

Gradient boosting vs. deep learning. Possibilities of using artificial intelligence in banking

Key Facts

- › Increasing competitive pressure for financial institutions through digital and more individual offers
- › Customer churn as a manifest problem
- › Currently low analysis and reaction possibilities
- › New analysis and action options through machine learning
- › COREai demonstrates the advantage of gradient boosting over deep neural networks in the case of the churn case

Report

Retail banking is undergoing radical change. New players in the market and changes in the regulatory framework (e.g. PSD 2) mean that the market is being opened up as a result of market conditions and regulations, posing new challenges for traditional banks. Customer behaviour has also changed decisively. The willingness of customers to change accounts has doubled in the last four years, with the absolute number of accounts in the German market remaining constant over this period. The traditional instruments of financial institutions currently offer hardly any opportunities to react to the effects of this change.

Artificial Intelligence provides various solution models for the challenges described, two of which will be discussed below along a concrete use case. The focus of this blog post is the so-called "Churn Case", in which customers willing to migrate are identified on the basis of customer master data and transaction data as well as, if possible, the associated reasons in order to be able to take measures for sustainable customer loyalty.

Response options of the banks

In the course of the challenges described above, banks have various options for responding to a declining customer base.

a) Acquisition of new customers

The focus of the business strategy is on acquiring new customers. It does not counteract the loss of existing customers.

b) Addressing customers who have already been terminated

The second variant focuses on winning back customers who have already given notice but are still in a business relationship with the bank due to legal deadlines. This means offering former customers attractive new customer offers as an incentive to remain loyal to the bank.

c) Identification of customers willing to migrate

The third option anticipates potential redundancies and the underlying motivation and addresses them in advance through needs-based support. The basis for this is a model that enables the identification of dissatisfied customers on the basis of available data.

In practice, in most cases a combination of the options mentioned is chosen. This stands however in no contradiction to the fact that for option c) expenditure must be operated for the creation of an information basis and attainment of action ability, which represents the focus of this Blogpost.

Challenges

According to expert estimates, the acquisition of new customers is about five times as expensive as keeping an existing customer. However, banks face two key challenges in order to realize the option of identifying customers willing to migrate.

1) Correct and timely forecasting, i.e. anticipation of the termination request

2) Correct determination of the reasons for a change of account provider

Often the banks do not receive information about a termination until it has already been received. Knowledge about the motives loses value at a later point in time, since a subset of the options for action can no longer be realized. Examples of customer reasons for switching can be, in addition to the obvious reasons discussed in the media, such as switching to a cheaper provider, e.g. direct banks, lack of accessibility to branches, bad experiences with customer advice, too few ATMs in the vicinity or an inadequate product portfolio.

Retail banking market and especially traditional institutions suffer from the increasing number of customers changing banks

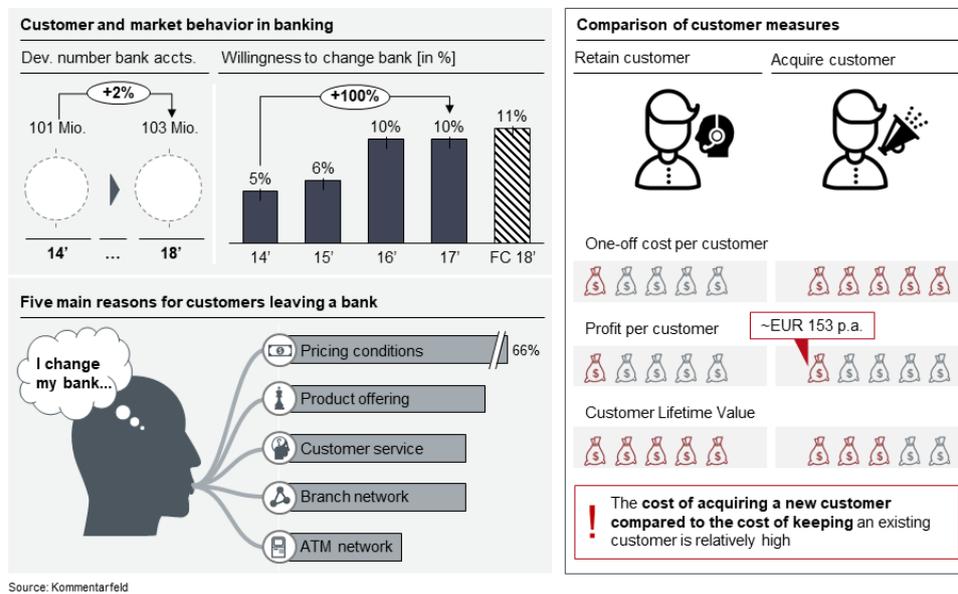


Figure 1: Retail banking market and especially traditional institutions suffer from the increasing number of customers changing banks

New options for action with the help of KI

Machine learning models can be used to utilize the data basis often already available in banks for regulatory reasons in order to train a forecasting model for customer churn. Classification models are used with increasing success to identify even high-dimensional patterns (patterns over a large number of variables), which are often too complex for human analysis capacities, and to recognize the connection with, for example, a termination event.

Banco Santander has also identified this problem as crucial. As early as 2016, for example, the Spanish bank published a public tender on Kaggle for data analysts worldwide. The data set, completely anonymous, was made available to 76,000 customers, 3,000 of whom terminated their accounts. The customers were characterised by 370 variables, with only Banco Santander being able to trace the data back to real circumstances due to anonymisation. The aim of the tender was to predict a customer's cancellation as accurately as possible, with the correct forecast of customers who cancelled being more important in the evaluation of the solution than the correct forecast of a customer remaining in the bank. The data set was analysed by more than 5,000 data analysis teams over a period of two months.

Comparison Deep Learning with Gradient Boosting

The COREai team also analysed the tendered data. As a basis for later discussion, the analysis strategy will be examined in more detail below.

1. Preparation: data cleansing and modelling

The preparation of each analysis provides for data cleansing and modelling. In a first step, using a Random Forest algorithm, the variety of variables was reduced to eliminate those with low information content. Further data cleansing, as would normally occur in data analysis projects, was not necessary as the data was already in a usable state. The two models Gradient Boosting with Decision Trees and Deep Learning were tested and their performance was compared. In order to improve the understanding of how the algorithms work, a rough explanation of the models is given below.

2. Gradient boosting and deep learning models in detail

Deep neural networks are so-called neurons (non-linear arithmetic operations) arranged in a layered architecture in which the information is abstracted layer by layer, whereby important features for classification are extracted from the data. This step-by-step abstraction makes it possible to process even very high-dimensional information. Thanks to their flexible architecture, these models are able to use different sources of information in one model and to recognize even highly complex patterns.

Gradient boosting is a strategy in which several performant individual models for a subarea of the solution space are combined to form a weighted overall model that performs over the entire solution space. The aim here is to exploit the cooperation in the ensemble, whereby the individual models are created in such a way that they compensate for the deficits of the already trained submodels. Decision trees are usually used as single models, which are weighted and included in the overall result.

Gradient Boosting is a meta model combining a set of weak classifiers to create a strong classifier

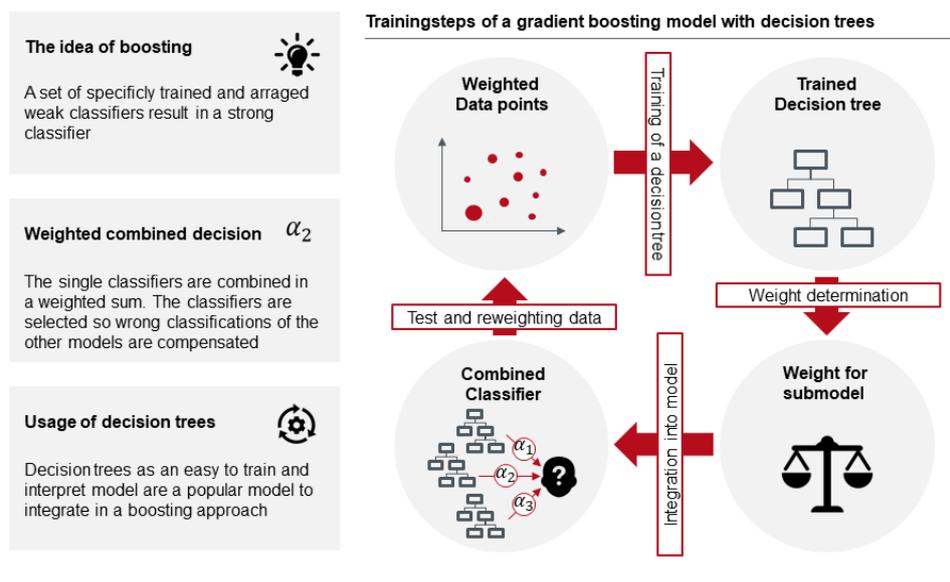


Figure 2: Gradient Boosting is a meta model combining a set of weak classifiers to create a strong classifier

3. Outcome 1: Predictive accuracy and performance

The two models presented were tested with regard to their predictive quality using the misclassification rate - the proportion of incorrectly predicted data points and customers respectively - whereby a correct classification of a cancelling customer was weighted higher than that of a remaining customer. In order to ensure a better intuitive interpretability of the models in this article, the results are given in unweighted form. In the test runs, the Deep Neural Network came to an erroneous prognosis in 23.0% of the cases. The gradient boosting model with decision trees only came to the wrong result in 3.3% of the cases.

Besides the better prediction accuracy, the gradient boosting model convinced with shorter learning and prediction times. In the test scenario considered, this circumstance did not lead to a decisive reason for selection due to the relatively small data set, but for larger data sets this aspect is of increasing importance (keyword Big Data).

4. Result 2: Interpretability

In view of the fact that, in the event of termination, not only the right forecast but also the right offer must be chosen, the interpretability of models represents a not negligible

advantage. The strength of deep learning models, with their scalable architecture in depth and breadth, unfortunately leads in return to some limitations in terms of interpretability - despite significant progress in this area. Gradient boosting, on the other hand, offers good interpretability of both the individual decision trees and the weighted ensemble, which is an advantage in the banking environment. By using an interpretable model, it may be possible to draw conclusions about the reasons for the termination in addition to forecasting terminations. Even though we would have liked to have done this in our test with Banco Santander's data set, no such findings could have been obtained due to the anonymisation.

Resume

Gradient boosting has proven to be the more promising solution for the described data set and use case than deep learning. In addition to the better analysis results, gradient boosting scores with better interpretability, which offers the user an improved possibility to understand the decision making of the algorithm and thus favours the ability to act in the event of termination.

A general statement on the dominance of one of the two algorithms cannot be made, since it depends on the patterns contained in the data. For example, Deep Neural Networks remain unchallenged in the field of image, speech and character recognition. In individual cases, it is now necessary to determine which information is actually contained in the existing data and thus whether, as with Banco Santander, a new type of customer service can be made possible.

Sources

Hadden J., Tiwari A., Roy R., Ruta D. (2005) "Computer assisted customer churn management: State-of-the-art and future trends", Computers & Operations Research 34, pp. 2902-2917

<https://www.wiwi-treff.de/Fuehrung-and-Strategie/Kundenmanagement/Beziehungspflege-zahlt-sich-aus-Neukundengewinnung-fuenfmal-so-teuer/Artikel-4977>

<https://www.wiwo.de/unternehmen/banken/studie-warum-die-deutschen-ihre-bank-wechseln/11478036.html>

www.kaggle.com

<https://www.kaggle.com/c/santander-customer-satisfaction>

<http://konto-report.de/girokonto/#girokonto>

<https://de.statista.com/statistik/daten/studie/187567/umfrage/anzahl-der-girokonten-bei-deutschen-kreditinstituten-seit-2002/>

<https://yougov.de/news/2017/04/05/das-wechselkarussell-im-bankensektor-nimmt-fahrt-a/>

<https://de.statista.com/statistik/daten/studie/652802/umfrage/gruende-von-bankkunden-fuer-einen-wechsel-des-girokontos/>

<http://www.tagesspiegel.de/advertorials/ots/a-t-kearney-deutsche-privatbanken-bleiben-im-europaeischen-wettbewerb-weiter-zurueck/20010868.html>