

“Printing Money” with Operational Machine Learning

Thomas H. Davenport and Rich Masi

Organizations have made large investments in big data platforms, but many are struggling to realize business value. While most have anecdotal stories of insights that drive value, most still rely only upon storage cost savings when assessing platform benefits. At the same time, most organizations have treated machine learning and other cognitive technologies as “science projects” that don’t support key processes and don’t deliver substantial value.

However, there are a growing number of large but innovative companies that are driving measurable value through “operational machine learning”—embedding machine learning on big data into their business processes. They’re employing a new generation of software, skills, and infrastructure technologies to solve complex, detailed problems and deliver substantial business value. One company found the approach so successful that a manager said it was like “printing money”—a reliable, production-based approach to generating revenue.

An Example of Operational Machine Learning

Take, for example, an investments firm that needed to create personalized cross-channel customer experiences. In the past, the company used “decision management” technology to create offers based on scores computed from past investments and the company’s perceptions of net worth. Today, however, the problem is much more complex. The company had tried to create cross-channel versions of the same idea, but it had never been successful because both the available technology and the collaboration between marketing and technology groups were lacking.

Over the past year, the firm created a cross-channel approach to personalized customer offers. It uses data from the customer’s website clickstreams, investing behaviors, and call centers. It can create both emailed offers and personalized, optimized website content. Personalized offers can also be made in call center interactions.

The solution learns from the responses of customers and tunes offers over time. It includes machine learning models to customize offers, an open-source solution for run-time decisioning, and a scoring service to match customers and offers. It supports millions of customer offers a day, and customer response is improved significantly over the single-channel legacy system. In order to help create these capabilities, the company created both a Chief Data Officer and a Chief Loyalty and Analytics Officer within the marketing function.

How to Drive Value from Big Data

With the adoption of big data platforms, many companies are experimenting with machine learning as a means of dealing with all the data. Data scientists, who are typically key to making machine learning work for organizations, have been described as holding “the sexiest job of the 21st century.” With the prominence of machine learning and the data scientist, why isn’t there a continuous benefit stream of value that flows from big data?

Part of the reason is the labor-intensive nature of early machine learning initiatives. In practice, the majority of machine learning initiatives follow the traditional resource consuming process of discover, model, deploy, monitor, and update that has been used for decades. Today, modern data and analytics

architecture components can be used to infuse automation into each step of this process and embed scalable machine self-learning into operational processes.

Embedded business rules and predictive analytics that drive operational decisions is not new, and there have been product offerings in this space with robust functionality for years. However, this technology has gained limited adoption, due to both cost barriers and the complexity of deployment and support. Today's contemporary big data architecture and open source software may be the gateway to more widespread adoption. The data management vendor space in this brave new world of data and analytics is crowded, but the area of real-time decision management that allows for production scoring and learning within analytical assets is much less populated. There is a large opportunity for organizations to build these types of applications on top of their big data stack and an even bigger opportunity for vendors in the data management space to extend their offerings to address real-time decision management.

There are three core functional capabilities that need to be developed to support real-time decision management: a decision service, a learning service, and a decision management interface.

1. The *decision service* determines the array of possible outcomes of a process. It accepts decision requests from business processes, applies business rules to filter a decision set, scores predictive analytics for the decision set, arbitrates by a business defined strategy, and returns an optimized result back to the business process. This is typically a rules engine of some kind, either proprietary or open source.
2. The *learning service* improves statistical predictions or categorizations over time. It maintains analytical assets for the decision set, updates predictive assets when responses are available, and passes production ready predictive models to the decision service. This would be a machine or statistical learning offering, also available from both proprietary vendors and in several open source versions.
3. The *decision management interface* allows business to define and update a decision set and/or decision set metadata, define business rules, and define a segmented decision-making strategy that includes rules, predictive analytics, and other key decision metrics. This could be adapted from existing decision management tools or built from scratch.

Building these capabilities on top of a big data stack (including data lake storage and data transformation capabilities) is transformational in terms of the availability of information to support the decision, the performance of the decision request, and the performance of the learning service. We have seen cases where the data query run time to support a decision has been reduced tenfold (for example, from around fifty milliseconds down to less than five milliseconds per query). Applications that used to only consider one month of customer history due to performance constraints can now include all customer history. In other situations where the learning service previously choked on the volume of responses, but when moved to a Hadoop data cluster, the distributed nature of the environment is not overly taxed. With the potential for processing thousands of concurrent requests per second, these big data-driven benefits change the game in operational contexts.

Exploratory analytics and machine learning can certainly generate insights that may be turned into actions that may drive value. On the other hand, operational machine learning that can scale within an embedded business process can drive value without ongoing human intervention. While your company

may not feel it has become a money printing press, this capability does offer the potential to generate massive and ongoing business value.

Rich Masi heads NewVantage Partners' data science and analytics practice and its Charlotte, NC office.