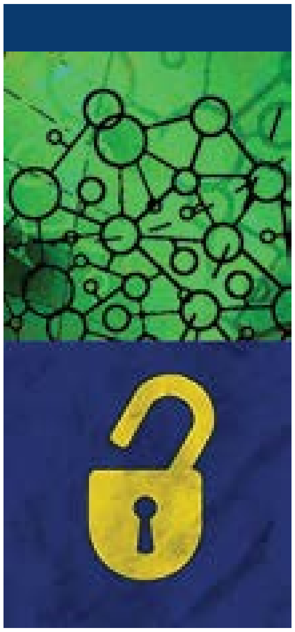
McKinsey&Company

McKinsey on Payments

Special Edition on Advanced Analytics in Banking

Volume 11 Number 28 August 2018



Using data to unlock the potential of an SME and mid-corporate franchise

Banks have long pondered the untapped value of the commercial segment but often lack the means to identify the precise needs of individual companies in this large and diverse population. This is changing, however. By mining huge reserves of customer data, banking analytics leaders are meeting the needs of hundreds of thousands of commercial customers—from small businesses to medium-size corporations—with new levels of convenience and cost efficiency. Several banks have achieved a ten-fold increase in the success rate of product recommendations, thus delivering highly relevant offers with clear economic benefit. This article highlights recent examples of how "next-product-to-buy" (NPtB) recommendation engines are identifying time-critical needs for their small- and medium-size enterprise (SME) and mid-corporate clients.

Ignacio Crespo
Carlos Fernandez
Huw Kwon

The problem: Most businesses are "invisible"

Many banks currently use rules-based models to generate recommendations for SMEs and mid-corporate companies (with annual sales up to $100 million), but with limited success. Relationship managers often view these recommendations with skepticism, as conversion rates typically range between three and five percent. They resort to general propositions designed for the consumer segment and devote most of their energy to those clients whose businesses they already know well and whose needs they can anticipate reliably. The result is that 25 percent of a bank's commercial customers usually account for 85 percent of the revenues, and the remaining 75 percent represents the "long tail" of untapped potential. These companies are effectively invisible to the bank's sales force (Exhibit 1).

The solution: Anticipate the client's next step

Banks are investing in building up predictive models globally: US Bank and TD Bank in North America; Itau and Banco do Brasil in Latin America; Barclays Bank and Lloyds Bank in the UK; ING, Banco Santander, and BBVA in Europe are just some examples of banks improving their commercial performance by leveraging machine learning. These advanced techniques have proven effective in diverse customer segments, from self-employed individuals to large corporate customers. SMEs and mid-corps are the sweet-spot for NPtB, as they generate massive amounts of data, which are typically underused. With the help of advanced analytics decisioning engines, banks have demonstrated that it is now practical to mine vast (and often messy) amounts of data, separating signal from noise, to arrive at precise recommendations for a client's next action. In addition, by broadening the types of data collected for the commercial segment, banks are also analyzing customer behaviors, transactions, and customer preferences across more extensive databases.

Successful implementation of NPtB engines has boosted new sales upwards of 30 percent and increased commercial segment revenues by between two to three percent. The impact on sales efficiency has been radical in some cases, with an increase of more than 50 percent in the number of leads offered per client and as many as six out of ten customers purchasing a new product in response to a sales call.

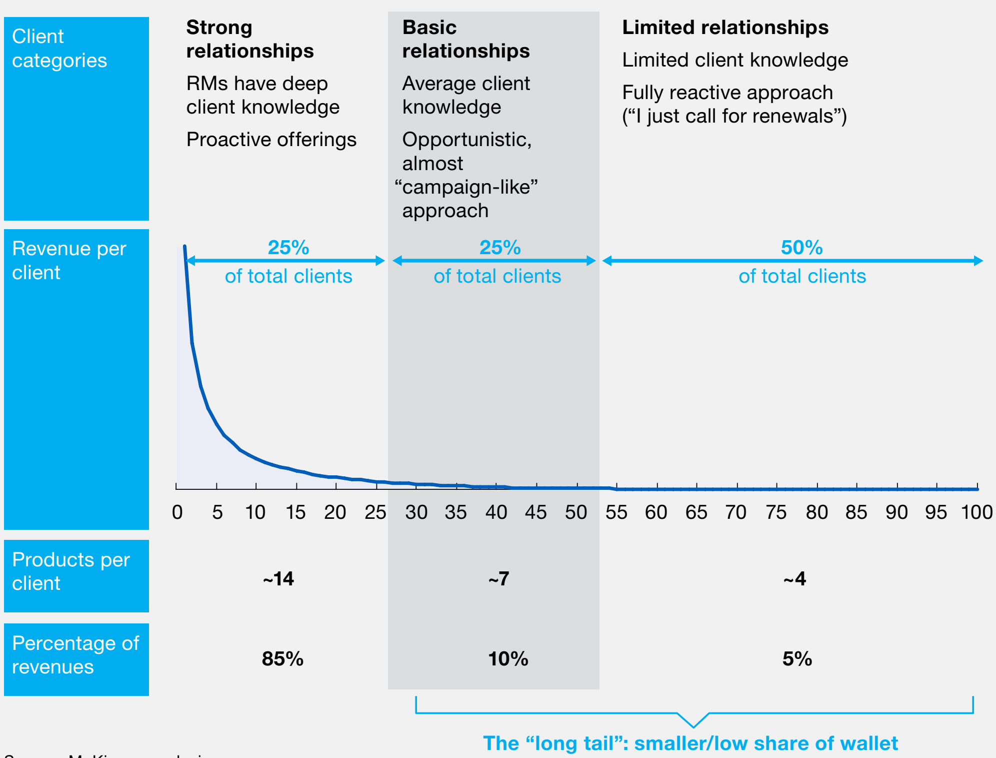
Leveraging data for NPTB recommendations

More than a decade ago, Amazon and Netflix began leveraging data and analytics to improve their cross-selling efforts. They started with simple analytics, dividing huge customer populations into several dozens of microsegments according to key behaviors (inputs). In order to achieve this new level of precision, they used singular value decomposition (SVD) to classify customers according to patterns in their purchase histories, each pattern culminating in a target output, that is, the "next product to buy." The number of inputs and the complexity of the algorithms used to analyze these inputs have been increasing in recent years, achieving outputs that have much greater precision than was possible with next-best-action (NBA) models. (See sidebar on page 33 for a summary of the evolution of NPtB from NBA.)

The NBA engines employed by Amazon (with more than 300 million customers reported in 2016) and Netflix (125 million subscribers reported in the first quarter of 2018) are not

Exhibit 1

Cross-selling efforts often focus on the companies that relationship managers know well.



Source: McKinsey analysis

entirely suited to banks serving SMEs and large corporations with products/services that address a relatively narrow range of business activity—domestic and cross-border payments, financing, documentary credit, investments, and insurance. By building NBA recommendation engines designed specifically for transaction banking, banks have increased service levels and profitability, improving their responsiveness to SMEs and helping large corporate clients cut through complex banking relationships and account structures to optimize liquidity.

To maximize the impact of each recommendation, decision engines should identify both customer needs and the preferred channel(s) for delivering the proposal and related communication. In some markets, companies tend to rely more heavily on direct communication with relationship managers, who play a key role in following up on recommendations. In other markets, such as the Nordics and the UK, the digital channel is the primary means both for alerting a customer to a recommended action and for delivering more detailed information about the opportunity.

Consolidate data for analysis in three waves

The data reserves required to power an NPtB engine are consolidated in three waves. As the volume and complexity of data increase across the three waves, analytical algorithms become progressively more sophisticated and accurate in predicting precise, time-critical needs of individual customers.

The first wave starts with the aggregation and analysis of internal structured data of various formats, including customer demographics, product usage, profitability, and transaction history. For example, one bank in Europe started by consolidating the information it had for 1.3 million SME customers, ranging from beauty salons, doctor's offices, and family-owned stores to small manufacturing companies and technology start-ups. This data set yielded 1,200 variables for analysis.

Continuing the focus on internal data, the second wave introduces algorithms capable of digesting unstructured data (e.g., call records, email communication), as well as a broader range of structured data from CRM systems (e.g., share of wallet, historical risk scoring, maturities, customer relationship lifecycle, company value chain, and suppliers). Fast-evolving algorithms augment the value of data already at hand by learning to recognize unanticipated clusters and associations in increasingly complex data sets. The algorithms generate actionable insights into a company's current needs, from payables management to financing for new equipment, based on information coming from transactions and payments along the customer value chain.

The third wave analyzes a broad range of data from point-of-sale transactions to industry news and comments on social media to generate ever more precise recommendations. As machine learning algorithms become more sophisticated, it is possible to produce recommendations from increasingly diverse types of unstructured data (including, voice, image, and video files) extracted from industry and company web sites, as well as news and social media (Exhibit 2).

How to build an NPtB engine: Design, develop, deploy

The preparation of an NPtB model moves through three phases: design, development, and deployment. Clear milestones mark the advance from analytics to proof-of-concept to implementation in the front line.

Design the NPtB engine

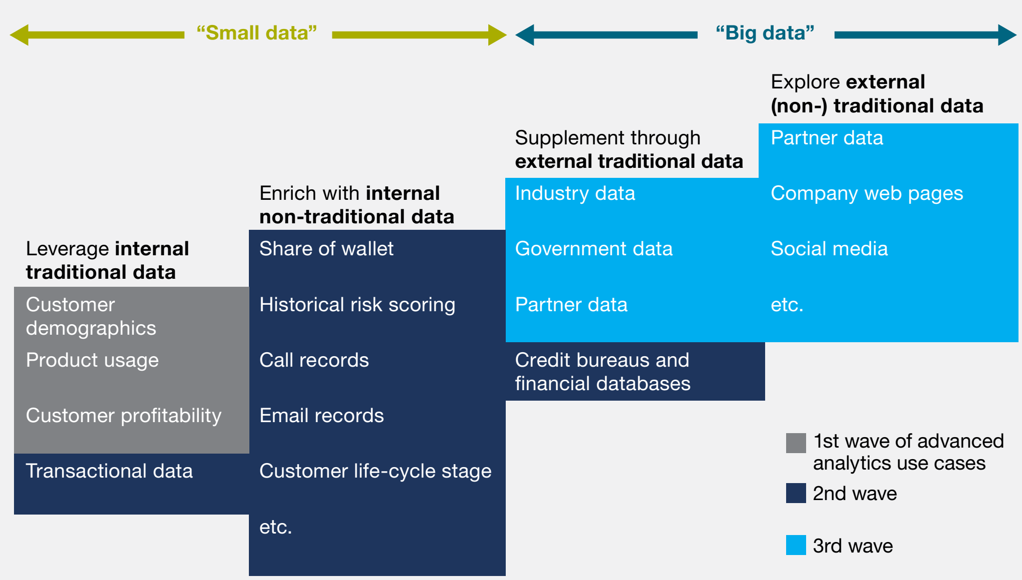
This first phase is the preparation and design of the NPtB engine—adjusting the scope, mapping pitfalls, and elaborating the business case to convince stakeholders throughout the organization of the value of the effort. This phase has five key activities:

1. **Prepare the data.** Looking across the three waves of data consolidation, the first step is to identify the types and sources of data and map the variables that can be analyzed. The goal is to locate data that can be combined with unique customer identifiers and provide sufficient record history (at least two years of data). It is also necessary during this phase to check data quality (that is, data consistency, format validation) to identify points for improvement and to restructure data ingestion (for example, defining treatment of unavailable values and primary keys, synchronizing time periods).

2. **Ensure the IT infrastructure meets the processing requirements to run the model.** It is important to design the recommendation engine so that it runs efficiently on the current IT infrastructure. Most banks have adequate processing and storage capacity to run basic NBA models that will generate actionable recommendations. For example, a CPU system with expanded memory capacity is adequate to run induction algorithms based on decision trees. However, deep learning algorithms require a GPU system to support complex artificial

Exhibit 2

Data from diverse sources are consolidated in successive waves of increasing complexity.



Source: McKinsey analysis

neural networks.

3. **Identify business sponsor and form multidisciplinary team.** The sponsor should be a business-line owner authorized to make binding decisions. The team should include data scientists, data engineers, business translators, UX designers, and data architects. The team identifies the variables to be analyzed and builds the analytical model based on precise understanding of business goals, including familiarity with the market segment to be addressed and the products to be targeted.
4. **Scope the effort and build the business case.** The executive team must reach a shared understanding of the problem to be solved and agree on a "back of the envelope" business case for the NPtB effort. It should also identify the main elements for evaluation (model performance, adoption rate, conversion rate improvement, etc.).
5. **Identify potential roadblocks.** It is crucial to follow proper legal and compliance procedures to ensure that the bank has the necessary consent/permission to use and merge the targeted data. In our experience, most of the data available for the corporate segment can be leveraged for NPtB analysis; however, it is important to identify early in the design phase any business limitations that may delay implementation. Such limitations may include, for example, business involvement, change management challenges, and workers' council policies.

Develop the analytical solution

In the development phase, data scientists, data engineers, and business translators collaborate to build the analytical solution of the NPtB model. The goal is to identify the products a customer is likely to buy, prioritize recommendations, and determine the most effective channels for delivering an offer. The team works toward these goals by building algorithms to answer three main questions:

1. **Which products does a particular company need or is willing to acquire?** The NPtB analytical engine identifies opportunities for cross-selling. For each specific company doing business with the bank, the engine ranks commercial leads for each product according to two criteria: probability (which leads are most likely to result in a transaction?) and value (which will be most profitable?).
2. **Which companies need a particular product?** Next, the engine also prioritizes clients according to their potential value/business priorities, propensity to buy, and more. (This step is particularly helpful for relationship managers, who must decide how to prioritize follow-up calls and visits.)
3. **Which channels should be used to optimize the success rate of the commercial opportunities?** Leads are distributed to digital and traditional channels based on company behavior and preferences, contact policies, and relationship managers' commercial activities.

To answer these three questions precisely, banks can analyze customer data, such as payments transactions and digital interactions. Machine learning algorithms can identify patterns in past customer behavior to predict future customer purchases. Data scientists build NPtB engines leveraging modeling environments such as R, Python, or Spark.

While deep-learning algorithms generate the most accurate predictions if data sources are

From NBA to NPtB

Next-product-to-buy (NPtB) models represent a significant advance over those using first- and second-generation next-best-action (NBA) methodologies. In the past decade, product recommendations for corporate transaction banking were generated through statistical analysis of historically observed behaviors across diverse sub-segments (usually between 10 and 20 in number). Each recommendation called attention to a particular area (e.g., disbursement services, accounts receivable and working capital management, investments, trade finance, foreign exchange) and targeted broadly companies sharing general characteristics, as defined by a limited range of profile data, for example, company size, geography, industry, supply chain position, financial behaviors.

With the proliferation of data points, second-generation NBA recommendation engines required more intensive work to describe the data, that is, to teach machines the data features to look for when analyzing a data set. This enabled banks to identify specific products (e.g., receivables financing, FX hedging, corporate procurement cards) that a customer would likely be considering. However, the analysis behind these recommendations focused on behaviors within specific cash-management functions, reinforcing the narrow scope of product "siloes," with little opportunity to optimize financial performance across the full value chain.

The goal of feature engineering, therefore, is to find the best combination of variables to enable a learning algorithm to recognize meaningful patterns in a particular data set. It is a key process in developing a conventional NBA model and until recently was a time-consuming manual process drawing on deep knowledge of business practices.

Drawing upon a much broader set of structured and unstructured data and taking advantage of recent gains in processing capacity, today's NPtB engines generate recommendations that are much more granular and precise than is possible with NBA models. The improvement in analytical sophistication and processing power comes thanks to deep learning, which can develop highly accurate predictive algorithms. Neural networks make it possible to automate the discovery of data features and the identification of the best combination of features to produce the targeted prediction. These calculations produce remarkably precise predictions and deliver actionable results.

complex and unstructured, gradient boosting machines, random forest algorithms, and even logistic regressions provide valuable commercial opportunities for the NPtB engine. In addition, gradient boosting machines have the advantage of generating reliable recommendations with smaller data sets or data sets where there are gaps. These models are trained with past data and statistically back-tested on an out-of-sample and out-of-time customer base to quantify model performance.

Deploy and pilot NPtB engine

The last phase is to embed recommendations in digital channels and relationship managers' interactions with clients. Banks test and refine the NPtB pilot in the field before rolling it out to the full market.

A European bank recently tested an NPtB engine in five branches to evaluate the precision of sales leads. A team of relationship managers, product specialists, and branch managers participated in the pilot, which included training on how to follow up on recommendations and testing the effectiveness of leads with clients. Over three months, the team tested the leads with companies and provided feedback. A sceptical relationship manager selected an offer for a letter of guarantee recommended for a particular client, and commented, “I don’t believe the customer will buy this, I know the company.” When the relationship manager asked the company owner if he needed a letter of guarantee this month, he answered, “How did you know? I am currently negotiating this product with another bank.”

In the course of a similar pilot with another financial institution, the model predicted that 4,500 companies, for which there was no indication in the data of previous international trade activity, would purchase international trade products in the coming month. As it happened one in five of these companies purchased an international trade product for the first time within 30 days. Based on the performance of the pilot, the analytics team updated the model before implementing it across the entire organization to target more than one million customers. The full launch included four weeks of coaching for more than 600 relationship managers. Within five months of starting the project, the NPtB was fully up and running, with the predictive model stabilized and relationship managers fully trained. Ultimately, this bank increased new sales by more than 30 percent, and relationship managers increased their interactions with commercial customers by more than 50 percent.

The pilot is an important opportunity to secure the endorsement of team members participating in the pilot, who then share information about the model with other colleagues. The pilot is also an opportunity to test metrics for evaluating the sales process, such as number of visits, percent of leads used by relationship managers, conversion rates, and the level of satisfaction among relationship managers participating in the pilot (versus control group). In addition, the pilot phase is the time to begin testing long-term performance metrics (in order to ensure sustainability in the front line), for example, hit rates for branch staff and relationship managers, customer profitability, and customer satisfaction.

In the transition from pilot to full roll-out, it is crucial to ensure that the organization is aligned around the NPtB use case and that a support team is assigned for the deployment.

Adapting an NPtB engine to serve large corporate clients

Banks have also been able to improve the relevance and timeliness of their recommendations to large corporate clients. At many institutions, relationship managers are thoroughly familiar with the general needs of their corporate customers, but sometimes they are at a loss to anticipate changes in these needs. This was once the case for a large European bank operating in diverse regions. It now draws on a broad range of data to understand general market trends and specific company behaviors. This requires not only applying advanced analytics to traditional types of information (annual reports, market conditions, competitor news) but also collecting publicly available data on social media (company-managed pages, customer comments, etc.). An NPtB engine extracts insights from available data to alert corporate treasurers to new opportunities, for example, to leverage complex banking relationships to improve cash flow and lower the cost of short-term financing. Identified leads include opportunities in various currencies, possibly triggering a change in cross-border pooling arrangements; letters of credit, domestic and international guarantees; or even investment banking products, e.g., debt capital management. NPtB engines can boost new sales among large corporate clients by as much as 15 percent.

Implementing NPtB for SMEs/mid-corporates

The lessons learned from banks that have implemented an NPtB engine can be summarized in five points:

1. **Design the NPtB engine according to the characteristics of the market segments served.** Consider first internal data, including company profiles, relationship characteristics, product granularity, and opportunities. Expand the data set to include external data, testing the relevance of the new variables in generating useful recommendations. In developing algorithms to generate predictions for the large corporate segment, it is important to test a broad variety of external data in order to build a robust data set that can produce insights with a new level of accuracy.

2. **Build the model around customer needs and interests.** One of the biggest impacts is shifting from a "product push" approach to interactions that address specific customer needs, as reflected in current transaction activity and financial performance. This shift enables relationship managers, service representatives, and product specialists to help customers weigh their options and choose the path that best serves the company's financial interests.

3. **Pilot the outcome of the NPtB engine to build confidence and secure buy-in.** Relationship managers must be confident in the opportunities identified by the NPtB engine; at the end of the process they will leverage leads to improve their sales effectiveness, but change management and internal buy-in are key for successful implementation.

4. **Focus on prototypes that create excitement.** Don't let IT and the complexity of legacy systems become the bottleneck, but start with a pragmatic "proof-of-concept" to demonstrate the model's potential. Quick test-and-learn prototypes have multiple purposes, including learning and improving but above all showing prompt impact to create enthusiasm.

5. **Ensure impact from multiple levers.** Better targeting based on analytics is crucial, but there are additional levers, including the timing of recommendations, framing recommendations within a broader value proposition, measuring the impact of recommendations (including the performance of relationship managers), which can also improve performance.

Ignacio Crespo is a partner in McKinsey's Madrid office, where **Carlos Fernandez Naveira** is an expert associate partner. **Huw Kwon** is a senior analytics expert in the London office.