Dr. Raphael Reinwald
Mai 2023

KÜNSTLICHE INTELLIGENZ IN DER FINANZ- UND BANKEN-BRANCHE – EINE ÜBERSICHT

ChatGPT und die Chancen und Risiken von aktuellen KI-Systemen

Nachdem am 30.11.2022 ChatGPT 3.5 von OpenAI für die Öffentlichkeit frei zugänglich gemacht wurde und schon im Januar 2023 – in präzedenzlosem Rekord – über 100 Millionen Nutzer hatte, sind die Möglichkeiten von künstlicher Intelligenz (KI) mittlerweile auch in das Bewusstsein einer breiten Öffentlichkeit gelangt. Eine weitaus präzisere und intelligentere Antwortfunktion als bisherige Chatbot-Systeme, Features wie mehrstufige Antworten und das „Erfinden" realistischer Anekdoten, Eigenschaften wie das „Erkennen" von Ironie oder von philosophischen Fragestellungen, „Kreativität" beim Schreiben von Gedichten oder gar von Programmcode, sowie sogenannte emergent abilities wie z. B. Mehrsprachigkeit – u. a. aufgrund der sehr großen Parameteranzahl des Modells – gingen über die Fähigkeiten bisher frei verfügbarer KI-Systeme deutlich hinaus. Diese Kompetenzen sind mit ChatGPT 4.0 (kostenpflichtig) im März 2023 sogar nochmals deutlich weiterentwickelt worden, das Modell besitzt nun schon ca. 100 Billionen Parameter [1].

Vor- und Nachteile des Einsatzes von KI-Software wurden in der Folge noch kontroverser diskutiert als zuvor, u. a. wurde bezüglich der Gefahr von Arbeitsplatzverlusten, der Möglichkeit des Kontrollverlustes (bspw. im Bereich autonomes Fahren oder Kriegstechnologie), von Datenschutz und Copyright-Problemen sowie des Einsatzes für Betrug, Deep-Fakes und Propaganda-Zwecke gewarnt (bspw. von Europol, siehe [2]). Das Future of Life Institute veröffentlichte am 23. März 2023 sogar einen offenen Brief, in dem eine Entwicklungspause solcher Systeme gefordert wurde – mit vielen berühmten Unterzeichnern wie Stuart Jonathan Russell, Elon Musk und Steven Wozniak [3].

Allerdings sind auch die positiven Einsatzmöglichkeiten und Chancen von künstlicher

Intelligenz in Bereichen wie Medizin (Molekularmedizin, Krebserkennung und -heilung, Bildanalyse), Mobilität (mit vrstl. deutlich geringeren Todesraten im Individual- wie auch ÖNV, Stau- und Emissionsreduktion), Kundenservice (effizientere Chatbots, Sprachanalyse und Synchronübersetzung), Rechtswesen (smart contracts im Vertragswesen), Konstruktion und Kunst (generative AIs) und im Finanzbereich (Portfoliooptimierung und Marktprognosen, Robo Advisory, AML/ATF[1] und Risikomanagement) sehr vielversprechend.

Wie bei jeglicher Technologie ist somit ein passender Gesetzes- und Regulierungsrahmen nötig, welcher im Bereich der KI (und zukünftig vrstl. gar „starker KI“) sicherlich auch Datensicherheitsfragen und ethische Aspekte umfassen muss.

KI in Verknüpfung mit weiteren Technologien (wie DLT)

Dies gilt umso mehr als KI in den meisten Fällen nicht wie ChatGPT in der Form einer general AI (oder eines allgemeinen question-answer-Chatbot-Systems) auftritt, sondern dezentral, eher spezialisierter und in anderen Geräten und Anwendungen gekapselt (ChatGPT selbst zukünftig in Office365, der Suchmaschine Bing, Microsoft Visual Studio/Azure DevOps etc.). Wie in anderen Branchen auch – betrachten wir bspw. die weitere Verschmelzung von künstlicher Intelligenz und Industrie 4.0 (5.0) / Industrie IoT im Maschinenbau – wird auch in der Finanzindustrie das Verschmelzen von künstlicher Intelligenz mit anderen Technologien wie bspw. der Blockchain/Distributed Ledger Technologie (DLT) stattfinden, welche derzeit schon digitale Emissionen, die Registrierung, den Handel sowie die Abwicklung und (Zentral-)Verwahrung von Kryptoassets ermöglicht und auf europäischer Ebene insbesondere derzeit durch die MiCAR und das EU DLT Pilotregime (flankiert durch das deutsche eWPG) reguliert wird [4].[2] Zukünftig wird zudem eine aufsichtliche Behandlung und Risikobetrachtung (RWA-Unterlegung) von Kryptoassets i. W. durch eine (europäischen) Umsetzung des Baseler BCBS 545 - Papiers erfolgen [5].

Künstliche Intelligenz und die verschiedenen Arten maschinellen Lernens

Dieser Fachbeitrag soll sich mit dem Einsatz von maschinellem Lernen im Finanzbereich und insbesondere in Banken beschäftigen. Während der Terminus Künstliche Intelligenz (KI, engl. AI: Artificial Intelligence) als allgemeiner Oberbegriff fungiert, ist maschinelles Lernen eine Teilmenge von KI, bei welcher eine große Menge an Daten i. W. durch Mustererkennung („pattern recognition“) verarbeitet, verallgemeinert und zur Prädiktion/Generierung genutzt werden kann. Maschinelles Lernen wird mittlerweile als eine Querschnittsdisziplin betrachtet, ist ursprünglich aber ein Teilgebiet der statistischen Interferenz.

Nach einer kurzen Einführung in die wichtigsten Arten von maschinellem Lernen sowie den regulatorischen Rahmenbedingungen und aufsichtlichen Sichtweisen werden einige aktuelle Anwendungsbereiche und Einsatzmöglichkeiten in diesem Fachbeitrag genauer vorgestellt.

Die verschiedenen Arten maschinellen Lernens

Im Bereich des maschinellen Lernens unterscheidet man im Wesentlichen zwischen drei Kategorien, eine Zwischenkategorie spielt zudem im Bereich der sogenannten

[1] AML: Anti-Money-Laundering (Geldwäschebekämpfung), ATF: Anti-Terror-Financing

[2] MiCAR: Markets in Crypto Assets Regulation (in Anlehnung an MiFIR: Markets in Financial Instruments Regulation, von welcher das Pilotregime auch gewisse Ausnahmen erlaubt)

© 2023 1 PLUS i GmbH

Large Language Models (Transformer-basierte Modelle mit einer großen Anzahl an Parametern wie bspw. auch Chat-GPT) eine wichtige Rolle.
Diese drei Kategorien (Klassen) sind:

Unsupervised Learning

Unsupervised Learning:
Unsupervised Learning bezeichnet das maschinelle Lernen ohne ex ante bekannte Labels (Sollwerte/Zielwerte) für die Daten und ohne Korrekturimpulse wie Belohnungen („Rewards“) während des Lernprozesses. Ziel ist es Muster (wie Cluster, Anordnungen, Hierarchien und andere Beziehungen) in den zunächst unstrukturierten Eingabedaten zu erkennen. Aufgrund der geringen Einschränkungen bezüglich der verwendenden Daten können mit Hilfe von Unsupervised Learning auch neue Strukturen in Daten identifiziert werden.
Unsupervised Learning lässt sich aus der Bayes’schen Statistik (Posterior Schätzer und eine folgende prediktive Verteilung) und informationstheoretischen Überlegungen (wie der Kullback-Leibler Divergenz sowie AIC, BIC[3]) motivieren.
Wichtige Formen des Unsupervised Learning sind die Clusteranalyse, Hauptkomponentenanalyse (beides latente Variablen Modelle), Zeitreihenmodelle, Anomalienerkennungen und Graphenmodelle.
Diese lassen sich bspw. wie folgt genauer unterteilen:

Clusteranalyse:
Die Clusteranalyse gruppiert Untersuchungsobjekte zu natürlichen Gruppen bzw. Ansammlungen anhand eines Proximitäts- bzw. Ähnlichkeitsmaßes mit Algorithmen wie K-means (nicht-hierarchisch), Linkage-Methoden, KNN (K-Nearest-Neighbor) oder der Ward Varianzmethode (alle hierarchisch, oft auch mit Hilfe von Dendrogrammen dargestellt).

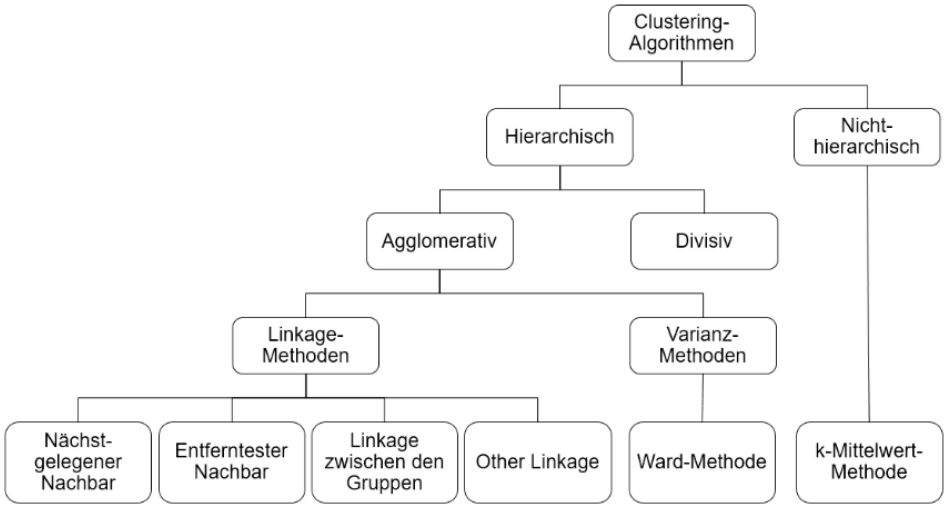

Abbildung 1: Hierarchische und nicht-hierarchische Clustermethoden. Quelle: UZH Universität Zürich https://www.methodenberatung.uzh.ch/de/datenanalyse_spss/interdependenz/gruppierung/cluster.html

Neben den *hierarchischen Verfahren* (auch Multi-Level Verfahren) existieren noch (zentrumsbasierte) *partitionierende Verfahren*, welche eine gegebene Partitionierung (mit festgelegter Clusterzahl) verwenden und entweder die Elemente durch eine Aus-

[3] AIC: Akaike Information Criterion, BIC: Bayesian Information Criterion (oder SIC: Schwarz Information Criterion)

© 2023 1 PLUS i GmbH

tauschfunktionen umordnen, bis die Zielfunktion optimiert ist, oder iterierte Minimaldistanzverfahren anwenden (bspw. die Methoden k-means++, k-median oder k-medoid).
Zudem werden im Bereich Unsupervised Learning auch häufig *dichtebasierte und gitterbasierte Verfahren* verwendet. Beim dichtebasierten Clustering sind die Cluster Objekte in einem mehrdimensionalen Raum, welche dicht beieinander liegen, getrennt von Gebieten mit niedrigerer Dichte (Dichteverteilung), beim gitterbasierten Verfahren wird der Datenraum in endlich viele Gitterzellen aufgeteilt und vermessen (bekannte Vertreter sind OPTICS [6] als dichtebasiertes Verfahren und CLIQUE als gitterbasiertes Verfahren [7][8]).

Hauptkomponentenanalyse bzw. Dimensionsreduktion:
Ziel ist es bei diesen Verfahren, die Information aus vielen einzelnen Variablen in wenige Hauptkomponenten (Basiselemente) zu bündeln bzw. die Dimension der Eingangsdaten zu reduzieren.
Die am meisten verwendeten Methoden in diesem Bereich sind die Faktoranalyse (FA), die Hauptkomponentenanalyse (Principal Component Analysis) und die unabhängige Komponentenanalyse (ICA), eine Erweiterung der Faktoranalyse, welche keine Gaußschen (normalverteilten) Faktoren erfordert. Eine bekannte Erweiterung der ersten beiden Varianten wäre aber bspw. auch das Gaußsche Mischmodell (Gaussian Mixture Model) mit EM (Expectation-Maximization) - Algorithmus.

Zeitreihenmodelle:
Zeitreihendaten liegen meist nicht unstrukturiert i. S. v. iid sondern autokorreliert vor. Sie werden traditionell oft durch (V)ARIMA[4] oder GARCH-Modelle dargestellt und bspw. durch das Box-Jenkins-Verfahren gelöst [9][10]. Oft sind im KI-Bereich bei Zeitreihendaten Hidden-Markov Modelle (generell sogenannte dynamische Bayes'sche Netzwerke; räumlich auch HMRF-EM[5]) oder State-Space-Modelle (SSM, im nichtlinearen Fall von SSM auch bspw. der EP-Algorithmus oder Partikelfilter) zur Mustererkennung im Einsatz.
Weitere Abhängigkeiten (bspw. räumliche „spatial structures") lassen sich oft mit den folgenden Ansätzen finden:

Probabilistic Relational Models oder Graphenmodelle, speziell Bäumemodelle (Tree-Based models):
Diese geben Relationsstrukturen anhand von (un-/gerichteten bzw. Faktor-) Graphen oder Baumstrukturen an.
Eingesetzte Algorithmen in diesem Feld sind Factor Graph Propagation, (Loopy) Belief Propagation oder der Junction Tree Algorithmus [11] – zur approximativen Lösung der oft hochdimensionalen Probleme können beispielsweise MCMC (Markov Chain Monte-Carlo), EP (Expectation Propagation) oder DPM (Dirichlet Process Mixtures) Verfahren Verwendung finden.

Supervised Learning

Supervised Learning:
Supervised Learning ist eine Form des maschinellen Lernens, welche vorher gelabelte (Ziel/Soll-) Datensätze verwendet, um Algorithmen zu trainieren, und dann Ergebnisse

[4] VARIMA: Vector Auto-Regressive Integrated Moving Average, GARCH: Generalized Auto-Regressive Conditional Heteroscedasticity
[5] HMRF-EM: Hidden-Markov Random Field Expectation Maximization

© 2023 1 PLUS i GmbH

korrekt vorherzusagen. Während die Eingabedaten in das Modell eingespeist werden, werden die Modell-Parameter (Gewichtungen) angepasst, bis das Modell verglichen mit den Label-Daten angemessen präzise ist. Es leitet also eine Funktion aus gelabelten Trainingsdaten ab, die aus einer Reihe von Trainingsbeispielen bestehen. Beim Supervised Learning ist jedes Beispiel somit ein Paar, das aus einem Eingabeobjekt (typischerweise einem Vektor) und einem gewünschten Ausgabewert/Zielwert besteht. Das Ziel des Lernalgorithmus ist es von den Trainingsdaten in angemessener Weise, also ohne ein over- oder underfitting der Trainingsdaten, funktional auf neue (out-of-sample and out-of-time) Daten verallgemeinern zu können.
Verwendete Verfahren sind:

Lineare oder Quadratische Diskriminanzanalyse (LDA/QDA):
Die LDA ist wie die lineare Diskriminante von Fisher ein Verfahren, um eine lineare Kombination von Merkmalen zu finden, die zwei oder mehr Klassen von Objekten charakterisiert bzw. trennt. Dies geschieht durch Maximierung der Zwischen-Gruppen-Varianz und Minimierung der Inner-Gruppen-Varianz. Die LDA ist eng mit der bekannten linearen Regression verbunden. Die QDA erlaubt zur Trennung der Objekte unterschiedliche Kovarianzmatrizen bei quadratischer Funktion zur Zuordnung der Klassen [12].
Weitere bekannte Methoden sind die linearen Regularisierungsverfahren LASSO (Least Absolute Shrinkage and Selection Operator) und Ridge Regression (von Tikhonov), oder auch der stochastische Gradientenabstieg (Stochastic Gradient Descent).

Naïve Bayes-Schätzer:
Naïve Bayes Verfahren sind einfache probabilistische Klassifikatoren, die direkt auf der Anwendung des Bayes-Theorems zur Berechnung des a posteriori Schätzers mit (naiven) Unabhängigkeitsannahmen zwischen den Merkmalen basieren.

Decision- / Random-Trees und Random-Forests:
Ein Entscheidungsbaum ist eine Struktur, in der jede Vertex-Formation eine Frage darstellt und jede Kante, die von diesem Scheitelpunkt absteigt, repräsentiert eine (probabilistische) mögliche Antwort auf diese Frage. Random Forests kombinieren die Ergebnisse verschiedener Entscheidungsbäume, um anschließend ein einziges aggregiertes Ergebnis zu erzeugen. Beide Methoden können zur Klassifikation verwendet werden und die Ausgabeklassen der Blätter durch Vergleich mit den gelabelten Ausgaben sukzessive angenähert. Random Forrests werden häufig als Benchmark im Vergleich mit anderen Modellen genutzt.

Die am häufigsten verwendeten Verfahren des Supervised Learning: Künstliche Neuronale Netze

Künstliche Neuronale Netzwerke (engl. ANNs), speziell mehrschichtige sogenannte Deep Neural Networks (DNN) mit einer Vielzahl verborgener Schichten (hidden layers):
ANNs sind Netzwerksysteme, die von den biologischen neuronalen Netzen inspiriert sind, aus denen menschliche Gehirne bestehen. Ein ANN basiert auf einer Sammlung verbundener Einheiten, die als künstliche Neuronen bezeichnet werden und die Neuronen in einem biologischen Gehirn modellieren - jede Verbindung kann, wie die Synapsen in einem biologischen Gehirn, dabei ein Signal an andere Neuronen übertragen. Wann ein Neuron aktiviert wird, wird hierbei durch eine Aktivierungsschwelle und eine Aktivierungsfunktion verschiedener Form (wie z. B. einer Sigmoid-, ReLU-

© 2023 1 PLUS i GmbH

oder Softmax-Funktion) bestimmt. Es gibt die Neuronen der Eingabeschicht (input layer), der Ausgabeschicht (output layer) und der zwischenliegenden Verarbeitungsschicht wie auch verborgener (Zwischen-)schichten (hidden layers). Auch bei künstlichen neuronalen Netzwerken werden die Eingabedaten in das Modell eingespeist und die Modell-Gewichtungen so lange angepasst, bis das Modell verglichen mit den Label-Daten hinreichend genau approximiert.

Support Vector Machines (SVM):
SVM sind Klassifikationsalgorithmen, welche Eingangsdaten linear (oder nichtlinear, per Transformation in einen höherdimensionalen Raum mithilfe des sogenannten Kernel-Tricks) in zwei oder (iterativ) mehrere Zielklassen separieren und hierbei die Distanz zwischen den Klassenobjekten maximieren (sogenannte Large Margin Classifier) um später eine hinreichende Verallgemeinerungsfähigkeit zu besitzen (ein overfitting auf die Trainingsdaten zu vermeiden). Kernel(dichte)funktionen hierfür können bspw. lineare Funktionen, radiale Basisfunktionen oder Sigmoid-Funktionen sein. Auch im Falle der SVM werden die Modell-Parameter wieder so lange angepasst, bis die Modellzielklassen verglichen mit den vorgegebenen Label-Klassen hinreichend genau approximieren.

Neben Random Forests werden für komplexere Probleme in der Praxis insbesondere Support Vector Maschinen und Artificial Neural Networks eingesetzt. Auch Kombinationen verschiedener Supervised Learning Modelle (sogenannte Ensembles) finden in der Praxis häufig Verwendung (bspw. AdaBoost oder XGBoost als Boosting- oder Bootstrapping als parallelisierbare Bagging-Methode).

Reinforcement Learning

Reinforcement Learning:
Reinforcement Learning (RL) ist ein Bereich des maschinellen Lernens, der sich mit der Frage befasst, wie intelligente Softwareagenten in einer Umgebung Maßnahmen ergreifen/sich anpassen sollten, um die kumulative Belohnung (den Reward, deswegen auch Reward Based Learning oder verstärkendes Lernen genannt) zu maximieren – es wird also eine zielorientierte Strategie (Policy) erlernt, um erhaltene Belohnungen zu optimieren. Es müssen hierbei von außen – von einem Softwaresystem oder einem Menschen – Impulse oder Rewards gegeben werden. Da Reinforcement Learning eine Reihe optimaler Aktionen beinhaltet, wird es als sequenzielles Entscheidungsproblem betrachtet und kann mit dem Markov-Entscheidungsprozess (MDP) modelliert werden.
Die Bellman-Gleichung löst beim MDP die Frage nach dem maximalen Nutzen und leitet die optimale Strategie ab. Reinforcement Learning ähnelt somit dem Lösen eines – ggf. partiellen – MDP, aber hierbei sind die Übergangswahrscheinlichkeiten und die Belohnungsfunktion (Reward) unbekannt, der Agent muss somit Aktionen ausführen, um eine möglichst optimale Policy zu lernen.
Oft wird Reinforcement Learning, statt simples menschliches Feedback zu verwenden, auch mit Methoden der Spieltheorie kombiniert. Die Spieltheorie dient der Modellierung und Lösung von Optimierungsproblemen in Multiagentensystemen. Hierbei sind die Zielfunktionen („Wünsche“/„Ziele“) der einzelnen Agenten über ihre Entscheidungsvariablen in kooperativer oder nicht-kooperativer Form gekoppelt, die vorherigen Handlungen der Agenten beeinflussen ihre Strategien also gegenseitig. Meist werden sogenannte Nash Gleichgewichte als Lösung angestrebt.

© 2023 1 PLUS i GmbH

Reinforcement Learning wird häufig noch weiter unterteilt in: **Model-free, Model-based, Online Learning oder Offline Learning.**

	Model-free RL	Model-Based RL
Offline Learning (Passive)	Direct Utility Estimation Temporal-Difference Learning (TD Learning)	Adaptive Dynamic Programming (ADP)
Online Learning (Active)	Q-Learning (Active TD Learning) SARSA (Active TD Learning)	Exploration (Active ADP)

Abbildung 2: Formen des Reinforcement Learning. Quelle: Introduction to reinforcement learning terminologies, basics, and concepts (model-free, model-based, online, offline RL), Towards Data Science, Kay Jan Wong, 25.11.2022.

Einsatz und Aufgaben des maschinellen Lernens, kombinierte Verfahren und Self-Supervised Learning

Beispiele für Reinforcement Learning Algorithmen sind Direct Utility Estimation (DUE), Temporal-Difference Learning (TD) oder auch die klassischen Verfahren Value Iteration, Policy Iteration (beide modelfrei) im Offline Learning. Modellbasiert wird im Offline Learning meist eine spezielle Form der dynamischen Programmierung, die Methode Adaptive Dynamic Programming (ADP), eingesetzt.
Offline-Lernalgorithmen arbeiten mit einer großen Menge an Daten aus einem Datensatz. Strikt offline lernende Algorithmen müssen von Grund auf neu ausgeführt werden, um aus geänderten Daten zu lernen. Online-Lernalgorithmen hingegen arbeiten mit Daten, sobald sie verfügbar sind. Strikte Online-Algorithmen verbessern sich inkrementell mit jedem neuen Datenelement, sobald es eintrifft; sie verwerfen diese Daten danach und verwenden sie nicht mehr. Im Online Learning finden bevorzugt die Algorithmen Exploration (AADP), Q-Learning und SARSA Verwendung.
Aufgaben des maschinellen Lernens lassen sich prinzipiell grob unterteilen in Mustererkennung und Klassifizierung, Regression und das Erlernen von komplexeren Verhaltensmustern im Reinforcement Learning mit dem Ziel einer möglichst genauen Prediktion.
Natürlich können die oben genannten Methoden Unsupervised Learning, Supervised Learning und Reinforcement Learning und die dazugehörigen Modelle auch kombiniert sequenziell oder (teilweise) parallel genutzt werden (z. B. bei Klassenübergreifenden Boosting-Methoden die technisch i. W. einen verbesserten Gradientenabstieg in einem Funktionsraum unter Verwendung einer konvexen Kostenfunktion durchführen oder bei gewichteten Bagging-Ensembles).
Als Zwischenkategorien zwischen Unsupervised und Supervised Learning gelten zudem das Semi-Supervised Learning, bei welchem nur eine Untermenge der Lerndaten bereits gelabelt und bekannt sind, sowie Self-Supervised Learning, also selbstüberwachtes Training (bspw. mit Texten), welches für LLMs bedeutend ist.
Self-Supervised Learning erfuhr insbesondere durch das Paper [13] und seine Verwendung in aktuellen LLMs wie ChatGPT, BERT oder in die Suchmaschinen Bing Chat und BARD großer Beliebtheit.

Self-Supervised Learning:
Self-Supervised Learning ist wie erwähnt eine Zwischenkategorie von Unsupervised und Supervised Learning, meist umgesetzt mithilfe künstlicher neuronaler Netze, bei welcher keine durch Menschen ex ante gelabelten Beispieldaten benötigt werden. Es wird auch als prädiktives Lernen bezeichnet.

Das Modell trainiert sich hierbei zunächst selbst, um einen (späteren) Teil der Eingabe

von einem anderen Teil der Eingabe zu lernen. In diesem Prozess wird das unüberwachte (unsupervised) Problem in ein überwachtes (supervised) Problem umgewandelt, indem die Labels automatisch generiert werden. Es werden also keine expliziten Labels benötigt, stattdessen werden in die Daten eingebettete Strukturen (wie Korrelationen, Metadaten) selbstständig aus den Daten extrahiert (und labels erzeugt). Der erste Schritt ist auch als Pre-Training bekannt, da ein Modell erzeugt wurde, das zwar noch nicht die finale Aufgabe lösen kann, bei welchem die hieraus erhaltenen Parameter aber sehr günstige Startbedingungen für ein weiteres Fine-Tuning darstellten (downstream task). Danach wird das Modell per überwachtem (supervised) Lernen auf die eigentliche Aufgabe trainiert („getuned“).

Am Ende wird das Modell oft noch per Reinforcement Learning (bspw. per Proximal-Policy-Optimization) with Human Feedback RLHF für menschliche Interaktionen angepasst.
Self-Supervised Learning findet beispielsweise in den Bereichen automatische Bild-, Sprach- (Natural Language Processing/NLP) und Videoerkennung bzw. Generierung, oder auch in den Feldern autonomes Fahren und Robotik Verwendung [14].
Früher häufig verwendete Arten neuronaler Netze wie RNN (Recurrent Neural Networks) oder LSTM (Long Short-Term Memory) – beide wurden bspw. insbesondere in der AI gestützten Zeitreihenanalyse (Time-Series Analysis) und Signaltheorie eingesetzt – werden in vielen Anwendungsbereichen durch die zuvor genannten Transformer-basierten LLMs (mit „attention“-Verarbeitung) abgelöst. Dies ist insbesondere für auf Text oder Zahlen basierende Zeitreihen der Fall, durch in den Bereichen Bild, Musik und Video sind derzeit Generative Adverserial (Neural) Networks (GANs) die Modelle der Wahl.
Streng genommen handelt es sich bei einem GAN um zwei neuronale Netze: eines, das zum Labeln/Kategorisieren und Bewerten entwickelt wurde, und das andere, um Werke von Grund auf neu zu erstellen. Indem man sie miteinander koppelt, ist es möglich eine KI zu kreieren, die auf Befehl (realistische) Inhalte wie Bilder oder Videos generieren kann.

Viele der bekannten KI-Softwareanwendungen tragen Teile ihrer Lernmethode bereits im Namen bspw. ChatGPT (Generative Pre-trained Transformer) oder BERT (Bidirectional Encoder Representations from Transformers) bei den Such-/Chatbots auf LLM-Basis.
Im Bereich Bildbearbeitung und Generierung werden v. a. die Programme Stable Diffusion, Midjourney, Part-I oder Dalle-2 eingesetzt. Zur Erkennung von Deep Fakes finden ebenfalls (geschachtelte) GANs wie der Microsoft VideoAuthenticator Verwendung.
Es ist zu beachten, dass viele dieser Modelle, insbesondere diejenigen welche im Bereich Spracherkennung (NLP) und für Chatbots Verwendung finden, i. W. zur Textgenerierung („erzeugen eines nächsten sinnvollen Wortes bzw. Satzes“) trainiert sind und im Bereich (formaler) Logik oft schlecht performen (hier sind einfachere KI-Modelle derzeit meist überlegen).
Auch im Finanzbereich werden KI und die verschiedenen, damit verbundenen Lernmethoden bereits genutzt, zukünftig ist noch mit einem deutlich verstärkten Einsatz zu rechnen.

© 2023 1 PLUS i GmbH

Kritische Punkte beim Einsatz von KI (im Finanzwesen)

Vor diesem Hintergrund erfährt die Nutzung von KI, insbesondere in sensitiven Bereichen wie bspw. den internen Risiko- und Steuerungsmodellen, eine zunehmende kritische Würdigung und Regulierung von aufsichtlicher Seite und gesetzgeberischer Seite in der EU.
Da viele der Modelle eine „Blackbox"-Funktionalität - welche sich mit den Daten selbst ändert - darstellen und oft auch aufgrund der Modelltiefe (Anzahl bspw. der hidden layer oder Modell(hyper)parameter, Menge der Eingabedaten) keine komplett deterministisch nachvollziehbare Verarbeitung jedes einzelnen Eingabetokens mehr möglich ist, ist ein relativ strenger Regulierungsansatz zu erkennen. Allgemein bevorzugt die Aufsicht, wenn möglich, den Einsatz von sogenannter Explainabe AI (XAI) oder eine gute Approximierbarkeit der Methodik der eingesetzten KI-Software.
Weitere kritische Punkte betreffend des KI-Einsatzes sind (noch) nicht vorhersehbare Seiteneffekte (side effects) und emergent abilities, sowie in einigen Fällen ein mögliches verstärkendes Verhalten bzw. „Herdenverhalten" der Systeme was – auch in Verbindung mit anderen Technologien wie High Frequency Trading (HFT) oder Instant Payment Infrastructure – die Systemstabilität des Finanzsektors gefährden könnte [15].
Auch Liquiditätsrisiken und Risiken aus der Fristentransformation, welche vor kurzem durch den Fall der Silicon Valley Bank, sowie der Signature Bank und der First Republic Bank in den USA wieder in das Bewusstsein der Öffentlichkeit rückten, können durch digitale Einlagen, welche per Smartphone instant übertragbar sind (künftig aufgrund DeFi und CDBC wie den E-EURO noch weiter verstärkt), in deutlich verschärfter Form als bisher auftreten. KI kann solche Krisen also verschärfen, oder aber auch auf aufsichtlicher Seite helfen, um diese früher zu erkennen.
Schon bisher notwendige Anforderungen an die Sicherheit der eingesetzten Informations-technologien und Systeme, den Datenschutz, an Redundanzen und BCM/Notfallmanagement sowie die Systemverantwortlichen werden mit dem Einsatz von KI-Technologie noch höher (vgl. die Regulierungen durch BAIT, DORA und seit einiger Zeit MiCAR im Bereich Kryproassets).
Die qualitative Validierung des Designs, der Use Cases (sin diese operativ, datentechnisch wie auch personell angemessen für den KI-Einsatz) und v.a. der Datenqualität (häufig ist Bias in den KI-Trainingsdaten vorliegend) von Modellen erhält noch größere Bedeutung.
Am Ende muss die Entscheidung bzw. (Gesamt-)Verantwortung über den Einsatz und die Ergebnisse von KI und den spezifischen KI-Modellen in Menschenhand in den jeweiligen Instituten liegen. Hierfür sind völlig andere Fähigkeiten und Kenntnisse als bisher vonnöten.

Regulierung von künstlicher Intelligenz

Auf europäischer Ebene ist der Einsatz von KI bisher konkret wie folgt reguliert:
Übergreifend sind zunächst EU-weit die Harmonisierten Vorschriften zum Umgang mit KI von 2021 [16] zu beachten, welche Branchenübergreifend ausgestaltet und recht allgemein gehalten sind.
Ziel der Verordnung ist es einen einheitlichen und harmonisierten Rechtsrahmen für den Einsatz künstlicher Intelligenz (KI) im europäischen Binnenmarkt zu schaffen und sicherzustellen, dass die Entwicklung und Anwendung von KI in Europa ethischen und gesellschaftlichen Standards entspricht sowie die Grundrechte und -freiheiten der Menschen weiterhin beachtet werden. Dies soll die Schaffung eines vertrauenswürdi-

Regulierung von KI in Europa

gen und innovativen Umfelds für den Einsatz von KI für Verbraucher wie Unternehmen fördern und europäischen Unternehmen einen Wettbewerbsvorteil verschaffen. Interessant ist die Klassifikation von Kreditwürdigkeitsbeurteilungs- bzw. Scoringverfahren als sogenannte High Risk AI.

Auch die EZB hat sich zu dieser Verordnung in Bezug auf die Finanzbranche und ihrer aufsichtlichen Tätigkeit hierin – im Rahmen einer ECB Opinion – geäußert [17].
Die EZB konzentriert sich hierbei auf die Potenziale und Risiken der KI und gibt Empfehlungen für die Umsetzung der Verordnung in der Praxis. Sie betont die Bedeutung von KI für die Wirtschaft und Gesellschaft in der EU, hebt allerdings auch hervor, dass transparente und für alle zugängliche Datenstrukturen, sowie vertrauenswürdige Dienstleistungen notwendig sind, um das richtige Verständnis für die Funktionsweise von KI-Systemen sicherzustellen. Sie fordert insbesondere Kohärenz mit der Kapitaladäquanzrichtline CRD. Zusätzlich erläutert die Europäischen Zentralbank (EZB) ihre eigene Rolle im Zusammenhang mit KI-Anwendungen im Finanzsektor. Die EZB bezeichnet sich dabei selbst als einen wichtigen Akteur, der regulatorische Richtlinien umsetzt und Bedarf für spezifische KI-Lösungen identifiziert – dabei verpflichtet sie sich wie üblich der Technologieneutralität. Sie sieht sich weiterhin nicht als Marktüberwachungsbehörde, aber sieht Aufsichtsbefugnisse im Bereich der Konformitätsbewertung von KI-Hochrisikosystemen, wie bspw. zur Kreditwürdigkeitsprüfung, im Rahmen des SREP[6]. Sie fordert weitere Festlegungen zur Qualität der Trainings-, Validierungs- und Testdaten sowie zur automatischen Protokollierung entsprechender KI-Systeme über ihre gesamte Produktlaufzeit und zur Bedienung durch natürliche Personen via Mensch-Maschine-Schnittstelle. Zudem müssen Robustheit, Genauigkeit und Cybersicherheit stets gewährleistet werden. Die EZB schlägt auch konkrete Anwendungen von KI im Finanzsektor vor wie eine Kreditwürdigkeitsprüfung oder Kreditpunktebewertung, eine Echtzeitüberwachung von Zahlungen oder die Erstellung von Kunden- oder Transaktionsprofilen zum Zwecke der Bekämpfung von Geldwäsche und Terrorismusfinanzierung. Sie hebt hervor, dass weitere Forschung und Entwicklung notwendig ist, um die Potenziale von KI im Finanzwesen vollständig auszuschöpfen.
Die Europäische Zentralbank beschäftigt sich auch weiterhin fortlaufend mit den Auswirkungen und Einsatzmöglichkeiten von KI in der Finanzbranche, insbesondere im Rahmen ihres SupTech Hubs[7] zunehmend auch in Bezug auf eigene aufsichtliche Verwendungszwecke wie Überwachungs- und Auswertungsmöglichkeiten [18].
Zudem legten die European Banking Authority (EBA) und die deutsche BaFin für sie essenzielle Prinzipien und Einschätzungen zum Umgang mit großen Datenmengen und künstlicher Intelligenz (Big Data and Artificial Intelligence – BDAI) bzw. fortgeschrittenen Analysetechniken (BD&AA – Big Data and Advanced Analytics) dar:

EBA Report on BD&AA, 2020 [19]:

Inhalt des Berichts ist die Identifikation von Risiken und Chancen, die sich aus der Nutzung von Big Data- und Advanced Analytics-Anwendungen im Bankensektor ergeben, sowie die Empfehlung von Maßnahmen zur Vermeidung von Risiken und zur Maximierung der Chancen. Zu den Potentialen gehören eine verbesserte Kundenanalyse und

[6] SREP: Supervisory Review and Evaluation Process
[7] Supervisory Technologies, in Analogie zu RegTech, den Regulatory Technologies im Bankenbereich bspw. des Meldewesenanbieters Regnology

© 2023 1 PLUS i GmbH

EBA und BaFin: Leitlinien und Anforderungen zum Einsatz von BD&AA respektive BDAI

Produktentwicklung, die Schaffung neuer Geschäftsmöglichkeiten, eine effiziente Risikobewertung und -kontrolle sowie eine höhere Betriebseffizienz. Die Risiken sind hauptsächlich Datenschutz- und Informationsrisiken, die eine Gefahr für die Vertraulichkeit und Integrität von Kundendaten darstellen können.
Der EBA-Report nennt die folgenden vier Säulen, welche beim Einsatz von BD&AA essenziell und von Instituten somit stets zu reviewen sind:

(i) Datenmanagement
Ein effektives Datenmanagement ermöglicht die Kontrolle und Sicherheit von Daten für Unternehmenszwecke unter Berücksichtigung von verschiedenen Datentypen und Datenquellen, Datenschutz (vgl. DGSVO) und Datenqualität. Ein erfolgreiches Datenmanagement, welches nebenbei Vertrauen schafft und die rechtlichen Anforderungen erfüllt, kann zu einer verbesserten Entscheidungsfindung, höheren betrieblichen Effizienz und einem besseren Verständnis der verwendeten Daten führen.

(ii) Technologische Infrastruktur
Die technologische Infrastruktur umfasst die Verarbeitung der Daten, die Datenplattformen und die Infrastruktur, welche die notwendige Unterstützung für die Verarbeitung und den Betrieb von BD&AA bieten. Sie muss zuverlässig und sicher sein.

(iii) Organisation und Verwaltung
Geeignete interne Führungsstrukturen und organisatorische Maßnahmen sowie die Entwicklung ausreichender Fähigkeiten und Kenntnisse unterstützen die verantwortungsvolle Nutzung von BD&AA in allen Banken und gewährleisten eine solide Aufsicht über ihre Nutzung.

(iv) Analyse-Methodik
Insbesondere betont das Dokument die Bedeutung von fortgeschrittenen statistischen Analysemethoden, um die wertvollen Informationen in großen Datensätzen zu extrahieren und zu interpretieren.

Der Bericht stellt also insgesamt fest, dass eine umfassende Governance und ein effektives Risikomanagement notwendig sind, um die Risiken im Zusammenhang mit Big Data-Anwendungen zu minimieren. Die Banken sollten ihre Datenerfassung, -speicherung und -verarbeitungspolitik überprüfen. Eine transparente Informationspolitik gegenüber Kunden und anderen Stakeholdern ist erforderlich, um Vertrauen und Zustimmung aufrechtzuerhalten. Die Regulierungsbehörden müssen zudem auch in der Lage sein, die Systeme und Prozesse der Banken zu überprüfen, um ein angemessenes Risikomanagement zu gewährleisten.

Die vier Säulen sollten zudem von einer Reihe von Vertrauenselementen („trust elements“) begleitet sein, hier nennt die EBA:
Ethischen Einsatz, Erklärbarkeit und Interpretierbarkeit, Rückverfolgbarkeit und Überprüfbarkeit, Fairness und Verhinderung/Erkennung von Bias, Datenschutz und -qualität, sowie Verbraucherschutzaspekte und Systemsicherheit.

Abschließend hebt der Bericht hervor, dass der Einsatz von Big Data-Anwendungen und fortgeschrittenen Analysetechniken (wie KI) im Bankwesen weiterhin zunehmen

© 2023 1 PLUS i GmbH

wird und fordert eine aktive Zusammenarbeit zwischen den Akteuren im Bankenbereich und den Regulierungsbehörden, um die Vorteile hieraus zu nutzen, ohne dabei die Sicherheit und Vertraulichkeit von Kundendaten zu gefährden.

Bafin Prinzipienpapier BDAI, 2021 [20]:

Dieses Prinzipienpapier befasst sich mit den Grundsätzen, die von Instituten bei der Anwendung von Künstlicher Intelligenz (KI) beachtet werden sollten. Im Folgenden sind die wichtigsten Punkte des Dokuments zusammengefasst:
Die BaFin fordert die Betrachtung des gesamten Prozesses bei einem Einsatz von KI, ob diese dann verwertbare Ergebnisse erzeugt, „hängt davon ab, wie beaufsichtigte Unternehmen sie in Entscheidungsprozesse [insgesamt] einbetten“. Die BaFin selbst will im Rahmen von Prüfungen wie üblich risikoorientiert, proportional und technologieneutral Vorgehen und bestehende Methoden und Vorschriften sukzessive unter Einbeziehung von Überlegungen zur KI weiterentwickeln.

- Einhaltung von Rechtsvorschriften: Banken und Finanzdienstleister müssen auch bei der Verwendung von KI selbstverständlich die relevanten Rechtsvorschriften einhalten. Insbesondere sollten sie die Datenschutzbestimmungen, die Anti-Diskriminierungsgesetze (auch in Bezug auf möglichen Datenbias) und die Bestimmungen zur Informationspflicht erfüllen.

- Verantwortung und Transparenz: Banken und Finanzdienstleister müssen sichergehen, dass die Verwendung von KI transparent ist. Dies umfasst die Veröffentlichung bzw. Dokumentation von Details zur Funktionsweise der KI und die Offenlegung der verwendeten Datenquellen. Berichtslinien und Berichtsformate müssen laut BaFin so gestaltet sein, „dass eine risikoadäquate und adressatengerechte Kommunikation gesichert ist – und zwar von der Ebene des Modellierens bis hin zur Geschäftsleitung“. Die Gesamtverantwortung liegt – MaRisk konform – weiterhin bei der Geschäftsleitung.

- Datenstrategie und Governance: Banken und Finanzdienstleister müssen über ein überprüfbares Verfahren (Datenstrategie) verfügen, welches eine fortlaufende Datenbereitstellung sicherstellt. Es muss zudem klar definiert sein, welche Voraussetzungen jeweils an die Qualität und die Quantität der Daten erfüllt werden müssen. Die Datenstrategie muss in ein Data-Governance-Konzept eingebunden sein.

- Risikomanagement: Banken und Finanzdienstleister sollten ein geeignetes, regelmäßig überprüftes Risikomanagement-System implementieren, um potenzielle Risiken bei der Verwendung von KI zu minimieren. Das System muss Regeln zur Entscheidungsfindung und zur Überwachung der Ergebnisse umfassen. Außerdem muss es geeignete Notfallprozesse (und somit i. Z. wohl auch Ersatzmodelle) für einen möglichen Anwendungsausfall geben.

- Kontinuierliche Überwachung und Anpassung: Banken und Finanzdienstleister müssen Sorge tragen, dass KI-Systeme kontinuierlich überwacht und angepasst werden, um sicherzustellen, dass sie effektiv und korrekt arbeiten.

© 2023 1 PLUS i GmbH

- Menschliche Aufsicht: Banken und Finanzdienstleister sollten sicherstellen, dass auch eine angemessene menschliche Aufsicht bei der Verwendung von KI existiert. Ein intensiver Freigabe- und Feedback-Prozess für KI-Anwendungen muss eingerichtet sein. Insbesondere muss die Verantwortung für Entscheidungen, die auf Datenanalyse durch KI basieren, bei Menschen liegen.

- Qualifikation und Schulung: Banken und Finanzdienstleister sollen stets beachten, dass ihre Mitarbeiterinnen und Mitarbeiter über die notwendigen Qualifikationen und Kenntnisse in Bezug auf KI verfügen müssen.

Insgesamt betont das Bafin Prinzipiendokument die Bedeutung von Verantwortung, Transparenz und Risikomanagement bei der Verwendung von Künstlicher Intelligenz. Darüber hinaus sollten KI-Systeme zielgerichtet überwacht und angepasst werden, um ein effektives und korrektes Arbeiten zu gewährleisten. Um eine angemessene menschliche Aufsicht zu gewährleisten, müssen Mitarbeiterinnen und Mitarbeiter entsprechend qualifiziert sein – was derzeit wohl noch eine Schwachstelle in vielen Instituten darstellt.

Während die Prinzipien und Reports zu BDAI (BD&AA) sich allgemein auf Anwendungen von KI in der Finanz- und Bankenbranche beziehen und hierbei die aufsichtliche Erwartungshaltung in Bezug auf striktes Datenmanagement und -sicherheit, Strategie und Governance, Risikomanagement sowie klare menschliche Verantwortlichkeiten festlegen, hat die EBA auch ein Diskussionspapier zum konkreten Einsatz von maschinellem Lernen im Bereich von internen Kreditrisikomodellen (IRB) veröffentlicht.

Vorausgegangen ist auf Seiten der Banken u. a. der IIF 2019 Report on Machine Learning in Credit Risk [21].
Dieser konstatierte: Die häufigsten Anwendungsfälle von maschinellem Lernen im Kreditrisiko liegen im Bereich des Kredit-Scorings und der Entscheidungsfindung. Finanzinstitute haben sich von der Verwendung von KI für regulatorische Bereiche wie Risikokapital, Stresstests und Rückstellungen abgewandt und konzentrieren sich auf die Anwendung von maschinellem Lernen in Bereichen wie der Kreditüberwachung und der Rückgewinnung.

EBA: Machine Learning for IRB models

Machine Learning for IRB models, 2022 [22]:
Das Papier zielt darauf ab harmonisierte Voraussetzungen zum Einsatz von Machine Learning (ML) Modellen im IRB-Bereich dazulegen und Einsatzbedingungen und Möglichkeiten aufzuzeigen.
Letztere sind – neben einem Einsatz als volles IRB-Modell (bspw. zur PD oder im A-IRB Fall auch zusätzlich der LGD und CCF/EAD-Berechnung):

- Modellvalidierung: Machine Learning Modelle werden verwendet, um Modellherausforderer zu entwickeln, die als Benchmark zum Standardmodell, das für die Berechnung des Kapitalbedarfs verwendet wird, stehen.
- Datenverbesserungen: Machine Learning Techniken können verwendet werden, um die Datenqualität von Daten zu verbessern, die für Schätzungen ver-

wendet werden. Sie sind zudem sowohl im Hinblick auf eine effizientere Datenaufbereitung als auch auf die Datenexploration eine Hilfe, Machine Learning kann dabei auch im Kontext von Big Data verwendet werden, um umfangreiche Datensätze (vor-) zu analysieren.

- Variablenauswahl: Machine Learning könnte verwendet werden, um erklärende Variablen und Kombinationen von Variablen zu selektieren mit nützlichen Vorhersagefähigkeiten innerhalb eines großen Datensatzes.
- Risikodifferenzierung: Machine Learning Modelle können als Modul für Risikozwecke verwendet werden zur Differenzierung des Probability of Default (PD)-Modells, bei dem das Modul bspw. per Modellierung durch Text-Mining Upgrades/Downgrades auf die PD-Klasse zulässt, die zuvor von dem "traditionellen" PD-Modell zugewiesen wurde

Aber gerade auch der Einsatz als „Full IRB-model“ soll unter gewissen Voraussetzungen ermöglicht werden.
Um dies sicherzustellen, müssen eine Reihe an Herausforderungen in den Bereichen Risikodifferenzierung, Risikoquantifizierung, Datenaufbereitung und Validierung gelöst werden. Dies wird in Sektion 3.1 des Dokuments erläutert.

Risikodifferenzierung:
Die Definitions- und Zuordnungskriterien zu Ratingnoten oder Pools (Artikel 171 Absatz 1 Buchstaben a und b CRR) können schwierig zu analysieren sein, wenn ausgefeilte Machine Learning Modelle als Hauptmethode zur Risikodifferenzierung verwendet werden. Das Ermitteln einer klaren ökonomische Erklärung/ Theorie hinter der Zuordnung und die Annahmen hinter dem Modell können eine Herausforderung darstellen, Hilfsmittel zur Interpretierbarkeit werden empfohlen.
Eine geforderte Ergänzung des Modells durch Expertenschätzungen und menschliches Urteilsvermögen erfordert ein genaues Verständnis des Modells und dieser Aspekt könnte die Verwendung von schwer zu interpretierenden Machine Learning Modellen behindern. Wenn ein Institut statistische Modelle für den Zuordnungsprozess für Ratingnoten oder Pools nutzt, sollten dies ergänzt werden durch Experteneinschätzungen (Artikel 172(3) CRR, i. V. m. Artikel 174(e) CRR)
In Bezug auf die Dokumentationsanforderungen sind Artikel 175 Absatz 1, Artikel 175 Absatz 2 und Artikel 175 Absatz 4 Buchstabe a der CRR zu beachten, welche das Institut, wenn es bei der Ratingvergabe ein statistisches Modell verwendet, zwingen die Modellierungsannahmen und die Theorie hinter dem Modell zu dokumentieren. Auch hierfür ist – insbesondere bei KI - Modellen – ein genaues Verständnis des Modells vonnöten.

Risikoquantifizierung:
In Bezug auf den Schätzprozess ist auch die Plausibilität und Intuition der Schätzungen nach Artikel 179 Absatz 1 Buchstabe a der CRR zu betrachten, Machine Learning Modelle können jedoch zu nicht intuitiven Schätzungen führen, vor allem, wenn die Struktur des Modells nicht leicht interpretierbar ist. Darüber hinaus kann es schwierig sein, richtige Beurteilungen bzw. Einschätzungen zu treffen, wie es in Artikel 180 Absatz 1 Buchstabe d CRR gefordert wird, wenn Resultate eine Kombination der Ergebnisse von Technik und der Vornahme von Anpassungen verschiedene Arten von Einschränkungen darstellen.

© 2023 1 PLUS i GmbH

Artikel 180 Absatz 1 Buchstabe a in Verbindung mit Artikel 180 Absatz 1 Buchstabe h CRR und Artikel 180 Absatz 2 Buchstabe a in Verbindung mit Artikel 180 Absatz 2 Buchstabe e CRR verpflichtet die Institute, die PDs nach Schuldnereinstufungen (Ratingnoten der Schuldner) oder Pools aus den langfristigen Durchschnittswerten (LTA) der einjährigen Ausfallquoten, unter Voraussetzung, dass die Dauer des zugrunde liegenden historischen Beobachtungszeitraums mindestens fünf Jahre für mindestens eine Datenquelle beträgt, zu schätzen. Dieser Zeitraum kann ggf. für Machine Learning Modelle schwierig darzustellen sein.

Validierung:
Schwierigkeiten bei der Interpretation der Ergebnisse der Validierung. Machine Learning Modelle können die Behebung von identifizierten Mängeln komplexer machen: Es kann z. B. nicht einfach sein, ein Absinken der Performance des Hauptmodells zu verstehen (gemäß Artikel 174 Buchstabe d der CRR), wenn der Zusammenhang zwischen Eingabedaten und Risiko-Parametern nicht richtig verstanden wird. Sensitivitäten und Risikotreiber müssen klar erkennbar sein. Außerdem kann die Validierung interner Schätzungen gemäß Artikel 185 Buchstabe b der CRR sonst schwieriger sein.
Schwierigkeiten im Zusammenhang mit den Validierungsaufgaben selbst: Im Hinblick auf die Bewertung der Inputs der Modelle, bei denen alle relevanten Informationen bei der Zuordnung von Schuldnern zu Klassen oder Pools gemäß Artikel 172 Absatz 1 CRR beachtet werden, kann es schwieriger sein, die Repräsentativität zu beurteilen sowie die operativen Datenanforderungen (z. B. Datenqualität oder Datenspeicherung und Data Maintenance) zu beurteilen. Auch das Modelldesign, Modellannahmen (nach Artikel 185 CRR) und passende Anwendungsfälle und deren Implementierung (Artikel 144 und 144 1b CRR) sind von der Validierungsfunktion bei komplexen Modellen schwerer zu beurteilen. Bei der Bewertung der Modellergebnisse muss aufgrund des hohen Risikos eines overfittings an die Trainingsdaten besonderes Augenmerk auf die Verwendung von Stichproben (samples) gelegt werden – diese müssen out-of-time und out-of-sample gewählt sein (wie bereits in Artikel 175 Absatz 4 Buchstabe b der CRR gefordert).
Einige der Funktionen von Machine Learning Techniken führen zu beiden o. g. potenziellen Herausforderungen, z. B. kann ein komplexeres Modell zu einer komplexeren Dokumentation führen, was wiederum die Validierung für die Validierungsfunktion erschwert.

Die Kategorisierung von Modelländerungen (wie in Artikel 143 Absatz 3 der CRR14 vorgeschrieben) und MCP (Model Change Policy) Anzeigen können eine Herausforderung darstellen, wenn Modelle mit hoher Frequenz aktualisiert werden und Variablen mit zeitlich unterschiedlichen Gewichtungen verknüpft sind.
Auch eine genaue Kenntnis der Hyperparameter eines Machine Learning Models sollte vorhanden sein und eine Bestimmung über das Validation Data Set motivierbar.
Natürlich müssen auch die Anforderungen der Corporate Governance nach Artikel 189 CRR bzgl. Interpretierbarkeit und Gesamtverantwortung des Vorstandes gegeben sein – auch dort muss somit ein gewisses Modellverständnis vorhanden sein. Vor übermäßiger Modellkomplexität wird gewarnt.

Die Interpretierbarkeit ist also der entscheidende Faktor: Neben Explainable AI (XAI)

können laut EBA auch graphische Tools, inverse Betrachtungen, Feature-Importance-Measures der Statistik (univariate oder multivariate wie fit-time oder predict-time measures) bspw. PPS, MIC, oder VAE, lokale Erklärungen wie LIME oder Ankerwerte oder auch die häufig verwendeten Shapley-Werte hierzu Verwendung finden. Eine andere Möglichkeit ist bspw. eine Linearisierung/Polynomisierung als Modellapproximation.
In Abschnitt 3.2 werden von der EBA nochmals potenzielle Stärken und Vorteile der Verwendung von Modellen des maschinellen Lernens dargestellt. In Sektion 4. des Berichts werden die o. s. Anforderungen, um einen Einsatz von Machine Learning Modellen zu ermöglichen, zusammengefasst.
Weitere Anwendungsbereiche – viele davon wurden auch schon im EBA-Report und vom FSB 2017 [15] erwähnt – sind die Folgenden [23][24]:

Anwendungen von KI in der Finanzbranche

- Allgemeine Prozessoptimierungen und Automatisierungen/RPA – Robotic Process Automation (noch die häufigste Anwendung)
- Kundenservice (Chatbots)
- KYC (Know Your Customer) Anwendungen
- Credit Scoring Applications (außerhalb Europas auch Verhaltensscoring mit alternativen Daten wie Amazon/Alibaba-Käufen, Social-Media Nutzung)
- Frühwarnsysteme und Kreditüberwachung
- Use for pricing, marketing and managing insurance policies
- Kapital- und RWA-Optimierung
- RegTech and regulatory compliance
- Modellrisikomanagement (Modellvalidierung) und Stress-Testing
- Risikomanagement (opRisk und Cyberrisk, Marktrisiko und Kreditrisiko)
- Portfoliooptimierung
- Green Finance Anwendungen (wie ESG-Ratings, Klimastresstest, Offenlegung SFDR Art. 8, 9)
- Trade and Market Order Execution
- Market Impact Assessment
- AML/TF, sowie Compliance und Fraud Detection

Auch immer mehr Versicherungen (bspw. zum Aufdecken von Versicherungsbetrug aber auch in der Schadensfallmodellierung) und sogenannte smart FinTechs nutzen KI in ihren Geschäftsmodellen (FinTechs verschiedener Bereiche wie bspw. Banking-Techs, LendTechs, WealthTechs, TradeTechs, PayTechs, InsurTechs, RiskTechs und RegTechs).
Zudem sind viele kleinere, öffentlich meist noch relativ unbekannte Firmen wie Dydon AI (bspw. Kooperationspartner vieler deutscher Landesbanken), Kensho Technologies, AlphaSense, Enova, Scienaptic AI, Socure oder Vectra AI im KI - Bereich der Finanzbranche tätig.
Mittlerweile existiert auch eine Vielzahl an KI-basierten Softwareanwendungen und Apps für die Bankenbranche, häufig zudem Cloud-basiert oder auch als SaaS/SaaP.
Die Software Feedzai® ist bspw. im Bereich Anti-Fraud und AML weit verbreitet, im Bereich Kreditrisiko und Kreditscoring ZestAI®, FIS Credit Intelligence®, Moody's Analytics RiskAuthority®, SAS Risk Management for Banking® oder ZenRisk®, hingegen im

Trading/Markets u. a. LOXM® von JPMorgan.

Die Verwendung von Methoden der künstlichen Intelligenz wird Sektorübergreifend [25] als auch speziell im Finanzbereich – wenngleich mit temporären Akzentverschiebungen – weiterhin zunehmen, insbesondere bei Projekten mit Schnittstellen zum Meldewesen und Risikomanagement können wir sie hierbei unterstützen. Sprechen Sie uns hierzu gerne an (info@1plusi.de).

Bibliographie:

[1] https://arxiv.org/pdf/2303.08774.pdf

[2] https://www.heise.de/news/Betrug-Fake-News-krimineller-Code-Europol-warnt-vor-Missbrauch-von-ChatGPT-8116268.htMachine Learning

[3] https://futureoflife.org/open-letter/pause-giant-ai-experiments/

[4] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32022R0858

[5] https://www.bis.org/bcbs/publ/d545.htm

[6] https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.129.6542

[7] https://list01.biologie.ens.fr/wws/d_read/machine_learning/SubspaceClustering/CLIQUE_algorithm_grid-based_subspace_clustering.pdf

[8] Campello, R. J. G. B., Kröger, P., Sander, J., & Zimek, A. (2020). Density-based clustering. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(2), [e1343]. https://doi.org/10.1002/widm.1343

[9] A First Course on Time Series Analysis — Examples with SAS, by Chair of Statistics, University of Würzburg, Prof. Falk. Version 2011.

[10] Box, G. E., Jenkins, G. M., and Reinsel, G. (1994). Times Series Analysis: Forecasting and Control. Prentice Hall, 3rd edition.

[11] https://www.merl.com/publications/docs/TR2001-22.pdf

[12] https://www.statistik.tu-dort-mund.de/fileadmin/user_upload/Lehrstuehle/Datenanalyse/Wissensentdeckung/Wissensentdeckung-Li-5_2x2.pdf

[13] Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., Soricut, R. (2020). ALBERT: A Lite BERT for Self-supervised Learning of Language Representations, ICLR 2020. https://arxiv.org/pdf/1909.11942.pdf

[14] https://towardsdatascience.com/the-quiet-semi-supervised-revolution-

edec1e9ad8c

[15] https://www.fsb.org/wp-content/uploads/P011117.pdf

[16] Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für Künstliche Intelligenz (Gesetz über Künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206

[17] https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52021AB0040

[18] https://www.bankingsupervision.europa.eu/press/publications/newsletter/2019/htMachine Learning/ssm.nl191113_4.en.htMachine Learning

[19] https://www.eba.europa.eu/sites/default/documents/files/document_library/Final%20Report%20on%20Big%20Data%20and%20Advanced%20Analytics.pdf

[20] https://www.bafin.de/SharedDocs/Downloads/DE/Aufsichtsrecht/dl_Prinzipienpapier_BDAI.pdf?__blob=publicationFile&v=4

[21] https://www.iif.com/Portals/0/Files/content/Research/iif_Machine Learningcr_2nd_8_15_19.pdf

[22] https://www.eba.europa.eu/sites/default/documents/files/document_library/Publications/Discussions/2022/Discussion%20on%20machine%20learning%20for%20IRB%20models/1023883/Discussion%20paper%20on%20machine%20learning%20for%20IRB%20models.pdf

[23] https://builtin.com/artificial-intelligence/ai-finance-banking-applications-companies

[24] https://arxiv.org/ftp/arxiv/papers/2107/2107.09051.pdf

[25] https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review

© 2023 1 PLUS i GmbH

© 2023 1 PLUS i GmbH