

Köln, 29.09.2021

Stellungnahme der Deutschen Aktuarvereinigung zum Konsultationspapier *Maschinelles Lernen in Risikomodellen – Charakteristika und aufsichtliche Schwerpunkte*

Generelle DAV Kommentare und Argumente

Die DAV sieht AI & DS als relevante technologische Entwicklungen im Zusammenhang mit der Digitalisierung, die datengestützten Tätigkeiten insb. in Versicherungen innerhalb aktuarieller Prozesse unterstützt und weiterentwickelt. Die zugehörige Forschung und Entwicklung ist für die DAV ein wichtiges Instrument zur Aus- und Weiterbildung von Aktuaren, da neue Techniken und Verfahren heute und in Zukunft die Arbeitsgebiete von Aktuaren bestimmen.

Die DAV bietet interessierten Aktuaren die Ausbildung zum Certified Actuarial Data Scientist. Die Aus- und Weiterbildung der DAV mit ihren hohen fachlichen Standards trägt dazu bei, dass Aktuare AI & DS Methoden und Daten Grundlagen in der Modellierung adäquat einsetzen und den Output der Modelle angemessen bewerten. Dies war schon in der Vergangenheit eine Hauptaufgabe der aktuariellen Arbeit und wird sich mit diesen neuen Werkzeugen der AI & DS analog fortsetzen. Dort wo es nötig und angemessen ist, wird die DAV in der Zukunft - wie in anderen Fachgebieten auch - fachliche Standards (z.B. Hinweise oder Richtlinien) für die fachliche Arbeit von Aktuaren in AI & DS erlassen.

Insoweit AI & DS Methoden bei der Modellierung im Zusammenhang mit Solvency II eingesetzt werden, sollte auch die Versicherungsmathematische Funktion und / oder der Verantwortliche Aktuar / die Verantwortliche Aktuarin entsprechendes Methoden-Wissen aufweisen, um eine adäquate Beurteilung der Modellierung vornehmen zu können. Die DAV unterstützt mit entsprechender Weiterbildung: so steht der diesjährige Akademietag für VA's unter dem Motto: "Data Scientist - Die VA's von morgen".

Eine direkte Übertragbarkeit von Algorithmik, Bewertung sowie Regulierungsmethodik von Banken auf Versicherungen sieht die DAV nicht unkritisch. Hauptgrund sind die unterschiedlichen Perspektiven, Anwendungsfälle und Prozesse. Hier gibt es maximal bei speziellen Versicherungssparten (Beispiel Kreditversicherung), bei Zahlungsprozessen (Prämien, Schäden) sowie bei der Kapitalanlage gewisse Ähnlichkeiten zur Banken-Logik. Speziell die versicherungsmathematischen Modelle (z.B. innerhalb der Risikobewertung) sind häufig sehr andersartig in Algorithmik und Methodik als Bewertungsmodelle von Banken.

Typischerweise gibt es zwei Aspekte, die Versicherer von den Banken unterscheiden: einerseits die Frequenz / Automatisierung und zum zweiten der Risikotransfer. Versicherer sind nicht im Hochfrequenz-Handel tätig – ganz im Gegenteil sind Geschäftsbeziehungen von Versicherern auf beiden Seiten der Bilanz noch oft mindestens auf einen einjährigen Zeitrahmen ausgerichtet. Daher müssen hier nicht so hohe Anforderungen an die Automatisierung gestellt werden. Versicherer hebeln das Geschäft nicht durch den starken Risikotransfer an Dritte / externe Investoren. Mit dieser ungehebelten Geschäftsstrategie verhalten sich Risiken insgesamt linearer und kontrollierbarer.

BaFIN Fragen und DAV Antworten

Diskussionspapier, Kapitel II:

Auszug „Charakteristika von ML-Methoden“ – Tabelle:

AI/ML-Szenario	Charakteristika
<i>1 Methodik und Datengrundlage</i>	<i>Komplexität und Dimension des Hypothesenraums</i> <i>Komplexität des Trainings</i> <i>Adaptivität</i> <i>Datenquellen</i> <i>Datentypen</i> <i>Datenmenge</i>
<i>2 Nutzung des Outputs</i>	<i>Bedeutung im Modell</i> <i>Anwendungsbereich</i> <i>Grad der Automatisierung</i>
<i>3 Auslagerung und IT</i>	<i>Auslagerung</i> <i>IT-Infrastruktur</i>

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?

Antwort a) DAV:

- Regulierungsrelevante Kriterien bestehen aus Sicht der DAV darin, die strukturelle Eignung der verwendeten Methoden im Hinblick auf Performance, Stabilität und Reproduzierbarkeit zu betrachten, aber nur insofern als dass die gleichen Maßstäbe wie an bestehende/etablierte Methoden anzusetzen sind. Konzeptionell besteht hierbei ein Unterschied, ob die Herstellung von Erklärbarkeitsansätzen über Sampling- und Marginalisierungsbetrachtungen lediglich prohibitiv ineffizient aber grundsätzlich zweckmäßig sind (vgl. bspw. Ensemble-Methoden/Boosting etc.) oder aufgrund struktureller Eigenschaften prinzipiell nur schwache menschenlesbare Einblicke bieten kann (vgl. bspw. bestimmte tiefe KNN).
- Eine explizite bzw. strikte Definition von ML-Methoden/Gruppen könnte es ermöglichen, Erklärbarkeit mit Blick auf Diskriminierung zu gewährleisten. Dies sollte für die Aufsicht im Sinne der Solvenz- und Risikobetrachtung allerdings nachrangig sein, da eine mögliche Diskriminierung unabhängig von der verwendeten Methode zu betrachten ist und sich schon aus dem Input ergibt. Es ist in diesem Sinne ein Thema für Datenschutzrechtler (bei nachträglicher Profilbildung aus augenscheinlich nicht personenbezogenen Daten) und dem Gesetzgeber (insofern als dass weder die Aufsicht noch die Versicherer zu entscheiden haben, welche Merkmale als diskriminierend einzuordnen sind). Hierbei besteht ein offensichtliches Spannungsfeld zwischen VAG und AGG, das von BDAI unabhängig ist und deshalb nicht Gegenstand der Betrachtung sein kann.
- Ein wesentlicher Aspekt des Charakteristikums 1 ist die Unterscheidung von BDAI-basierten Datenquellen, die als (statische) Merkmale in "klassische" Modelle eingehen, also nur mittelbar der Definition folgen. In diesem Falle ist die Validierung der Inputs so zu betrachten wie für nicht-BDAI-Merkmale in klassischen Modellen auch, es ist also keine weitere Regelung notwendig.
- Charakteristik 2 hat pragmatisch höheren Einfluss, um für Charakteristik 1 die Validierungs-Tiefe/Notwendigkeit zu bestimmen. Insofern ist eine Trennung zwischen Input/Methode und Output nicht immer sinnvoll möglich, da spezifische Methoden auch spezifische Betrachtungen erfordern. Beispielhaft sei hier die Analyse von Residuen in generalisierten linearen Modellen genannt: Während dies dort mit vorliegender Modellannahme eine Betrachtung der (bspw. quantilsweisen oder quantilsortierten) Abweichungen erlaubt, ist bspw. Homoskedastizität bei den meisten verteilungsfreien ML-Methoden statistisch nicht definiert (oder gar nicht *definierbar*), sodass klassische Gütemaße nicht anwendbar sind. Auch Konfidenzintervalle sind für Prädiktoren ohne geschlossene funktionale Darstellung nicht trivial zu bestimmen. Während diese Gütemaße in der Praxis auch für ML-Methoden durchaus funktional sind, bleibt formal festzuhalten, dass es in diesem Falle

(also im Gegensatz zu bspw. GLMs) ratsam ist, nicht statistische Sicherheit mit heuristischer Glaubhaftigkeit zu verwechseln.

- Abgesehen von Fragen der IT-Sicherheit und des Risikomanagements im Allgemeinen (also der Governance) gibt es im Bereich der Charakteristik 3 (intern/extern) aus Sicht der DAV keinerlei Regelbedarf.

b) Welche weiteren Charakteristika von ML-Methoden können aus Ihrer Sicht für die Aufsichtspraxis oder auch die interne Modelle-Governance von Bedeutung sein?

Antwort b) DAV:

- Die Frequenz des Lernens und der Anpassung stellt zwar möglicherweise einen praktischen Paradigmenwechsel dar, ist aber aufsichtsrechtlich irrelevant: Wo eine feste Prüffrequenz oder Testatpflicht besteht, muss diese auch mit ML-Methoden sichergestellt werden. Wenn sich daraus ergibt, dass komplexe Modelle deshalb einen erhöhten Pflegebedarf entwickeln, ist die Abwägung ökonomischer, nicht aufsichtsrechtlicher Natur, ob und wie man damit umgeht.
- Ein relevanter Punkt ist die (teilweise) mangelnde Analytizität von verteilungsfreien ML-Methoden: Ohne geschlossene Darstellungsform und Verteilungsannahme können Vorhersagen nicht (programmatisch) mit Konfidenzen versehen werden, sondern allenfalls in spezifischen, begrenzten Anwendungsfällen. Dies führt dazu, dass in Bereichen, wo es um Stabilitäts- und Langzeitprognosen geht, die Anwendung grundsätzlich kritisch zu sehen ist, wenn für das verwendete Modell keine kuratierte und durch aktuarielle Fachkenntnisse validierbare Konfidenzbetrachtung vorliegt. (Anm.d. Verf.: Man spricht dabei in der Wissenschaft davon, dass (möglicherweise kritische) Rand- und Spezialfälle in sog. "non-linear pockets" des Zustandsraumes verborgen liegen können.)

c) Welche Charakteristika gehören aus Ihrer Sicht nicht in diese Übersicht?

Antwort c) DAV:

- Die Charakteristik 3 ist eine generelle Eigenschaft, die für ML Methoden keinen besonderen Einfluss haben sollte.
- Die Auslagerung von Prozessen ist ein generell mögliches Verfahren, welches schon heute stark durch regulatorische Anforderungen beeinflusst ist. Hierbei spielen Transparenz und Governance eine relevante Rolle, die unabhängig vom Einsatz von ML-Methodik bewertet werden muss. Dieses Thema wird deshalb in der weiteren Beantwortung ausgeklammert.

- Auch IT-Infrastruktur ist ein Kriterium, welches heute wichtig für verschiedenste Datenverarbeitungs-Prozesse von Versicherungsunternehmen ist. Dieses ist unabhängig von ML-Methodik relevant für die Prozesse und Verfahren, bei denen ML-Algorithmen bisherige Ansätze ergänzen oder ersetzen könnten.

d) In welchen relevanten Anwendungsbereichen kommen ML-Methoden bei Ihnen zum Einsatz bzw. wo planen Sie deren Umsetzung?

Antwort d) DAV:

- Als DAV sorgen wir für die Ausbildung unserer Mitglieder in AI und DS und legen fachliche Standards und Richtlinien für die aktuarielle Arbeit fest. Zum konkreten Einsatz von ML-Methoden in den Unternehmen oder möglichen Planungen kann die DAV keine Auskunft geben.

Diskussionspapier, 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.

Antwort e) DAV:

- Aktuell sehen wir keine Regelungen, welche den Einsatz von ML-Methoden grundsätzlich verhindern würden.
- Allerdings ist beispielsweise bei Stresstesting im Sinne von Solvency II bei der Nutzung von II bei der Nutzung von ML Vorsicht geboten, weil Marginal- und Sektionalbetrachtungen durch die weitgehende Verteilungsfreiheit keinen nachweislich monotonen Zustandsraum bilden, bzw. eine Überprüfung von Sicherheitsmerkmalen anhand bspw. Spezialfällen exponentiell aufwändig ist.

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?

Antwort f) DAV:

- Die EBA-Leitlinien EBA/GL/2020/06 richten sich an Banken und nicht an Versicherungen. Die DAV kann hierzu keine Aussage treffen.

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?

Antwort g) DAV:

- Validierungen von ML-Verfahren erfordern gewisse Standards, insbesondere bei nicht spezifisch erklärbaren Entscheidungen – aber indirekt schon in bisheriger Aufsichtspraxis enthalten – sowie indirekt auch in aktuariellen Richtlinien.
- Wir sehen das eher durch Ausbildung/Fachkenntnisse/Standard und Richtlinien abgebildet. Unseres Erachtens sind wir als Aktuare gut positioniert und adressieren das Thema durch unser neues Ausbildungsformat "Actuarial Data Science".

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?

Antwort h) DAV:

- Aus regulatorischer Perspektive existieren schon heute Anforderungen an IT, Implementierung und Auslagerungsmanagement, die unabhängig vom Tool-Einsatz (Algorithmik, Komplexität) auch Ansprüche an zukünftigen Einsatz von ML-Methoden stellen. ML-Methoden stellen ihrerseits intrinsische Anforderungen an IT und Implementierung, die möglicherweise tiefgreifender sind als alternative traditionelle Methoden - je nach Komplexität der angewandten Verfahren.
- Generell kann sich das systemische Risiko bei Auslagerung von Prozessen und Datengenerierung erhöhen, insbesondere bei existierender Monopol-/Oligopol-Stellung des zugehörigen Auslagerungs-Unternehmens.
- Adversarielle Angriffe sind bei allen verteilungsfrei trainierten Methoden grundsätzlich vorstellbar (Wobei das Risiko je nach Methode differiert). Rückfallebenen über explizit erklärbare Surrogate Modelle sind prinzipiell vom VAG gedeckt, aber aktuell nicht verbreitet und erfordern zudem hohen Pflegeaufwand.

Diskussionspapier, Kapitel III.2:

i) Welche Herausforderungen sehen Sie bei der Auswahl der Daten und bei der Sicherstellung der Datenqualität von ML-Methoden?

Antwort i) DAV:

- Grundsätzlich nimmt die Bedeutung des sog. *Curse of Dimensionality* zu: Die Wahrscheinlichkeit dafür, in hochdimensionalen (und dynamisch gewonnenen) Daten zufällige, nicht von Signalen(=Kausalen Zusammenhängen) unterscheidbare Korrelationen zu finden, wächst exponentiell mit der Input-Menge. Dies führt dazu, dass im Zuge der Digitalisierung die Hypothesenvalidierung komplexer und zugleich wichtiger wird. Das Thema ergibt sich aber aus der Datenmenge, nicht der Wahl der Methode, ist also genauso auch für klassische Methoden relevant.
- Ansonsten gibt es die "klassischen" Fragestellungen wie zum Beispiel: Passen die Daten zur gewählten Methodik, zur Fragestellung bzw. zum geplanten Einsatz der Modellergebnisse? Wird sichergestellt, dass kein Bias in den Daten vorliegt und keine Diskriminierung erfolgt? Daten müssen immer mit Fach- und Sachkenntnis im Kontext bewertet werden.
- Diese Herausforderungen im Umgang mit Daten sind seit jeher Bestandteil der aktuariellen Arbeit. Aufgrund der aktuariellen Aus- und Weiterbildung sehen wir uns als Aktuare als besonders qualifiziert, diese Fragestellungen auch bei ML-Methoden im Versicherungskontext angemessen zu bewerten und zu berücksichtigen. Neben den entsprechenden regulatorischen Anforderungen (SII, VAG, AGG, ...) gibt es für Aktuare zusätzlich noch eigene Standesregeln und zu befolgende fachliche Standards (DAV-Richtlinien).

j) Welche Aspekte der Datenqualität werden durch die Anwendung von ML-Methoden aus Ihrer Sicht erleichtert?

Antwort j) DAV:

- Die typischerweise bei ML-Methoden benutzten Open Source-Tools und -Frameworks bieten prinzipiell automatisierte Datenprozessierung inkl. Analyse, Visualisierung und Transformation der Daten. Dies kann die Erkennung und Bereinigung von z.B. Artefakten oder Bias in Daten erleichtern, erfordert aber tiefgehende Fachkenntnisse über die Erfordernisse der Datenqualität in Bezug auf die jeweiligen regulatorischen Vorgaben und Prüfungen. Eine Vollautomatisierung ist deshalb nicht absehbar und gegenwärtig auch nicht vorstellbar. Im Umkehrschluss bedeutet dies indes, dass der Anwender zuzüglich zu den erforderlichen aktuariellen und regulatorischen Kenntnissen vertiefte methodische Kenntnisse aufweisen muss.

Diskussionspapier, Kapitel III.3:

Auszug/Quote Seite 15 oben:

„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....“

k) Welchen Einfluss hat die „Blackbox“-Eigenschaft Ihres Erachtens auf die Validierung der Verfahren?

Antwort k) DAV:

- Siehe Antwort b): XAI-Methoden stellen keinen Ersatz für plausibilisierte Hypothesen dar, insbesondere, wenn der Prädiktor verteilungsfrei ist. Da Edge Cases komplexer und teilweise schwieriger aufzufinden werden, ist es unerlässlich, sich nicht auf Methoden allein zu verlassen, sondern mit Fach- und Sachkenntnis die dargestellten Zusammenhänge zu bewerten.
- Grundsätzlich sind zwei unterschiedliche strukturelle Eigenschaften der Grund für die Blackbox-Eigenschaft: Während einerseits die Komplexität der Lerner eine (kombinatorische) Evaluation des Zustandsraumes erforderlich machen kann, steht dem die bereits vorstehend mehrfach erwähnte Eigenschaft der Verteilungsfreiheit einiger Methoden entgegen, weil Marginalschnitte und -projektionen die nicht-lineare Eigenschaft nicht abzubilden in der Lage sind, aber andererseits keine analytische Form der Prädiktorfunktion existiert, bzw. gefunden werden kann. Daraus ergibt sich, dass für Ensemble-Methoden wie Boosting XAI uneingeschränkt sinnvoll ist, sofern der Rechenaufwand im günstigen Verhältnis zur durch die Methode gegebenen Präzision steht, wohingegen bei tiefen neuronalen Netzen der Stand der aktuariellen Forschung dies aktuell nicht hergibt.
- Nach der Auswahl der grundsätzlich geeigneten Methoden sind methodenagnostische XAI-Ansätze zu bevorzugen, da sie es erlauben können, unterschiedliche Methoden hinsichtlich ihrer qualitativen und quantitativen Eignung gegenüberzustellen. Spezifische XAI-Ansätze können nachgelagert helfen, ein präferiertes Modell auszuvalidieren.

- Grundsätzlich lässt sich beobachten, dass in der Praxis oft kein großer Unterschied in der Validierung von z. B. linearen Modellen und Machine-Learning-Modellen existiert. Insofern können viele der derzeitigen Validierungsansätze im Bereich z. B. der internen Modelle (Proxymodellierung) oder der Projektionsmodelle bei einem (ggf. teilweisen) Einsatz von ML-Methoden weiterhin angewandt werden. Methodenspezifische Validierungsansätze, die beispielsweise spezifisch für lineare Regressionen entwickelt wurden, müssen durch andere Methoden ersetzt werden.

l) Welche Bedeutung messen Sie dem Trade-Off zwischen Performance und Erklärbarkeit bei?

Antwort l) DAV:

- Aktuarielle Berechnungen verlangen als Grundlage für Prüfung und Validierung einen gewissen Grad von Erklärbarkeit. Insbesondere aus Risikomanagement-Perspektive wird Nachhaltigkeit mit Sicht auf langfristige Stabilität des Unternehmens verlangt, insbesondere sollen proaktiv zukünftige Trends und neuartige Risiken und deren Einfluss auf die Finanzlage des Unternehmens bewertet werden. Solche Berechnungen benötigen die Bewertung unternehmensübergreifender Prozesse und zugehöriger Datengrundlagen für mögliche Re-Evaluierungen von Einflüssen und Trends auch in Folgejahren. Für diese und andere Arten von Nachhaltigkeit (stabile Preisberechnungen für langjährige Verträge, ...) bietet ein Minimum an Erklärbarkeit sowie eine Nachhaltigkeit von Algorithmus-verändernden Einflüssen eine solide Grundlage, und setzt damit möglicherweise eine Schranke bzgl. der Performance.
- Strukturell unterschiedliche Methoden sind generell unterschiedlich gut erklärbar. In Bezug auf sowohl die Komplexität als auch Vorhersagegüte verschiedener ML-Methoden spricht man vom sog. Performance-Erklärbarkeit-Tradeoff-Korridor, womit ausgesagt ist, dass besonders genaue Methoden (bspw. des Deep Learnings) im Allgemeinen größere Hürden an die Erklärbarkeit zeigen, sie sind "schwärzere" Boxen. Insbesondere gilt dies für verteilungsfreie Methoden. Je nach Anwendungszweck ist also eine Abschätzung zu treffen: Es ist davon auszugehen, dass "schwärzere Boxen" umso größeren Aufwand der Validierung und Plausibilisierung mit sich bringen. Der Trade-Off ist damit letztlich ein betriebswirtschaftliches Kosten-Nutzen-Thema.
- Ein zunehmend relevanter Aspekt der BDAI-Thematik wird zukünftig auch der Energieverbrauch von komplexen Modellen, insbesondere des Deep Learnings sein. Während die Genauigkeit sich weiter steigern lässt, geht dies auf Kosten zusätzlichen Energieverbrauchs im Vergleich zu einem simpleren, aber nachhaltigeren Modells. Wenngleich

dies aktuell kein spezifisch aktuarieller oder auch aufsichtsrechtlich relevanter Aspekt ist, ist doch vorstellbar, dass die Betrachtung der Energieeffizienz zukünftig breitere Beachtung finden wird. Deshalb ist wichtig, dass alle an der Auditierung beteiligten Parteien sich verdeutlichen, dass die Modellkomplexität im Allgemeinen sinkenden Grenzerträgen unterliegt, was sich insbesondere darin zeigt, dass die Grenzerträge mit Blick auf den Energieverbrauch im Speziellen besonders ungünstig skalieren.

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?

Antwort m) DAV:

- Grundsätzlich sind XAI-Techniken eine Möglichkeit, Licht auf die „Blackbox“ eines Modells zu werfen. Ob das ausreichend möglich ist, ist im Einzelfall insbesondere auch in Abhängigkeit von der jeweiligen Anforderung an das genutzte Modell und seiner Erklärbarkeit abhängig.
- Eine konkrete Nennung von Verfahren scheint nicht zielführend, nicht zuletzt auch, da sich die XAI-Techniken aktuell rasant entwickeln und Antworten daher schnell veralten können.

n) Wie sollte aus Ihrer Sicht eine der Methode nachgelagerte XAI in die Validierung einbezogen werden?

Antwort n) DAV:

- Ebenso wie die Validierung ist die Modellerklärung als ein Teil eines Modells zu betrachten und ist somit integraler Bestandteil des Modellierens (und gehört zum Modellkonzept).

Diskussionspapier, Kapitel III.4:

o) Welche Fragen ergeben sich aus Ihrer Sicht zur Aufsichtspraxis in Bezug auf Modellanpassungen bei ML-Methoden?

Antwort o) DAV:

- Kritisch zu bewerten ist hier in erster Linie der sog. Model Drift, also eine schleichende „Entvalidierung“ durch kleine, aber sich kumulierende Entfernung von den (geprüften) Grundannahmen. Dies gilt aber gleichermaßen für „klassische“ Ansätze und ist kein ML- sondern ein Datenthema.

p) Sehen Sie für bestimmte ML-Methoden die Notwendigkeit sehr häufiger Retrainings?

Antwort p) DAV:

- Der Bedarf eines Modell-Retrainings rührt aus den fachlichen Anforderungen und ggf. der Verfügbarkeit der Daten, nicht aber aus dem Modell.

q) Werden ML-Methoden Ihres Erachtens eine Anpassung der Modell-Governance notwendig machen? Wie arbeiten klassische Modellierungseinheiten, Validierer und neue „Data Science“-Einheiten zusammen?

Antwort q) DAV:

- Grundsätzlich ist die existierende Modell-Governance vorerst unabhängig vom Einsatz unterliegender Tools und Verfahren. Die klassische Verkettung zwischen Modellentwicklung, Prüfung/Validierung und Modelleinsatz kann weiterhin funktionieren und bedarf möglicherweise zusätzlicher Elemente, um eine erhöhte Frequenz von Anpassungen zuzulassen.

ML-Verfahren, die zusätzlich zu Anpassungen von Kalibrierungen auch methodische Änderungen vornehmen (z.B. Erhöhung/Reduktion von eingesetzten Parametern), benötigen neuartige Ansätze auch in Modellgenehmigungs-/Anpassungsverfahren.

Die Modell-Governance erfordert wie bisher eine entsprechende Zusammenarbeit aller relevanten Beteiligten (Modellierer, Validierer, Data Scientists, etc.).

Die Deutsche Aktuarvereinigung (DAV) ist die berufsständische Vertretung der Aktuare in Deutschland. Sie schafft die Rahmenbedingungen für eine fachlich fundierte Berufsausübung ihrer Mitglieder und steht im ständigen Dialog mit allen für sie relevanten nationalen und internationalen Institutionen, um im Interesse der Aktuare und zum Nutzen von Verbrauchern und Unternehmen ihren Sachverstand in gesetzgeberische Prozesse einzubringen.