McKinsey&Company

FIG – Banking
Advanced analytics in banking – a sneak peek

Leading in a disruptive world

October 2018

#10

#10

Advanced analytics in banking – a sneak peek

By Philipp Koch, Max Flötotto, Christoph Sporleder, Ranja Reda-Kouba
Additional contributors: Mario Thaten

The German banking sector has been under sustained pressure. Low interest rates have been depressing revenues. In response, many banks are in the process of transforming their business models, but quite a few players have not yet achieved a substantial reduction of their cost base. As a result, the cost-to-income ratio (CIR) is deteriorating across all banking segments. In a three-year period ending in 2016, the CIR was 75 percent, compared to 69 percent in the three years before (2010 to 2013). Advanced analytics may not be able to put a big dent in a cost base driven by expensive branch networks, but it can be a source of competitive advantage, which will help banks secure revenue streams, even under adverse conditions. Use cases exemplifying this exist in many segments of the banking industry. In this article, we will look at the real-life application of advanced analytics in a retail banking use case before examining potential applications in other segments.

Exhibit 1
Advanced analytics value tree

Value tree			Categories of use cases	Typical impact	
Value to business	Revenue growth	Client/financial profitability	Micro segmentation	+30%	EVA
			1-to-1 pricing	+20 bps	Deposit margin
			Transactional analytics	40%	Lead conversion
		Client acquisi-tion and retention	Churn reduction analytics	-20%	Churn rate
			Customer experience analytics	+5 PPT	CSAT score
	Risk control	Risk losses	Digital credit assessment	-80%	Manual underwriting
			Advanced early warning systems	-25%	Gross NPL inflow
			Credit collection analytics	-10%	Collection expenses
		Fraud losses	Fraud detection analytics	+15%	Fraud detection
	Efficiency	Operational costs	Geospatial analytics	-10%	Location costs
			Productivity optimization	-20%	Back-office costs
		HR management	Predictive HR	2x	Motivation score
		Financial control	Advanced financial analytics	+15 PPT	Market-to-book ratio

PROVEN IMPACT ALONG THE ENTIRE BANKING VALUE CHAIN

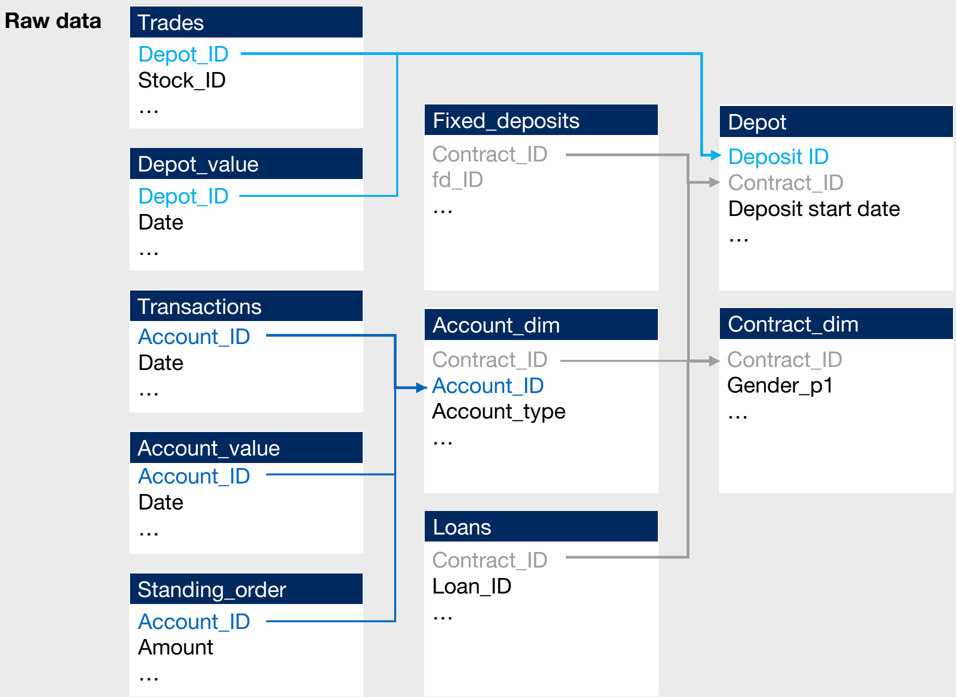
Advanced analytics holds the potential to grow revenue, control risks, and increase efficiency across the entire banking value chain. Impact has been proven in real cases, e.g., value-based client segmentation, eligibility assessment for debit cards or loans, and customer retention. For details and typical impact ranges, see Exhibit 1. It is time for banks to determine the most promising areas of application and reap the benefits before others beat them to it. But how does this play out in practice? In what follows, we will provide an actual use case to illustrate the challenges and benefits of implementing advanced analytics at scale.

DEEP DIVE: CHURN REDUCTION ANALYTICS USE CASE

A global bank was seeking to prevent churn in its consumer retail division. Prior to the effort, the checking account cancellation rate was in the range of 5 percent per year. While the bank had win-back offers in place for certain cancelling customers, there was no proactive retention program. The bank's goal was to identify potential churners and reach out to them before they canceled their accounts. In the given use case, the proactive churn prevention program helped the bank sustain ten times the revenue it had managed to reclaim with its previous win-back scheme. At the same time, better prioritization helped double the efficiency of the bank's churn reduction efforts.

To achieve this goal, the bank implemented an advanced analytics program that helped predict churn and take appropriate measures to retain valuable customers. The core module is an analytical model that predicts who is most likely to churn within the next three months, based on comprehensive customer data from the last twelve months. The most important lesson learned was that excellence in analytics is not only about the best algorithms. Data quality, diligent interpretation of the corresponding outcomes, and effective operational execution turn out to be equally important.

Exhibit 2
Comprehensive customer view construction



Cleansed data

1 2 3

Contract_ID	Target	Variable_1	Variable_2	Variable_3
40032	0	M	53	0
93467	1	W	85	0
93306	0	W	26	0
20489	1	M	51	0
84544	1	W	72	0
93704	0	W	66	0
65452	1	M	72	0
98891	0	M	34	0
79931	0	M	41	0

1 **1 line for every customer** contains all the data you need to develop your model

2 **Target variable –** this is the variable your model needs to learn how to predict

3 **Predictors –** variables your model uses to predict the target variable

Step 1: data cleansing

Working with unreliable data is dangerous. If the data fed into a predictive algorithm is inconsistent or incorrect, one can easily arrive at inaccurate or downright erroneous conclusions. This will result in conducting churn prevention measures on customers who are loyal or missing potential churners and losing their business. Data cleansing and alignment are key prerequisites to any successful application of advanced analytics. In our case, the dependent variable of the model was defined by asking, "Is the customer likely to churn?": "Yes" (1) or "No" (0)? Examples of predictors derived from comprehensive data included:

- Products (credit card, insurance, custody account, mortgage lending, etc.)
- Demographics (age, gender, profession, number of persons in household, ZIP code, etc.)
- Customer information (segment, "customer since," etc.)
- Volume (sum of balances of various products, etc.)
- Income (deposits, loans, securities, etc.)
- Interactions (adviser contacts, online banking log-ins, cash withdrawals, etc.)
- Rating (credit standing, credit, etc.).

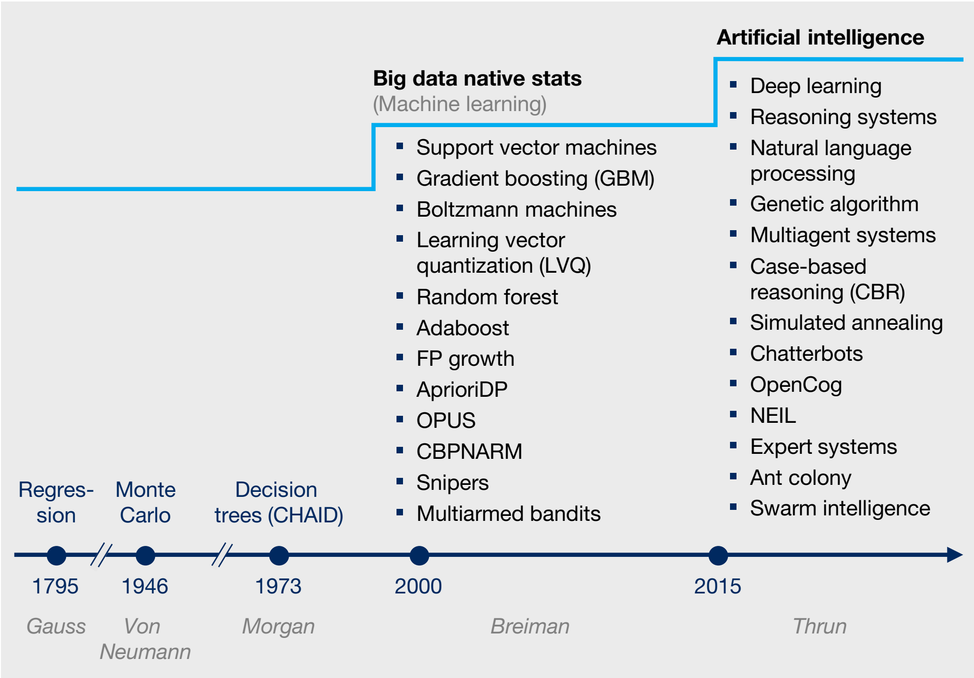
This data was not found in a single source, but was spread throughout multiple tables and databases. To build and develop the churn prediction model, it was necessary to construct a comprehensive customer view, using an identifier to integrate all data on a given customer and consolidate it accordingly (see Exhibit 2). Also, data points were missing in many sources. Conversely, not all data points contained in a given source were relevant to the analysis at hand. Generally, no data is ever clean when at the beginning. It takes special data engineering capabilities to retrieve, access, manage, select, clean, and split the data to prepare it for data scientists to work their magic. While it is not necessary to have the perfect data architecture in place to reap the benefits of advanced analytics, the quality of the output is closely correlated with the quality of the data from the model on which it is built. Smart preparation will give a model an edge before running the first algorithm.

Step 2: modeling

The algorithms used to gain relevant insights from data are evolving fast. While simple regression analysis dates back to the late eighteenth century, the past few years have seen the advent of powerful tools, such as deep learning and genetic algorithms (see Exhibit 3). For details, see our article "Artificial intelligence – how advanced analytics and smart machines will change the way we work."

All approaches have their limitations. Some algorithms will produce clear results but tend to oversimplify the relation between predictors and target variables. Others are more accurate but tend to be overly dependent on historical data ("overfitting"). This is why we recommend combining multiple algorithms to overcome the shortfalls and apply the individual strengths of each one ("ensemble learning"). In the given use case, the bank used three different types of algorithms, including decision tree and random forest modeling, to predict churn. The combined model relied on more than 200 data fields to identify churners, and it turned out to be much more powerful than a logistical regression model in terms of differentiation between likely churners and loyal customers. As a result, far fewer customers needed to be contacted to achieve the same impact as compared with a less discriminative approach. To master this kind of advanced modeling, the bank relied on the services of a team of highly specialized data scientists. The capabilities of a data scientist exceed those of a typical business analyst. Successful data scientists excel at data wrangling, mathematics, and coding. Therefore, modeling is only one of the building blocks of any successful advanced analytics application.

Exhibit 3
Evolution of statistical learning



Step 3: interpreting

A good forecast is valuable but not enough to turn the tide. To generate actual impact, effective churn prevention is just as important. Total impact is a product of three factors:

- Value of the customer at risk
- Churn forecast precision
- Cancellation prevention effectiveness.

Exhibit 4
Example of customer segmentation

○ Low priority ◐ Medium priority ● High priority

Customer segment	Actions: Advanced customer care	Actions: Cross-sell	Actions: Change product/price
Empty nester			
▪ Older customers ▪ Decreasing debt ▪ Increasing savings	● Contact through personal advisor	● Offer of further investments	◐ Reduce price or improve conditions
House builder			
▪ Starting to finance property ▪ Only partial credits at this bank	● Contact through personal advisor	● Offer of follow-up financing	
Transactor			
▪ Using their checking account only for very few transactions ▪ No other product holdings		◐ Offer for investments and other products	
Creditor			
▪ Many consumer credits ▪ Bad credit rating	◐ Contact through personal advisor		
Student			
▪ Young people and students ▪ Low credit and asset volume		◐ Digital offers	○ Increase overdraft
Senior			
▪ Very old ▪ High savings, little activity ▪ Not churning, but passing away?			

The more valuable a given customer is, the higher the return on investment to retain them will be. Precise forecasting allows for an accurate identification of potential churners, while effective cancellation prevention drives the retention rate. Once the bank in our example identified high-risk customers, it set out to determine the drivers of churn. The first step was to use statistical clustering to create segments of customers with comparable behavioral patterns. In the second step, these segments were analyzed to determine the root causes of churn. In the given case, life events turned out to play a big role. Examples include:

- Young customers switching banks upon starting their first job
- Married couples setting up a joint account and cancelling individual accounts
- Empty nesters consolidating their pension accounts.

In the third step, segments of likely churners were enriched with additional information, such as interaction preferences (e.g., in person vs. online) from market research and other sources. Based on these segment profiles, the bank devised a set of actions to prevent cancellation. For example, valuable empty nesters were contacted by an advisor who offered them improved conditions to prevent them from cancelling their accounts. Using a "test and learn" approach, the bank identified the kinds of actions that were most effective for various types of customers at risk (see Exhibit 4 for examples). Many banks use similar approaches to combat not only churn, but also other types of unwanted behavior, such as credit default or fraud.

Lessons learned

Using advanced analytics to reduce churn is not as straightforward as some experts assert. What is needed is the right architecture to handle large quantities of data, the right people to choose and implement the most potent algorithms, and the right experts to make sense of the output. Throughout the entire process, executives are needed who have a solid command of both the analytics and economics involved in reducing complexity and making trade-off decisions. While a cutting-edge artificial intelligence model might increase forecast precision to a certain extent, a team might be better served with a simpler, more robust model to inform decision making. Also, it is important to foresee and manage the implications of decisions made in the early process stages, e.g., which data to use at later stages to ensure reliable output and effective actions once a bank has successfully implemented a particular use case, such as the one at hand. The momentum and the knowledge gleaned from the pilot effort can also help build a case for rolling out advanced analytics to other parts of the organization.

APPLICATIONS IN RETAIL BANKING, WEALTH AND ASSET MANAGEMENT, AND CORPORATE BANKING

Churn reduction in consumer retail banking is only one of many applications of advanced analytics in banking. In this section, we look at opportunities to create value using advanced analytics in three other areas: customer vetting in retail banking, wealth and asset management, and corporate banking.

Retail banking

For now, advanced analytics can help improve basic analytics and maximize revenue per customer, e.g., by using a recommendation engine to propose additional products and services to a retail bank's current customers. In the next three years, microsegmentation, enabled by nontraditional data, will provide fast movers with competitive advantages when it comes to revenue generation and collections. Specifically, advanced analytics can help increase the accuracy of customer vetting. Customers who would have otherwise failed traditional eligibility tests may well be granted debit cards or loans thanks to advanced analytics of their credit history and other factors. This is because machine learning can be used to identify complex patterns without being constrained by a predetermined set of criteria. While the application of advanced analytics in other areas of retail banking, such as customer retention and customer experience management, can also generate high impact, it is less urgent from a competitive perspective (see Exhibit 1 for details).

Wealth and asset management

In wealth and asset management, immediate priorities include lead generation and relationship management, especially pricing based on insights gained through advanced analytics of a customer's history. These applications hold the potential to create substantial value over the course of the next three years. For example, advanced algorithms can help wealth managers focus their acquisition efforts on the most valuable leads. Predictive modeling can be used to estimate a customer's lifetime value with unprecedented accuracy and

Exhibit 5
Wealth and asset management

! Urgent use cases

Value tree			Categories of use cases	Typical impact	
Value to business	Commercial effectiveness	Customer acquisition and retention	Insights-based investment advice	+20%	More leads
			! Optimize purchase funnel leakage	+200%	Increased conversion
			Retention/churn reduction analytics	-20%	Churn rate
		Customer profitability	! Dynamically recommend next conversation/trade	+15%	Incremental trade activity
			Customer behavioral segmentation	+30%	EVA
			! 1-to-1 pricing	+10 bps	Revenue margin
		Investment performance	Debias investment decisions	+1 - 3	Alpha
			Generate investment ideas using big data	+2 - 3	High-potential ideas/day
	Risk control	Risk losses	Advanced early warning systems	-25%	Gross NPL inflow
			Credit collection analytics	-10%	Collection expenses
		Fraud losses	Fraud detection analytics and KYC	+15%	Fraud detection
	Operational efficiency	Operational costs	Automation and productivity optimization	-20%	Back-office costs
		Financial control	Advanced financial analysis and reporting	+15 PPT	Market-to-book ratio

help prioritize a bank's activities accordingly. Additionally, advances in analytical capabilities also present several no-regrets opportunities to mitigate risk (see Exhibit 5 for examples). Applications geared toward relationship management and customer value maximization in wealth should be a priority for banks in the near future. In parallel, banks should explore risk mitigation use cases to safeguard the mid-term viability of their wealth services.

Corporate banking

In corporate banking, there is an emerging set of high-value advanced analytics use cases focused on revenue. The next "product to buy" systems will increase revenue while price optimization and leakage fee reduction will boost profitability. Robotic process automation and optimized service models can improve RM productivity by up to 20 percent (see Exhibit 6 for details). For now, banks should focus on pricing and performance management use cases. Also, they should proceed to capture low-risk, high-yield opportunities, such as savings derived from automation.

Exhibit 6
Corporate banking

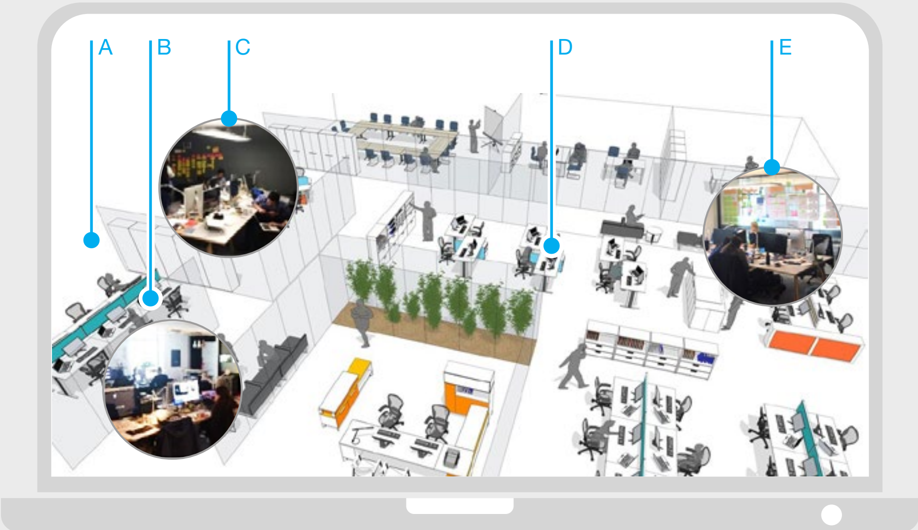
Value tree			Categories of use cases	Typical impact	
Value to business	Revenue growth	New customer identification	More effective customer targeting	+75%	More relevant leads
		Churn reduction analytics	Spot clients who are likely to switch to another bank	+10%	Increase in profit
		Next product to buy	Identify commercial opportunities per single client by looking at internal data and external information	+10 - 15%	Increase in revenue
		Pricing analytics	Leakage identification Price optimization	+10 - 15%	Increase in revenue
	Risk control	Digital credit assessment	Support RMs/credit offices via digital workflow Decision engine and/or RM recommendation tool	-20 - 25%	Reduction of gross inflow to NPL
		Advanced early warning system	Identify clients who are likely to become riskier/problematic	-15 - 20%	Reduction in RWA
		Anti-money-laundering control	Identify suspicious client transactions/behavior Improve detection rate (OCR)	+25%	Cost savings of total losses
		Fraud detection analytics	Identify potential fraud cases through patterns of past fraud	TBC	Reduction in false positives
	Efficiency	RM production optimization	Maximize frontline productivity defining the optimal service model	+10 - 20%	Improved RM productivity

IMPLEMENTING ADVANCED ANALYTICS AT SCALE

To create sustainable competitive differentiation, banks should apply advanced analytics not only in selected areas, but throughout their entire business. The prerequisite for advanced analytics at scale is agile organization. Agility is about balancing stability with

Exhibit 7
A "sneak preview" into the near future of scaled-up analytics capabilities

Your own "shop floor" of an analytics lab and factory, closely integrated with all banking business functions



A | Data science: use AI techniques like machine learning and deep learning to drive insights and actions

B | Data engineering: get, prepares and quality assure data for analytics

C | Data lab: approachable unit that helps drive a continuous series of use cases

D | Design and integration: operationalize advanced analytics by embedding it into existing work flows and processes or into new applications with intuitive UX

E | Concept sprint: advanced analytics will be fully agile flexibility in a way that is conducive to sustainable success in a dynamic environment (see "The five trademarks of agile organizations.")

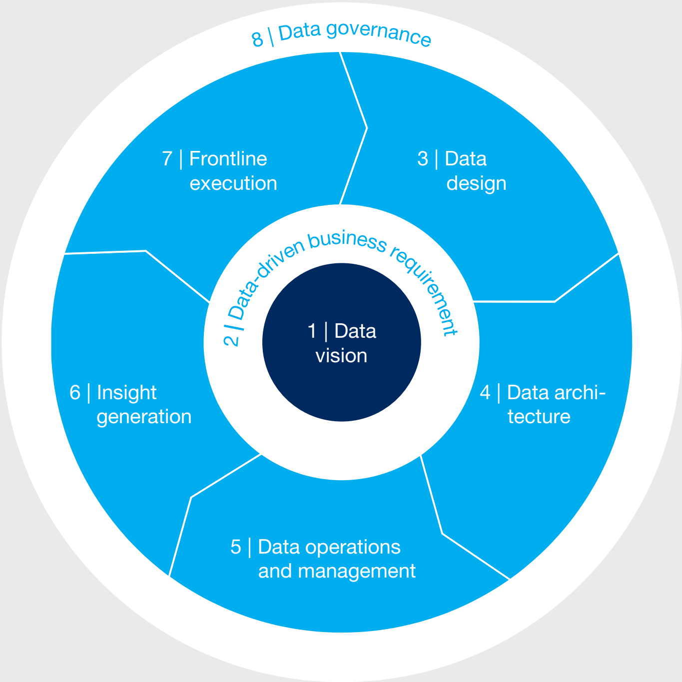
To implement advanced analytics at scale, banks should take a step-by-step approach and:

- Define the bank's advanced analytics ambition and strategy
- Design and conduct a series of use cases to generate buy-in of key stakeholders
- Build a scalable foundation (see Exhibit 7 for a glimpse of the future).

Key enablers of advanced analytics at scale include a comprehensive data strategy (see Exhibit 8) and the introduction of new roles and professions, such as data scientists and solution architects. While these roles may be new, the concept of getting the finest minds to collaborate on competitive differentiation is an age-old success factor within the banking industry. Jakob Fugger, for example, relied on the services of accounting wizard Matthäus Schwarz to secure a competitive edge for his bank. 500 years ago, Schwarz published his "Musterbuchhaltung," a study of accounting principles that were revolutionary at the time and covered not only debit and credit, but also included the concept of a "secret" account detailing profit and loss. Jakob Fugger proceeded to become the richest man in history. At the peak of his career, his wealth was equivalent to about USD 400 billion.

Banks that take advanced analytics seriously will treat data as an asset, make analytics a core competence, establish an appropriate governance model, use data and analytics to drive innovation, and implement agile principles to enable the implementation of advanced analytics at scale.

Exhibit 8
Successful advanced analytics requires a data strategy



1 | Aligned goals and business value aspiration from data

2 | Business perspective on needed capabilities

3 | Data definition and layered models to describe attributes and relationships

4 | Technology to organize, Extract and analyze data

5 | Manage data entry, quality, access, security, and business use

6 | Use data to generate new insights

7 | Processes and tools to translate business insights into actions

8 | Organizational models, tools, and artifacts to manage all data-related activities

KEY TAKEAWAYS

- Because of high cost pressure and disruptive market conditions, banks should take advantage of advanced analytics to generate competitive differentiation and protect their market share.

- Advanced analytics is more than algorithms. Thorough data preparation, diligent interpretation of the corresponding outcomes of application, and effective actions to monetize resulting insights are equally important.

To take full advantage of advanced analytics, a bank should treat data as an asset, implement agile principles throughout the organization, and make analytical capability a core competence across all its functions.

"Leading in a disruptive world"
October 2018

Visual Media Europe
www.mckinsey.de