

# KÜNSTLICHE INTELLIGENZ IM BANKING: DEEP LEARNING VS. GRADIENT BOOSTING

---

Eine Gegenüberstellung

Bastian Berghaus, Julius Heitmann

30.10.2018

COREtechmonitor | Blogpost

Copyright © CORE GmbH

## Key Facts

- › Datenanalyse und Künstliche Intelligenz von zunehmender gesellschaftlicher wie wirtschaftlicher Relevanz
- › Deep Learning als hoch performantes Modell der Künstlichen Intelligenz häufig im Mittelpunkt der Diskussion
- › Gradient Boosting auf Basis von Entscheidungsbäumen als performantes (Funktion) und interpretierbares (Transparenz) Modell als Alternative in der Finanzindustrie

## Report

„Künstliche Intelligenz“ ist eines der aktuell meistdiskutierten Themen in der Öffentlichkeit. Geschehnisse wie die Auswertung von Wählerdaten durch die Cambridge Analytica für die US-Präsidentenwahl 2016 oder die Erzeugung einer durch viele Beteiligte als menschlich eingeschätzte Konversation mit der neuen Generation des Google Assistenten schüren zwar Befürchtungen wie Begeisterung für die neue Technologie; jedoch erlaubt erst die Unterscheidung einzelner Methoden der Künstlichen Intelligenz sinnvolle Anwendungsfälle und Nutzenaspekte kritisch zu diskutieren, d.h. ihre Leistung und Grenzen in wirtschaftlicher Dimension einzuschätzen und vor dem Hintergrund sozialer und rechtlicher Kontexte zu beurteilen.

### Einsatz Künstlicher Intelligenz in der Wirtschaft

Die Geschäftsmodelle der Technologieriesen Google, Apple, Facebook und Amazon (GAFAs) genauso wie die von Baidu, Alibaba und Tencent (BATs) werden durch Künstliche Intelligenz erst ermöglicht. Die Analyse kundenseitiger Daten, z.B. Stamm-, Nutzungs- oder Transaktionsdaten erlaubt die passenden Interaktionen, Empfehlungen oder Suchergebnisse kundenindividuell zu ermitteln. Mit ihren Produkten im Bereich Personal Assistance (Siri, Alexa und Google Home) bringen sie Künstliche Intelligenz in den Alltag der Kunden und unterstützen sie

bei täglichen Problemen. Huawei und Baidu auf dem chinesischen Markt kündigen Konkurrenzprodukte noch für 2018 an. Die Folge ist eine stetig wachsende Datenbasis dieser Unternehmen; dies wiederum ermöglicht ihnen die kontinuierliche Präzisierung von Aussagen über künftiges Kundenverhalten.

Auch in der Finanzbranche gewinnt Datenanalyse zunehmend an Bedeutung und wird als zukünftig differenzierender Wettbewerbsfaktor angesehen. Beispielsweise führte JPMorgan Chase jüngst die Anwendung „Contract Intelligence“ ein. Damit ist eine sekundenschnelle Analyse von 12.000 kommerziellen Kreditverträgen zur Extraktion wichtiger Datenpunkte und Klauseln möglich. Das manuelle Vertragsreview würde etwa 360.000 Stunden in Anspruch nehmen. Unternehmen wie Scalable Capital nutzen Algorithmen, um das Trading von ETF-Portfolios zu automatisieren und manuelle Prozesse zu minimieren. Zusätzlich profitiert Scalable Capital von den Fähigkeiten der Algorithmen, aus großen Mengen an Informationen systematisch und im Sekundenbruchteil zu abstrahieren. Dadurch werden schnelle Reaktionszeiten bei komplexer Informationsbasis ermöglicht, um so die Rendite zu steigern.

### Deep Learning- vs. Gradient Boosting-Modell

Die Technologie, die bei dieser Diskussion am stärksten im Fokus steht, ist das Deep Learning, ein künstliches neuronales Netz,

bei dem Funktionselemente – auch als Neuronen bezeichnet – in einem geschichteten Netzaufbau dazu genutzt werden, Informationen sukzessive zu verarbeiten. Seit dem Jahr 2006 sind diese Modelle auf dem Gebiet der hochdimensionalen Klassifikation – der Zuordnung von Datenpunkten zu einer begrenzten Anzahl an Klassen bei einer viele Variablen umfassenden Problemstellung – nicht mehr wegzudenken. Besonders auf dem Gebiet der Bild-, Sprach- und Handschrifterkennung gibt es aktuell kaum Alternativen zu Deep Learning. Ihre Stärke hinsichtlich der Lösung von Problemstellungen beziehen diese Modelle aus ihrem geschichteten Aufbau. Hierbei wird die komplexe Information Schicht für Schicht abstrahiert, bis schließlich in der letzten Schicht eine Entscheidung bzgl. einer vorher definierten Problemstellung getroffen wird.

Warum aber wird diese Technologie nicht bei jeder Künstlichen Intelligenz implementiert? Der komplexe Aufbau eines Deep Learning-Netzes birgt auch Nachteile. Um den schichtenbasierten Aufbau zu trainieren, wird in Abhängigkeit der Tiefe des Modells, d.h. der Anzahl der implementierten Schichten, eine vergleichsweise große Menge an Daten und ein hoher Einsatz an Rechenleistung benötigt. Zudem kann es abhängig vom Anwendungsfall erforderlich sein, die Gründe für eine durch das Netz getroffene Entscheidung nachvollziehen zu können. Gerade in gesetzlich stärker reglementierten Kontexten, wie Banken oder Versicherungsunternehmen, kann die Nachvollziehbarkeit eine zwingende Voraussetzung sein. Hier trifft es sich, dass die BaFin in ihrer aktuellen Publikation aus der Reihe BaFin Perspektiven 1/2018 unmissverständlich klargestellt hat, dass die Aufsicht eine Argumentation, wonach die Funktionsweise eines Algorithmus aufgrund

seiner Komplexität nicht mehr nachvollziehbar sei und daher auch nicht mehr aufsichtlich geprüft werden könne, nicht akzeptiert.

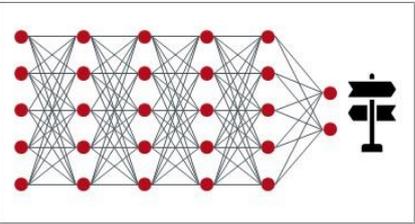
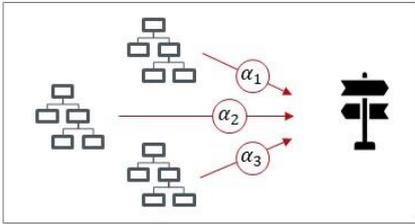
Deep Learning-Netze bieten mit ihrer hohen Anzahl an Neuronen und Verknüpfungen eine hohe Lernkapazität. Doch genau dieser Vorteil verhindert, dass eine Entscheidung eines Netzes leicht bzw. überhaupt nachvollzogen werden kann.

Eine vielversprechende Alternative zum Deep Learning-Ansatz bildet das Ensemble Learning, in dessen Rahmen verschiedene Methoden und Algorithmen kombiniert werden und das insbesondere im Kontext von Entscheidungsbäumen eingesetzt wird. In den letzten Jahren hat sich als Unterart das Gradient Boosting in verschiedenen öffentlichen Ausschreibungen durchsetzen können. Statt die Entscheidung entlang eines einzelnen, großen Entscheidungsbaums zu treffen, vereint Gradient Boosting eine Menge kleinerer Entscheidungsbäume, die jeweils trainiert werden, die Schwächen anderer Bäume zu mitigieren. Die Ergebnisse des jeweiligen Baumes werden gewichtet und mit den Ergebnissen der weiteren Bäume zu einer gemeinsamen Entscheidung zusammengeführt.

Metaphorisch gesprochen bildet diese Methode eine Gruppenentscheidung von Experten ab, wobei jeder Teilexperte für einen bestimmten Bereich des Lösungsraums optimal entscheidet und seine Entscheidung gewichtet in die Gesamtentscheidung eingeht. Dies führt im Gegensatz zu Deep Learning-Modellen dazu, dass eine Entscheidung des Modells nachvollzogen werden kann, indem die Bäume mit dem höchsten Beitrag zur Entscheidung identifiziert werden, wobei die Entscheidungen der Teilbäume leicht nachvollziehen sind. Zudem wird die Komplexität der Einzelbäume im Training oft beschränkt, wodurch die Nachvollziehbarkeit weiter verbessert wird. Diese beschränkte

Komplexität bildet einen weiteren Vorteil gegenüber dem Deep Learning-Ansatz, da weniger Daten zum Training benötigt werden. Zudem weist das Gradient Boosting einen relativ geringen Trainingsaufwand zu anderen hochdimensionalen Klassifikatoren auf. Im Ergebnis verspricht Gradient Boosting, mit einem kleinen Datensatz zu einem konvergierten Trainingsergebnis zu führen.

### Je nach Kontext kann Gradient Boosting bei hoch-dimensionalen Klassifikationsproblemen eine passende Alternative darstellen

Deep Learning	Dimensionen	Gradient Boosting
	<p><b>Graphische Darstellung</b></p>	
<ul style="list-style-type: none"> <li>▪ Neuronen sind in Schichten angeordnet</li> <li>▪ Schichtweise Abstraktion der Eingangsinformation, genannt feature extraction</li> <li>▪ Gestaltungsfreiheit in Netzwerkarchitektur und Typ der gewählten Neuronen</li> </ul>	<p><b>Technische Beschreibung</b></p>	<ul style="list-style-type: none"> <li>▪ Ensemble mehrerer Entscheidungsbäume</li> <li>▪ Entscheidung als gewichtetes Scoring der Einzelentscheidungen der Bäume</li> <li>▪ Gestaltungsfreiheit in Hyperparametern des Ensembles sowie der Entscheidungsbäume</li> </ul>
<p>Kapazität um hochkomplexe Muster zu lernen <span style="float: right;">+</span></p> <p>Flexible Modellarchitektur <span style="float: right;">+</span></p> <p>Open source community für Handschrift-, Sprach- und Bilderkennung <span style="float: right;">+</span></p>	<p><b>Vorteile der Modelle</b></p>	<p><span style="float: left;">+</span> Leicht interpretierbare Modelle und Ergebnisse</p> <p><span style="float: left;">+</span> Niedrigerer Trainingsaufwand</p> <p><span style="float: left;">+</span> Open source community für hochdimensionale Klassifikationsprobleme</p>

---

## Resümee

Eine Aussage zu Dominanz eines der beiden angesprochenen Modelle auf einem bestimmten Gebiet ist nicht möglich. Beide Modelle leisten gute Ergebnisse bei der Klassifikation hochdimensionaler Daten, wobei die Performance abhängig ist von zugrundeliegenden Daten und den darin enthaltenen Mustern. Einen Spezialfall stellt dabei die Aufgabe der Sprach-, Bild- und Handschrifterkennung dar, bei dem in nahezu allen Systemen heutzutage Deep Nets implementiert werden.

Eine weitere zu beachtende Entscheidungsdimension - insbesondere in der Finanzwelt - stellen Nebenbedingungen des Use Cases dar in Form von Nachvollziehbarkeit der Modellentscheidung oder Konvergenzgeschwindigkeit (Menge der Daten, die benötigt wird bis zum trainierten Modell), die zu

einer Differenzierung der Modelle neben der Vorhersagequalität führen können. Hier hat die BaFin möglicherweise bereits eine Marktvoentscheidung erzwungen.

Abhängig von Use Case und verfügbarer Daten kann Gradient Boosting zu präferieren sein, da es für spezifische Anwendungsfälle und hinsichtlich Transparenz zu besseren Ergebnissen als Deep Learning führt. Anhand verschiedener Use Cases sollen daher die speziellen Vorteile und die Leistung des etwas außerhalb des Rampenlichts stehenden Gradient Boosting-Modells aufgezeigt und anhand von Analysen quantifiziert werden. Der nächste Blogpost wird den Santander Customer Satisfaction Use Case vertiefen, einem im Jahr 2016 öffentlich ausgeschriebenen Wettbewerb zur Prognose von Kundenabwanderung, auch bekannt als Churn Case.

---

## Quellen

### Cambridge Analytica

[German only] "Facebook informiert nach Datenskandal die betroffenen Nutzer", 10. April 2018:  
<https://www.wiwo.de/unternehmen/it/cambridge-analytica-facebook-informiert-nach-datenskandal-die-betroffenen-nutzer/21160858.html>

### Engel, C.

"Artificial Intelligence: The Weapon of Choice in Banks' Fight for Survival", 07. March 2018:  
<https://thefinancialbrand.com/71108/artificial-intelligence-banking-machine-learning/>

### Kaggle: Santander Customer Satisfaction

"Santander Customer Satisfaction - Which customers are happy customers?":  
<https://www.kaggle.com/c/santander-customer-satisfaction>

### Bundesanstalt für Finanzdienstleistungsaufsicht (BaFin)

"BaFin Perspektiven. Ausgabe 1/2018":  
[https://www.bafin.de/SharedDocs/Downloads/DE/BaFinPerspektiven/2018/bp\\_18-1\\_digitalisierung.pdf;jsessionid=88AC19BE9AD943AE14AB8A3B8BDE0EBA.1\\_cid372?\\_blob=publication-File&v=17](https://www.bafin.de/SharedDocs/Downloads/DE/BaFinPerspektiven/2018/bp_18-1_digitalisierung.pdf;jsessionid=88AC19BE9AD943AE14AB8A3B8BDE0EBA.1_cid372?_blob=publication-File&v=17)



**Bastian Berghaus** ist Transformation Associate bei CORE und unterstützt Projekte im Bereich Banking. Der studierte Betriebswirt sammelte erste Berufserfahrung bei verschiedenen Beratungsunternehmen und war bei CORE bereits als Student Professional mit dem Schwerpunkt auf Authentifizierungstechnologien tätig. In seiner neuen Position unterstützt Bastian Klienten der Finanz- und Versicherungsindustrie in der Entwicklung innovativer Lösungen für komplexe IT-Vorhaben.

**Mail: [Bastian.Berghaus@core.se](mailto:Bastian.Berghaus@core.se)**



Als Expert Manager unterstützt **Julius Heitmann** bei CORE im Schwerpunkt die Konzeption komplexer IT-Transformationen. In der Automobilindustrie hat er umfangreiche Erfahrungen im Umgang mit Big Data und Data Analytics sammeln können. Diese bringt er insbesondere in Projekten zur Analyse und Optimierung großer Datenmengen ein und entwickelt passgenaue Steuerungssysteme für die Datenverarbeitung vor allem im Finanzsektor.

**Mail: [Julius.Heitmann@core.se](mailto:Julius.Heitmann@core.se)**

COREinstitute  
Am Sandw erder 21-23  
14109 Berlin | Germany  
<https://institute.core.se>  
Phone: +49 30 26344 020  
[office@coreinstitute.org](mailto:office@coreinstitute.org)

COREtransform GmbH  
Am Sandw erder 21-23  
14109 Berlin | Germany  
<https://www.core.se>  
Phone: +49 30 26344 020  
[office@coretransform.de](mailto:office@coretransform.de)

COREtransform GmbH  
Limmatquai 1  
8001 Zürich | Helvetia  
<https://www.core.se>  
Phone: +41 442 610 143  
[office@coretransform.ch](mailto:office@coretransform.ch)

COREtransform Ltd.  
One Canada Square  
London E14 5DY | Great Britain  
<https://www.core.se>  
Phone: +44 203 319 0356  
[office@coretransform.co.uk](mailto:office@coretransform.co.uk)

COREtransform MEA LLC  
DIFC – 105, Currency House, Tow er 1  
P.O. Box 506656 Dubai | UAE  
<https://www.core.se>  
[office@coretransform.ae](mailto:office@coretransform.ae)