

# ARTIFICIAL INTELLIGENCE IN BANKING: DEEP LEARNING VS. GRADIENT BOOSTING

---

## A COMPARISON

Christoph Sonn, Julius Heitmann

30.10.2018

COREtechmonitor | Blogpost

Copyright © CORE GmbH

## Key Facts

- › Data analysis and artificial intelligence of increasing social and economic relevance
- › Deep learning as a high performance model of artificial intelligence often at the centre of discussion
- › Gradient boosting based on decision trees as a performant (function) and interpretable (transparency) model as an alternative in the financial industry

## Report

"Artificial intelligence" is one of the most discussed topics in public today. Events such as the evaluation of voter data by Cambridge Analytica for the US presidential election in 2016 or the generation of a conversation with the new generation of Google Assistants, which many participants regard as humane, do fuel fears and enthusiasm for the new technology; However, it is only by distinguishing individual methods of artificial intelligence that meaningful application cases and benefit aspects can be critically discussed, i.e. their performance and limitations can be assessed in an economic dimension and assessed against the background of social and legal contexts.

### Use of Artificial Intelligence in the Economy

The business models of the technology giants Google, Apple, Facebook and Amazon (GAFAs) as well as those of Baidu, Alibaba and Tencent (BATs) are only made possible by artificial intelligence. The analysis of customer-side data, e.g. master data, usage data or transaction data, allows the appropriate interactions, recommendations or search results to be determined on a customer-specific basis. With their Personal Assistance products (Siri, Alexa and Google Home), they bring artificial intelligence into the everyday lives of customers and support them in their daily problems. Huawei and Baidu on the Chinese market announce competition

products for 2018. The result is a constantly growing database of these companies, which in turn enables them to continuously refine statements about future customer behavior. Data analysis is also becoming increasingly important in the financial sector and is seen as a differentiating competitive factor in the future. For example, JPMorgan Chase recently introduced the "Contract Intelligence" application. This makes it possible to analyze 12,000 commercial credit agreements in a matter of seconds to extract important data points and clauses. The manual contract review would take about 360,000 hours. Companies such as Scalable Capital use algorithms to automate the trading of ETF portfolios and minimize manual processes. In addition, Scalable Capital benefits from the algorithms' ability to systematically abstract large amounts of information in a fraction of a second. This enables fast response times to complex information bases and thus increases returns.

### Deep Learning- vs. Gradient Boosting Model

The technology that is most in focus in this discussion is deep learning, an artificial neural network in which functional elements - also known as neurons - are used in a layered network structure to successively process information. Since 2006, these models have been indispensable in the field of high-dimensional classification - the assignment of data points to a limited number of classes in a

---

problem that encompasses many variables. Especially in the field of image, speech and handwriting recognition, there are currently hardly any alternatives to deep learning. These models derive their strength in solving problems from their layered structure. Here, the complex information is abstracted layer by layer until a decision regarding a previously defined problem is finally made in the last layer.

But why is this technology not implemented in every artificial intelligence? The complex structure of a deep learning network also has disadvantages. In order to train the layer-based structure, a comparatively large amount of data and a high level of computing power is required, depending on the depth of the model, i.e. the number of layers implemented. Depending on the application, it may also be necessary to understand the reasons for a decision made by the network. Especially in more strongly regulated contexts, such as banks or insurance companies, traceability can be a mandatory prerequisite. Here it is true that the BaFin has made it unmistakably clear in its current publication from the series BaFin Perspectives 1/2018 that the supervisory authority does not accept an argument according to which the functionality of an algorithm is no longer comprehensible due to its complexity and can therefore no longer be checked by the supervisory authority.

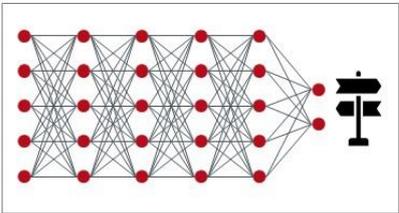
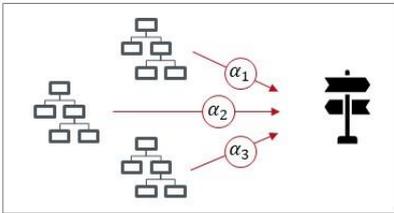
Deep learning nets, with their high number of neurons and links, offer a high learning capacity. However, it is precisely this advantage that prevents a network's decision from being easily or even comprehensible.

A promising alternative to the deep learning approach is ensemble learning, in

which various methods and algorithms are combined and which is used particularly in the context of decision trees. In recent years, gradient boosting has been able to assert itself as a subspecies in various public tenders. Instead of making the decision along a single, large decision tree, gradient boosting unites a number of smaller decision trees, each of which is trained to mitigate the weaknesses of other trees. The results of the respective tree are weighted and combined with the results of the other trees to form a joint decision.

Metaphorically speaking, this method represents a group decision by experts, whereby each sub-expert decides optimally for a certain area of the solution space and his decision is weighted in the overall decision. In contrast to deep learning models, this means that a decision of the model can be traced by identifying the trees with the highest contribution to the decision, whereby the decisions of the subtrees are easy to understand. In addition, the complexity of the individual trees in training is often limited, further improving traceability. This limited complexity is a further advantage over the deep learning approach because less data is required for training. In addition, gradient boosting requires relatively little training compared to other high-dimensional classifiers. As a result, gradient boosting promises to lead to a converged training result with a small data set.

**For the task of high-dimensional classification given the context gradient boosting can be a suitable alternative to deep learning**

Deep Learning	Dimensions	Gradient Boosting
	<p><b>Graphical illustration</b></p> 	
<ul style="list-style-type: none"> <li>▪ Neurons arranged in a layered architecture</li> <li>▪ Successive abstraction of the input information layer by layer called feature extraction</li> <li>▪ Degrees of freedom in definition of network architecture as well as types of neurons</li> </ul>	<p><b>Technical description</b></p> <ul style="list-style-type: none"> <li>▪ Ensemble of multiple decision trees</li> <li>▪ Decision as a weighted voting of individual decision of the decision trees</li> <li>▪ Degrees of freedom in definition of hyper parameters of ensemble of and individual trees</li> </ul>	
<p>Capacity to learn highly complex patterns </p> <p>Flexible model architecture </p> <p>Open source community for handwriting, speech and picture analysis </p>	<p><b>Resulting advantages</b></p> <ul style="list-style-type: none"> <li> Easily interpretable models and results</li> <li> Lower training effort</li> <li> Open source community for high dimensional classification tasks</li> </ul>	

Artificial intelligence in banking: Deep Learning vs. Gradient Boosting – a comparison

**Resume**

A statement about the dominance of one of the two models in a particular area is not possible. Both models provide good results in the classification of high-dimensional data, where the performance depends on the underlying data and the patterns it contains. A special case is the task of speech, image and handwriting recognition, in which deep nets are implemented in almost all systems today.

A further decision dimension to be considered - especially in the financial world - are the constraints of the use case in the form of traceability of the model decision or convergence speed (amount of data required up to the trained model), which

can lead to a differentiation of the models in addition to the prediction quality. Here, BaFin may already have forced a market prediction.

Depending on the use case and available data, gradient boosting may be preferable, as it leads to better results than deep learning for specific use cases and in terms of transparency. The specific benefits and performance of the gradient boosting model, which is somewhat out of the limelight, will therefore be demonstrated using various use cases and quantified using analyses. The next blog post will deepen the Santander Customer Satisfaction Use Case, a public competition in 2016 to predict customer churn, also known as the Churn Case.

---

## Sources

### 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)



**Christoph Sonn** is a Transformation Manager at CORE. His main focus is on developing innovative products and digitization solutions for banks, insurances and biotechnology companies. From strategic conception to the go-live of innovative IT transformations, Christoph develops solutions for companies to position themselves in the market.

**Mail: [Christop.Sonn@core.se](mailto:Christop.Sonn@core.se)**



As Expert Manager at CORE, Julius Heitmann supports the conception of complex IT transformations. In the automotive industry, he has gained extensive experience in dealing with big data and data analytics. He contributes this experience in particular to projects for the analysis and optimization of large amounts of data and develops tailor-made control systems for data processing, especially in the financial sector.

**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)