

Use Case First - AI Second

Welche AI ist die beste für mein Unternehmen? Und was möchte ich damit erreichen?

von Sebastian Petry

Eine Künstliche Intelligenz ist in erster Linie ein Werkzeug, eine Technologie. Um sie gewinnbringend einzusetzen, muss sie mit einem oder mehreren wertschöpfenden Produkten verbunden werden. Dabei gilt: Use Case First – AI Second!

Künstliche Intelligenz gehört zu den wichtigsten Triebfedern der digitalen Revolution. Schon heute verändert sie die Art, wie Organisationen agieren und Entscheidungen treffen, und sie hilft Unternehmen dabei, schneller und effizienter zu agieren. Nähert man sich dem Thema Künstliche Intelligenz, trifft man schnell auf die Unterscheidung zwischen starker und schwacher KI. Während das Ziel einer starken Künstlichen Intelligenz darin besteht, intellektuelle Fertigkeiten wie intelligentes und flexibles, nicht reaktives Handeln oder die Transferleistung von Handlungsstrategien zu erlangen oder sie zu übertreffen, stellt eine schwache Künstliche Intelligenz das Nachahmen oder Automatisieren menschlichen Verhaltens in den Mittelpunkt. Zwei wesentliche Aspekte sind für sie charakteristisch:

1. Sie verfügt über einen zum Teil selbstständigen Lerneffekt als Resultat aus sich ändernden äußeren Bedingungen.
2. Ihr Algorithmus kann mit äußerst komplexen und unvorhergesehenen Daten umgehen.



Man unterscheidet zwischen vier Erscheinungsformen von AI. Welche davon kommt für das Unternehmen infrage?

Die gegenwärtige Debatte unterscheidet im Wesentlichen zwischen vier verschiedenen Erscheinungsformen von AI, die nicht alle diesen beiden Kriterien gerecht werden, die aber je nach Anwendungsfall und Umfeld im Unternehmen ihre Berechtigung haben.

Zum einen gibt es da die echte AI: selbstfahrende Autos, Bots, die komplexe Spiele wie Schach, Go oder Dota2 besser spielen als jeder Mensch und dabei selbstständig neue Spielstrategien entwickeln, oder auch Sprachassistentenprogramme, die im täglichen Leben durch verbale Interaktion Fragen des Alltags beantworten und sich von Tag zu Tag verbessern. Dabei handelt es sich um Algorithmen (oft komplexe Formen von tiefen neuronalen Netzen und deren Kombination in komplexen Frameworks wie Reinforcement Learning), die in einem bestimmten Umfeld Aktionen ausführen und fest an ein Produkt gekoppelt sind, wie auch in den Marketingplattformen von Facebook und Google. Die echte AI hat die schwache Intelligenz als Zielbild und einen eigenen Anspruch.

Der echten AI gegenüber steht die Marketing-AI. Hierbei handelt es sich um „alten Wein in neuen Schläuchen“. Typischerweise werden dabei Verfahren wie Regressionen oder Ensemble-Methoden (Random Forests, Boosted Trees, Support Vector Machines etc.) eingesetzt. Durch die Algorithmen werden Handlungsempfehlungen ausgesprochen oder Aktionen direkt ausgelöst. Eine klassische Anwendung ist die Prognose von Kundenverhalten (etwa Kündigungsprognose) mit logistischen Regressionsmodellen [Hof18] und der daraus resultierende automatische Versand einer passenden E-Mail an „gefährdete“ Kunden. Im Allgemeinen sind die Datensituationen nicht hochkomplex, und das selbstständige Lernen ist nicht Bestandteil der Anwendungsfälle.

Verwandt mit der Marketing-AI ist die Methoden-AI. Hier wird Deep Learning oft mit AI gleichgesetzt. Beim Deep Learning werden sogenannte „tiefe neuronale Netze“ eingesetzt, die ihrerseits auch als Methode des Machine Learning eingeordnet werden könnten. Es gibt verschiedene Architekturen und Varianten von neuronalen Netzen, die bereits für diverse Anwendungsfälle wie Image und Speech Recognition, Weblog Analytics, Text Mining etc. entwickelt und bereitgestellt wurden. Andererseits besteht die Möglichkeit, Architekturen selbst zu entwickeln. Tools wie Keras ermöglichen einen guten und intuitiven Zugang zu dieser mächtigen Modellklasse. Erst neue technische Möglichkeiten und Big-Data-Technologien machen es möglich, tiefe neuronale Netze in praxisrelevanter Zeit zu rechnen und somit in produktive Prozesse einzubinden. Tiefe neuronale Netze sind sehr datenhungrig, und es bedarf einer großen Menge an Daten, um diese Netze zu trainieren. Nicht ohne Grund kommen viele Deep-Learning-Frameworks und Architekturen von Technologiekonzernen wie Microsoft (CNTK), Facebook (Torch) oder Google (TensorFlow). In vielen Fällen ist die Methoden-AI eine Vorstufe der echten AI.

Die vierte Erscheinungsform ist die Baukasten-AI. Dabei handelt es sich um vortrainierte Bausteine, die man auf Produkte adaptieren kann, wie etwa Sprach- oder Chatbots im Kundenservice oder zur kognitiven Steuerung von Dashboards. Sie werden häufig für den jeweiligen Zweck als vorgefertigte Lösungen und On-Premise oder in der Cloud bereitgestellt. Diese können dann in eigene Produkte integriert werden und sind fester Bestandteil von Produkten. Gute Baukasten-AI-Elemente kommen der menschlichen Handlungsweise sehr nahe. Die Qualität derartiger Bausteine und Produkte ist jedoch sehr unterschiedlich und muss von Fall zu Fall eingehend geprüft werden.

Für Unternehmen mit einem geringeren digitalen Reifegrad wird es in den meisten Fällen auf die Einführung von Marketing- oder Baukasten-AI hinauslaufen. Eine Baukasten-AI lässt sich leicht durch externe Experten in ein Unternehmen und dessen Produkte integrieren. Eine Marketing-AI hingegen wird oft durch eigene Data-Science-Abteilungen im Unternehmen bedient. Ist die Marketing-AI bereits etabliert, besteht die Möglichkeit, zusätzlich die Methoden-AI einzuführen, was jedoch den Aufbau von technischem und methodischem Expertenwissen verlangt. Entweder verschmelzen Marketing- und Methoden-AI, oder die Methoden-AI wird an die Produkte des Unternehmens angegliedert und nicht an die Data-Science-Abteilungen. Inwieweit die jeweilige AI den beiden Kriterien von schwacher Intelligenz entspricht, hängt vom Nutzen und ihrer Implementierung ab. „Echte AI“ zielt meistens auf ein hoch spezialisiertes Produkt ab und zählt oft zum Kerngeschäft eines Unternehmens, wie Amazons Alexa, selbstfahrende Autos oder Gesichtserkennung im Security-Umfeld.

Einführung einer AI setzt ein hohes Maß an Vertrauen voraus

Eine AI wird in Unternehmen eingesetzt, um bestehende Prozesse oder Produkte zu automatisieren und zu optimieren oder neue Produkte oder Prozesse ins Leben zu rufen oder zu gestalten. Die Ausarbeitung des einzelnen Use Case ist dabei Aufgabe des Fachbereichs. Mögliche Use Cases sind die Personalisierung einer Website, die Einführung eines kognitiven KPI-Dashboards, die Implementierung von Chat- oder Sprach-Bots im Customer-Service oder eine intelligente Kampagnen-Lösung.

Um aus dem Use Case ein Produkt zu entwickeln, sollten Nutzen und Erwartung an das Produkt (und damit an die AI) von allen Stakeholdern abgefragt und berücksichtigt werden. Die Einführung einer AI setzt ein hohes Maß an Vertrauen in die jeweilige AI-Lösung voraus. Denn oft geht es direkt um das jeweilige Produkt (etwa eine Website), in dem die AI-Lösung konkrete Aufgaben und damit auch Verantwortung übernimmt. Das Vertrauen in die Lösung muss vorab gegeben sein. Eine AI-Lösung ermöglicht es eben nicht, Entscheidern einen Einblick in die wichtigsten Einflussgrößen einer Vorhersage, geschweige denn bildliche Beschreibungen der Wirkzusammenhänge zu geben. Leichte, intuitive Bedienbarkeit des Produkts und einfache Integrierbarkeit sind beliebte Verkaufsargumente, die jedoch nicht davon ablenken dürfen, wie gewissenhaft man die richtige Entscheidung treffen sollte.

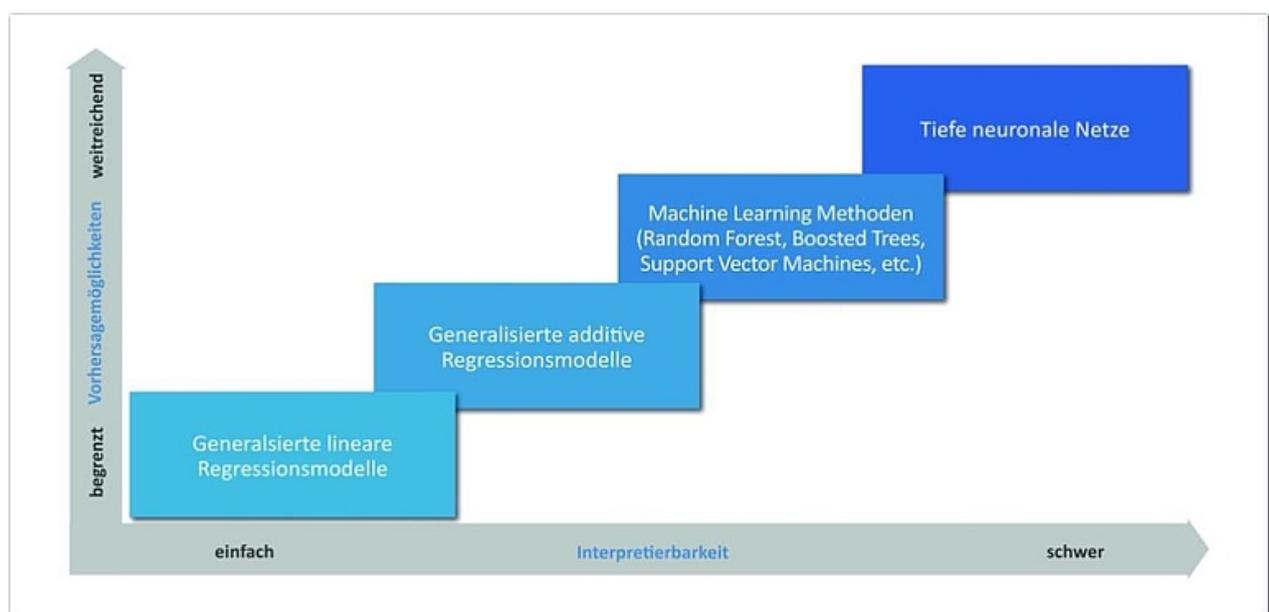
Schlussendlich sollte bei der Implementierung einer AI immer ein möglichst neutrales Überwachungskonzept umgesetzt werden, um die Effekte der AI auf das Geschäft kontrollieren und optimieren zu können: Vertrauen ist gut – Kontrolle ist besser.

Realistische Einschätzung des digitalen Reifegrads berücksichtigen

Hilfe bei der Auswahl der AI-Komponenten holt sich der Fachbereich am besten von internen Abteilungen wie Data Science, BI und IT. Bei einem derart innovativen Thema empfiehlt es sich, zusätzlich auch neutrale Experten hinzuzuziehen, um fehlende Expertise auszugleichen. Die externe Beratung ist vor allem notwendig, wenn der Reifegrad des Unternehmens in diesen Themen gering bis mittel ist. Oft handelt es sich um eine Make-or-Buy-Entscheidung zwischen Baukasten-AI und Marketing- bzw. Methoden-AI. Darüber hinaus fließen dabei auch Aspekte wie Time-to-Market, das Abhängigkeitsverhältnis zu Anbietern, Implementierungsaufwand und externe Kosten mit in die Bewertung ein.

Wie wichtig ist mir das Wissen und Verständnis über Prozesse und Entscheidungen, die zukünftig eine AI in meinem Unternehmen trifft? Möchte ich verstehen, warum Kunde A genau das Angebot B erhalten hat? In den meisten Fällen wird eine AI direkter Bestandteil der Wertschöpfungskette und damit stark in das Produkt integriert sein. Es gilt deshalb, das Vertrauen in eine Blackbox gegen eine Inhouse-Entwicklung abzuwägen, sofern eine solche möglich ist. Bei diesem Entscheidungsprozess, aber auch bei der Festlegung von Nutzen und Erwartung an die AI ist eine realistische Einschätzung des digitalen Reifegrades unbedingt zu berücksichtigen. In einem solchen Entscheidungsprozess darf der Hype um AI die Sinne nicht trüben. Schließlich geht es in erster Linie um Ihr Geschäft.

Wann welche Methode einsetzen?



Je flexibler die Methode, desto weitreichender die Vorhersagemöglichkeiten und desto schwerer die Interpretierbarkeit

Möchte man vorrangig mithilfe datenanalytischer Methoden wie AI, Data Science oder Machine Learning den „digitalen Wandel“ in seinem Unternehmen vorantreiben, stellt sich die Frage der Methodik erst später und man klärt für sich zuerst folgende Fragen:

Use Case: Wie lauten die genaue Definition, und was ist ein klares, abteilungsübergreifendes einheitliches Verständnis des Use Case bzw. des Produkts?

Daten: Welche Art und Menge von Daten werden anfallen?

Technik: Auf welche Technologien wird (nicht) vertraut?

In vielen Fällen geht es in der Methodendiskussion fachlich um den Trade-off zwischen Interpretierbarkeit und Vorhersagegüte. Die Vorhersagegüte hängt dabei primär von der Datensituation ab. Zur Lösung von Vorhersageproblemen bei komplexen Datensituationen benötigt man komplexe Verfahren, die per se schwerer zu interpretieren sind. Oft ist der geringe Performance-Gewinn durch die Verwendung komplexerer Verfahren bei einfacheren Datensituationen eher gering und geht auf Kosten der Interpretierbarkeit.

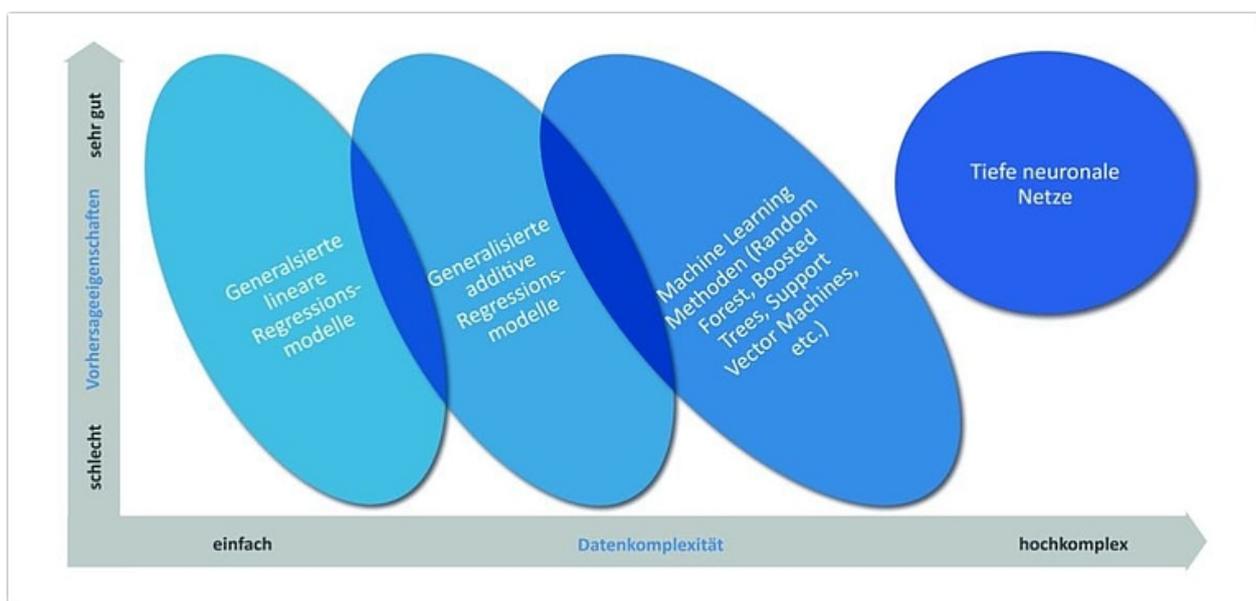
Viele klassische statistische Methoden (zum Beispiel Regressionsmodelle, generalisierte [additive] Regressionsmodelle mit Regularisierung [Büh07; Woo17]) haben den großen Vorteil der guten Interpretierbarkeit. Bei Machine-Learning-Methoden (zum Beispiel Random Forests, Gradient Boosted Tree, Support Vector Machine etc. [HTF08]) ist dies nur eingeschränkt möglich und bei tieferen neuronalen Netzen [bte18; HiH18; Lan16; Sch15] nur noch mit gesonderten Methoden und in speziellen Fällen (zum Beispiel LIME – Local Interpretable Model-Agnostic [RSG16]). Generell gilt: Mit steigender Flexibilität der Methode steigt, insbesondere bei komplexen und hochkomplexen Datenstrukturen, die Prognosegüte, was zum Verlust der Interpretierbarkeit führt.

Wie ändert sich die Arbeitsweise eines Data Scientist?

Die Nutzung schwer interpretierbarer Methodenklassen ohne signifikanten Performance-Gewinn verschenkt viele wertvolle Erkenntnisse gegenüber statistischen Regressionsmodellen. Dadurch, dass bei Methoden des Machine Learning das Feature Engineering oft von den Methoden übernommen wird, geht die direkte Arbeit mit den Daten und somit auch das Gefühl für die Daten verloren. Das wird besonders bei neuen Datenquellen deutlich, die nicht mehr so genau erschlossen werden, wie es mit klassischen Methoden notwendig war, um sie bestmöglich zu nutzen. Die Suche nach den kleinen und großen Nuggets im Datenschatz wird nicht mehr so stark durch die Datenexploration und -analyse im Feature Engineering unterstützt. Andererseits sind die modernen Methoden sehr mächtig und geben uns die Möglichkeit, völlig neue Use Cases umzusetzen.

Der Data Scientist rückt näher an das Produkt

Insbesondere durch die tiefen neuronalen Netze wird die Klaviatur des modernen Data Scientist deutlich vergrößert, wobei sich auch die Arbeitsweise wandeln wird. Der Schwerpunkt wird sich stärker auf das Algorithmenvverständnis sowie auf die stetige Validierung der Ergebnisse und Methoden verschieben. So wird er sich schnell neue Architekturen von Netzwerken aneignen müssen, um diese dann optimal für seine Use Cases zu nutzen.



Bei tieferen neuronalen Netzen ist eine Interpretierbarkeit nur noch mit gesonderten Methoden möglich.

Der Data Scientist muss nicht mehr alles selbst umsetzen, sollte aber gut verstehen, was er selbst nicht entwickelt hat. Zusätzlich wird sein Aufgabenbereich näher an das Produkt heranrücken, sodass er sich stärker mit technischen und anderen Nebenbedingungen auseinandersetzt. Wie bereits oben erwähnt, sind Bilder, Filme, Weblogs, geschriebene und gesprochene Sprache Datenarten, für die sich tiefe neuronale Netze bei Prognosemodellen hervorragend eignen und die Benchmark stellen. Allein die Anzahl von Use Cases, die durch die Nutzung von Bildern, Text und Sprache möglich sind, machen tiefe neuronale Netze zu einer zentralen Methode der AI und der modernen Data Science, geht es bei AI doch idealerweise um das Nachahmen menschlichen Verhaltens und Lernens, das fraglos durch Bilder, Text und Sprache stark bestimmt ist.

Das Verständnis für Algorithmen wird wichtiger

Neue Methoden stellen Unternehmen vor neue technische Herausforderungen hinsichtlich Performance und Maintenance, die wieder eine Make-or-Buy-Diskussion aufkommen lassen werden. Auch neue AI im analytischen Bereich wie Watson von IBM, Einstein von salesforce sowie eine Vielzahl von AI-Plattformen in unterschiedlichen Cloud-Umgebungen (Microsoft Azure, Google Cloud, AWS etc.) werden die Arbeitsweise eines Data Scientist nachhaltig verändern. Die Arbeitsgeschwindigkeit wird höher werden, da vieles vor dem Open-AI-Gedanken schnell verfügbar gemacht wird und somit nicht selbst entwickelt werden muss. Das Verständnis für die Algorithmen wird wichtiger werden, um nicht selbst entwickelte Algorithmen optimal zu nutzen. Ein weiterer Schwerpunkt wird die Ergebnisdarstellung und Plausibilisierung werden, mit der man die Fachbereiche überzeugt, ohne ihnen den Algorithmus oder die Wirkungsweise im Detail verständlich machen zu müssen. 80:20-Standardlösungen werden mit diesen AI-Plattformen schnell bereitgestellt werden können. Optimierte Lösungen werden aber weiterhin viel manuellen Aufwand bedeuten.

Was muss ein Unternehmen tun, um AI-ready zu werden?

Erfolgreiche AI muss wie Data Science, Reporting oder auch Kampagnenmanagement zum integralen Bestandteil einer BI-Strategie werden. Bisher stellt die Business Intelligence eines Unternehmens vor allem die Datengrundlage und die Strukturen bereit, auf deren Basis der Mensch Handlungen ableitet. Zukünftig wird die BI auch Datengrundlagen und andere Strukturen schaffen, um automatisierte Handlungen zuverlässig auszuführen und nachzuhalten, wie dies in Unternehmen mit hohem digitalem Reifegrad bereits heute der Fall ist. Dies wird auf adäquaten Plattformen und Applikationen der IT fußen, als fachliche Basis eine strukturierte Reporting-Landschaft besitzen und aufgabengerechte Business-Datenlayer nutzen.

AI ist neben aller Faszination, die von den Algorithmen ausgeht, ein sehr technisches Thema, das es im Unternehmen zu verankern gilt und das mit Big-Data-Technologien bearbeitet werden muss. Gerade das Deployment und der produktive Betrieb von modernen Methoden sind für viele Unternehmen eine Herausforderung.

Was nicht in den Daten steckt, kann kein Algorithmus finden

Wenn ein Unternehmen die Themen AI und Digitalisierung auf seine Agenda hebt, braucht es eine solide Datenstrategie. Dazu gehört insbesondere das Aufbrechen von Datensilos unter Einhaltung der Datenschutzbestimmungen. Oft gilt zwischen Abteilungsleitern von Unternehmen die Losung: „Wer die Daten hat, hat die Macht.“ Diese Denkweise hat in einem erfolgreichen Digitalisierungsprozess keinen Platz mehr. Einzelne Daten-Owner haben oft nur ein geringes Interesse an der generellen Verfügbarkeit ihrer Daten. Dieses Thema sollte top-down angepackt werden. Langwierige Entscheidungsprozesse, ein starres Projektmanagement und Deploymentzyklen sowie eine geringe Erfahrung mit moderner Fehlerkultur bremsen das aus, was für eine erfolgreiche AI am wichtigsten ist:

1. Mut zur Nutzung neuer Möglichkeiten mit realistischer Erwartungshaltung – auch ohne Erfolgsgarantie
2. Die Kreativität, aus den Möglichkeiten der AI schnell neue Produkt-Features und Produkte zu entwickeln
3. Die Möglichkeit, der Erste am Markt zu sein und sich damit einen Wettbewerbsvorteil zu sichern

Wo liegen die Grenzen von AI? Stephan Wolfram beantwortet diese Frage wie folgt: „What will AI allow us to automate? We'll be able to automate everything that we can describe. The problem is: it's not clear what we can describe.“ [Fri18] Diese Aussage bringt die großen Möglichkeiten von AI zum Ausdruck, die wir nutzen sollten, ruft aber auch ganz klar dazu auf, dass wir uns erst über den Use Case klar werden, bevor wir nach einer AI rufen. Das impliziert auch, dass ein Use Case konsequent zu Ende

gedacht werden muss und hierbei auch ethische und rechtliche Aspekte unbedingt zu berücksichtigen sind.

AI bietet viele Chancen für neue Geschäftsmodelle, die wir nutzen sollten. Man übergibt der AI Aufgaben, die zuvor ein Mensch ausgeführt, beeinflusst oder beobachtet hat. Dadurch werden menschliche Entscheidungen oder Handlungen aus einem Prozess entfernt und durch eine Künstliche Intelligenz ersetzt. Ihre Effizienz und Qualität hängt in hohem Maße von den richtig ausgeführten Schritten ab, die uns an diesen Punkt führen und in unserer technischen, fachlichen und ethischen Verantwortung liegen.

Literatur

- [bte16]** b.telligent Blog: (2016), Deep Learning im Scheinwerferlicht. März 2016, <http://www.btelligent.com/blog/filter/deep-learning-im-scheinwerferlicht/>, abgerufen am 18.3.2018
- [BüH07]** Bühlmann, P. / Hothorn, T.: Boosting algorithms: regularization, prediction and model fitting (with discussion). In: Statistical Science 22, 2007, S. 477-505
- [Fri18]** Fridman, L.: Artificial General Intelligence. MIT 6.S099, März 2018, <https://agi.mit.edu>, abgerufen am 29.5.2018
- [HiH18]** Higham, C. / Higham, D.: Deep Learning: An Introduction for Applied Mathematicians. 2018, arXiv:1801.05894v1
- [Hof18]** Hoffmann, C.: Ode an die logistische Regression. 2018, <http://www.btelligent.com/blog/ode-an-dielogistische-regression/>, abgerufen am 18.3.2018
- [HTF08]** Hasti, T. / Tibshirani, R. / Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer 2008
- [Lan16]** Lang, T.: Deep Learning for Understanding Consumer Histories. 2016, https://jobs.zalando.com/tech/blog/deep-learning-for-understanding-consumer-histories/?gh_src=4n3gxh1, abgerufen am 18.3.2018
- [RSG16]** Ribeiro, M. T. / Singh, S. / Guestrin, C.: „Why Should {I} Trust You?": Explaining the Predictions of Any Classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, August 2016, S. 1135-1144
- [Sch15]** Schmidhuber, J.: Deep learning in neural networks: An overview. In: Neural Networks, Vol. 61, 2015, S. 85-117
- [Woo17]** Wood, S. N. et al.: (2017) Generalized Additive Models for Gigadata: Modeling the U. K. Black Smoke Network Daily Data. In: Journal of the American Statistical Association, 112:519, 2017, S. 1199-1210



Dr. Sebastian Petry

ist Data Scientist bei b.telligent mit langjähriger Erfahrung bei der erfolgreichen Umsetzung von Projekten in diversen Branchen. Als promovierter Statistiker verfügt er über fundiertes Wissen in der Methodik und Algorithmik klassischer Machine-Learning-Verfahren, wobei Deep Learning für ihn eine natürliche Erweiterung der Data-Science-Werkzeugkiste ist.

E-Mail: [sebastian.petry\(at\)btelligent.com](mailto:sebastian.petry(at)btelligent.com)

Bildnachweise:

b.telligent

[Online Themenspecial](#)

[Impressum](#)

|

[Kontakt & Anfrage](#)