Q&A: Scale AI With Automated Machine Learning

Answers To Seven Key Questions That Unlock Enterprise Outcomes With AutoML

by Kjell Carlsson, Ph.D.
March 18, 2020

Why Read This Report

There's good news for enterprises that want to transform using machine learning (ML) and AI. Automated machine learning (AutoML) helps data scientists, and everyone else, build predictive models, derive advanced insights, and infuse AI into applications at scale. But AutoML must be used with care and caution to unlock promised business value. This report helps application development and delivery (AD&D) professionals and analytics leaders answer the most important questions about AutoML and understand how to use it to effectively drive business outcomes.

Key Takeaways

Use AutoML To Democratize, Accelerate, And Scale AI

AutoML boosts the productivity of your data scientists and empowers everyone in your organization to build predictive models. Have your developers use it to embed AI capabilities in your apps and to rapidly detect and extract business insights across your organization.

Better Together: AutoML And Data Scientists

Yes, you'll still need data scientists. Indeed, you'll want more of them, because they now drive more AutoML-enabled business value and provide training, support, and adult supervision for all your newly created "citizen data scientists."

Watch This Space

AutoML is on a tear, with vendors across the spectrum embedding capabilities into everything from machine learning platforms to business applications and continuously expanding the range of machine learning problems it can tackle. Don't expect offerings to look the same in six months, let alone a year.

Q&A: Scale AI With Automated Machine Learning

Answers To Seven Key Questions That Unlock Enterprise Outcomes With AutoML

by Kjell Carlsson, Ph.D.
with Srividya Sridharan, Mike Gualtieri, Brandon Purcell, Boris Evelson, Leslie Joseph, Sami Kaipa, Tienan Li, Jeremy Vale, Daniel Weber, and Madison Bakalar

March 18, 2020

Table Of Contents

Related Research Documents

AI Unlocks The Business Intelligence In BI

The Forrester New Wave™: Automation-Focused Machine Learning Solutions, Q2 2019

The Forrester Wave™: Multimodal Predictive Analytics And Machine Learning Solutions, Q3 2018

Shatter The Seven Myths Of Machine Learning

FORRESTER®

Forrester Research, Inc., 60 Acorn Park Drive, Cambridge, MA 02140 USA
+1 617-613-6000 | Fax: +1 617-613-5000 | forrester.com

AutoML Helps Scale AI Across The Enterprise

Whether you're a tech titan or an enduring industrial giant, AI ambitions are shackled by too few data scientists who take too long to develop production-grade predictive models or deliver advanced insights. Thankfully, new AutoML solutions promise to:

- **Endow data scientists with productivity superpowers.** Data scientists spend large amounts of time on manual, repetitive tasks, such as testing innumerable permutations of features and algorithms. This results in too few, usually suboptimal, models and, ultimately, unhappy data scientists. AutoML automates this grunt work for a broad range of ML use cases, freeing data scientists to use their unique talents on more projects, with faster and better results. For example, data scientists at Hortifrut, which ships a quarter of the world's blueberries, used H2O's Driverless AI solution to reduce the time it took to build models predicting blueberry spoilage from three to five months to three to five days.
- **Empower anyone to build predictive models.** No data science skills? No problem. AutoML lets business-savvy workers and leaders build ML models that can predict everything from a customer's creditworthiness to an organization's staffing needs with the right data and a touch of a button. Even senior execs can use it to get predictions and insights fast. For example, the CEO of the Merrow Sewing Machine Company used Aible's AutoML solution, by himself, to find an additional $3 million in high propensity-to-convert sales leads.
- **Embed AI capabilities like vision, text, and speech in your applications.** Even among data scientists, the skills to build deep learning (artificial neural network) models used for computer vision tasks (like image recognition or object detection) or natural language understanding tasks (like intent recognition) are rare. AutoML solutions enable anyone to build a custom deep learning model for a wide range of use cases and usually provide features that make it easy to create the training data needed. For example, shopDisney used Google's Cloud AutoML to create models that would recognize Disney characters to help customers search and get recommendations for Disney products.[1]
- **Discover new, valuable business insights faster.** Many business users don't want a predictive ML model per se. They want data-driven insights about how to improve the business that you can get using ML methods. AutoML helps business users discover the drivers of important patterns in their data (like sales, Net Promoter Score, or customer attrition) and proactively alerts them to data anomalies on which they should take action.[2] For example, marketers at Vodafone used the Albert platform's AutoML capabilities to derive proactive insights about marketing spend effectiveness. Over a three-month period, Vodafone reduced the cost of acquiring each customer by 40% while increasing sales by 45%.

Get Smart About How To Use AutoML Effectively

Abracadabra! Not! AutoML doesn't automatically lead to business value. There's an ever-growing array of AutoML solutions, and none address all the different ways in which companies can use AutoML to drive business value. Companies hoping for the wrong type of business value from the wrong AutoML tool are setting themselves up for failure. Even with the right solution, pitfalls abound. You can dramatically improve your chances of success with answers to these seven questions:

1. What is AutoML, and what are the different types of offerings today?

AutoML is about automating the steps that a data scientist performs when building ML models (see Figure 1).[3] Specifically, Forrester defines AutoML as:

> *Capabilities and platforms that partially or fully automate the machine learning model development lifecycle — especially feature engineering, feature selection, algorithm selection, model training, hyperparameter tuning, and model validation — but also extend to aspects of data discovery, extract-transform-load, data prep, model deployment, model monitoring, and automated model retraining, as well as the automatic generation of visualizations and text explanations.*

There are four main types of AutoML offerings, each with dramatically different capabilities that make them suited for different business use cases (see Figure 2).

FIGURE 1 AutoML Automates Each Step In Developing A Machine Learning Model

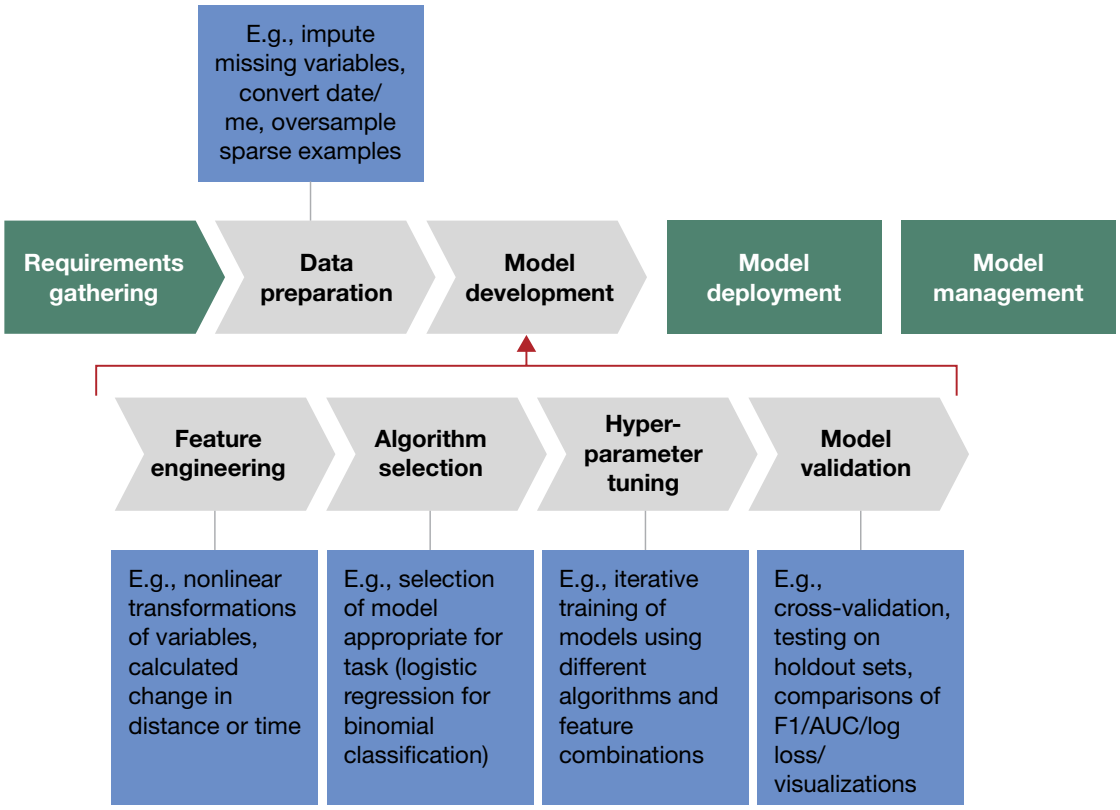


FIGURE 2 The Four Different Types Of AutoML Offerings

Type	Description	Human effort	Example use cases	Sample vendors
Automation-focused machine learning platforms 	Automate the process of developing a predictive ML model as much as possible	Designed to minimize the need for human involvement	• Lead scoring • Attrition risk • Credit risk • Forecasting	• Aible • DataRobot • dotData • EdgeVerve • H2O.ai • Lityx • Squark
Multimodal machine learning platforms 	Offer the widest range of advanced analytics and ML capabilities, including automatic training models	Active human participation in model development	• Attrition risk • Credit risk • Forecasting • Clustering • Market basket analysis • Text analytics	• Dataiku • IBM • KNIME • RapidMiner • SAS
Deep learning-focused AutoML solutions 	Customize pretrained deep learning models (artificial neural networks), especially for vision or text data	Human labeling and annotation of training data	• Visual search • Product recognition • Content moderation • Intent recognition	• Chooch AI • Clarifai • Google AutoML • IBM PowerAI Vision • Microsoft • Neurala
Augmented business intelligence solutions 	Surface insights (e.g., anomalies, opportunities, or drivers) based on ML models packaged with business intelligence platforms	Automatic initial insights guiding or augmenting human follow-up analyses	• Anomaly detection • Key influencers and root-cause analysis • Trends and forecasting	• Albert.ai • Anodot • Avora • Microsoft • Oracle • Outlier.ai • Salesforce • Sisense • Tableau • Tellius • TIBCO

2. What are the limitations of AutoML?

Do you even need data scientists now that AutoML exists? Yes. However, the role and skills of your data scientists will shift away from repetitive tasks to higher, value-added activities, thanks to AutoML. Even though your data scientists will be far more productive, you'll still want more of them to scale your ML and AI operations. This is because your data scientists will now be generating more value, but you'll also need them to:

- **Do everything else in advanced analytics and machine learning.** AutoML tools tackle a limited (albeit valuable) set of standard ML tasks but don't help with most advanced analytics and ML tasks — mining association rules (e.g., market basket analysis) and most unsupervised machine learning (e.g., clustering). These tools often don't support specialized use cases involving unusual data structures and complex or use-case-specific feature transformations. Few do anomaly detection or allow for custom model performance metrics, and many lack any significant capabilities in the other parts of the ML model lifecycle, such as data discovery, model monitoring, or the ability to create and manage pipelines of different models. Further, most offerings lack related capabilities — like optimization and simulation or business rules engines for suggesting next-best actions — that you need to create full solutions.
- **Provide adult supervision, training, and support.** Creating a bad model is easy, and even easier with AutoML. Selecting tools with built-in guardrails helps, but you'll need data scientists to ensure that your new legions of "citizen data scientists" are leveraging their new ML powers appropriately. Have them train and assist business users as well as validate and sign off on models before they're put into production. For example, Wunderman Thompson engaged IBM's Data Science Elite team to train its employees to use their AutoAI tools to leverage over 17,000 features in an automated pipeline to create new customer acquisition models spanning 270 million customers.
- **Tackle the most important or complicated projects.** Building models for the organization's most mission-critical applications — those that require the highest accuracy, entail the greatest risk, and are the most complicated or most valuable (for example, assessing risk in finance and insurance) — and the organization's most important strategic analyses requires trained, experienced data scientists. While AutoML offerings accelerate individual steps, they cannot, on their own, provide the in-depth analysis of hypothesis, develop the custom algorithms, align across the data, model and business requirements, and constraints, or provide the trust needed for the company's most important use cases.

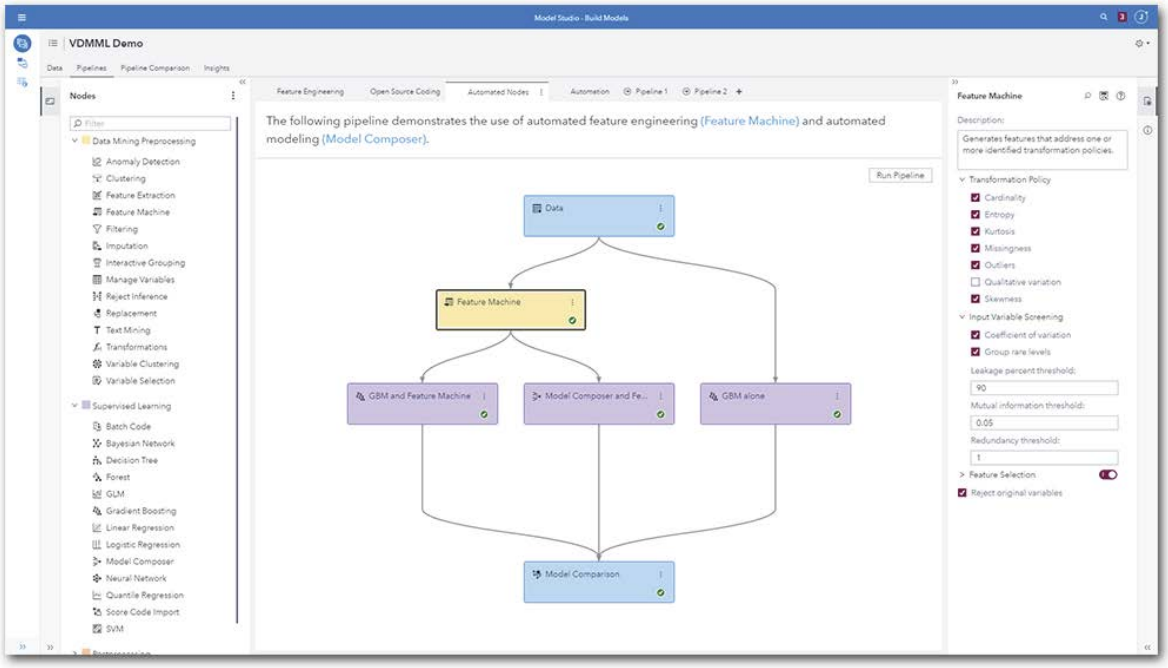
3. How do I boost the productivity of data scientists and the data-savvy with AutoML?

Far from being the sexiest job of the 21st century, data scientists spend the majority of their time on repetitive, manually intensive tasks across every part of the model development lifecycle — not just in acquiring, prepping, and cleaning data, but in creating, testing, validating, and comparing innumerable

permutations of features and algorithms to create a single model. To boost the productivity, quality of work, and morale of your data scientists — and everyone else in your organization who currently builds ML models:

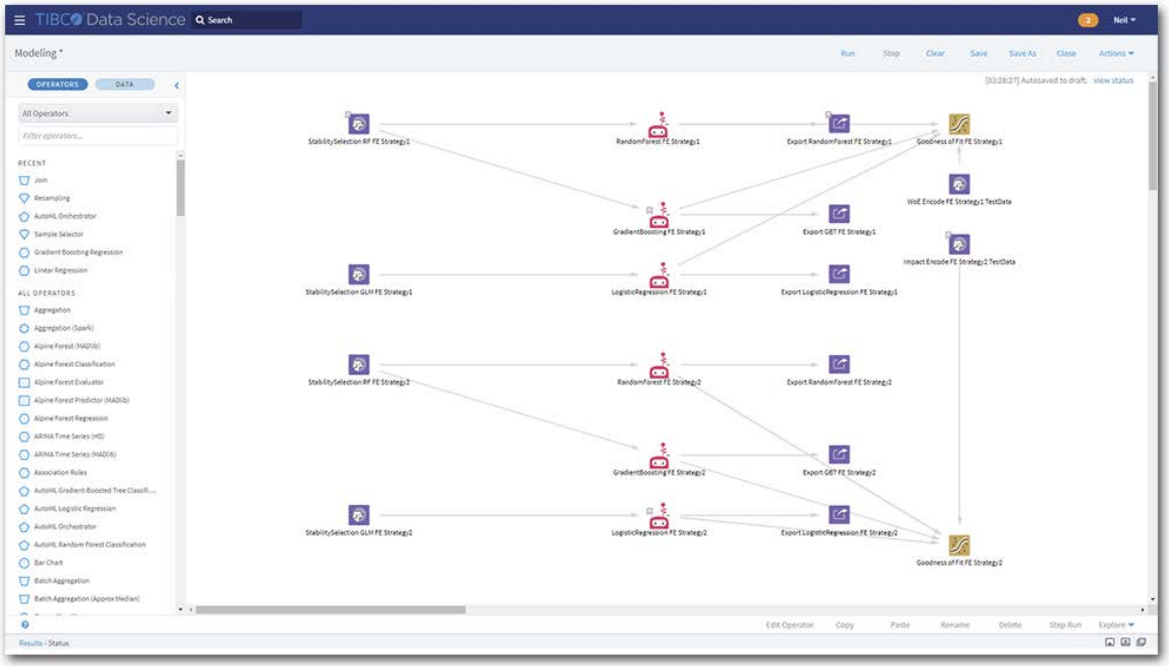
- **Unleash AutoML capabilities within your existing multimodal ML platforms.** The tools that your data scientists are using today have (or will soon have) AutoML features. However, they might not have access to them, or they may be unaware of how to use them effectively. Spring for upgraded licenses, if needed, as well as scalable infrastructure (AutoML is compute-intensive) so that your data scientists can use these features. Further, invest in training and adapting your analytics processes to get the full value from these tools. Better yet, open these tools to your quantitative users (e.g., data analysts and engineers) who want to grow their data science skills as a gateway to the full array of advanced analytics and ML methods. Chemical engineers at Sappi, the leading global provider of sustainable wood fiber products, adopted RapidMiner's Auto Model to create and test thousands of models to predict pulp viscosity for reduced quality variation and variable cost reduction.
- **Supplement with an automation-focused ML solution to drive speed and scale.** An automation-focused ML solution will never be the be-all, end-all tool for a data scientist — it's too limited in the tasks and methods it provides. However, automation-focused ML solutions reduce the time and effort needed for every iteration of the ML model development lifecycle. This enables data scientists to create more and better models faster. Steward Health Care used DataRobot's AutoML solution to develop over 1,100 models it needed to predict staffing needs across the shifts at its 38 hospitals.[4] Since data scientists can't be experts in all ML methods and feature transformations, these tools help them tackle problems that go beyond their training and experience.
- **Select tools with maximum capability, scalability, and reusability.** For non-data scientists, less is more. For data scientists, more is more. Choose solutions that support the widest range of use cases (e.g., time series and anomaly detection), ML algorithms (e.g., neural networks and ensembles of models), and feature transformations (e.g., for geolocation and text data). Ensure that the tools provide as many options as possible and support the greatest variety and volume of data possible. Even though these tools will be harder to use, they'll allow your data scientists to eke out every performance improvement and tackle the widest set of use cases. To really make your data scientists happy, choose tools that output model pipelines or code that they can edit, reuse, and build on (see Figure 3).

FIGURE 3 Multimodal Predictive Analytics And ML Solutions Deliver Editable Pipelines Or Code



Source: SAS website

FIGURE 3 Multimodal Predictive Analytics And ML Solutions Deliver Editable Pipelines Or Code (Cont.)



Source: TIBCO website

4. How do I democratize predictive analytics and ML for non-data scientists with AutoML?

Every part of your business can benefit from having more predictive models. With AutoML, your domain experts can now become directly involved in the process of developing these models without any knowledge of ML. The result? Models that use the right data and are more closely tied to the business need, better leverage of your data scientists, and more data-driven decisions across the business. To democratize the power to build predictive models:

› **Make an automation-focused ML platform widely available to drive experimentation.** Automation-focused ML platforms specialize exclusively in streamlining the end-to-end process of developing and, increasingly, deploying an accurate predictive model. They abstract away complexity and can be used with little to no knowledge of data science, while embedding many ML model development best practices. This enables non-data scientists to do the initial data discovery and experiment with different algorithms, and even complete the end-to-end process by deploying and monitoring the ML model on their own (see Figure 4). For example: One & All, the fundraising agency for nongovernmental organizations like The Salvation Army, brought the creation of donor

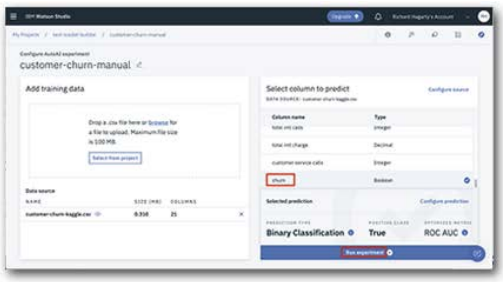
selection models in-house with the help of Lityx, enabling the agency to reduce costs by 70%, increase donor giving by 8%, and increase turnaround from months to hours, while also opening up the use of data-driven targeting to a host of smaller nonprofits.

› **Select solutions that make it easy to show business impact.** All AutoML tools make it easy to build models, but the best ones help build models that are aligned closer to business needs. Look for tools that solicit and leverage users' business domain knowledge, have templates or recipes for common use cases, have features for explaining the models and their predictions, and have built-in guardrails that prevent users from making obvious mistakes. For example, Aible asks users for the business impact generated by correct vs. faulty predictions and business constraints to create models tailored to the business need (see Figure 5).

› **Create a sustained, virtuous cycle of business alignment.** The data, ML method, and business use case are usually in flux at the beginning of an ML project and require iteration. The exact business requirements are only clear once a model is deployed, requiring new data and a new ML model. Accelerate this iterative process with AutoML tools that support as much of the ML lifecycle as possible, from data discovery and acquisition through deployment of the model and beyond, with features that help monitor models in production and make it easy to retrain them with new data.

FIGURE 4 Many AutoML Offerings Reduce The Model Development Process To Three Steps

1. Select data and variable to be predicted

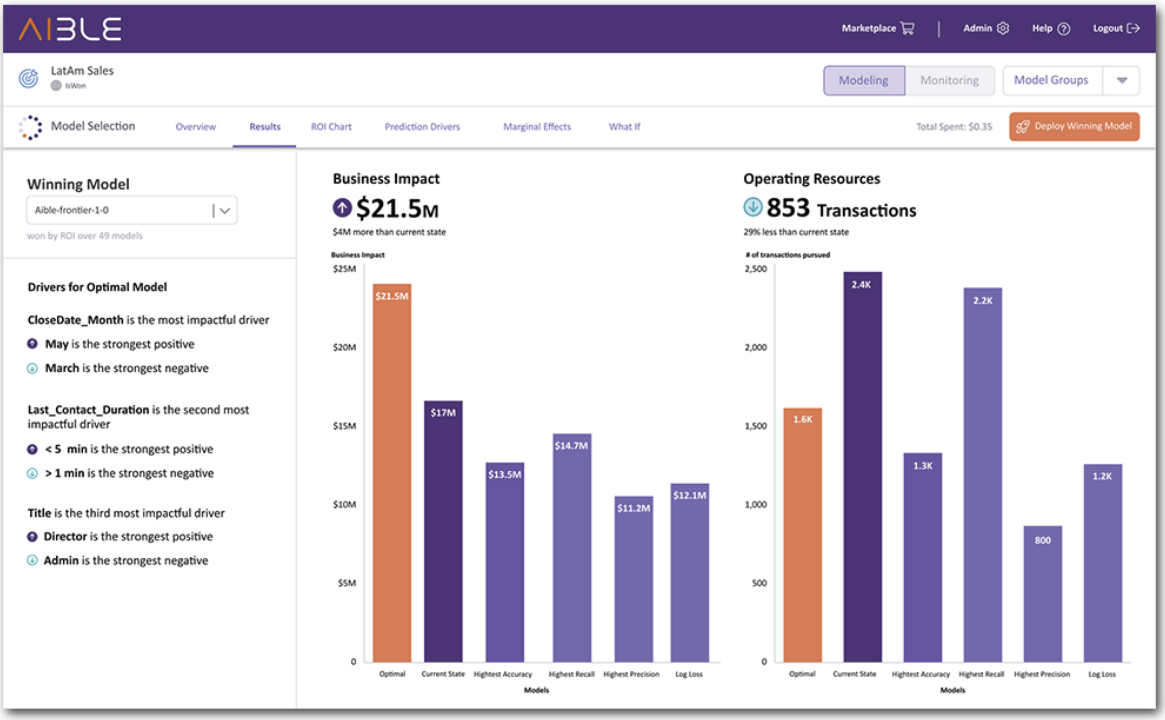


2. Choose a few desired model characteristics

3. Push a big button (and wait)

Source: IBM, H2O.ai, and DataRobot websites

FIGURE 5 An Example Comparing The Business Impact Of Different Models



Source: Aible website

5. How do I use AutoML to add AI capabilities like vision, text, and speech to my apps?

Few data scientists have the skills and experience to build production-grade models from scratch for computer vision applications (like recognizing your products in images and video) or natural language processing applications (like understanding the different parties in legal documents). They haven't had the training, the data, or the need to build these models. Thankfully, for a wide range of use cases, that's not necessary. You can add customized AI capabilities to your applications if you:

› **Consider deep-learning-focused AutoML solutions.** Offered by most of the major cloud vendors as well as an array of focused computer vision and natural language processing providers, these solutions enable you to develop custom deep learning models with no knowledge of deep learning. They either enable you to modify an existing model that has been pretrained on a large corpus of data using your new training data (an approach known as transfer learning) or go the extra step of identifying which of a range of pretrained models is most suited for your data and optimizing that model (known as neural network search). For example, robots in more than 500 Stop & Shop supermarkets detect spills (and thus lower the stores' insurance premiums) with spill detection models built using Neurala's Brain Builder platform (see Figure 6).

› **Select offerings that support each step of model development and deployment.** For most applications, more important than training a deep learning model are the steps before and after — acquiring the training data, deploying your model, incorporating it into the application, and continuously improving the model over time. The best solutions offer a rich workbench for labeling and annotating your images or text, with features that make it faster to annotate, reduce the amount of data you need to annotate, and ensure the quality of the annotation. When using drones to inspect wind turbines, AES used Google's AutoML service to reduce the number of images that engineers needed to review by half.[5] As deep learning models typically require a lot of computing power and are often deployed on edge devices, you want solutions that scale and help you monitor and manage your models in deployment.[6]

› **Have the right features in your automation-focused ML platform.** For some use cases, you might not need a deep-learning-focused AutoML solution. The most established automation-focused ML platforms have support for extracting topics and other predictive features from unstructured data like text. The real estate marketing firm G5 used H2O.ai to build a model using call transcripts to score leads, leading to a five times improvement in qualified leads for their agents.[7]

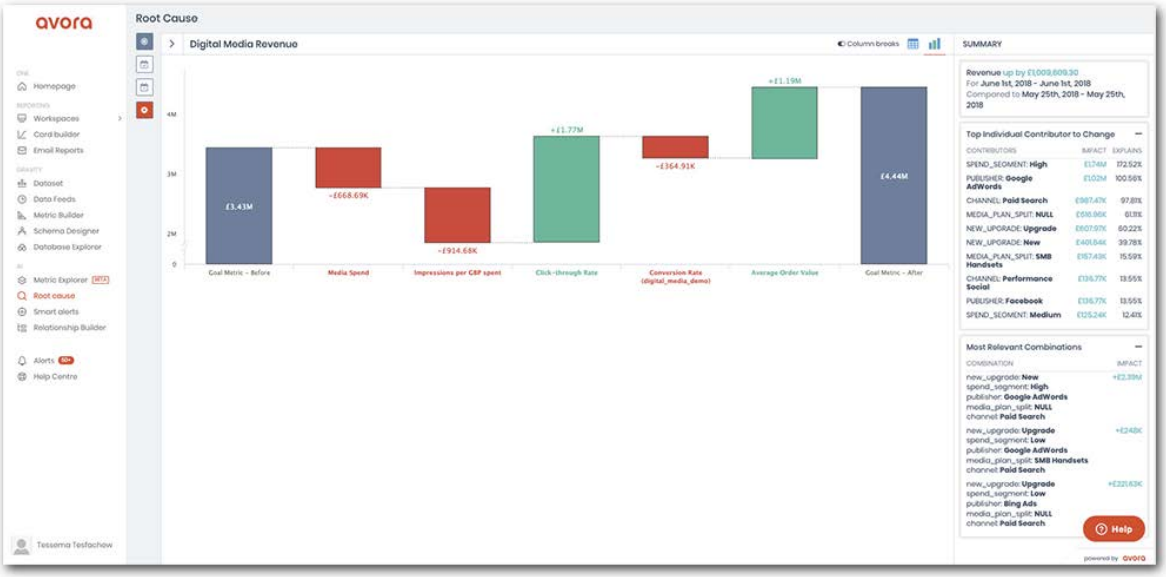
FIGURE 6 Detecting Spills (To Reduce Insurance Premiums) In Stop & Shop Supermarkets

6. How do I accelerate and scale new business insights from data with AutoML?

Contrary to popular belief, not everyone's looking to make predictions. More business users are looking for insights from their data to make better decisions and react faster to changes in their business. However, they don't have the skills, time, or enough talented data analysts. To empower these individuals to leverage the power of ML for data-driven decision making:

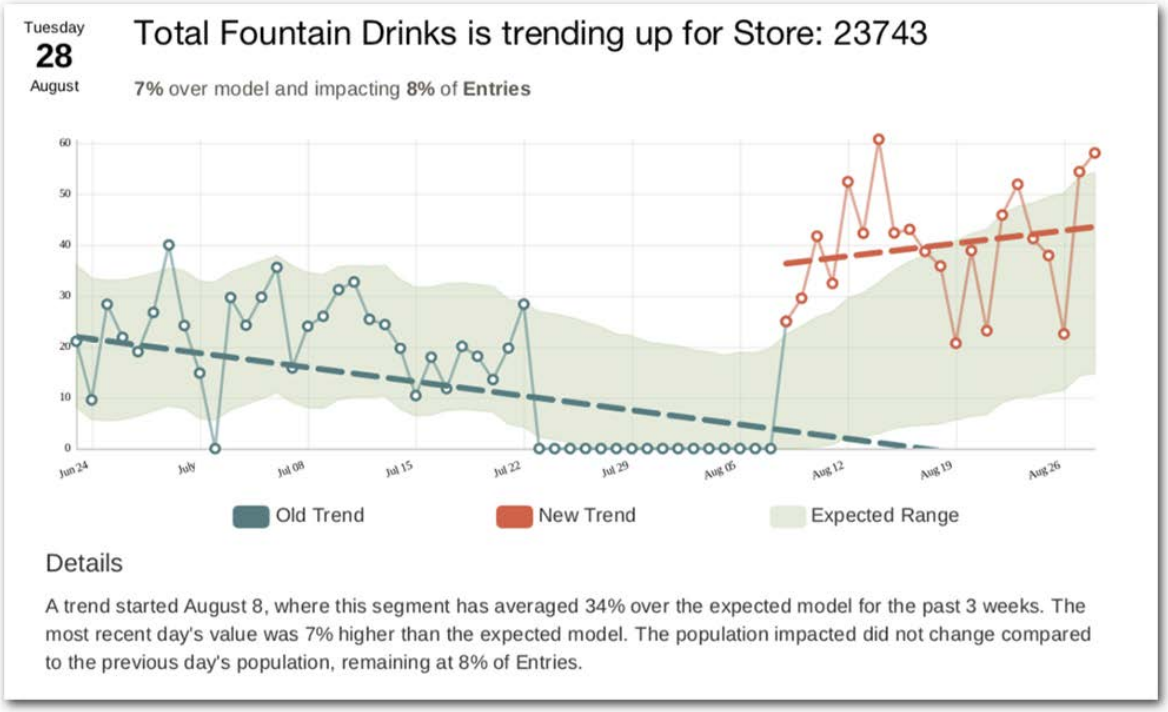
- **Implement augmented business intelligence (BI) solutions.** Data scientists and other data-savvy users will understand how to use automation-focused and multimodal ML solutions to rapidly extract business insights. However, everyone else will struggle to figure out how they're supposed to interpret variable importance scores and partial dependency plots. Augmented BI solutions use AutoML behind the scenes to identify valuable anomalies or potential root causes and suggest either next-best actions or additional data to explore (see Figure 7).[8] A large fast-food chain used Outlier.ai's solution to learn that soft drink sales had dramatically increased in one of their over 25,000 restaurants thanks to a change in the positioning of the soda dispenser, leading them to reconfigure all their restaurants (see Figure 8).
- **Choose tools that deliver a rich, complete analysis.** While every BI vendor is launching features that make it easier to extract insights from data, the best offerings present a rich analysis with autogenerated visualizations and natural language explanations. Where applicable, they also suggest either further areas for analysis or next-best actions (see Figure 9).[9] Typically, when a solution is preconfigured, or can be configured for your particular horizontal use case, it's better able to deliver actionable business insights to your users.
- **Consider a converged offering.** Vendors like H2O.ai and Lityx are going beyond their capabilities for building predictive models and are piloting features that help business users rapidly extract insights from data. Similarly, firms are using dotData's ability to search for new explanatory features across multiple data sets for new business insights.[10] Augmented BI platforms like Microsoft and Tellius are offering features similar to those of the automation-focused ML solutions to automate the creation of predictive models. Dutch professional services firm Macaw used the automated machine learning for dataflows feature in Microsoft's Power BI to build predictive models for inventory management.[11]

FIGURE 7 Example Of An Augmented BI Solution Suggesting Actionable Root Causes



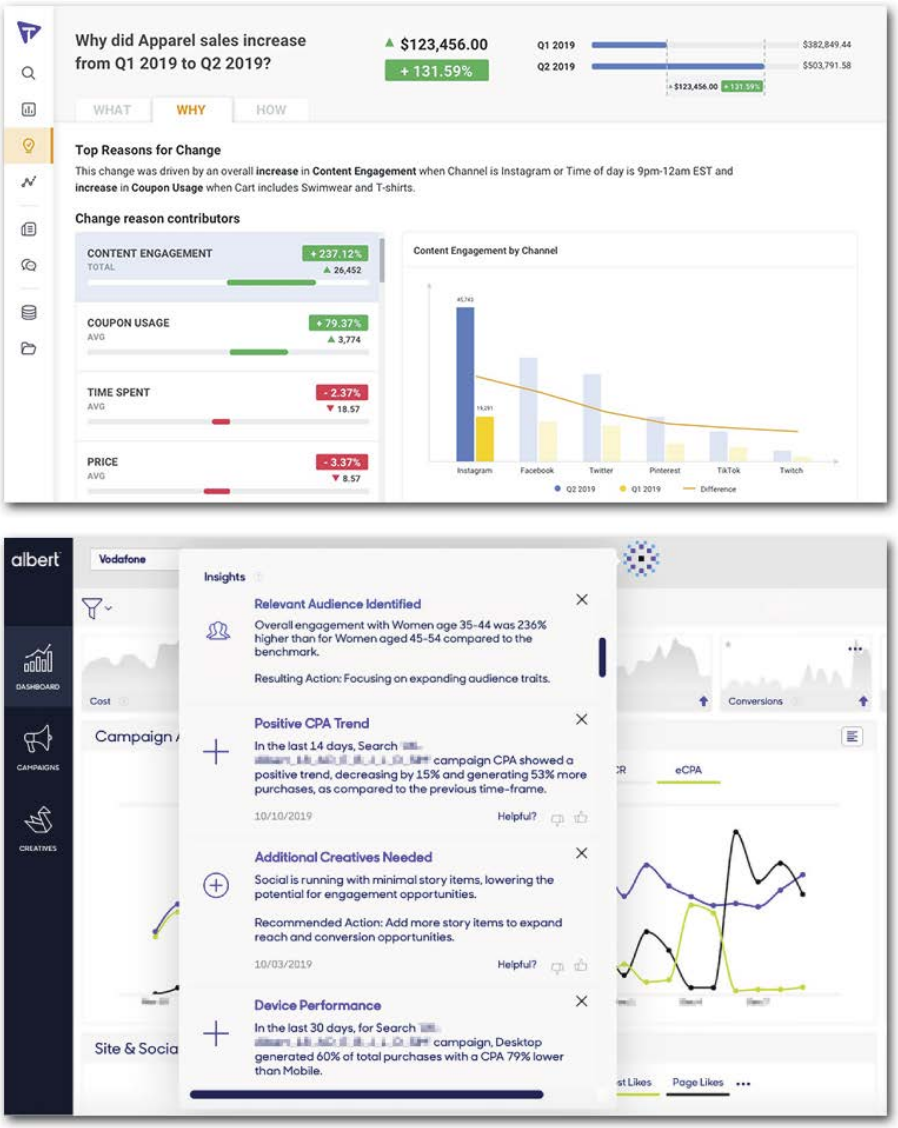
Source: Avora website

FIGURE 8 Example Of Finding A Valuable Business Insight Needle In A Haystack Using AutoML



Source: Outlier.ai website

FIGURE 9 Using Natural Language Querying, AutoML, And Natural Language Generation For Business Insight



Source: Tellius and Albert.ai websites

7. What is the future of AutoML?

AutoML won't spawn an army of self-thinking machines.[12] And you'll still need more data scientists. However, AutoML is developing at a breakneck pace, dramatically increasing its range of use cases, features, availability, and impact. Be on the lookout as:

- **AutoML automates more ML.** New features abound and are being rapidly disseminated across offerings. Time series support, geolocation, topic mining, anomaly detection, support for ensembles, model deployment, and monitoring features that were once the differentiator of individual vendors are becoming common. Expect the arms race to continue with all vendors supporting more and more feature transformations, algorithms, and larger data sets, as well as new battlefronts in the fields of AutoML for data discovery and automated business insights.
- **AutoML makes more impact.** Much like data scientists, AutoML offerings are rapidly learning the importance of understanding the business and are evolving toward solutions that are aligned to the business need and business users. In addition to covering more of the steps needed to implement a business solution based on AutoML, vendors are implementing recipes and templates for common business use cases, making it easy to embed model predictions into business apps. They're also experimenting with new interfaces to make them easier for business users (such as conversational interfaces) and new explainability features to drive trust in the models.
- **AutoML disappears as a standalone product (in the distant future).** AutoML offerings are converging as automation-focused ML platforms build out their capabilities to resemble traditional multimodal ML platforms (e.g., DataRobot's acquisition of ParallelM and Paxata) and vice versa — as witnessed by the proliferation of AutoML features in multimodal platforms in 2019.[13] Simultaneously, expect AutoML features to be embedded in business apps (similar to Salesforce Einstein) and in focused horizontal and vertical solutions. Don't expect any of these to be effective substitutes in the near term though.

Supplemental Material

Companies Interviewed For This Report

We would like to thank the individuals from the following companies who generously gave their time during the research for this report.

Aible
Albert.ai
Avora
dotData
Google
H2O.ai
IBM
Lityx
Microsoft
Outlier.ai
RapidMiner
SAS
Squark
TIBCO

Endnotes

[1] "Cloud AutoML's technology is helping us build vision models to annotate our products with Disney characters, product categories, and colors. These annotations are being integrated into our search engine to enhance the impact on guest experience through more relevant search results, expedited discovery, and product recommendations on shopDisney," said Mike White, CTO, SVP, Disney consumer products and interactive media. Source: "Cloud AutoML," Google Cloud (https://cloud.google.com/automl).

[2] Net Promoter and NPS are registered service marks, and Net Promoter Score is a service mark, of Bain & Company, Inc., Satmetrix Systems, Inc., and Fred Reichheld.

[3] Specifically, most AutoML tools focus on four main tasks: binary classification (e.g., predict whether a customer will buy or leave), multinomial classification (e.g., determine which customer segment they belong to), regression (e.g., estimate customer lifetime value), or anomaly detection (e.g., detect fraudulent purchases).

[4] Source: Erin Sullivan, "Reducing Costs with DataRobot at Steward Health," DataRobot (https://www.datarobot.com/casestudy/reducing-costs-with-datarobot-at-steward-health-care/).

[5] Source: "How Google's AutoML Vision helps AES fight climate change," YouTube video, April 12, 2019 (https://www.youtube.com/watch?v=lF-u7j1x0C0).

[6] For an evaluation of the capabilities of different computer vision platforms, see the Forrester report "The Forrester New Wave™: Computer Vision Platforms, Q4 2019."

[7] Source: "Real estate marketer improves sales lead quality using H2O.ai and AWS," H2O.ai, March 2019 (https://www.h2o.ai/wp-content/uploads/2019/03/H2O.ai-Case-Study-G5.pdf).

[8] See the Forrester report "AI Unlocks The Business Intelligence In BI."

[9] For evaluations of business intelligence offerings, see the Forrester report "The Forrester Wave™: Enterprise BI Platforms (Client-Managed), Q3 2019" and see the Forrester report "The Forrester Wave™: Enterprise BI Platforms (Vendor-Managed), Q3 2019."

[10] For example, Macromill, the Tokyo-based global market research and digital marketing solutions provider, used dotData to identify a new customer profile that wouldn't have been readily uncovered with traditional analysis. The profile was composed of late-night-shopping, health-agnostic, wedding-planning young couples across web access log, POS, and survey data.

[11] Source: "Dutch full-service digital firm tailors a fast AI solution for fashion retailer," Microsoft, June 28, 2019 (https://customers.microsoft.com/en-us/story/724164-macaw-partner-professional-services-power-bi).

[12] For more on what machine learning does and doesn't do, see the Forrester report "Shatter The Seven Myths Of Machine Learning."

[13] Source: Ron Miller, "DataRobot is acquiring Paxata to add data prep to machine learning platform," TechCrunch, December 13, 2019 (https://techcrunch.com/2019/12/12/datarobot-to-acquire-data-prep-startup-paxata/); Sudipto Ghosh, "Latest Acquisition: DataRobot Ushers into New Enterprise AI Horizon with ParallelM," AiThority, June 24, 2019 (https://www.aithority.com/ait-featured-posts/datarobot-ushers-into-new-enterprise-ai-horizon-with-parallelm/); Steve Brooks, "RapidMiner increases Data Scientist efficiency with Auto Model," Enterprise Times, October 15, 2018 (https://www.enterprisetimes.co.uk/2018/10/15/rapidminer-increases-data-scientist-efficiency-with-auto-model/); and "IBM Watson Studio's Latest AutoAI Tool to Automate Processes," Yahoo Finance, June 13, 2019 (https://finance.yahoo.com/news/ibm-watson-studios-latest-autoai-131901291.html).

We work with business and technology leaders to drive customer-obsessed vision, strategy, and execution that accelerate growth.

PRODUCTS AND SERVICES

- Research and tools
- Analyst engagement
- Data and analytics
- Peer collaboration
- Consulting
- Events
- Certification programs

Forrester's research and insights are tailored to your role and critical business initiatives.

ROLES WE SERVE

Marketing & Strategy Professionals

CMO
B2B Marketing
B2C Marketing
Customer Experience
Customer Insights
eBusiness & Channel Strategy

Technology Management Professionals

CIO
- Application Development & Delivery

Enterprise Architecture
Infrastructure & Operations
Security & Risk
Sourcing & Vendor Management

Technology Industry Professionals

Analyst Relations

CLIENT SUPPORT

For information on hard-copy or electronic reprints, please contact Client Support at +1 866-367-7378, +1 617-613-5730, or clientsupport@forrester.com. We offer quantity discounts and special pricing for academic and nonprofit institutions.