Allokationsverfahrens werden Mitarbeiter unter Berücksichtigung von Urlaub, persönlichen Präferenzen, Arbeitspensum sowie fachlicher Kompetenz (vgl. ISQC 1.31) den vorhandenen Prüfungsmandaten zugeordnet. Ziel der Zuordnung ist es, Reisezeiten zu minimieren, wodurch Kosten reduziert werden sowie die Mitarbeiterzufriedenheit zu erhöhen. Maschinelle Lernverfahren können zudem (je nach den Bedürfnissen der Prüfgesellschaft) auf Präferenzen (z. B. konstantes Prüfteam für den Mandanten) Rücksicht nehmen.

3.2 Prüfungsdurchführung. Nach erfolgter Prüfungsplanung erfolgt in einem Folgeschritt die eigentliche Ausführung der Prüfungshandlungen. Nachfolgend werden die Einsatzmöglichkeiten maschineller Lernverfahren für die verschiedenen Prüfverfahren und -techniken erläutert.

3.2.1 Funktionsprüfung. Ein Anwendungsgebiet maschineller Lernverfahren findet sich in der Funktionsprüfung interner Kontrollen. Im Zuge dieser Prüfungshandlung werden Belege und Dokumente gesichtet, um die Wirksamkeit der internen Kontrollen (IKS) zu validieren. Unter der Voraussetzung, dass das jeweilige physische Dokument digitalisiert wurde, z. B. durch eine automatisierte Texterkennung (engl. Optical Character Recognition [OCR]), können prüfungsrelevante Informationen extrahiert und automatisiert ausgewertet werden. In diesem Kontext lassen sich maschinelle Lernverfahren zunehmend auch einsetzen, um beispielsweise die Manipulation physischer Rechnungen automatisiert zu erkennen [20].

In gewissen Prüffeldern (z. B. Materialwirtschaft, Forderungen) wird der Prüfer aufgrund der Vielzahl von Buchungen auf die Prüfung einer repräsentativen Stichprobe zurückgreifen (ISA 530). Auch dabei können unüberwachte maschinelle Lernverfahren angewandt werden, um die repräsentative Stichprobenauswahl zu unterstützen und so die Prüfungssicherheit zu erhöhen.

Maschinelles Lernen findet sich auch im Bereich Process Mining. Process Mining ermöglicht durch die Visualisierung sog. «Event Logs», den tatsächlichen Prozessdurchlauf eines Unternehmens darzustellen. Der Ist-Prozess kann dabei dem Soll-Prozess gegenübergestellt werden, wodurch IKS-Verstöße schneller erkennbar und transparenter sind. In einer Feldstudie testeten Jans, Alles, Vasarhelyi (2014) den Einkaufsprozess eines Unternehmens mittels Process Mining. Dabei wurden Zahlungen ohne Genehmigung, Verstöße gegen die Funktionstrennung und Abweichungen von Soll-Prozessen aufgedeckt. Bemerkenswert ist, dass keine dieser Abweichungen zuvor von der internen Revision des Unternehmens aufgedeckt wurde.

3.2.2 Analytische Prüfungshandlungen. Auch im Bereich der analytischen Prüfungshandlungen finden zunehmend maschinelle Lernverfahren Anwendung. In diesem Kontext können überwachte Lernverfahren eingesetzt werden, um prädiktive Erwartungswerte für die Plausibilisierung von Unterneh-

menskennzahlen oder Trendanalysen vorherzusagen (vgl. KPMG Clara)[21]. In den Vereinigten Staaten finden sich zudem für ausgewählte unternehmerische Kennzahlen erste öffentlich zugängliche Prognosen, z. B. über Case Ware AI [22]. Ob der Wirtschaftsprüfer KI-Prognosen im Rahmen von analytischen Prüfungshandlungen berücksichtigt, liegt in seinem Ermessen bzw. professionellen Urteilsvermögen.

3.2.3 Einzelfallprüfungen. Im Bereich der Einzelfallprüfung kann maschinelles Lernen auch die Prüfung von Dokumenten und Belegen sowie die Stichprobenauswahl unterstützen. Diese Anwendungsbeispiele wurden bereits im Absatz «Funktionsprüfung» vorgestellt.

Bei der erstgenannten Prüfungshandlung kann maschinelles Lernen für die Auswertung von Verträgen, Rechnungen, Bankbestätigungen usw. eingesetzt werden [23]. Mithilfe von Natural Language Processing kann die Semantik der Verträge erkannt und aus den Verträgen extrahiert werden. In gewissen Bereichen würde diese Einsatzmöglichkeit dem Prüfer vor allem erlauben, eine Vollprüfung statt einer Stichprobenprüfung durchzuführen, um hierdurch eine erhöhte Prüfungssicherheit zu erlangen.

3.3 Dokumentation. Während der gesamten Prüfung dokumentiert der Prüfer sein Vorgehen sowie die erlangten Nachweise. In Bezug auf die bereits vorgestellten Einsatzmöglichkeiten im Bereich der Prüfung von Belegen kann maschinelles Lernen auch die Dokumentation unterstützen. Durch die Möglichkeit, semantische Informationen zu extrahieren, kann in Kombination mit Verfahren der RPA die Dokumentation der prüferischen Erkenntnisse unmittelbar innerhalb des Prüfprogramms erfolgen. Ein solcher Effizienzgewinn erlaubt es dem Prüfer, mehr Zeit für die Analyse der gewonnenen Informationen aufzuwenden. Erste Lösungsansätze finden sich auch in diesem Bereich in der Praxis, wie z. B. Deloitte Argus Software.

3.4 Prüfungsabschluss. Wie in *Abbildung 2* dargestellt, umfasst der Prüfungsabschluss die Beurteilung der Unterneh-

mensfortführung, die Bildung des Prüfurteils sowie die finale Berichterstattung.

In der Forschung finden sich erste Ansätze, um Fragen zur Unternehmensfortführung und prüferischen Urteilsbildung mittels maschineller Lernverfahren zu beantworten [24]. Dabei werden Modelle mit historischen Unternehmenskennzahlen trainiert. Anhand der Trainingsdaten lernt das Verfahren die jeweiligen Datenausprägungen bzw. Muster, welche eine Unternehmensfortführung bzw. Insolvenz determinieren. Nach erfolgreichem Abschluss der Trainingsphase kann das erlernte Modell verwendet werden, um Wahrscheinlichkeiten für eine Insolvenz von «neuen» bzw. unbekanntem Unternehmen zu prognostizieren.

Zusammenfassend lassen sich mögliche Anwendungen für maschinelle Lernverfahren in nahezu allen Phasen der Prüfung erkennen. Es zeigt sich, dass maschinelle Lernverfahren das Potenzial bergen, sich zu einem ergänzenden Werkzeug des Prüfers in der Prüfungspraxis zu entwickeln. Um dieses Potenzial jedoch vollumfänglich ausschöpfen zu können, ist eine Reihe von Herausforderungen zu bewältigen, die im nächsten Abschnitt skizziert werden.

4. HERAUSFORDERUNGEN FÜR DIE PRÜFUNGSPRAXIS

Die prüferische Datenanalyse birgt (z. B. aufgrund des Journal Entry Testing) bereits bekannte Herausforderungen, welche u. a. die Datenextraktion, die Datenaufbereitung und den Datenschutz betreffen [25]. Die Prüfungsgesellschaften sehen sich infolge des Einsatzes maschineller Lernverfahren indes aktuell mit einer Reihe zusätzlicher Herausforderungen konfrontiert. Die hiermit verbundenen Fragestellungen sind neuartig und verlangen den Prüfungsgesellschaften deshalb oftmals ab, selbst einen empirischen Erfahrungsschatz im prüferischen Kontext zu sammeln [26]. Die nachfolgenden exemplarischen Aufzählungen sollen einen Impuls zur Diskussion darstellen.

4.1 Aktuelle Herausforderungen

4.1.1 Fachwissen. Fraglich ist derzeit, durch welche Mitglieder des jeweiligen Prüfungsteams maschinelle Lernverfahren angewandt werden. Für einen Nutzen stiftenden Einsatz bedarf es im Idealfall der Kombination von ausgereiften prüferischen und technischen Fachkenntnissen. Die Gewinnung von Mitarbeitern mit einem solchen Schnittstellenwissen gestaltet sich aktuell jedoch schwierig [27]. Eine kurzfristige Lösung bietet die Einbindung technischer Experten unter Anwendung von ISA 620. Diese Experten finden sich in Berufsgruppen wie Data Scientists, Data Engineers und Data Analysts, welche aktiv in die Prüfungsplanung und -durchführung miteinbezogen werden können. Mittelfristig werden Prüfungsgesellschaften jedoch zunehmend selbst in die Ausbildung künftiger Prüfergenerationen investieren müssen, um Expertise an der Schnittstelle von IT-technischem und prüferischem Wissen aufzubauen.

4.1.2 Infrastruktur. Der Einsatz maschineller Lernverfahren stellt zudem neue Anforderungen an die Verfügbarkeit und Leistungsfähigkeit der im Rahmen einer Prüfung verwendeten

technischen Infrastruktur. Das Training tiefer Lernverfahren erfolgt heutzutage auf speziell hierfür entwickelten leistungsstarken Prozessoren. Die Notwendigkeit des Einsatzes solcher Prozessoren ist auf den Umstand zurückzuführen, dass tiefe Lernmodelle nicht selten mehrere Millionen Parameter umfassen. Erst durch das Training derart umfangreicher Modelle lassen sich komplexe Zusammenhänge in hochdimensionalen Daten in angemessener Zeit erlernen bzw. extrahieren (z. B. aus Buchungen der SAP-Finanzbuchhaltung [28]). Somit wird es zunehmend an Bedeutung gewinnen, dass Prüfgesellschaften in entsprechende technische Infrastrukturen investieren. Diese technischen Werkzeuge ermöglichen es Prüfern, auch zukünftig die angestrebten Potenziale zu erschliessen.

4.1.3 Trainingsdaten. Im Kontext der Anwendung maschineller Lernverfahren gilt es zudem, ausreichend Trainingsdaten in hinreichender Qualität zu akquirieren, da diese unmittelbar die Güte des erlernten Modells und der ableitbaren Ergebnisse determinieren. Dies gilt insbesondere für die Anwendung überwachter Lernverfahren, da solche Algorithmen basierend auf den durch die Experten vorgegebenen «Kennzeichnungen» lernen. Liegen diese Kennzeichnungen nicht für jedes mögliche Resultat in ausreichender Anzahl vor, können Klassifikationsmodelle nicht bzw. nur mit Einschränkungen erlernt werden.

4.2 Zukünftige Herausforderungen

4.2.1 Rechtssicherheit. Die rechtlichen Herausforderungen umfassen u. a. den Schutz des Urheberrechts von erlernten ML-Modellen. Da diese Modelle oftmals auf Grundlage von Daten der geprüften Unternehmen erlernt werden, ist es aktuell unklar, welche Partei an einem erlernten Modell einen Rechtsanspruch geltend machen kann [29]. Zudem können sich durch das Inverkehrbringen von ML-Modellen im Kontext der Prüfung neuartige Fragestellungen der Verantwortlichkeit und Haftung ergeben, bspw. durch fehlerhafte Entscheidungen eines ML-Modells oder die missbräuchliche Modellanwendung [30].

4.2.2 Compliance. Die Compliance-Herausforderungen betreffen die Erarbeitung und Anwendung geeigneter Richtlinien für die Risikobewertung, das Training sowie das Testen von ML-Modellen für Prüfungszwecke. Aufgrund des iterativen Trainings können sich die durch die Modelle getroffenen Entscheidungen im Zeit- bzw. Trainingsverlauf verändern [31]. Auch wird das Risikomanagement von ML-Modellen Kriterien und Prozesse umfassen müssen, die es ermöglichen, einen gegebenenfalls vorhandenen Bias [32] in den Trainingsdaten zu erkennen und diesen entsprechend zu korrigieren.

4.2.3 IT-Sicherheit. Eine weitere Herausforderung betrifft die Sicherheit von ML-Modellen. Durch den Einsatz von maschinellen Lernverfahren im Kontext der Prüfung werden sich zukünftig neue Möglichkeiten des wiederum ML-gestützten Angriffs auf ML-Modelle, sogenannter Adversarial Attacks [33], ergeben (z. B. um Buchungen zu generieren, die betrugsbehaftete Buchungen gezielt verschleiern [34]). Es

bedarf deshalb eines definierten Sicherheitsprozesses, durch welchen Risiken antizipiert und minimiert werden können. Von zunehmender Bedeutung wird auch die Sensibilisierung der Prüfer sein, um auf diese neue Art von Bedrohung und etwaige Manipulationen angemessen reagieren zu können.

Die kontinuierliche Bewältigung der genannten Herausforderungen und damit verbundenen Fragestellungen werden dazu beitragen, den Einsatz maschineller Lernverfahren in der prüferischen Praxis weiter zu etablieren. Es wird sich zeigen, inwieweit sich hierdurch die Versprechen eines effektiven und effizienten Audit 4.0 umsetzen lassen.

5. FAZIT UND AUSBLICK

Im Bereich der Abschlussprüfung befindet sich die Prüfungspraxis in einer Phase der Transformation von der IT-gestützten Prüfung, welche gekennzeichnet ist durch den Einsatz von Prüfsoftware zur zielgerichteten Datenanalyse, hin zu einer zunehmend durch maschinelle Lernverfahren unter-

stützten Prüfung [35]. Abschliessend bleibt jedoch anzumerken, dass der vollständig autark prüfende KI-Wirtschaftsprüfer noch nicht in greifbare Nähe gerückt ist. Maschinelle Lernverfahren entwickeln sich zunehmend zu einem unterstützenden und wertvollen Werkzeug, welches in den verschiedenen Phasen des Prüfprozesses eingesetzt werden kann. Parallel hierzu werden für digitalaffine Prüferinnen und Prüfer die Hürden für die Anwendung solcher Verfahren zunehmend kleiner.

Vor mehr als 30 Jahren stellte Microsoft die Anwendung von Excel für die tabellarische Datenverarbeitung vor. Eine Anwendung, welche sich heutzutage nur noch schwer aus dem Arbeitsalltag der Prüfung wegdenken lässt. Vielleicht lässt sich durch maschinelle Lernverfahren die nächste «Lingua Franca» in der Prüfungspraxis etablieren. Die Anwendungspotentiale hierfür sind in ihren Grundzügen deutlich erkennbar und werden heute bereits punktuell genutzt. ■

- Fussnoten:** 1) Vgl. Klauser, Herzog, 2017. 2) Vgl. Rausenberger, Prenrecaj, 2017. 3) Vgl. Dai, Vasarhelyi, 2016. 4) Vgl. Eberle, 2017. 5) Vgl. Karpathy, 2017. 6) Vgl. Bishop, 2006. 7) Vgl. Duda, Hart, Stork, 2012. 8) Vgl. Borth, 2020. 9) Vgl. LeCun, Bengio, Hinton, 2015. 10) Vgl. Rosenblatt, 1958. 11) Vgl. Dickey, Blanke, Seaton, 2019. 12) Vgl. Issa, Ting, Vasarhelyi, 2016. 13) Vgl. IAASB, 2019; Expertsuisse, 2015. 14) Vgl. Expertsuisse, 2015. 15) Vgl. Expertsuisse, 2015. 16) Journal Entry Testing per se wird erst im Zuge der Prüfungsdurchführung abgewickelt. Bei den genannten ISA-Standards handelt es sich jedoch um Standards, welche eher dem Bereich der Risikobeurteilung zuzuordnen sind als der Prüfungsdurchführung. Deswegen haben sich die Autoren entschieden, diese Einsatzmöglichkeit in diesem Prüfungsabschnitt einzuordnen. 17) Vgl. Schreyer, Sattarov, Schulze et al., 2019a; Lahann, Scheid, Fettke, 2019; Westermann, Spindler, 2017. 18) Vgl. Tatu, Schreyer, Hagelauer et al., 2014. 19) www.aspaara.com. 20) Vgl. van Beusekom, Shafait, Breuel, 2013; van Beusekom, Schreyer, Breuel, 2010. 21) Vgl. Koskivaara, Back, 2007. 22) https://ai.caseware.com/. 23) Vgl. Lin, Hazelbaker, 2019. 24) Vgl. Jiang, Jones, 2018; Kirkos, Spathis, Nanopoulos et al., 2007. 25) Vgl. Bönner, Riedl, Wenig, 2011; Klauser, Herzog, 2019. 26) Vgl. van Giffen, Borth, Brenner, 2020. 27) Vgl. Salijeni, Samsonova-Taddei, Turley, 2019. 28) Vgl. Schreyer, Sattarov, Borth et al., 2018. 29) Vgl. SBFi, 2019. 30) Vgl. EU-Kommission, 2020. 31) Vgl. Unterthiner, Keyzers, Gelly, 2020. 32) Vgl. Barocas, Hardt, Narayanan, 2019. 33) Vgl. Goodfellow, Shlens, Szegedy, 2017. 34) Vgl. Schreyer, Sattarov, Reimer et al., 2019b. 35) Vgl. Eberle, 2017.
- learning, in: Nature, 2015, 521, S. 436–444. ▶ Lin, P., Hazelbaker, T., Meeting the Challenge of Artificial Intelligence: What CPAs Need to Know, in: The CPA Journal, 2019, 89, S. 48–52. ▶ Rausenberger, R., Prenrecaj, K., Audit 4.0 – Digitale Wirtschaftsprüfung, Der Einsatz innovativer Technologien verändert Abschlussprüfung und Prüfer, in: Expert Focus, 2017/11, S. 779–783. ▶ Rosenblatt, F., The perceptron: a probabilistic model for information storage and organization in the brain, in: Psychological review, 1958, Vol. 65, Nr. 6, S. 386–408. ▶ Salijeni, G., Samsonova-Taddei, A., Turley, S., Big data and changes in audit technology: Contemplating a research agenda, in: Accounting Business Research, 2019, 49, S. 95–119. ▶ Schreyer, M., Sattarov, T., Schulze, C., Reimer, B., Borth, D., (2019a), Detection of accounting anomalies in the latent space using adversarial autoencoder neural networks, https://arxiv.org/pdf/1908.00734.pdf, abgerufen am 29. Juni 2020. ▶ Schreyer, M., Sattarov, T., Reimer, B., Borth, D., (2019b), Adversarial learning of deepfakes in accounting, https://www.alexandria.unisg.ch/258090/1/NeurIPS_2019_RAIFS_final.pdf, abgerufen am 29. Juni 2020. ▶ Schreyer, M., Sattarov, T., Borth, D., Dengel, A., Reimer, B., Künstliche Intelligenz in der Wirtschaftsprüfung – Identifikation ungewöhnlicher Buchungen in der Finanzbuchhaltung, WPg, 2018, 11, S. 674–681. ▶ Staatssekretariat für Bildung, Forschung und Innovation (SBFi), Herausforderungen der künstlichen Intelligenz, 2019, https://www.sbf.admin.ch/sbf/de/home/bfi-politik/bfi-2021-2024/transversale-themen/digitalisierung-bfi/kuenstliche-intelligenz.html, abgerufen am 7. Juli 2020. ▶ Unterthiner, T., Keyzers, D., Gelly, S., Bousquet, O., Tolstikhin, I., Predicting Neural Network Accuracy from Weights, https://arxiv.org/abs/2002.11448, abgerufen am 7. Juli 2020. ▶ Van Beusekom, J., Shafait, F., Breuel, T., Text-line examination for document forgery detection, in: IJDAR, 2013, Vol. 16, No. 2, S. 189–207. ▶ Van Beusekom, J., Schreyer, M., Breuel, T., Automatic counterfeiting protection system code classification, in: Media Forensics and Security II, 2010, Vol. 7541, S. 75410–75418. ▶ Van Giffen, B., Borth, D., Brenner, W., Management von Künstlicher Intelligenz in Unternehmen, in: HMD Praxis der Wirtschaftsinformatik, 2020, 57, S. 4–20. ▶ Westermann, C., Spindler, C., Anomalie-Erkennung in der Wirtschaftsprüfung mithilfe von Machine Learning, Vorstellung einer technologiebasierten Methode mittels automatisierter Lernverfahren, Expert Focus, 2017/11, S. 791–797.
- Vol. 13, S. 1–15. ▶ Dickey, G., Blanke, S., Seaton, L., Machine learning in auditing: Current and future applications, https://www.cpajournal.com/2019/06/19/machine-learning-in-auditing/, abgerufen am 29. Juni 2020. ▶ Duda, R., Hart, P., Stork, D., Pattern Classification, New Jersey, 2012. ▶ Eberle, R., Die Zukunft der Wirtschaftsprüfung, Droht eine Wirtschaftsprüfung der unterschiedlichen Geschwindigkeiten?, in: Expert Focus, 2017/11, S. 751–755. ▶ EU-Kommission, Weissbuch, Zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen, https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_de.pdf, abgerufen am 7. Juli 2020. ▶ EXPERTSuisse, Schweizer Handbuch der Wirtschaftsprüfung, Zürich, 2015. ▶ Goodfellow, I., Shlens, J., Szegedy, C., Explaining and Harnessing Adversarial Examples, https://arxiv.org/abs/1412.6572, abgerufen am 7. Juli 2020. ▶ Issa, H., Ting, S., Vasarhelyi, M., Research ideas for artificial intelligence in auditing: The formalization of audit and workforce supplementation, in: Journal of Emerging Technologies in Accounting, 2016, 13, S. 1–20. ▶ Jans, M., Alles, M., Vasarhelyi, M., A field study on the use of process mining of event logs as an analytical procedure in auditing, in: The Accounting Review, 2014, 14, S. 1751–1773. ▶ Jiang, Y., Jones, S., Corporate distress prediction in China: a machine learning approach, in: Accounting and Finance, 2018, 58, S. 1063–1109. ▶ Karpathy, A., Software 2.0, https://medium.com/@karpathy/software-2-0-a64152b37c35, abgerufen am 9. Juni 2020. ▶ Kirkos, E., Spathis, C., Nanopoulos, A., Manolopoulos, Y., Identifying qualified auditors' opinions: A data mining approach, in: Journal of Emerging Technologies in Accounting, 2007, 4, S. 183–197. ▶ Klauser, M., Herzog, D., Digitalisierungsgrad der Treuhandbranche, Ergebnisse einer repräsentativen Umfrage bei Expertsuisse-Mitgliedunternehmen, in: Expert Focus, 2017/11, S. 832–837. ▶ Klauser, M., Herzog, D., Digitalisierungsstudie 2019, Die Branchentrends zeigen Handlungsbedarf, in: Expert Focus, 2019/12, S. 960–968. ▶ Koskivaara, E., Back, B., Artificial neural network assistant (ANNA) for continuous auditing and monitoring of financial data, in: Journal of Emerging Technologies in Accounting, 2007, 4, S. 29–45. ▶ Lahann, J., Scheid, M., Fettke, P., Utilizing Machine Learning Techniques to Reveal VAT Compliance Violations in Accounting Data, in: 2019 IEEE 21st Conference on Business Informatics (CBI), 2019, Vol. 1, S. 1–10. ▶ LeCun, Y., Bengio, Y., Hinton, G., Deep