dotData

Case Study:

Sumitomo Mitsui Banking Corporation Accelerates Data-Driven Colture With Automated Feature Engineering

dotData Enterprise helps accelerate AI/ML development by 48X

48X FASTER DEVELOPMENT PROCESS	**2M** FROM 2,000 FEATURES TO 2,000,000 FEATURES	**30%** IMPROVEMENT IN MODEL ACCURACY	**10X** MORE DATA SCIENCE PROJECTS PER YEAR

Data science is a major area of investment for banks due to its proven impact on operations such as fraud protection, risk mitigation, customer relationship management, and more.

But while investments in AI are growing, banks are often finding that their existing analytics and business intelligence technology and talent aren't capable of meeting their current and expanding needs. Challenges in resources, technology infrastructure, and the ability to operationalize models quickly and efficiently can prevent financial institutions from fully leveraging AI and data science to drive business impact.

These challenges, paired with the need to remain competitive in a quickly evolving market, compelled the Sumitomo Mitsui Banking Corporation to work with dotData to help it maximize AI and machine learning (ML) investments.

"dotData makes it increadibly simple for our data scientists to explore a multitude of use-cases quickly and without the headaches typically associated with data science."

- Sumitomo Mitsui Banking Corporation

https://dotdata.com

ABOUT SUMITOMO MITSUI BANKING CORPORATION:

Sumitomo Mitsui Banking Corporation was formed in 2001 by the merger of Sumitomo Bank and Sakura Bank. With over $1,775.14B in assets, Sumitomo Mitsui Banking Corporation is the world's 14th largest bank and provides offerings across a broad spectrum of financial services including consumer banking, corporate and investment banking, international banking, and more.

THE CHALLENGE:

In early 2016, Sumitomo Mitsui Banking Corporation's IT Planning department was tasked with addressing a growing concern: while the bank had begun using machine learning in several of their business divisions -- for operations, such as enhancing customer product upsell and cross-sell opportunities, managing customer attrition and identifying default risk, Sumitomo Mitsui Banking Corporation's nascent data science team was facing a shortage of talent faced with an excess of demands.

While building ML and AI models were feasible, it was a 100% manual operation that required a lot of coding and data manipulation. Building a single model typically took two to three months with as much as 80% of the time spent on the process of creating the complex multi-dimensional flat tables required by ML and AI models. This process, known as feature engineering, coupled with the complexity and time-consuming nature of ML model selection and optimization, was hampering the ability of the team to deliver on all the projects required of them.

The combination of not enough talent, the complexity of models, and the time-consuming nature of feature engineering prevented Sumitomo Mitsui Banking Corporation's team from scaling their data science practice, and restricted its output to only five new ML models in any given calendar year, with bandwidth to update an additional five models in that same timeframe.

THE SOLUTION

The Sumitomo Mitsui Banking Corporation team, led by Akinobu Funayama and Tomohiro Oka, decided that AutoML technology was a possible solution to the broad lack of talent, and one that might help them accelerate development lifecycles for their data science projects. Critical goals

were the ability to analyze and optimize business models quickly and automatically, as well as the ability to automate as much of the data science lifecycle as possible. During the evaluation process, Sumitomo Mitsui Banking Corporation identified the automated creation of features, also known as "automated feature engineering" as a critical requirement for their project. Automating this manual and time-consuming process would enable Sumitomo Mitsui Banking Corporation to optimize resources and shorten project timelines. An additional crucial element for the use of these AutoML platforms was the need for transparency, to enable their data science team to provide a higher level of transparency to business units that were asking for ML and AI applications.

To identify the best possible technology providers, Sumitomo Mitsui Banking Corporation explored more than 300 platforms, ultimately short-listing 50 providers to evaluate in more detail. Sumitomo Mitsui Banking Corporation ranked the short-listed platforms by their ability to meet the essential requirements of automated feature engineering and AutoML as well as ease of use for less experienced users. Several of the short-listed vendors were then tested in purpose-specific proof of concepts.

After their extensive evaluation cycles, Sumitomo Mitsui Banking Corporation decided that a combination of two separate types of platforms, used together, would best fit their needs. The first was a smart interactive data preparation platform that would help cleanse the master data. The second one was dotData's Data Science Automation (AutoML 2.0) platform to help automate the full-cycle of feature engineering, the selection, and optimization of ML models and also provided vital capabilities to help explain features associated with ML models to non-technical users. dotData was selected because of its advanced feature set, its AI-powered feature discovery, creation, and evaluation engine, as well as dotData's automated machine learning evaluation and optimization capabilities.

RESULTS

Once implemented, the benefits achieved as a result of investing in dotData's AutoML 2.0 platform were immediate and substantial. Prior to implementing dotData, it took two months for data scientists to explore 2,000 features for each project during their development process. Through data science automation, Sumitomo Mitsui Banking Corporation can now examine more than two

RESULTS

- Sumitomo Mitsui Banking Corporation can now generate 2,000,000 features for each project by using AI-based feature engineering.

- Development times have been accelerated 48X from 2 months to 2 days per project

- Increased capacity means Sumitomo Mitsui Banking Corporation can develop over 100 models per year instead of 10.

- Model accuracy has been improved on average by 30%.

LEARN MORE

https://dotdata.com

https://dotdata.com/products

million features for each project. The benefits of automating both the feature engineering as well as the machine learning have also allowed the team to reduce development times dramatically. Sumitomo Mitsui Banking Corporation has cut its data science development times from an average of two to three months per project to less than ten hours per project -- a 48X acceleration in development times. The ability to explore the huge number of feature hypotheses has also improved the accuracy of models by as much as 30%, providing additional data insights that augment Sumitomo Mitsui Banking Corporation's core domain knowledge.

The adoption of dotData has provided significant benefits to multiple business units. Data scientists at Sumitomo Mitsui Banking Corporation no longer struggle with managing data and creating features, and can instead focus on feature analysis and on identifying models that are likely to be most beneficial to their business units. Sumitomo Mitsui Banking Corporation has also increased the number of models built per year to 100 models from 10 -- a growth of 10X over the previous manual process. Finally, more transparency in feature engineering is allowing the team to explain ML models more succinctly to business units and to accelerate the feedback loop to continue to improve models over time.

As a result of their investment in data science and AutoML 2.0, Sumitomo Mitsui Banking Corporation has been able to expand the number of supported use-cases for data science and ML applications across the organization. In addition to customer management and marketing and sales, Sumitomo Mitsui Banking Corporation now uses AI and data science in numerous other divisions within the institution, including compliance, risk management, and financing and loans. This expansion has happened without having to add additional headcount to the data science team and has allowed Sumitomo Mitsui Banking Corporation to broaden the use of data science while also improving the accuracy of their models.

Sumitomo Mitsui Banking Corporation's adoption of AutoML 2.0 has enabled them to automate many aspects of its data science process that were formerly time and resource-consuming. As a result, Sumitomo Mitsui Banking Corporation has been able to rapidly scale their AI/ML initiatives to drive transformative business changes

About dotData

dotData is the first and only company focused on full-cycle data science automation. Fortune 500 organizations around the world use dotData to accelerate their ML and AI projects and deliver higher business value. dotData's automated data science platform speeds time to value by accelerating, democratizing, augmenting and operationalizing the entire data science process, from raw business data through data and feature engineering to machine learning in production. With solutions designed to cater to the needs of both data scientists as well as citizen data scientists, dotData provides unmatched value across the entire organization.

dotData's unique AI-powered feature engineering delivers actionable business insights from relational, transactional, temporal, geo-locational, and text data. dotData has been recognized as a leader by Forrester in the 2019 New Wave for AutoML platforms. dotData has also been recognized as the "best machine learning platform" for 2019 by the AI breakthrough awards and was named an "emerging vendor to watch" by CRN in the big data space. For more information, visit www.dotdata.com, and join the conversation on Twitter and LinkedIn.