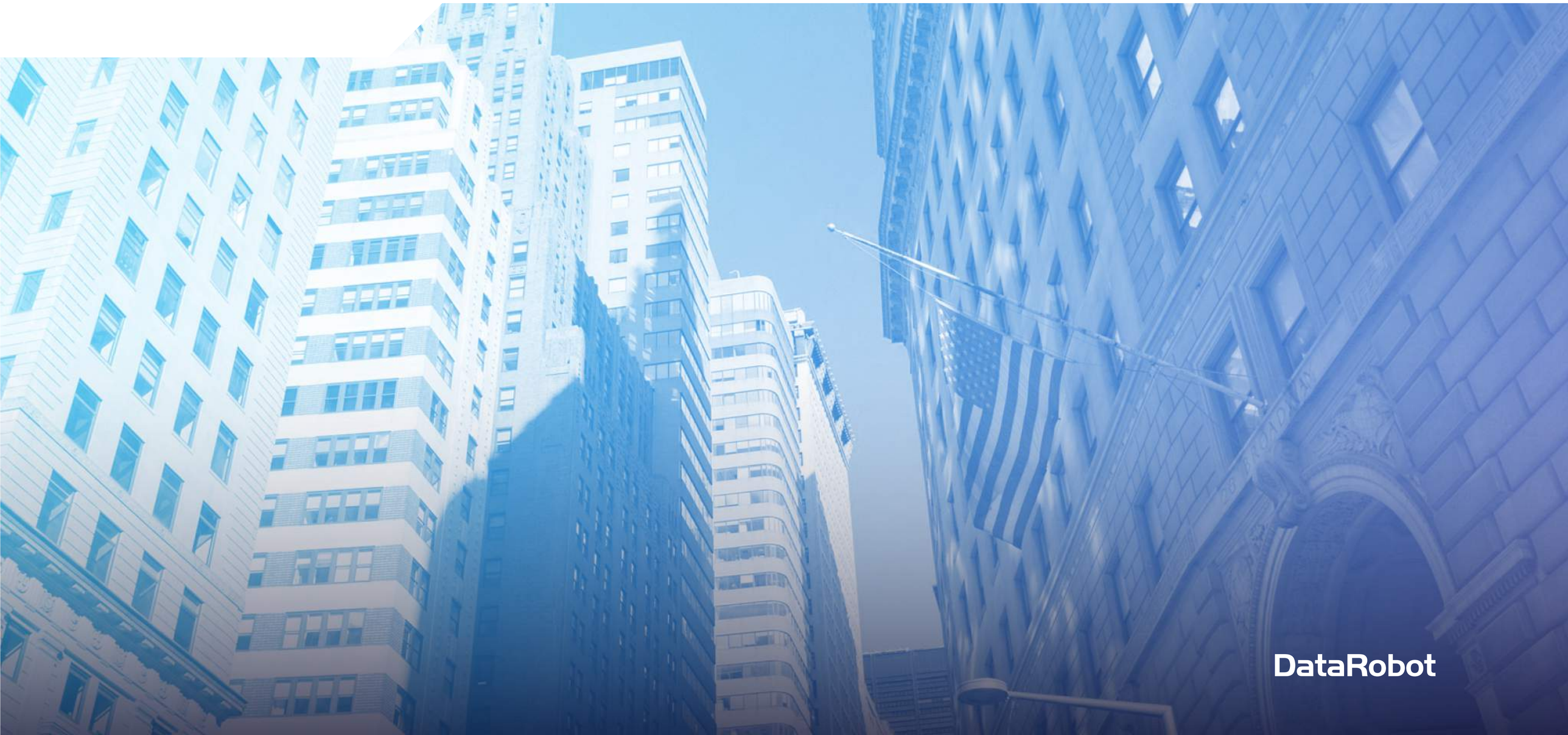
Automated Machine Learning Drives Banking Innovation

BANKING OVERVIEW



Executive Summary

Prediction has been a part of the banking sector for a long time. The art of underwriting and pricing loans has always required some mechanism for estimating and evaluating the credit risk associated with a borrower. Years ago, this was as simple as a conversation and good solid handshake, but today's financial institutions are far more savvy. Sophisticated default and loss modeling is second nature to many banks.

There remains, however, a significant and costly gap between the common practice in banks today — enlightened as it may be — and the state of the art. Most commercial banking businesses have yet to embrace machine learning. And while they're farther ahead in putting machine learning to good use, consumer banks are still leaving opportunities on the table. At the same time that technology has caught up to financial institutions' burgeoning need to parse through and make the most of the loads of data they possess, they are dealing with the competitive heat of fintech upstarts, ever-increasing regulatory requirements, and pressure to progress as efficiently and effectively as possible.

Machine learning offers many ways to meet current demands across the organization — not just with the initial assessment of a customer. Prospecting, loan prepayment, customer churn, and relationship deepening all have the potential to be dramatically enhanced by machine learning.

Executive Summary (continued)

On the regulatory side of the house, banks employ armies of model validators who spend hours checking the hand-built, bespoke models churned out by data science shops. These custom- built models are time-consuming to build, prone to error, and vary substantially in quality.

Similarly, banks spend tremendous sums on regulatory compliance related to Comprehensive Capital Analysis and Review (CCAR) stress testing and anti-money laundering efforts. It's not unusual for the biggest banks to spend more than $100 million annually on stress testing alone. Even with that expenditure, many of the largest banks in the US have failed, or been forced to revise, their CCAR submissions over the last five years. In other words, the current process is ripe for disruption.

At a broader level, robo-advising is replacing the traditional advisor, and even though quantitative analysts (or "quants") abound, the buy-side has yet to embrace machine learning for much of the work that they do.

Machine Learning Drives Banking Innovation

This paper outlines the opportunities provided by machine learning and describes how automated machine learning can be the difference between soaring revenues and defeat at the hands of the fintech invasion.

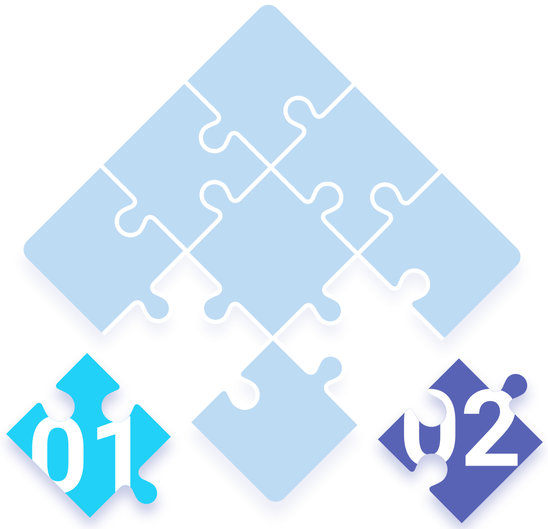
Commercial banks need machine learning to survive

In consumer banking, machine learning is commonplace among the larger players and often supported by large, well-funded data science teams. By contrast, most commercial banking shops have yet to embrace machine learning outside of traditional credit risk models, even though machine learning provides a substantial competitive advantage.

One example is the use of technology to identify the best potential customers based on existing data. Although third-party sources can provide data about a business, the information is often untrustworthy, dirty, incomplete, and off-target. By combining in-house data with external data sources, however, banks are discovering that they can use machine learning to accurately predict potential commercial customers who have high growth potential and high credit quality. These sorts of models help sales teams focus their efforts on the right customers, resulting in better new business at a lower cost.

Another example of low-hanging fruit for commercial banking is relationship deepening. The more products that a customer purchases, the stickier and more profitable they become. Using machine learning to develop models that predict which products to pitch to which clients is a simple and highly effective approach that reliably increases revenue and retention.

The chief constraint in expanding the use of machine learning in commercial banking is the lack of data scientists, both due to the cost of retention for these “unicorns” and the scarcity of the resource. By investing in advanced tools, like automated machine learning, commercial banks empower more users throughout the organization to create real value, without needing to add more data scientists.



Identification of the best potential customers

Relationship deepening

Consumer banks need a simpler path to implementation

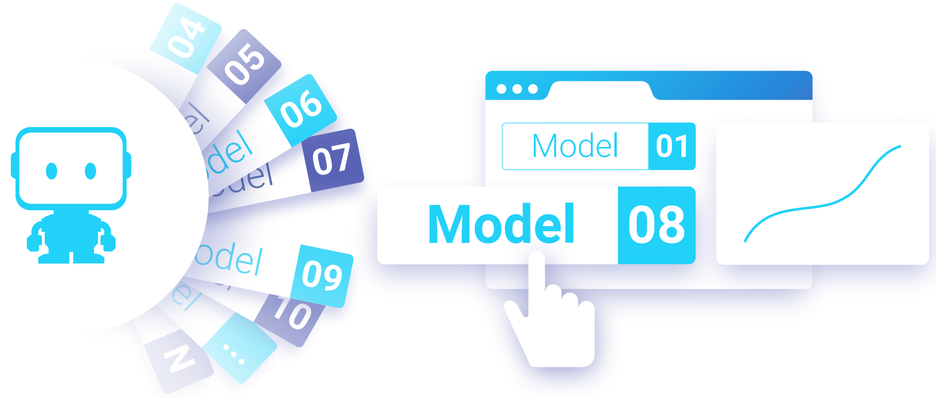
The power of machine learning in the consumer banking space cannot be overstated. From detecting fraud and credit risk to optimizing rewards programs, the sheer quantity of data means that even smaller banks can optimize their business through artificial intelligence and see significant results. Unfortunately, most banks struggle with speed to market — both for technical and regulatory reasons.

On the technical side, the tools that data science teams use to build models — tools like R or SAS — rarely lend themselves to easy implementation in back-end systems. Code rewrites, error checking, and the incredibly slow, painful, and expensive process for maintaining these valuable models challenges even the most sophisticated organization.

The resistance stems from how they begin. Today, nearly all predictive models are custom-built. Because hand-coded models are more error prone and risky, model risk management teams spend months on model governance, monitoring, and validation to ensure that these models are safe to use and that all regulatory requirements have been satisfied.

Automated machine learning solves these problems and instills confidence in the outcome. Because automation transparently and systematically builds dozens of models, regulators and validators get models that are built in a standardized way along with dozens of independently built challenger models that greatly reduce the amount of time spent in model validation.

Automated machine learning also offers options for deploying models: drag-and-drop interfaces; API-based approaches that require only a few clicks; distributed systems to enable scoring massive amounts of data in minutes; and downloadable scoring code for greater transparency. These options minimize implementation risks and dramatically reduce both the time and cost to deploy, monitor, maintain, and refresh key models.



Regulators and validators need a systematic approach

Because most predictive models are built by hand — an iterative and complex process — an entire industry has sprung up around the audit, governance, and validation of these models. High-priced consulting firms, model validation teams, and federal regulators have scores of people whose only job is to replicate and validate results to ensure that no inadvertent errors have been introduced.

Given the sheer number of models that will be built in the next decade, it is imperative that these validation teams, whether internal or external, find ways to increase their efficiency. Automated machine learning tools will not only facilitate the building of more models than ever before, but the best tools will enforce best practices, transparency, and thorough documentation in a way that reduces both errors and total validation time.

Model risk officers must embrace these new tools to ensure that the high standard of risk management is maintained without slowing innovation.

Buy-side and sell-side opportunities

Although sell-side players are losing ground to robo-advisors and automation, they still have a huge opportunity to provide value to clients — and machine learning plays a key role in that future. From identifying the right investors to making contact about an upcoming IPO to selecting appropriate research reports for clients, market makers will use machine learning to be smarter about how they help their clients.

On the buy-side, asset managers, hedge funds, and other players will supercharge their quants with machine learning. Predictive models built with machine learning will be used to design trades in ways that couldn't have been conceived of even a few years ago; Kaggle, Quantopian, and others have demonstrated the value of using machine learning to rapidly backtest such strategies.

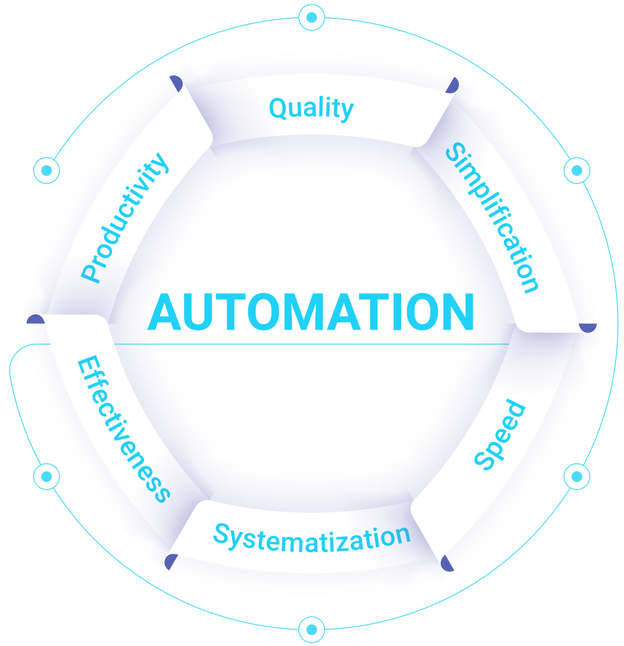
Automation is the Solution

Automation improves productivity and delivers better quality output. Factories are increasingly automated. Engineers have successfully automated complicated activities like surgery and anesthesia. Even fully autonomous driverless vehicles are no longer science fiction. The same principle holds for machine learning. Automated machine learning not only directly addresses the key challenges of limited data science talent and long project timelines, it also simplifies the provisioning problem.

By delivering the power of machine learning to business domain specialists, automation expands the pool of people in an organization who effectively contribute to data science projects. Plus, with expertise built into the software, automated machine learning tools are simple to use. Guardrails and best practices make it safe to engage more people in projects, with the assurance that users can't "forget" critical steps.

Of course, surgical robots do not replace surgeons; they make it possible for surgeons to do more demanding and productive work. By the same token, automated machine learning does not replace the data scientist. Rather, it handles the thousands of routine and repetitive tasks that consume valuable time. With automated machine learning, data scientists no longer need to spend their time manually provisioning server instances, distributing jobs to servers, or maintaining complicated software stacks in virtual machines. Automated machine learning software handles these tasks behind the scenes, with a single administrative interface. By automating such low-level tasks, expert data scientists handle more projects and perform more high-value work — like engaging with their internal clients to understand the business problem or to explain results.

Finally, automation means systemization. Individual, hand-coded models are not only difficult and time-consuming to build, they're also error-prone and introduce implementation risk to the modeling process. By systematizing the technical aspects of the process — for example, out-of-sample validation, model tuning, variable selection, model selection, and so on – implementation risk is greatly reduced. As a result, organizations more consistently adhere to best practices, streamline model validation, and, ultimately, dramatically reduce speed-to-deployment.



DataRobot: The Leader in Automated Machine Learning

DataRobot, a global data science and machine learning company, helps banking institutions around the world leverage the power of machine learning. DataRobot provides a platform for users of all skill levels to produce accurate predictions in a fraction of the time that would be needed using conventional tools and methods. To leverage the most innovative techniques, DataRobot uses open source tools like Apache Spark, H2O, Python, R, TensorFlow, Vowpal Wabbit, and XGBoost.

Empowers business users. DataRobot's automated machine learning platform perfectly complements business users' domain knowledge and acumen. It combines an intuitive web interface and graphical model assessment tools with flexible programming APIs so users work with the tools they know. Plus, DataRobot supports advanced tuning and custom extensions to provide the most sophisticated analysis available.

Enables collaboration. Fraud investigators, credit risk analysts, insurance underwriters, model validators, and other business analysts collaborate on machine learning projects with DataRobot. Collaboration moves machine learning out of the skunk works and into the front office where, with a clear definition of the business problem and direct feedback from stakeholders, it delivers even better results.

Provides interpretable and transparent results. With DataRobot, every step in the modeling process is visible and reproducible. DataRobot offers extensive facilities for understanding the behavior of the predictive models it builds. It also provides aids to the business user or stakeholder seeking to understand and interpret the model. DataRobot explains its predictions by highlighting the most influential factors so that your team can share and explain results to customers, executives, validators, and regulators.

Solves many problems. DataRobot is a general-purpose automated machine learning platform that transforms a business across different products, functions, and regions. With DataRobot, organizations avoid the integration headaches of using many different "point" solutions in various parts of the business. For example, banks may use DataRobot's automated machine learning to increase insights across business functions, including credit risk, fraud, marketing, and underwriting projects.

Reduces time-to-value. DataRobot is fast in building, testing, and deploying predictive models. Its modern software architecture uses distributed computing to run many experiments in parallel. It also automates routine tasks like data discovery, feature engineering, and model assessment. Once a model is approved, DataRobot offers multiple options for deployment, including native scoring, exportable prediction code, and prediction APIs.

Arrives enterprise-ready. DataRobot works with data wherever it resides, including relational databases, Hadoop, text files, and many other sources. It runs on free-standing clusters, in Hadoop under YARN, or as a managed service in the cloud. In Hadoop, DataRobot uses native security, data provenance, and application management services. For low maintenance and easy integration, DataRobot uses the most current software development practices, including microservices and containerization.

Automated Machine Learning Drives Banking Innovation

Machine learning is transforming the banking industry across a wide range of functional areas. The future of banking depends on how quickly the industry embraces automated machine learning to transform and innovate. DataRobot's automated machine learning platform is the solution. Better predictions, faster. That's DataRobot.

For more information on DataRobot, or to schedule a demo, visit **www.datarobot.com.**

Contact Us

DataRobot
One International Place, Fifth Floor
Boston, MA 02110
www.datarobot.com | info@datarobot.com

DataRobot