mobey forum

Predictive Analytics Workgroup

PREDICTIVE ANALYTICS IN THE FINANCIAL INDUSTRY

THE ART OF WHAT, HOW AND WHY

PART 1

TABLE OF CONTENT

EXECUTIVE SUMMARY

This first white paper of the new series discusses the value of predictive analytics for the financial industry and answers the question why this is the right time to start with predictive analytics and how to empower entire organisations to use it. As mobile technology evolves and everything around us – not just our mobile devices– is becoming connected we are entering a new era of connected experiences. The customer journey in the financial industry is completely digitized. This exponentially increases the number of interactions between a financial service company and its customers. Customers expect banks to understand their context and the challenge for financial industry is to be relevant at all these interactions. Given the gold mine of data that banks have access to, the field of predictive analytics offers a range of untapped opportunities in doing so. When implemented successfully, predictive analytics will lead to vast improvements of existing static business rules and achieve progress like reducing cost, increasing revenues and improving customer experience. Mobey Forum expects that predictive analytics skills will soon be essential for banks to keep their position in the market against non-banks but also other banks that will be using predictive analytics as a competitive weapon. That's why building skills that go beyond the conventional descriptive analytics and focusing on the question "what will happen?" has to be top priority for every financial institution.

To understand the concept of predictive analytics the paper discuss the difference between descriptive and predictive analytics. The majority of business analytics at financial institutions is currently still focussed on the 'rear view mirror', resulting in descriptive analytics. The questions more concerned with the future mainly remained unanswered from an analytics point of view. These decisions are often made intuitively and are seldom fact based, often resulting in sub-optimal decisions. With predictive analytics we do identify and address questions concerned with the future. We learn from our experience and predict future behaviour in order to drive better decisions.

Used effectively, there are several areas of application for predictive analytics where financial institutions could make profitable investments while at the same time improving the attractiveness of their services. Mobey Forum is mentioning examples like card linked offers, next best action, pricing, claim handling, risk assessment to mention few.

But as always, where are many opportunities, there are also many challenges to be solved. Effective usage of predictive analytics is only possible in a data-driven organization. It all comes down to having access to right "past data" and using right skills, techniques and tools to find business relevant patterns that can be used to solve similar problems in the future. The challenges in the different phases of the predictive analytics lifecycle are discussed in the paper.

Considering the challenges and opportunities ahead, the final advice is to start now. But start small, create successful examples and iterate towards leveraging predictive analytics in to your organization.

This paper will help you to start. It explains the most important components, challenges and key application areas of predictive analytics in the financial industry.

1.
BUSINESS ANALYTICS AT FINANCIAL INSTITUTIONS STILL FOCUSES ON THE ‘REAR VIEW MIRROR’

The days in which customers developed a lifelong relationship with their financial institutions are long gone. Digital banking has made customer service more cost efficient but, when improperly implemented, has also made the relationship between the bank and the customer more distant. On the other hand, digital and mobile banking caused an increase in the number of interactions between a customer and its bank. These interactions offer numerous opportunities to make banking more personal again. Smart usage of data is becoming key in achieving this goal.

Financial industry must invest in understanding and anticipating the complete context in which each customer consumes their services. Financial industry has a competitive advantage here over most other industries thanks to the data they possess. Due to customer identification requirements and strict regulations, bank customer databases have potentially very high quality data, provided that master data management processes and systems are up to standards. Financial institutions also have transaction data that retailers can only dream about. This data allows banks to have a very detailed view into each customer's income, spending patterns, credit profiles, and so on. This 360 degree view of a customer is the starting point for numerous interesting analyses, creating actions to anticipate the future needs of a customer.

In reality, however, the majority of business analytics at financial institutions over the last two decades has been synonymous with transaction reporting. Historical data-oriented reporting systems and better visualization tools have been focusing on 'rear view mirror' organizational activities such as sales, transactions, risk management, customer satisfaction measures or various operational issues.

The questions more concerned with the future prospects of businesses given a certain course of actions, mainly remained unanswered from an analytics point of view. In practice however, these questions are answered every day. Think of the daily number of decisions made about the future within financial industries. Which offers are going to be most desirable to which of my customers, via which channels should I offer my products or where should we be careful with a price increase since this might detract customers? These decisions are often made intuitively and are seldom fact based, often resulting in sub-optimal decisions. Additionally, sometimes a decision maker might not even be aware of urgent decisions. In these cases he should be triggered, otherwise an opportunity to influence the result will be lost.

Even more important, data driven decision support could improve the customer experience throughout the (mobile) journey of a customer, adding value for the customer and the bank. Based on among others the behaviour of peers, a bank could provide continuous digital assistance for their customers. The assistance will directly help customers to answer questions like "do I need some kind of additional insurances?", "how much money can I save every month to fulfil my saving goals" or "can I afford this purchase (without creating balance problems at the end of this month)"? This kind of assistance which is focussed on future actions and value will definitely be more helpful for a customer than simply applyingoutdated static business rules from the past.

In both cases the decision maker (the manager or the customer) probably rather bases their decision or financial product choice on fact based advice, than on gut feeling.

1.1 PREDICTIVE ANALYTICS IS BECOMING ESSENTIAL FOR DIGITAL FINANCIAL SERVICES

This is why predictive analytics is so interesting for Financials. It is: "a technology that learns from experience (data) to predict the future behaviour of individuals in order to drive better decisions". [1] In many cases these individuals are customers, but it can also be companies, debtors, employees, products, machines, locations, etc. Predictive analytics does not stop at taking note of past events, but it goes on to predict a chosen target variable's future value. As such, predictive analytics leverages machine learning and statistical methods to arrive at predictions at scale.

Considering today's context and data-driven era, predictive analytics can be a great contributor to making banking personal again. As with online and mobile technologies before, understanding predictive analytics is becoming essential for digital financial services professionals in order to remain relevant. Offering decision support throughout the mobile journey of a customer will add value, both for the customer and the company. Through predictive analytics, the company can predict changes in customer needs, target customers with well-timed and appropriate product offerings, and build loyalty by offering contextual information and advice that improves the customer experience.

The retail industry has used predictive analytics for many years as a means to increase campaign effectiveness, build customer loyalty and reduce shopping cart abandonment in e-commerce. Telecommunications companies have applied predictive analytics to reduce customer churn and to predict customer life time value. The financial industry has traditionally used predictive analytics to identify potential fraudulent transactions and to evaluate credit risk, but the time has come for them to go far beyond these areas and keep on going. Four perspectives explain in more detail why 'now is the time'.

1 Siegel, Eric (2013). Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die

1.2 WHY THIS IS THE RIGHT TIME FOR PREDICTIVE ANALYTICS WITHIN FINANCIALS?

The customer: Growing expectations regarding service and relevance

In today's data economy being able to apply powerful data-gathering and analysis is no longer good enough. Customers' expectations about services in general are increasing and expectations in the customer experience field are set here by companies like Google and Apple. Google, the company that coined the "mobile-first" expression is moving to "AI-first" and is evolving from an information retrieval company to informing and assisting. Think of the Google Now service, which tells you when to leave for your appointment. Customers expect this kind of added value, and expect it to be neither annoying nor inaccurate. In information retrieval, anything over 80% recall and precision is pretty good. When assisting a customer, customer expectation rises to almost 100% accuracy (see also chapter 4). This is something that is only possible when the data quality is high and predictive analytics capabilities are outstanding.[2]

Financial institutions should therefore empower customers by creating real-time insights into their financial situation and offer personalized insights, services, and advice throughout their mobile decision journey. Predictive Analytics is becoming competitive ground and fundamental for these kind of innovations in financial services.

The smartphone: The perfect communication channel is already created

The perfect communication channel is already created By having an app on virtually all their customers' smart phones, the banks have created an ideal distribution channel for their products and market communication; It is always within arm's reach and offers a variety of options for real-time, two-way dialogue. Simultaneously, the smart phone is also an important source of data (for instance geolocation), which makes it possible to offer the customer advice, information and products that are relevant to his location or purchase situation.

Globally, the base of mobile banking customers is set to rise from 0.8 billion people in 2014 to 1.8 billion by 2019* (Juniper Research, KPMG analysis).[3] In Western countries, around 40% of all people are using mobile banking apps. In China and India, these numbers are around 60-70%.[4] Adding the introduction of other wearables, the amount of data that can be collected for even more accurate personal contextual messaging will increase dramatically in the near future.

2 Quora, Peter Norvig Session

3 Mobile Banking 2015, KPMG

4 Statista.com

The data: Digitization increased the number of potential data sources

Digitization of every aspect of our life has increased the number of potential data sources and the amount of data that can be captured and stored, some structured but mostly unstructured data. The wave of technologies that were originally developed by Silicon Valley heavyweights such as Google, Facebook, Amazon and Yahoo emerged as viable advanced data storage and processing alternatives to existing commercial Relational Database Management System and Business Intelligence tools. The costs of storing data have fallen below the costs of deciding which data to keep and which to delete. Computational processing power increases exponentially allowing us to do analytics at scale and in nearly real-time. Suddenly, Terabytes, Petabytes and even Zettabytes were part of the conversation for small start-ups like Airbnb and Uber looking to disrupt previously untouched markets by putting data-driven advanced analytics capabilities at the very core of their value propositions. Similar to these companies, a future-proof bank stores and matches all data which is gathered during the complete online and offline journey of its customers. This results in a real time 360-degree customer view which creates a perfect foundation for predictive analytics.

The competition: Upcoming changes in regulation will create new areas of competition

Financial institutions are constantly looking for new sources of differentiation, revenue and customer loyalty. The next few years provide great opportunity for financial institutions to prepare for an environment of open competition on the usage of customer data. Upcoming changes in data regulation in the EU could radically change the market environment for financial institutions when it comes to predictive analytics. In Europe regular financial services such as payments may be reduced to a commodity where due to the Payment Service Directive anyone with a payment licence can compete for customer payment data. So financial institutions need to leverage their use of data in order to remain relevant to their customers. Once the customer can freely decide with whom to share their transactional data a new competitive area will emerge. In this competitive environment banks will compete with non-banks and with each other for the right to use their customers' data. So if we look back to the previous Mobey Forum reports on the competitive aspects of mobile wallets we could state that not payment itself but payment data based services are becoming "the last mile" in the war of wallets. The better financial institutions use the remaining time until the existing regulation is in place to improve their skills, implement new tools and change the processes and requirements to their data analysis output, the easier they will be able to transit to the new reality. The current window of opportunity is closing. Banks had better start today.

2. PREDICTIVE ANALYTICS EXPLAINED

2.1 FOUR TYPES OF BUSINESS ANALYTICS

To explain the concept of predictive analytics it must be viewed in relation to other common types of business analytics. Gartner distinguishes four types of business analytics; descriptive, diagnostic, predictive and prescriptive analytics.

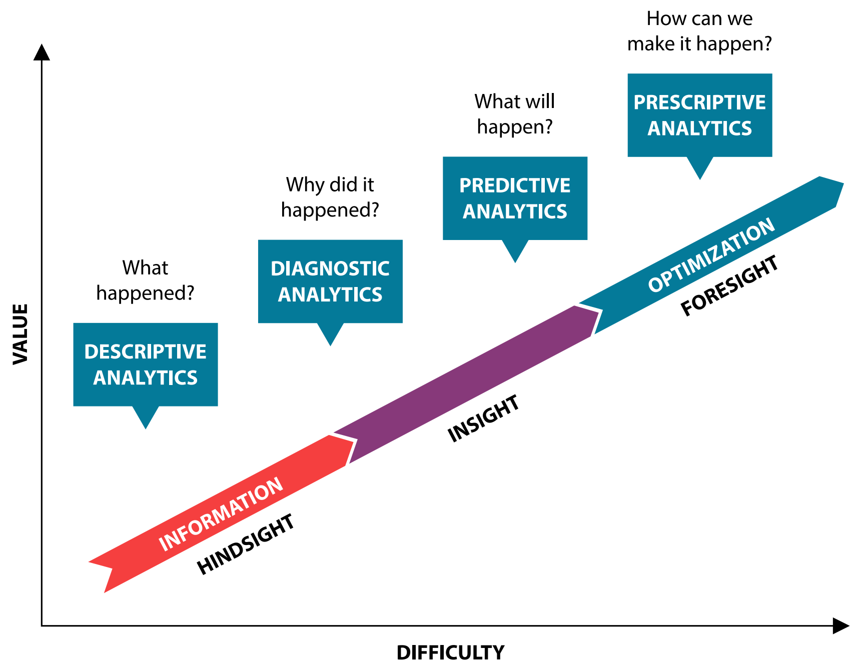


Figure 1: Gartner (March 2012). Gain value from higher levels of analytic maturity, moving from information and hindsight to optimization and foresight.

Traditional business intelligence which is used to understand the past typically includes descriptive and diagnostic analytics. It answers questions like 'what happened and why?'.

After descriptive and diagnostic analytic capabilities are in place, the next step is predictive analytics. This involves shifting from a historical orientation to become more forward-looking. This is a crucial difference, the aim is not just to describe and understand the past but also to explore 'what is likely to happen'. For this level of difficulty, we are making decisions about the future which we can only express with limited certainty. Diagnostic and descriptive analytics can be precise. Predictive analytics is, by definition, based on probability, and must be accepted as delivering 'degree of confidence' only.

In personal terms, think of the weather and questions such as: 'Will I need a jacket tonight?' or 'Should I bring an umbrella?'. Only using descriptive analytics a weather man could say: 'It rained yesterday'. But this does not help you with your current decision. Fortunately, there are weather forecasts, where you generally hear something like 'the chance of rain is 90% in the upcoming 24 hours'. Still, it is your own decision whether or not to take an umbrella, but the prediction gives you the opportunity to make a fact-based decision. In this example, you will probably take an umbrella with you when you go outside. Actually, your decision was backed up by predictive analytics.

Finally, a fourth type of analytics is prescriptive analytics, which tries to answer the question 'How can we make it happen?' or 'What should we do?'. Consider fresh food supply at a supermarket. A predictive model is used to predict demand and supply for fresh food on a daily basis. Together with a set of business rules, which will involve cost and risk considerations of the management, a prescriptive case can be constructed where the outcome tells the management (or a live system) how much additional stock is needed to reach the wanted cost-value ratio. If we think of our umbrella example once more, the final recommendation will depend on the costs we assign to getting wet or carrying an umbrella with you the whole day. According to Gartner, only 3% of companies are currently using prescriptive analytics. [5]

Financial institutions are likely to derive exponentially more value as they ascend the analytics ladder. In general, the difficulty of the analysis also increases in the same manner. For the financial industry, the current challenge is to move from descriptive (and / or diagnostic) to predictive analytics and continuing toward prescriptive analytics once the technical and cultural capabilities are in place.

5 http://www.enterpriseappstoday.com/business-intelligence/gartner-taps-predictive-analytics-as-next-big-business-intelligence-trend.html

2.2 PREDICTIVE OR DESCRIPTIVE?

Managers (and even analysts) often wrongfully think that they are already using predictive analytics. In many cases predictive analytics is used as a synonym for data analytics as a whole. The majority of the use cases within financial sector is still in the descriptive area. Let us consider three examples.

> ***Example A:*** *A Sales Dashboard with monthly sales per product shows that sales of product A is dropping month after month. Management concludes that the product is no longer interesting enough for the company and decides to remove the product from their portfolio.*

> ***Example B:*** *A cross-sell analysis shows that 30% of the customers who bought product A, bought product B the month after. The management decides to directly offer product B whenever a customer buys product A.*

> ***Example C:*** *Based on the sales of last year an analyst concludes that especially young customers are interested in product A. The management decides to call all young customers in their customer base and offer them product A.*

Note that in each of these examples a prediction is made based on an implicit assumption: what happened in the past will happen in the future in exactly the same way. This is a prediction, it may even be a fact-based one, but as is often stated in the investment management industry, 'past performance is no guarantee of future results'. Moreover, nothing is stated about the prediction value of the assumptions and there is no evaluation dataset used to create a confidence level around the decision. Let's consider the future decision in the first example; drop product A. The implicit model in the mind of the decision maker is a forecast trend analysis. The underlying assumptions are a linear model, stable data quality, no seasonal influence and no interference with any other new or competitive product. Ceteris paribus, sales of product A will continue to decline with a high level of certainty.

These examples illustrate the descriptive character common for the majority of analytics within financial institutions. Since there is no fact-based prediction supporting the management decisions, none of these examples should be considered as being predictive analytics use cases.

Despite the fact that no actual fact-based prediction is made, the analysis in example C presents a clear trigger for a predictive analytics case. Building a model to predict the probability of a customer buying product A in the future, where age is included as one of the variables.

This example illustrates the fact that descriptive analytics is crucial to understand what drives sales, business and customers' behaviour. It creates the first insights for the next level; a successful predictive analytics use case.

2.3 A PREDICTIVE ANALYTICS CHECKLIST

To understand if your company already has a (successful) use case of predictive analytics we present a checklist. Only if the following questions can be answered positively, a use case can be thought of as belonging to the field of predictive analytics:

- ✓ Is there a fact-based prediction of a future value of a variable (action or event)?
- ✓ Is this prediction obtained by relevant past information in similar contexts?
- ✓ Is a rational decision made based on the fact-based prediction of the future value?

The most straightforward example is targeted direct marketing. In this case one predicts the probability that a certain customer will respond positively to an offer given all the relevant past information on similar customers. A follow up of this prediction could be to contact all customers with a probability of positive response above a certain threshold. This completes the predictive analytics use case, since the management decided which customers to contact based on a fact-based prediction.

Examples of predictive analytics are found in business and finance, but also in sports, politics and social media. Facebook predicts which of 1,500 candidate posts (on average) will be most interesting to you and uses these predictions to personalize your news feed. Netflix sponsored a $1 million competition to predict which movies you will like in order to improve movie recommendations. The Obama 2012 presidential campaign used predictive analytics to influence individual voters. The campaign hired over 50 analytics experts to predict which voters will be positively persuaded by a particular type of political campaign message and type of contact.

These examples show that the number of operational predictive analytics examples which increase business value are growing and are no longer limited to one or two sectors. It is time to understand the main components behind it.

2.4. THE MAIN COMPONENTS OF PREDICTIVE ANALYTICS

As a manager in any industry, it helps to understand the main components of predictive analytics. The Harvard Business Review article "A Predictive Analytics Primer" lists the three basic components that underlie predictive analytics: [6]

The Data:

The confidence level of every prediction is directly linked to the underlying richness of the data and data quality, more than it is linked to the chosen model (or size of the dataset!).

The Statistics:

The set of mathematical techniques, ranging from basic to advanced, which are applied to the data to derive predictions.

The Assumptions:

The main assumption in predictive analytics is that the future will continue to be like the past, but not necessary in a linear way.

Machine Learning and Big Data are terms heard often in the field of predictive analytics. Neither of which are necessary for a predictive analytics use case, but they do offer fast-growing opportunities. The framework above enables us to address these terms and explain how they relate to predictive analytics.

2.5 THE DATA: IT'S ABOUT USING HIGH QUALITY DATA, NOT ABOUT SIZE

It is important to emphasize that predictive analytics is not about the size of the data, but about using the data to improve decision support. A regular customer database including variables like age, postal code, product purchases and contact history could be a great starting point for building a powerful predictive model. Often, organizations already face challenges in ensuring the quality of these types of data. The current lack of predictive analytics use cases within financial institutions is an incentive to start small, rather than directly incorporating 'Big Data' in your predictive analytics cases. Said differently, more low quality data is not likely to improve predictive capabilities. If you succeeded in creating an understandable model and you want to increase predictive power or robustness you can start incorporating more and different data sources in your model, slowly building towards a "big data" environment.

Big data is a popular term used in many different ways, but mainly referring to data that is too complex to be managed by traditional analytics methodologies and tools. Although there are many different definitions, the common phenomenon people have been using to further describe big data are the four V's. There is an increase in scale (Volume), forms (Variety), speed of generation (Velocity) and uncertainty (Veracity) of data sources when it comes to big data. In the end, opportunities of predictive analytics will obviously increase when other and larger amounts of data are stored. But the fourth V (Veracity) will give additional challenges when including big data sources in predictive analytics.

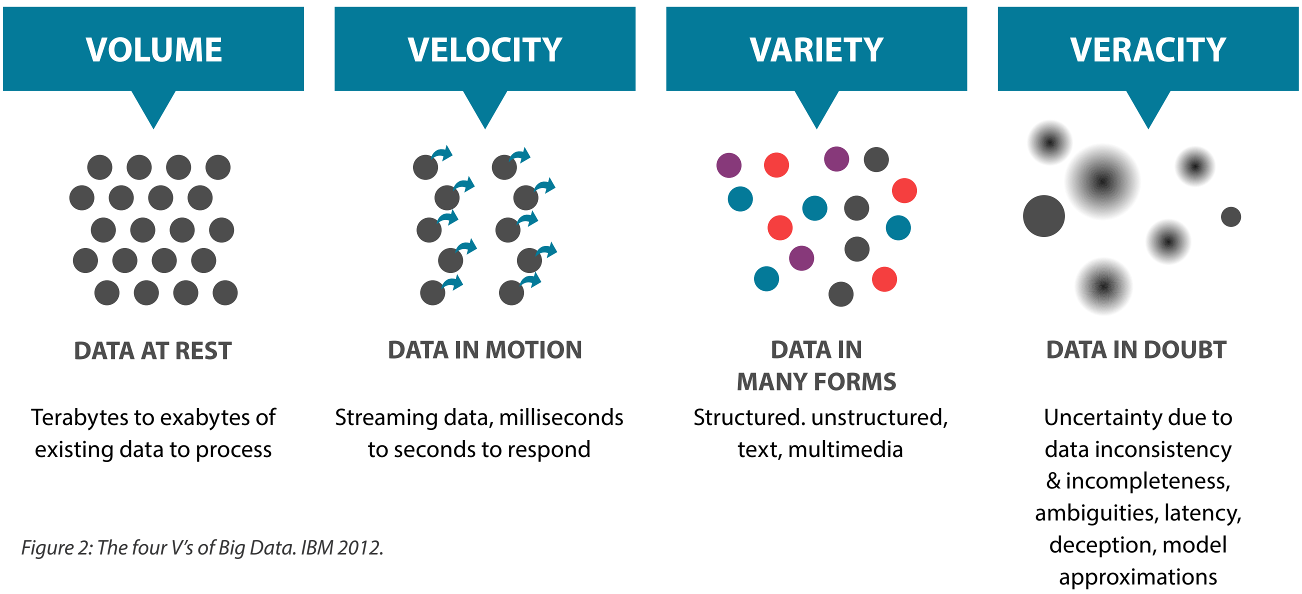


Figure 2: The four V's of Big Data. IBM 2012.

6 https://hbr.org/2014/09/a-predictive-analytics-primer

2.6 THE STATISTICS: CLASSICAL METHODS VERSUS MACHINE LEARNING

There are different techniques available which can be applied to a dataset in order to derive predictions. For banking professionals, the first thing to understand is the significant difference between classical statistical methods and machine learning, or who formulates the relationships between variables when creating a model? Second, since the number of data sources, variables and models keeps on growing, the question arises how to perform iterations on the created models in a more automated way.

Who formulates the relationships between variables?

In the case of classical statistical methods a human expert In the case of classical statistical methods a human expert formulates and tests relationships between different variables. For instance, what is the influence of age, gender or postal code on sales? When such a relationship is established, the relationship (or obtained formula) can be used to derive predictions.

Machine Learning is often used to invert this process. It uses "past data" to find patterns that can be used to solve similar problems in the future. In this case, the starting point is not the different explanatory variables but the outcome (sales). Next, a computer is taught to automatically uncover the factors that are driving sales. So machine learning is a 'field of study that gives computers the ability to learn without being explicitly programmed' [7]. If done properly this results in powerful predictive models. Learning from data is an iterative process that requires multiple tasks to be performed within each iteration (e.g., training multiple predictive models with different parameters, evaluating them to select the best performing predictive model and deploying it in production in order to generate predictions).

Who performs iterations of machine learning tasks?

Data quality, as described above, is the ultimate boundary for what can or can't be done with classic statistics and machine learning. Moreover, in both cases, updating models might be time consuming. Customer behaviour (and hence the data) is changing over time, making predictions based on non-updated models inaccurate. Iterating on these models to keep them up to date is a human labour intensive activity. Apart from applying machine learning techniques instead of classical statistics, there is an important difference regarding who performs the iterations on the models.

Industrialisation of machine learning enables automation of the machine learning tasks. It reduces the amount of associated human labour by simultaneously increasing the number of feasible predictive use cases. The technology giants today are already capable of increasing the granularity of predictive models to the point where they daily automatically generate a unique predictive models per customer or product. Moreover, performances of models are constantly tracked. Triggers are in place which will automatically update a model if performance falls below a certain threshold. This could remove current barriers for organisations to apply predictive analytics on the level and complexity needed to deliver individualized and contextualised customers experience.

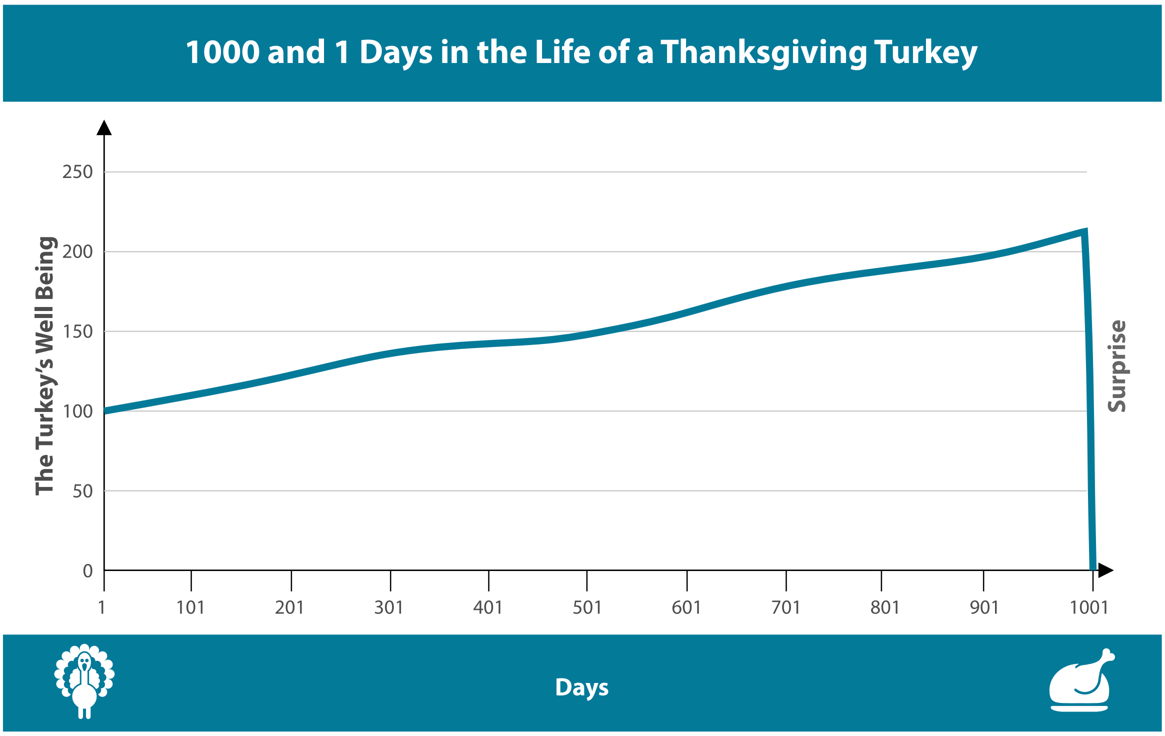
7 Arthur Samuel, 1959

Classical Statistics	Machine Learning	
Human formulates the relationships between variables	Algorithm formulates the relationships between variables	
Human tests the relationships between variables	Human performs iterations of machine learning tasks	Machine performs iterations of machine learning tasks

Figure 3: Technology inclusion in statistics, BigML 2016

2.7 THE ASSUMPTIONS AND ITS LIMITATIONS

The number one reason making the assumptions of a predictive model invalid is time. Has the new input data changed in comparison to the historic data used to create the predictive model? Is it necessary to create a new predictive model because the old model doesn't work well with the new data? Are all the key variables used in the old model still relevant? As Nicolas Taleb describes in his book Black Swan[8], the model that predicts the Turkey's well-being for the first thousand days of its life doesn't predict the fatal thousandth and first tragic day. It looks like some key variables were missing in its well-being model which relevance for model's accuracy became evident two days before Thanksgiving.



Finally, even if the assumptions still apply to the current situation predicting the right outcome doesn't necessarily mean that this outcome can in every case be influenced. Some things will happen no matter if we were able to predict them upfront or not. So predicting that 'the chance of rain is 90% in the upcoming 24 hours' like mentioned before, won't enable us to prevent the rainfall. Other future predicted outcomes can directly be influenced. When a bank is able to predict which customers are likely to churn, action can be taken and the customer behaviour might change.

8 Nicolas Taleb, The Black Swan

3.
KEY AREAS TO CREATE VALUE WITH PREDICTIVE ANALYTICS FOR FINANCIAL SERVICES

Used effectively, there are several areas of application for predictive analytics where financial institutions could make profitable investments while at the same time improving the attractiveness of their services. From a high-level business perspective, value for business breaks down to commercial effectiveness, risk control, operational efficiency and financial control. This so-called 'Value Tree' can be used to structure the different use cases of predictive analytics for financial industry.

When focussing on commercial effectiveness, examples of predictive analytics are card linked offers or next best action marketing. If operational efficiency is the goal, predictive analytics can for instance be useful in a predictive maintenance case. The one to one relationships between the example and the created value for the business will not always be so straightforward. If predictive analytics is used to decrease the down time of ATM's (predictive maintenance), this might also increase customer satisfaction and retention.

In this white paper the highlighted cases are discussed in more detail, the majority of these serve multiple goals and create value for the bank and the customer.[9]

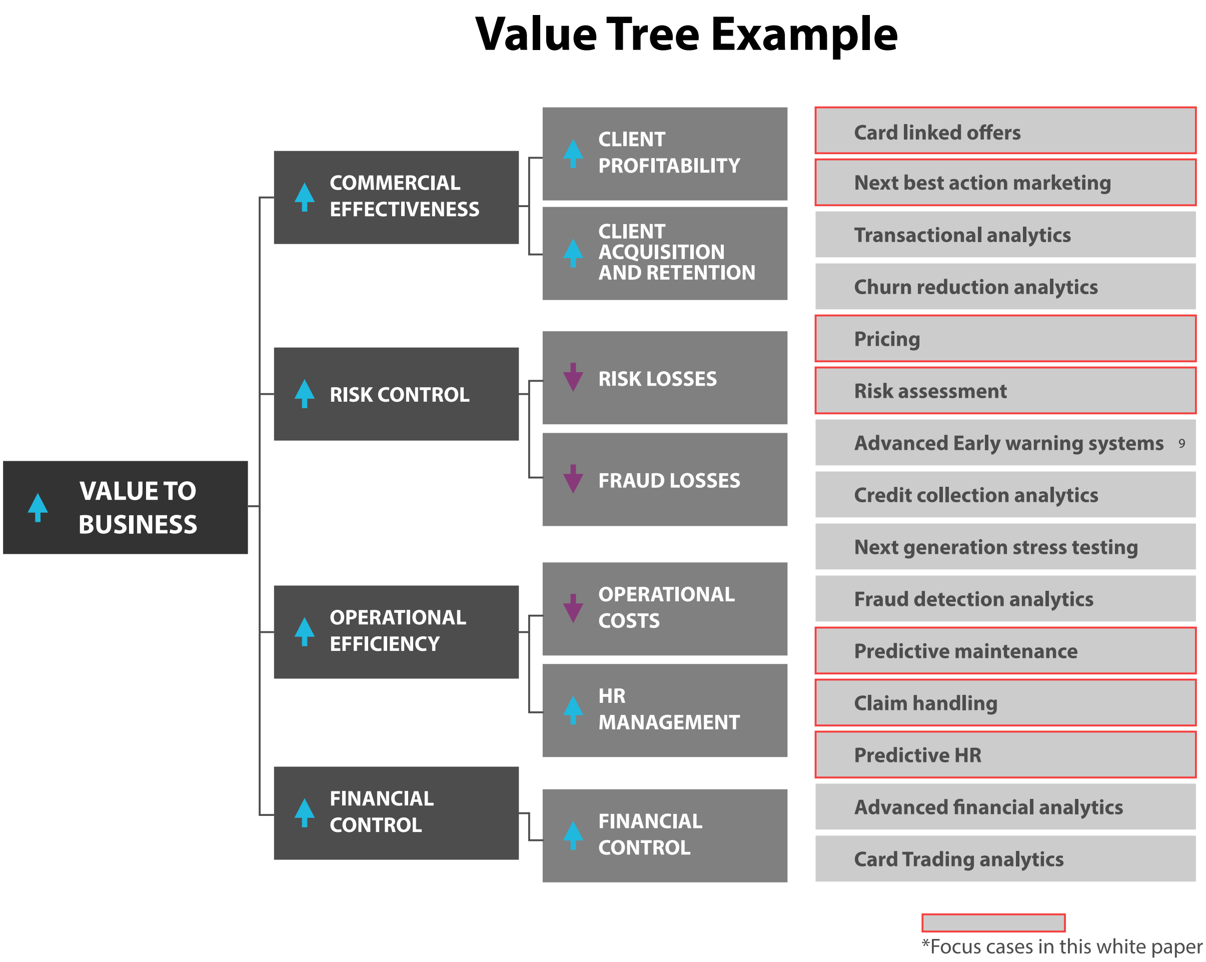


Figure 4: Value Tree.

9 Early warning systems: analytics used to identify certain events (the time needed to sign in or the pressure of how a pin code is entered on a touch-pressure sensitive device) which do not conform to an expected pattern.

Card linked offers

Several financial institutions are looking into new business models, where their customers are offered discounts from retailers, provided that they use the banks' card. So called 'card linked' offers have been around for a while, but has great potential when combined with analytics to ensure the offers are well timed and contextually appropriate from a consumers' point of view. And when this 'card' is placed in the context of a mobile application and/or mobile payment capability, the opportunity for the delivery of content-rich, timely, and relevant messaging is increased dramatically. Aggregating payment information across different payment service relationships of a single customers is also one of attractive value propositions for non-banks (see PSD2 above).

Next best action marketing

Card linked offers are already an example of next best action marketing, but the range of actions is mainly focused on making a new offer. The increased amount of interactions in the digitized journey offers opportunities to add value for customers in various ways, more than simply presenting offers. Financial institutions have extensive knowledge of their customers financial and lifestyle habits. The banks that are able to use this information to offer their customers better financial advice and support them in making decisions could earn both the customers' loyalty and increase sales. During all the interactions with a customer, banks should constantly evaluate the next best action. This range of actions is not limited to offers, but can include service, financial advice, etc. Algorithms constantly evaluate the different actions that can be taken for a specific customer. The one which is considered the 'best' in terms of timing and relevance is presented to the customer. Just imagine what the opportunities are for the next best action concept in the mobile customer journey. Since they are prosperous, they are discussed in more detail in the next chapter.

Pricing

In many countries, comparison sites are gaining popularity and are dominant in the client acquisition process. For products like savings accounts a rational customer should always compare interest rates before selecting their product and offering bank. This evokes a game theoretical question: Given the interest rates offered by my competitors, what interest rate should I set to balance customer acquisition and profitability? And if I take a certain action, how are my competitors likely to respond? Based on past customer behaviour, a bank can model the in- and outflow of money. Predictions made using these models provide a fact based instrument and improve the ability of setting the right interest rate.

Risk assessment

More accurate predictions about future behaviour would give the financial institutions the chance to assess risk even better. This would allow for lower risk premiums or higher profits by reducing both false positives and false negatives. From a customer point of view, this would on average be well received when it is provided with a rational explanation, since no one benefits from allowing customers to take on debt they are not able to service.

Predictive Maintenance

Customers expect banks to operate 24/7, so the downtime of a bank's infrastructure needs to be minimized. For example, an ATM that does not work frustrates customers. If a bank could predict these types of equipment failures, maintenance could be scheduled in order to prevent failure or minimize downtime. This technique is known as predictive maintenance and this approach is preferred over routine or time-based maintenance, because tasks are only performed when needed.

Claim handling

Claim management can be a time consuming and costly task for back offices of insurance companies. Based on components of a claim, such as the value of the claim, nature of injury, place, characteristics of the claimant and the customer's insurance history it is possible to predict which claims are likely to be fraudulent. These cases clearly need attention and can be prioritized and assigned to experienced claim handlers. On the contrary it is also possible to predict which cases are not complex and have a close to zero probability of being fraudulent. It could be interesting to not assign these cases to a claim handler, but settle them automatically and transfer the claim amount to the customer. This decreases claim duration and therefore positively influences the customer experience.

HR Analytics

HR departments of financial institutions are continuously searching for potential employees. But which profile to hire? Techniques for using predictive analytics are available, but not widespread among HR departments of financial institutions. By analysing data of current employees, HR departments could determine which variables (age, experience, grades, etc.) contribute the most to performance and productivity. When hiring new employees this makes it possible to predict if a certain profile is likely to be successful in the company

.

4. A CLOSER LOOK AT NEXT BEST ACTION: OPPORTUNITIES OF PREDICTIVE ANALYTICS IN THE MOBILE JOURNEY

Mobile is a still growing channel for financial services, both in terms of sales and services. In this section we will discuss the opportunities of predictive analytics within the mobile journey. Of particular importance is the introduction of the 'next best action' concept. NBA could create a personalized customer experience, while at the same time boosting sales and profit.

As mentioned before, mobile banking apps allow customers to do their banking when and wherever they want. In all these interactions a bank can build knowledge about every customer to better serve that customer, constantly adjusting service and offers based on the context of that customer. This topic will be explored more deeply in future papers, where Mobey Forum will focus on the path from next best action to mobile digital assistants in financial services.

From static offers to real-time recommendations

In the past, banks have relied on basic segmentation and static offers delivered through standard scripts from their service representatives (mostly descriptive analytics). These have often focused on retention and product promotion campaigns, and static decision trees that are used for all customers regardless of an individual customer's past history with the bank. A predictive analytics based NBA concept takes into account all the known information about the customer, including interactions or events, to arrive at optimal next best actions in the form of real-time recommendations, or real-time automated actions. New style NBA concepts also consider the optimal channel for the offer or interaction based on current channel, channel preferences, and geolocation, be it the branch, Web, contact center, ATM or smart phone.

If only business ruling is used, a marketing automation system will constantly fire different rules at customers. This can potentially lead to an enormous number of service and offers, which in turn can give the customer the feeling of being spammed. A predictive analytics based NBA concept tries to increase the value of interaction for the customer in such a way that the "annoyance factor" stays low. Only if the models predict that it is likely enough that the offer or information serves the customer, the action will be executed.

Continuous ranking of possible actions

To successfully develop a complete NBA concept, a 360-degree customer view is required based on historical client data and all real-time events (triggers) a customer experienced. Based on this customer view, the experienced triggers and the business KPI's, a probability of success is created for every possible action. Basically, behind every action is a predictive model. If a new trigger arrives, all models are evaluated again. This real-time evaluation creates a continuous ranking of all possible actions. When a customer logs in to his mobile app, the best possible action based on the probability of success is presented. It is wise here to include some kind of threshold. If none of the listed actions is likely to create success, it is better to do nothing. Remember, the best possible action can therefore also be to present 'no action at all'!

The 'predictive' explained

To further explain the predictive modelling within a next best action engine, consider the examples in the framework below.

In the first case a predictive model evaluates the probability of churn for customers with a savings account. The probability of churn depends on some general customer characteristics (like age, gender, number of products), but the model also identifies and includes a number of triggers which increase the probability of churn. An example could be a customer reading the conditions of the savings accounts. The probability of churn for this specific customer now increases above a determined threshold. Because of this trigger, all other models and corresponding actions are also evaluated (think of cross-sell, up-sell or service). For this customer, the action corresponding to churn ranks number one on the list of actions. The customer opens his mobile app again and finds an offer for a new savings account, including a gain calculation. The gain convinces the customer and he transfers his balance to the new product. The bank supported the customer in his mobile journey and boosted sales and customer experience.

The other examples have a similar structure. The service example shows that a next best action is not necessarily an offer. The model in this case predicts a negative account balance in future months. Compared to all other actions, simply notifying the customer is the best thing to do.

Using next best action in the mobile journey is therefore not just focussing on sales, but it is a way to constantly help customers in their financial mobile journey. When the timing and context are right, the personalized customer experience will boost customer loyalty and profit in the long term.

	RETENTION	SERVICE	CONTACT
360 DEGREES' CUSTOMER VIEW	Customer with a savings account has a high probability of churn.	Customer has a positive account balance for a period of two years.	The customer appreciates direct and fast service, he used the chat functionality multiple times.
REAL TIME TRIGGER	Customer reads conditions of his savings account in the mobile app.	The account balance of the customer decreases to below 1.000 for the first time in two years.	Customers spends more than two minutes within the mobile app on different pages.
NEXT BEST ACTION	The bank places an offer in the mobile app for a special savings product, including gain calculation.	The bank notifies the customer on a possible negative balance and explains the consequences.	An option for direct chat appears: 'Can I help you?'
DESIRED EFFECT	The gain convinced the customer and he transfers his balance to the new savings account.	Customer appreciates the service and transfers additional money to avoid the negative balance	Positive customer experience by proactive serving the customer via his prefered channel.

Figure 5: Example of Next Best Actions concept using Predictive Analytics. VODW, 2016.

5.
THE 10 CHALLENGES ON THE PREDICTIVE ANALYTICS LIFECYCLE

When embarking on the journey of predictive analytics there are challenges that need to be considered and preferably taken care of before they become non-overcoming hurdles. Overall, a first important task to complete when considering predictive analytics usage is to determine whether or not your organization is prepared and equipped to work effectively in a data driven environment. It is only within such a culture that a company will fully benefit from the opportunities of predictive analytics. Predictive Analytics should be viewed within a broader context of how the Financial collects, stores, accesses and uses customer and transactional data. This context has both ethical and cultural considerations.

Building a data driven organization should not be considered as an IT and/or intelligence department project. Effective usage of data will mean a culture change in your organization that's driven by using analytics and data input in practically every step of your decision-making processes.

Creating this data driven organization and at the same time enabling the power of predictive analytics is a continuous process. Let's examine the ten different challenges on the predictive analytics lifecycle to frame some very basic initial key questions about your organization's readiness for successful use of predictive analytics.

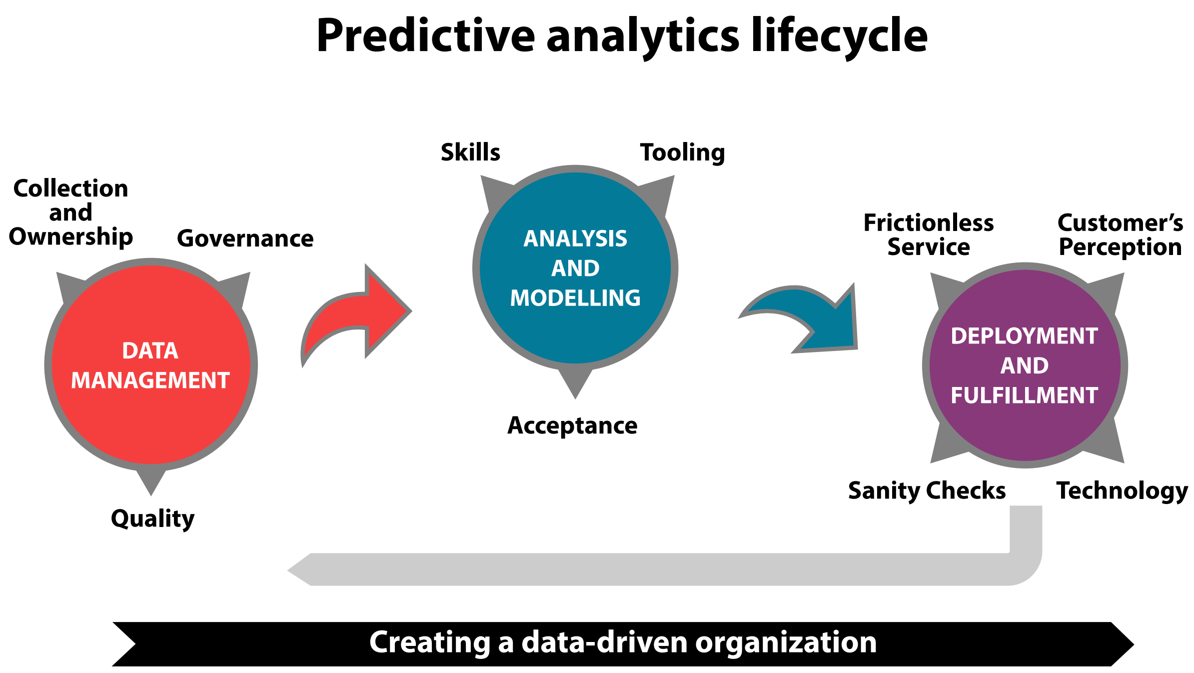


Figure 6: The predictive analytics lifecycle. VODW, 2016.

5.1 DATA MANAGEMENT

Data collection and ownership

When performing data management, the first question is to address is: 'What data should we collect?'. How can we find the meaningful data nuggets in a world where the amount of data is doubling every 18 months[10]? To answer this question, companies should inform themselves regarding questions like: 'How much may be stored, used and shared?' and 'May we combine data from other data sources?'. Data Ownership is another crucial topic. Who actually owns the data? Is it the customers' data or banks data we are dealing with? How will the 'right to be forgotten' influence this? Imagine that aggregated customer data is used for narrowing various outcomes and some of the customers wants to be forgotten. May financial institutions still use the aggregated data or will this dilute the data that the predictive analytics uses? There are many unanswered questions and the answers will vary from jurisdiction to jurisdiction. It is important to (at least try) to answer these questions, so the analytics can be created on a sound legal basis.

10 IBM. http://www-01.ibm.com/software/data/demystifying-big-data/

Data governance

Combining and managing multiple sources and types of data to create a 360-degree customer view is a major challenge for financial institutions. Historically, data from the transactional world (sales, payments etc.) has been available and is reliable. Customer interaction data however is a new world for financial institutions. Online analysis and marketing automation tools enable them to track customers in their entire online journey, creating an interesting new data source. Linking these new data types with the more traditional BI data sources is a challenge for financial institutions. These different data sources are typically stored in different silos, owned by different departments, each with different approaches to data management. These disconnected systems and incompatible data leads to analysts and IT staff spending too much time and effort in combining or aggregating data, leaving little time left for actual usage of the data to add value for the customer. When collecting data across different products and channels, organizations should work towards a synchronized data storage and collection entity.

Data Quality

Quality is the main determinant of a model's performance. The confidence level of every prediction is directly linked to data quality, more than it is linked to the chosen model or technique. The quality of data is influenced by issues like incompatible definitions, inconsistency, incompleteness and duplication. When these issues are too dominant, there will be a justified lack of confidence in presented analyses and predictions. Making them worthless. Wrong predictions might eventually even drive customers away or cause wrong investments.

5.2 ANALYSIS & MODELLING

Skills

If the data is sufficient enough to start modelling, new skills and capabilities are required. So far most skills in financial institutions have been used to look at business intelligence in hindsight. Now analysts need to make predictions about the future. Reporting skills are no longer sufficient, advanced modelling skills are required. The majority of financial institutions might have a first team of data scientists in place, which created comprehensive models in the past for credit scoring or churn. But is this team future proof? Are they able to deal with new challenges like unstructured data and machine learning? Moreover, are these 'rocket scientists' capable of translating their ideas to value added business concepts? These could be reasons to hire new competences, enter into a partnership or even use crowdsourcing.

Tooling

Based on its value proposition to the end-user the tools for predictive analytics can roughly be divided in three groups:

1. Open source tools like R and Python are widespread in the data science community. Proficiency in programming is required to derive value from them. Despite the right skill sets and the potential savings in licensing costs companies must also factor in the heightened need for support and services as well as the lack of enterprise-grade requirements such as maintainability and security out of the box.

2. Traditional licensed enterprise solutions of well-known brands. Those solutions rely on the close relationships between the providers and their installed base of customers and cover a wide range of specific business related use cases, but come at the highest Total Cost of Ownership among the choices. These solutions are under pressure to adapt to the newest developments in technology in the cloud era.

3. Cloud-based machine learning as service platforms complement data science teams by lightening workloads, automating machine learning tasks and helping the users to prototype ideas faster but they also enable non-data scientists (such as business analyst and developers) to learn and utilise predictive analytics in their business context. It has yet to be proven how fast these types of solutions will gain ground in the financial industry.

A recent trend toward hybridisation in the market for predictive analytics has begun to emerge. Open source tools and traditional licenced enterprise solutions are more and more trying to embrace a cloud and API-driven approach to keep up with the end-users' requirements. At the same, cloud-based machine learning as service platforms are maturing quickly, providing the users with the flexibility of the open-source solutions and reliability of the traditional licensed enterprise solutions.

Acceptance

The affected organizations (e.g. marketing, security) should be able and empowered to deploy effective actions based on data insights. Frequent interactions between analysts and managers are needed to ensure that analytics support the most important business decisions. When only the analyst understands the power of predictive analytics, it is doomed to fail. On the other hand, model designers need to understand the types of business decisions managers make. This will increase the relevance and acceptance of their analysis.

5.3 DEPLOYMENT & FULFILMENT

Frictionless service

An action based on predictive analytics can only be effective if it is delivered in the right channel and context and ends in a measurable result that is, ideally, frictionless for the customer. Making the perfect prediction for your customer will only turn into a negative experience if the customer can't act and close the given opportunity, right there and then, in the same context.

Customer's Perception

When applying predictive analytics in a service environment it is paramount that the customer's service perception be carefully considered. Predictions that are too accurate can unnerve customers, induce paranoia and deter them from engaging with the brand and service. Indeed, service providers may find that their customers' 'creepiness threshold' is lower than they anticipate. Here it is very easy to accidentally push customers away. One good, albeit descriptive, example of this was when banks introduced pie charts illustrating where customers had used their payments cards. Some customers were surprised that banks knew that much, even though it was just a graphical illustration of their account statement. This ethical consideration is an especially sensitive topic for financial institutions. While people are generally starting to get used to all the privacy sacrifices they make in social media services, large groups of people still want to maintain a selectively opaque relationship with their bank in order to maintain their image as a good customer.

Financial institutions need to consider carefully 'the creepiness effect' and find a solution. A first step could be to provide the customer with help and explanations, always appropriately matched to the customer's individual level of understanding, by addressing questions like: 'What do we do with your data?', 'Why should you allow this?', and 'What's in it for you?'.\

Sanity checks

Algorithms can sometimes 'be intelligent' in a very dumb way. Sanity checking or including some kind of restrictions on the outcome or action are therefore of crucial importance in many predictive analytics use cases. Two illustrative examples demonstrate why.

In 2015 a number of customers were banned from all services at their bank. The bank closed their accounts and deposits, leaving them unable to pay for food and other essentials. They were victims of the de-risking process of a large global financial institution, in which customers who were considered to be a potential threat to the bank were 'dumped' in order to avoid fines from authorities. When the consequences of an action are so severe, human sanity checks should be incorporated. Linking the algorithm directly to the execution of the action can severely damage customers and also the image of a bank.

An example from another company, Uber, shows why including a restriction might sometimes help in avoiding situations where the usage of an algorithm loses touch with reality. In December 2014, a hostage crisis took place in Sydney. As people sought to flee Sydney's business district, Uber's dynamic pricing algorithm responded to a big spike in demand, leading to single ride taxi costs of $80 U.S. There was a negative impact on the customer experience of Uber users, since they had to pay a huge fare to flee from a crisis situation. Uber could have avoided this situation by including some kind of restriction in its automatic pricing algorithm. Instead, a simple rule could have been applied, capping a short ride fare at AUS $50, for example.

Technology

Obtaining and processing data and delivering the service to the right location at the right time can be a challenge. Is the mobile device reachable and enabled (e.g. roaming)? Can your own systems process the data fast enough remotely to deliver real time interaction if needed or do you want to run the predictive model locally? Do you get the external data fast enough? How often do you want to automatically update the predictive model to ensure its as accurate as possible (every 3 months or every day)?

6.
THINK BIG, START SMALL, BUT START NOW

The previous chapter described the hurdles to overcome before predictive analytics can begin to deliver benefits. Since all of these challenges will become relevant sooner or later when embarking on the journey of predictive analytics, the advice is to be aware of the range of change ahead, and appreciate that the changes will not be incremental. At the same time, the best advice is to start small and simple, but start now.

Think big

If the digitization that the financial services industry has witnessed considered as a big step forward, then the shift caused by artificial intelligence will be immense. In the near future we will experience a level of service provided without any human interaction that we cannot imagine today. This new benchmark, similarly to mobile services in the past decade, will be set by the players outside of the financial industry. This will force the industry to adapt, the same reason that we see banks partner with fintechs and start calling themselves technology companies. This adaptation will be even less incremental than the one we experienced before and will force financial industry to master predictive analytics and put data at the core of their business. But this doesn't mean the banks should follow every technology hype around data and hire hundreds of data scientists.

Start small

Mobey Forum advises starting with a small, multi-disciplinary team and identify a simple goal which delivers value for the business and/or customer. This multi-disciplinary team should be able to complete the job from start to finish in a relatively short period of time. With respect to the described journey of predictive analytics this means the team should be able to collect and analyse the relevant data, build a model and finally use the model to execute some action towards the customer. The team should be allowed to experiment and work autonomously, enabling them to quickly deliver results which add value. Think, for example, of a predictive churn model for one single product. Start with a limited number of data sets for this purpose which are easy to synchronize. Create a simple predictive model and use the model to pro-actively contact customers with a high probability of churn and measure any churn reduction against a similar control group of customers. Start small, fail, learn and succeed. [9]

Start now

Traditionally IT and data projects have a linear and sequential approach. A typical project has a number of phases like target setting, requirements, design, implementation, verification and maintenance. All phases have fixed deadlines and the actual value of a project for the business or the customer is delivered when all phases are completed and tasks are done. If this principle is used to build a (big) data infrastructure which satisfies all needs of the business, it will probably take months or even years before the infrastructure is completed. Using predictive analytics to gain value from the data will only start after this period. A scenario which should be avoided since the delivered value for the business or customer is minimal during a long period of time.

Successful examples will create leverage for predictive analytics in the company and will create curiosity and questions from other department, generating new cases. Now we can move on and start incorporating new data sources. This iterative and incremental development approach allows you to work towards a 'big data' infrastructure on the long term, while constantly generating value for the business and the customer on the short term. The motto should be think big (data), but start small (data).

Mobey Forum's Workgroup for Predictive Analytics

CO-CHAIRS

Kasper Sylvest	Danske Bank
Amir Tabakovic	BigML

EDITOR

Nick Jetten	VODW

CONTRIBUTORS

Jonathan Atkinson	Fiserv
Couchon Benoît	Worldline
Bent Bentsen	DNB
Ivan Bracic	Infobip
Michel-Ange Camhi	Worldline
George Charalambous	HSBC
Joaquim Costa	SIBS
Lars Henrik Dahl	Nets
Janis Dirveiks	Nordea
Yuri Grin	Intervale
Douglas Hartung	Diebold
Markus Iofcea	UBS
Felipe Lopez	Tecnocom
Neil McDonald	Sequent
Anna Mialet	CaixaBank
Oriol Nunez	CaixaBank
Krisztian Otlakan	Erste Group Bank
Stefan Piel	UBS
Maurizzio Poletto	Erste Group Bank
Oscar Sala Palomes	CaixaBank
Ville Sointu	Ericsson
Jukka Yliuntinen	Giesecke&Devrient
André Zölch	PostFinance
Bastiaan van de Rakt	VODW

mobey forum

www.mobeyforum.org

@mobeyforum

#mobeyforum

/mobeyforum

Group: Mobey Forum