



DataRobot

EBOOK

5 WAYS AUTOMATION IS EMPOWERING DATA SCIENTISTS TO DELIVER VALUE



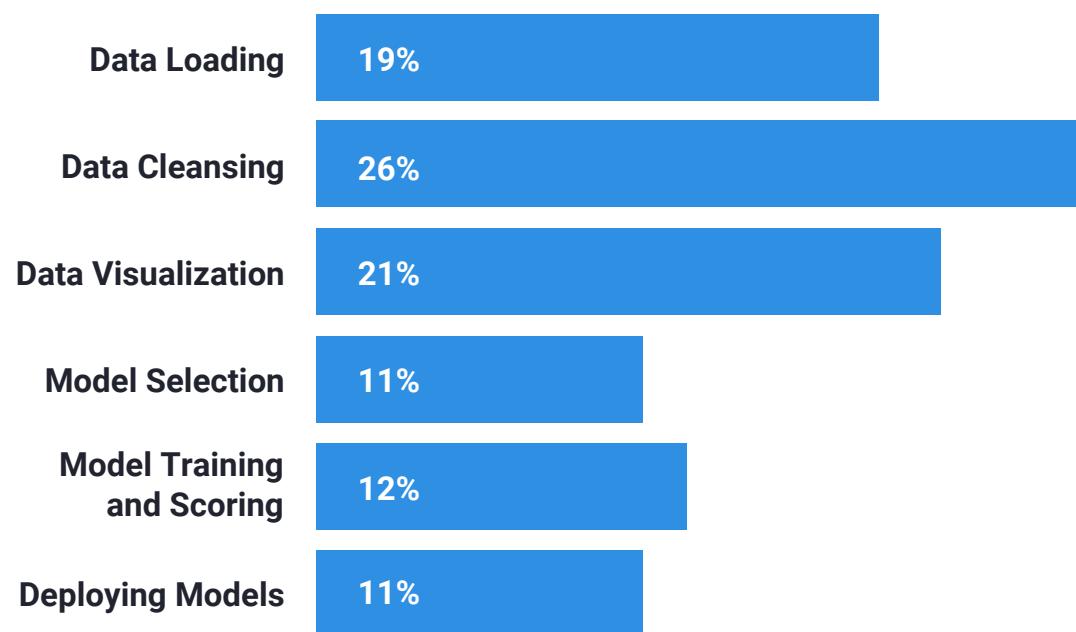


Automation Accelerates Data Science Development and Frees Data Scientists to Focus on Problems at a Higher Level

Data scientists are in demand now more than ever before. They are also busier than ever and increasingly under pressure to work rapidly, while also ensuring that their processes are transparent, reproducible, and robust. It's a situation that leads to some common pain points for data scientists.



The [2016 Crowdflower Data Science Report](#) infamously pointed out that three out of every five data scientists surveyed actually were spending the most time with data wrangling, or what the survey referred to as “digital janitor work.” By 2020, Anaconda’s report, [2020 State of Data Science: Moving From Hype Toward Maturity](#), explained that things had improved, but the efficiency disparity still looms large.



Besides the time spent in data prep, there are well known gaps around model deployment, with [a recent Forrester survey](#) finding that “less than 25% of respondents are very confident in their organization’s ability to implement key capabilities to operationalize machine learning models.”

This paper introduces several ways automation can be helpful to data scientists doing feature engineering, model development, and MLOps. These products and techniques allow data scientists to perform other tasks, such as building algorithms, exploring data, and doing predictive analysis.

Data prep often requires an extensive amount of data scientists’ time. This is where automation can really make a difference. DataRobot’s Feature Discovery automates joining datasets and performing automated feature engineering. With just a few commands, Feature Discovery can join dozens of datasets together, engineer new features, and identify which ones are likely to work best for your problem.

Data scientists can quickly bring value by adopting automation tools that accelerate the development of AI and operationalize machine learning models. For those that learned to manually code their own models, this does require a mindshift. The skills around using automation will develop over the next few years as organizations move to continual learning methods.



"We have transformed our business into one that is driven by data. We trained over **200 data scientists** on [DataRobot](#) and put into production **100 dashboards and 30 machine learning models**. These models have really transformed the business and helped decision-makers and executives change the way they do business."

— Dr. Megan Oftedal

Principal Product Manager, Data Scientist,
American Fidelity Corporation

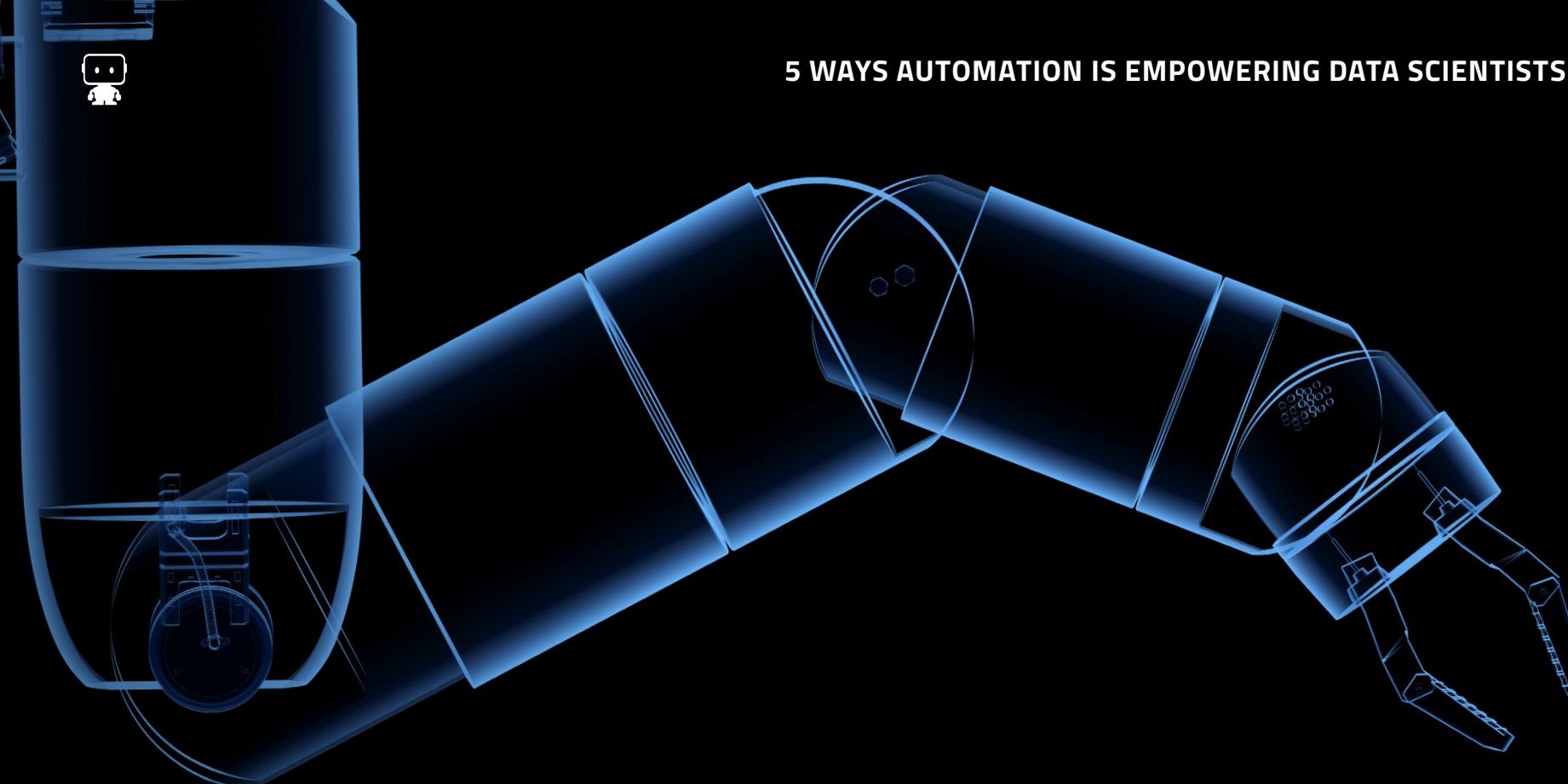


How Data Pipelines Train Models and How Image-Based Data Enhances Predictive Ability

After you have built your models, deploying to production and maintaining them requires time and an exhaustive examination of your data, predictions, and each model's accuracy. MLOps provides a structured framework to maintain proper governance over your AI projects, while Continuous AI can refresh your production model without unnecessary manual intervention.

Finally, as trust is essential in any relationship, you'll see the expertise and tools to test your systems across multiple dimensions of trust to maintain operational excellence and reflect your values.

Let's take a deeper look at these five best automation capabilities that empower data scientists to deliver real solutions faster.



1

TAKE AUTOMATED FEATURE
ENGINEERING TO THE NEXT LEVEL
WITH FEATURE DISCOVERY



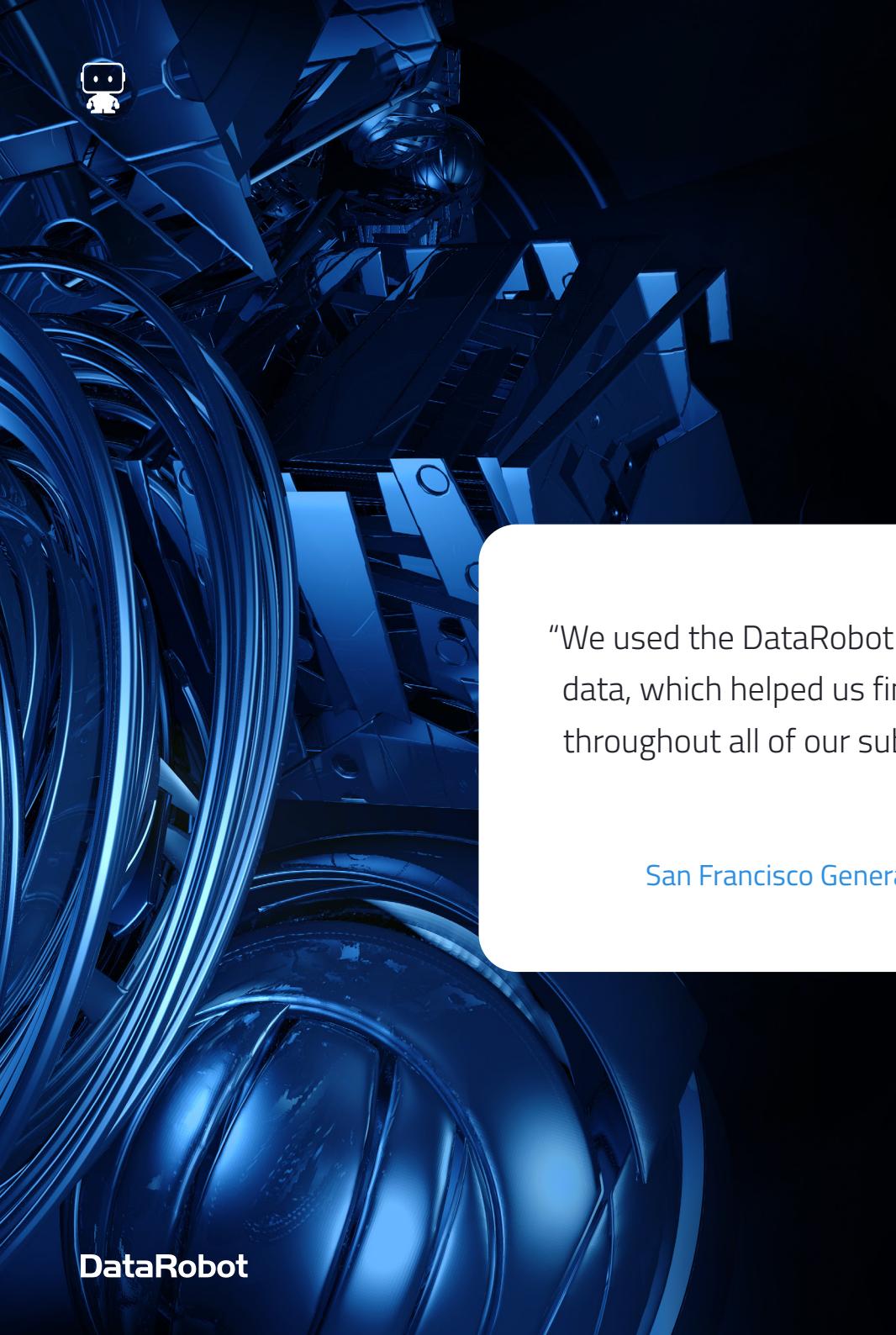
Take Automated Feature Engineering to the Next Level with Feature Discovery

Data scientists find that features are divided across multiple data assets. A substantial amount of data prep time is spent on joining these data assets together to run machine learning models on top of them. A typical issue that arises is that datasets are often stored across enterprises at different levels of granularity.

As part of data preparation, a data scientist may spend their time summarizing transactional level information to the level of a customer. This could involve a laborious process of aggregating the transactional data and creating new features, such as the average transaction amount, the max transaction amount, and so on. DataRobot's Feature Discovery automates this process.

"Joining datasets of different granularity is exhausting. With Feature Discovery, this process becomes a game. With little to no effort, I have hundreds of features automatically generated. The best part? **My models love them.**"

— Thodoris Petropoulos
a Customer-Facing Data Scientist at DataRobot,
illuminates the time-saving benefits of using Feature Discovery



"We used the DataRobot Feature Discovery tool on our high-frequency physiological data, which helped us find some new features that ended up being high impact throughout all of our subsequent analysis."

— **Dr. Austin Chou**
PhD, Data Scientist, UCSF Brain and Spinal Injury Center Zuckerberg
San Francisco General Hospital, has witnessed the power of Feature Discovery firsthand

Feature Discovery also provides users with flexibility, such as specifying how much history should be implemented when determining new features. Implicit guardrails allow you to specify how much history should be implemented when determining new features and ensure that future data is eliminated when generating new features.

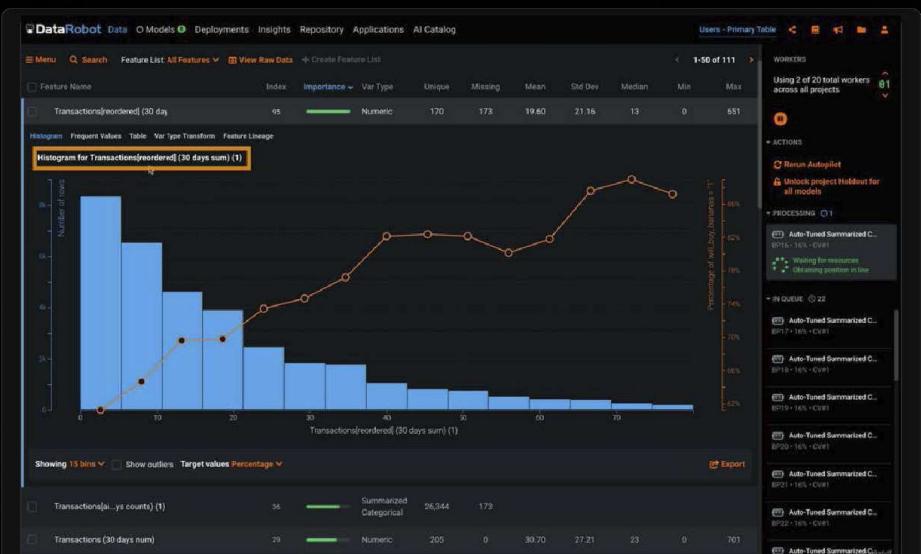
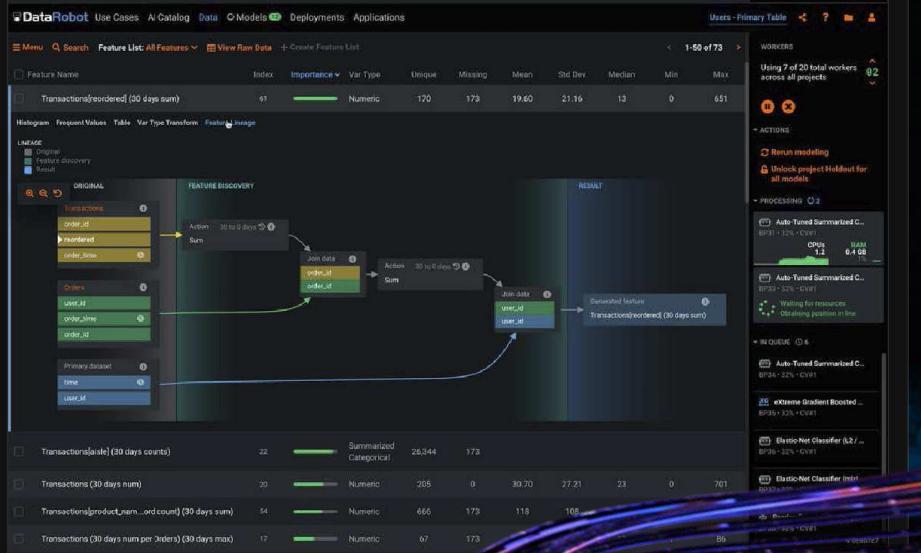


Figure 3. Histogram for 30-Day Sum of Reorders

Figure 4. Feature Lineage tab



Feature Discovery Exists to Automate Features and Remove Manual Procedures

To start using Feature Discovery, you first need to upload all of the datasets you want to use to the [AI Catalog](#), a centralized hub for datasets that can then be used for model training and batch predictions. After you load your data, DataRobot performs [exploratory data analysis \(EDA\)](#) and shows any [data quality issues](#) detected.

DataRobot will then automatically generate hundreds of features and eliminate the ones that do not make sense from the modeling procedure. Figure 3 shows an example of a generated feature.

This particular generated feature is the 30-day total of the reordered products found in the Transactions table. If you want to take a deeper dive into this feature, you can also visit the Feature Lineage tab (Figure 4).

Figure 5. Download Dataset and Feature Derivation Log

The screenshot shows the DataRobot Feature Discovery interface. At the top, there are tabs for 'Project Data', 'Feature Lists', 'Feature Associations', and 'Feature Discovery'. Below these, a summary shows 'Secondary Datasets' (2), 'Explored features' (309), and 'Generated features' (66). A prominent orange button labeled 'Download Dataset' is located in the center. To its right is another orange button labeled 'Feature Derivation Log'. On the left, three datasets are shown: 'Transactions' (1.3M rows x 6 features), 'Orders' (1.4M rows x 2 features), and 'Primary Dataset' (63.1K rows). Each dataset has a 'Days before time' slider. On the right side of the interface, there is a 'WORKERS' section showing 'Using 5 of 20 total workers across all projects' with a progress bar at 25%. Below this are sections for 'ACTIONS' ('Rerun modeling', 'Unlock project Holdout for all models') and 'PROCESSING' (which lists several machine learning models like 'Auto-Tuned Summarized C...', 'Extreme Gradient Boosted...', 'Elastic-Net Classifier (L2)...', 'Elastic-Net Classifier (Ridge...)', and 'RandomForest Classifier (...)' along with their performance metrics).

After DataRobot calculates the new features and removes the ones with no signal, these features are now available for model building. You also have the option to both download the dataset with the newest features created by Feature Discovery or see the details using the Feature Derivation Log. Both of these options can be accessed by clicking on the Feature Discovery tab under Data, as shown in Figure 5.

From intricate data schemas and datasets within different source systems, DataRobot automatically discovers, tests, and creates hundreds of valuable new features for your machine learning models, dramatically improving their accuracy, and deepening your overall understanding of the problem at hand.

DataRobot Feature Discovery can join and aggregate your dataset. After you have illustrated how the datasets will be united, you can leave the feature generation and modeling to DataRobot — and forget about many of the arduous and repetitive steps of feature engineering.



Assemble Feature Engineering and Model Training Tasks to Create Tailored Blueprints

While Feature Discovery exists to automate features and remove manual procedures, Blueprints automatically bring feature engineering and model training tasks together to generate training pipelines from the data at hand.

The Blueprints are then trained against the data to surface the best option for the problem at hand and a Leaderboard displays the list of resulting models, allowing easy exploration and iteration.

No matter the given dataset – and whatever the feature type – Blueprints can be created via one click or one line-of-code. The DataRobot Repository holds a list of created Blueprints, any of which can be run at any time.



2

DATA PIPELINES TRAIN
MODELS WITH ALL THE
NECESSARY NEW DATA



The Building Blocks of DataRobot Pipelines:

Prepare data and deploy pipelines for machine learning initiatives with a purpose-built application.

- **User-Defined Data Pipelines**

Build reproducible and reusable Data Science and Data Engineering Pipelines

- **Interactive Code-Centric User Experience**

Build faster using the interactive GUI to view data and edit code.

- **Intelligent Data and Schema Aware Engine**

Run pipelines on a highly optimized intelligent engine with smart caching of results.

- **Elastic Scalability**

Ingest and transform large data at scale utilizing a multi-node distributed architecture.

- **Scheduling and Operationalization**

Schedule, orchestrate, and manage your data pipelines.

Data Pipelines Train Models with All the Necessary New Data

Data scientists need to stay on top of machine learning data throughout every stage of model development and model deployment. Teams can start by collecting data from various sources, then cleaning and combining it.

Data science teams can then build a dataset at the unit of analysis by standardizing the values among other data preparation operations — and this is repeated throughout the lifecycle of a model. To facilitate the process, data scientists can lay a data pipeline — a set of connected data processing steps — to train models with all the necessary new data.

You can connect to data sources of various formats and implement data transformations to build and orchestrate your machine learning data flows with DataRobot Pipelines.



Create a Learning Dataset

Ingest raw data from various sources and use the distributed processing power of SparkSQL to join and clean data to arrive at a dataset that can be exported to the AI Catalog from where a Project could be created to start the training process.



Create Scheduled Prediction Pipelines

Create a Learning Dataset. Ingest raw data from various sources and use the distributed processing power of SparkSQL to join and clean data to arrive at a dataset that can be exported to the AI Catalog from where a Project could be created to start the training process.



Clean Up Outliers and Segments

For training datasets that contain outliers or data segments that you are not interested in training for, use the SparkSQL transformation module in a pipeline to remove unwanted rows and publish a clean dataset to AI Catalog.



Data Downsampling

With the power of SparkSQL, you can easily downsample your data using a simple WHERE condition or more complex operations like Window functions and Tablesample expressions. You can also apply the same approach to extract various segments of data from a full dataset.



3

AUTOMATE EXPERIMENTS
WITH A CLEAR PATH
TO PRODUCTION



Automate Experiments with a Clear Path to Production

Composable ML equips data scientists with a fully flexible approach to model building, allowing them to combine their own data science expertise with DataRobot.

With Composable ML, you can let DataRobot try out the initial approaches for you. With more than 200 transformers and estimators available out-of-the-box, you can then compose and edit your own training pipeline using your own expertise. In addition to this, you can:

- **Employ Python and/or R to define modeling logic.**
- **Stitch Python and R tasks together in a single blueprint and the platform will handle the data conversion.**
- **Install any dependency and – if required – bring your own Docker container.**

Once your blueprint is trained on DataRobot's infrastructure, you get instant access to the model Leaderboard, MLOps, model documentation, and model insights. Finally, you can share your Composable ML blueprints across your organization in just a few clicks.



For Code-First Data Scientists, Zepl Is Your Data Science Notebook Solution

For code-first data scientists, [Zepl](#) is your data science notebook solution. You can start by using Zepl's APIs for Python or R. Zepl's pre-built notebooks make it easy and productive to work with colleagues on your data science projects in one place and give you all the code you need to connect with DataRobot in the language of your choice.

Extensible at its core, Zepl integrates elegantly with countless other popular frameworks and libraries and allows you to easily explore datasets and do ad hoc analysis. Connect the data in Snowflake, S3, or other data sources in seconds. You can also visualize with Zepl's built-in Plotly Chart Editor or the visualization library of your choice.

Zepl provides data science teams all the tools they need to be successful analyzing data and supports the most widely adopted open-source notebooks on the market today, Jupyter and Apache Zeppelin. You can then easily share results with business users by using the publish feature. Zepl also requires no maintenance from you or your team.

With no software to download, install or maintain, you and your team can use Zepl across any platform and anywhere, evolving your questions and ideas into answers and outcomes more rapidly. Get all of the benefits of notebooks without the headaches with Zepl's fully-managed cloud solution.



Build, Manage, and Deploy Highly Accurate Forecasts Across Multiple Segments

Success in business and analytics – no matter the industry – depends on working with optimal partners that add the most value to your organization. Segmentation further divides the targets based on the most important features.

For example, [common customer segmentation models](#) range from simple to very complex and can be used for a variety of business reasons. Common segmentations include:

- **Demographic segmentation**
- **Recency, frequency, monetary (RFM)**
- **High-value customers (HVCs)**
- **Customer status**
- **Behavioral segmentation**
- **Psychographic**

[In most cases](#) related to machine learning, segmentation is used for data that is unlabeled, which means that only the inputs are given. The inputs are further segmented into different clusters or groups based on the relations between them.

DataRobot Segmented Modeling for Time Series empowers data scientists to build, manage, and deploy highly accurate forecasts across hundreds of segments. With Segmented Modeling, you can:

- **Automatically detect when a series will benefit from segmented forecast modeling**
- **Build highly accurate and optimized models for each defined segment**
- **Evaluate model performance on multiple levels**
- **Streamline management and fast-track deployment with intelligent model roll-up**



Ease of Use

- **Upload and Share.**

Add images to different folders to classify them or use a .csv file to define the mappings. Zip everything up, then upload to DataRobot's AI Catalog. From there you can create a new project or share your image dataset with individuals and teams across your enterprise.

- **Profile and Preview.**

DataRobot runs automated exploratory data analysis (EDA) on your image dataset. This allows you to preview sample records, see basic feature statistics, and assess overall data quality to identify issues like misclassified images, missing values, or duplicate images.

- **Run Autopilot for Image Models.**

Just pick the target and select Start. DataRobot automatically prepares, selects, trains, and blends a wide variety of traditional machine learning models with cutting-edge deep learning neural networks on your image data.

- **Bring Your Own Deep Learning Model.**

You can bring your own deep learning model via Composable ML to experiment with speed and scale.

Visual AI Enables You to Blend Images Alongside Multiple Features

The first commercial use of computer vision broke through [in the 1970s](#) as machines interpreted typed or handwritten text using optical character recognition — a development that was a major advancement for the blind.

Today, visual AI is helping the oldest and yet most [fundamental industry in human history](#): farming. John Deere's [Blue River See and Spray](#) machine accurately sprays herbicides only where weeds are present with a combination of computer vision and machine learning, thus reducing non residual, pre emergent herbicide use by 77% on average. Whether it's for agriculture, or [manufacturing and healthcare](#), visual AI is expected to be a [\\$5.2 billion market by 2026](#).

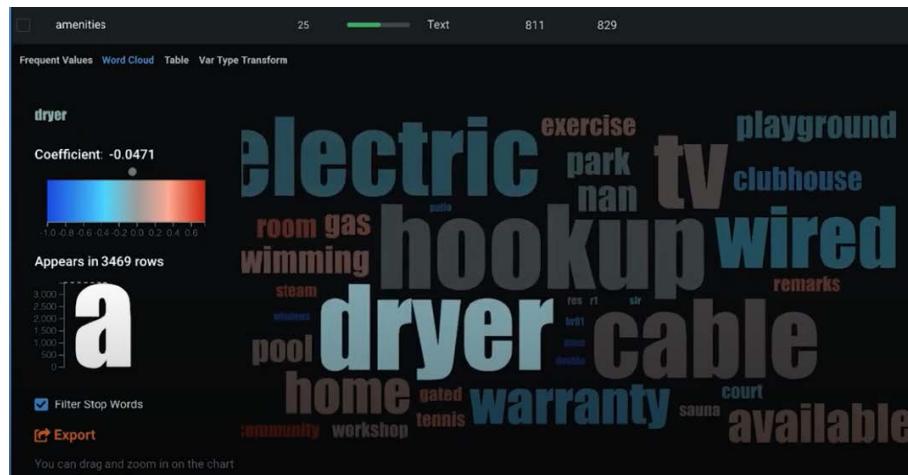
DataRobot Visual AI allows data scientists to quickly and easily build and deploy highly accurate and explainable machine learning models using image-based data. This technology unlocks classification, regression, clustering, anomaly detection, multilabel classification use cases with images in the DataRobot AI Cloud Platform.

DataRobot Visual AI also enables you to blend images alongside as many other feature types in your dataset as you like. You can mix images with numeric, categorical, date, and raw text features to provide much broader data perspectives to your AI models, allowing you to innovate in completely new ways.



Integrating Diverse Data Formats into a Single Model

Visual data also helps real estate developers predict listing prices. When combined with geospatial, text, and tabular data, a multimodal dataset can help make even more accurate predictions. This [House Listings Demo](#) analyzes a real estate dataset with 6k property listings and 32 features of various types, including text descriptions, the year built, zip code, latitude and longitude, as well as exterior and interior images.



Heat maps show areas with the highest and lowest square-foot locations, while word clouds summarize the key terms from freeform text features and show how they relate to the target.

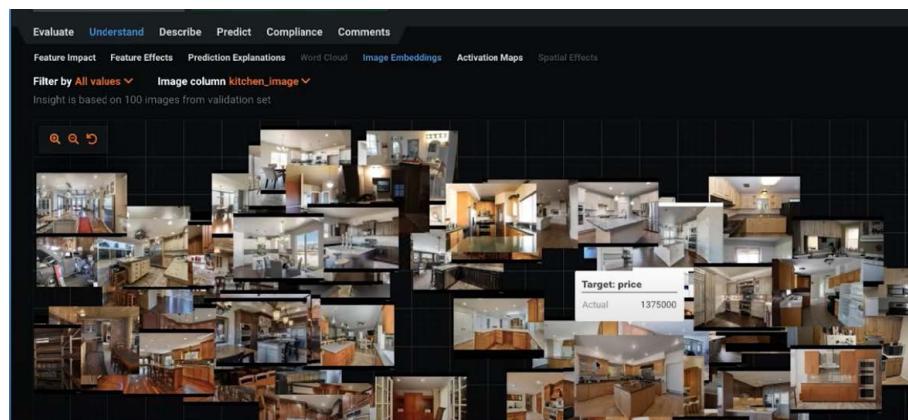


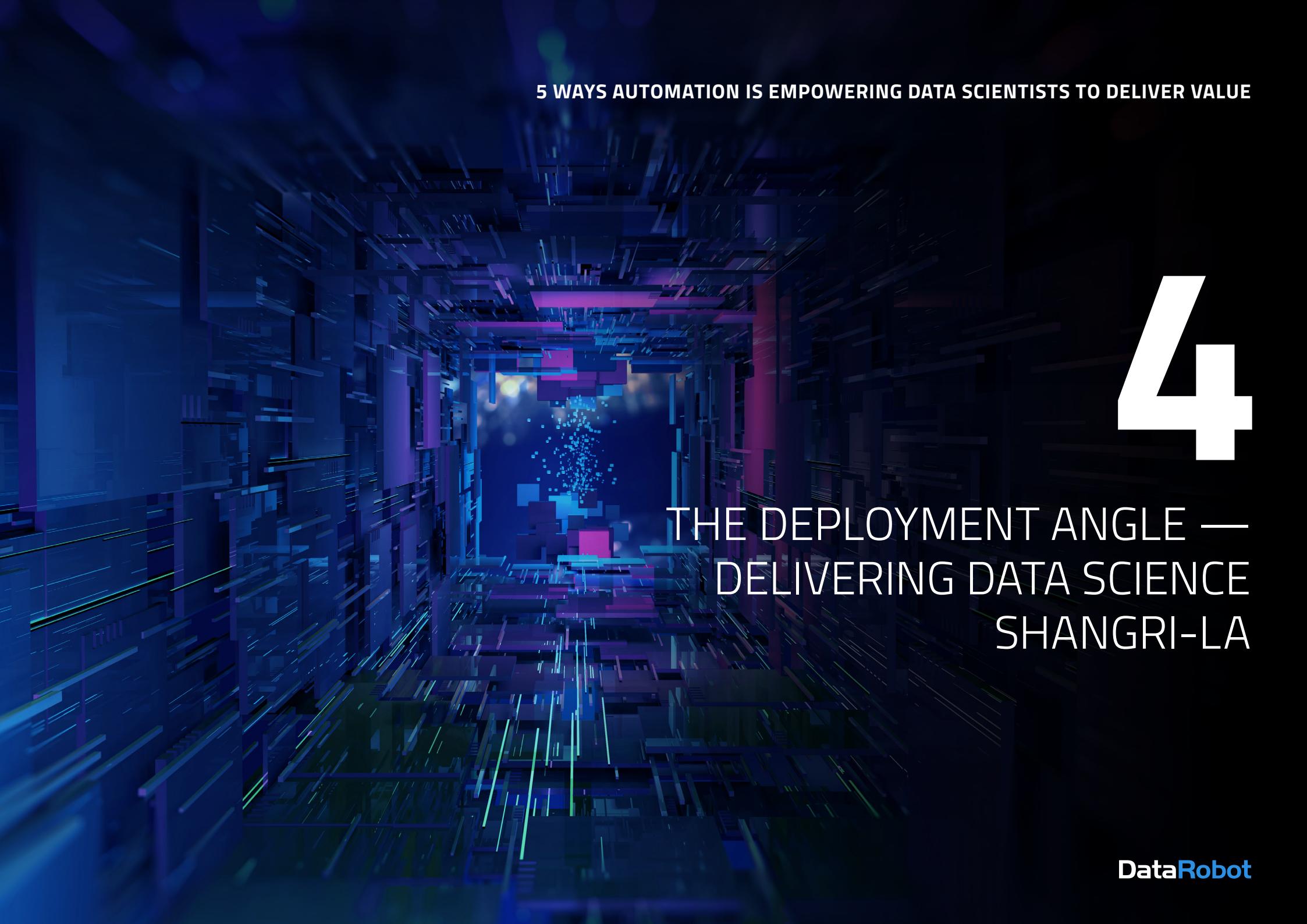


Image features, such as image embedding, performs unsupervised learning to cluster images and see unexpected patterns. For example, you can see clear differentiation between modern and more traditional kitchens with wooden cabinets. An image of a shabby-looking bathroom, for example, could be the key piece of data that helps accurately predict a low listing price.

Image activation maps then show exactly where the network looked for each and every prediction it gave—giving confidence that the model is looking at the right parts of the images when making predictions.

In addition to the power of implementing image data, data scientists also want to leverage text, categorical, and geospatial data to help answer many questions about their data, even when faced with a non-predictive problem. Multimodal Clustering provides users with a one-click/one-line-of-code experience to build and deploy clustering models on any kind of data — numeric, categorical, text, geospatial — not just images.

Multimodal Clustering also introduces the new Cluster Insights visualization, which allows you to understand, name, and explain each cluster for any model, allowing you to solve more complex use cases by generating insights from experimentation.

A dark, futuristic digital cityscape serves as the background. It features a complex network of glowing blue, purple, and green lines and rectangular nodes, creating a sense of depth and data flow. The overall atmosphere is high-tech and modern.

5 WAYS AUTOMATION IS EMPOWERING DATA SCIENTISTS TO DELIVER VALUE

4

THE DEPLOYMENT ANGLE — DELIVERING DATA SCIENCE SHANGRI-LA



The Deployment Angle – Delivering Data Science Shangri-La

Maintaining models requires time and energy-intensive inspections of your data, predictions, and each model's accuracy. Your team could waste loads of time doing manual work on dashboards. And when they stop working, they'll need to refresh and retrain them. Then the whole process starts again — overtaxing already strained data science teams and resources and potentially creating tension across your organization.

Even worse, you may not even know that one or more of your model's features or the target have changed in a significant way between training and inference. Before you can correct [data drift](#), or the change in model input data that leads to model performance degradation, you will also want to know if the feature in question is essential to your model's accuracy.

[Forrester](#) even went so far as to say that "AI models are like six-year-olds during quarantine: They need constant attention...otherwise, something will break."

Consequently, data scientists need to find a better way to deploy and manage the lifecycle of all their production models holistically across the enterprise.

DataRobot [MLOps](#) delivers this promise, empowering organizations to deploy, manage, monitor, and govern their machine learning models from one particular place, regardless of location or how they were created. MLOps also helps maintain consumer trust by ensuring accurate and reliable predictions — even after coming across scenarios not seen in the training data.



MLOps Delivers the Discipline to Maintain Control over Your AI

DataRobot MLOps allows stakeholders to work seamlessly on the mutual goal of scaling and managing trusted machine learning models in production. You can also deploy any model to any production environment that you like, whether that be a cloud provider, on-prem, or a hybrid result. MLOps [monitoring agents](#) also delivers monitoring in place of any deployed production model. Finally, you can explain your model degradation and keep an eye on service health, accuracy, and data drift by implementing automated best practices.

Undoubtedly, your deployed model will degrade over time. As the data used to train your model looks increasingly different from the incoming prediction data, prediction quality inevitably declines and becomes less reliable. Data scientists can do one of the following:

- **Build your own [challenger models](#)**
- **Use DataRobot's [Automated Machine Learning](#) product to build them for you**
- **Test these models against your current champion model**

This method of continuous learning and evaluation keeps you on your toes with model performance — something to keep in mind in our dynamic and [highly volatile world](#). MLOps delivers a disciplined framework where you maintain control over your AI projects across your entire organization as fully functional enterprise AI comes closer to meeting the end of the rainbow.



Improving Model Management to Ensure the Most Effective Business Adoption

Before choosing DataRobot, Stanley, Black & Decker had a data science team that used a wide range of open source code to solve a myriad of analytics problems across more than 30 brands. The team needed a tool that would quickly and accurately build the best machine learning models for a specific problem, including demand forecasting, which was previously handled using an outdated and extremely manual approach.

"DataRobot will help us to rapidly iterate on thousands of combinations of models, data preparation steps, and parameters that would take days or weeks to do manually. In addition, their unparalleled development, support, and variety in model management will enable us to quickly deploy the prediction models and ensure the most effective business adoption."

— Aleksandar Lazarevic,
VP of Advanced Analytics & Data Engineering,
Stanley, Black & Decker



Continuous AI Can Refresh Your Production Models Without Unnecessary Manual Intervention

The COVID pandemic changed so many things in life for 2020 and beyond. It changed the way people dress, how they work, eat, and work out, and altered their hygiene routines and standards, and even the lenses in people's glasses. Machine learning models trained on 2019 data, however, had no idea that the demand for Clorox wipes, athleisure wear, and anti-fog lenses would spike in demand — and they didn't know what to do. They simply panicked and started coughing up bad predictions.

Companies such as Zenni Optical were able to react quickly with fast-tracking innovations like anti-fog lenses for newly masked consumers.

[Courtney Fadio Biro](#), VP of Growth Marketing at Zenni, reports that pandemic-informed innovations helped [Zenni surpass its pre-COVID 2020 sales projections by 45%](#). Zenni, Clorox, and Thinx not only survived the pandemic, but thrived. [Other organizations, meanwhile, were less fortunate.](#)

Now, bearing the lessons of the pandemic in mind, imagine a production machine learning model that always stays accurate after it's deployed — all by itself. Imagine predictions becoming even more accurate as time passes without unnecessary manual intervention. Now, when accuracy drops below a predetermined threshold or data drift occurs, you can create multiple [MLOps](#) retraining strategies to refresh your production models based on the schedule of your choosing with [Continuous AI](#).

Continuous AI not only retrains your current production models for you, it also generates and tests a whole host of new models and presents the top ones as recommended challengers. Challengers are then replayed against historical prediction data for you (or the system) to decide if one of them should be promoted as the new champion.

Continuous AI combines the best of automated machine learning with the best of machine learning operations to continually improve models over their full lifecycle — and the entire process can be fully automated and scaled across your entire organization.

The background of the slide features a dynamic, abstract pattern of overlapping blue and black geometric shapes, including rectangles and triangles, creating a sense of depth and motion.

5 WAYS AUTOMATION IS EMPOWERING DATA SCIENTISTS TO DELIVER VALUE

5

BUILDING TRUST AND
REDUCING BIAS



Building Trust and Reducing Bias

Corporations have been coaxing customers to trust them in new and more intrusive ways, from obtaining personal information to allowing them to monitor online behavior. Simultaneously, news feeds are awash with technology-centered stories focusing on the misuse of personal data, hacks, unscrupulous surveillance, no transparency, algorithmic bias, and the dissemination of misinformation.

Even in instances where the code is working and the model looks to be generating adequate answers, it's possible for a model to encode fundamental data science mistakes that render its results invalid. The effects can be catastrophic when the model is implemented to make decisions in the real world.

Subtle examples of running code that led to [failing models](#) expose the Achilles' heel of machine learning. The following examples highlight the need for automation tools that offer guardrails and insights along the way:

- [Target Leakage in a fast.ai Example](#)
- [The SARCOS Dataset Failure](#)
- [The Poker Hand Dataset](#)

AI can malfunction, be intentionally corrupted, and acquire and even codify human biases in dubious ways (racial or gender discrimination) that may not be immediately recognizable. Hence, gaining the [trust](#) of regulators, the media, and customers is more important than ever for corporations across the board.

Companies that wish to ensure that their AI is trusted need to implement more secure safeguards. DataRobot's specialized [Trusted AI Team](#) understands the significance of trusted and ethical AI and is constantly developing statistical techniques and tools to assess bias, model confidence, and robustness across multiple dimensions and stages of the machine learning lifecycle.

You can also use [Bias and Fairness Testing](#) to identify bias in models with protected features, such as gender and ethnicity, and afford guidance to resolve upstream issues and prevent bias from reoccurring in the future.



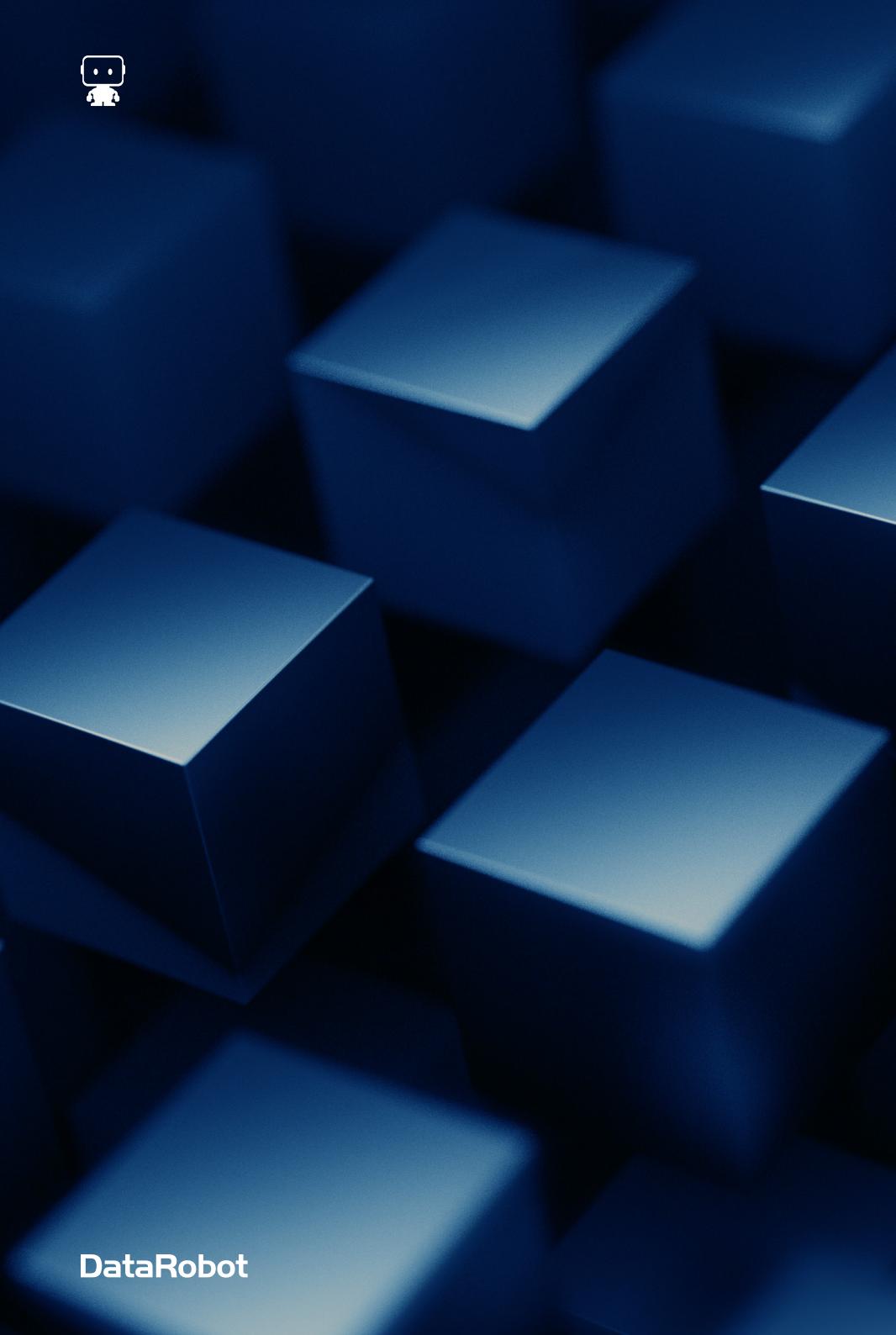
Automating Core ML Capabilities Is Crucial for Data Science Success

Firms expect data science teams and machine learning capabilities to bring business success, and organizations realize that adopting machine learning in business applications and processes will lead to stronger returns. However, these same firms are challenged to achieve machine learning at scale, and data scientists require the tools and processes to enhance cooperation and successfully automate and operationalize machine learning across the business.

The more automation capabilities that a data scientist has at their fingertips, the more strategic problems they can tackle head-on. In turn, they will be freer to use their coding skills to solve larger problems and find real solutions faster.

From data prep to model deployment, automation allows data scientists to be more agile. The more automation capabilities available to a data scientist, the more problems and concerns they can handle directly in each industry. By building better algorithms, exploring data, and experimenting with a clear path to production, they can deliver scalable solutions and repeatable, transparent success and value.

Even with the best automation weapons, most firms lack competency in key MLOps capabilities, and this inability to accomplish key machine learning capabilities points to a shortage of MLOps processes, tools, and technologies.



DataRobot [MLOps](#) empowers organizations to deploy, manage, monitor, and govern their machine learning models from one particular place, regardless of how they were created or where they are located, and finally gives data scientists the opportunity to solve a higher level of problems by using their cherished coding prowess and hard skills to implement real solutions faster. Of course, it means very little if your AI solutions are not trusted by your customers, partners, the media, and regulators.

Finally, while some may see automation as a threat, it's best to take what has been called an automation-first mindset and ask which parts can be automated. Think about the rise of free and open-source content management systems. This didn't lead to a lack of demand for HTML programmers, but rather resulted in an expansion of better-looking websites and a demand for experts to improve them and stand out among their peers and competition.

In the same way, the DataRobot AI Cloud Platform advances data science by accelerating the end-to-end journey from data to value and by automating a lot of the manual work that data scientists currently do. Ultimately, AI Cloud makes a data scientist more productive and helps businesses realize greater ROI.



DataRobot

DataRobot was founded in 2012 to democratize access to AI. Today, DataRobot is the AI Cloud leader, delivering a unified platform to accelerate delivery of AI to production for every organization. DataRobot's vision is to bring together all data types, all users, and all environments to deliver critical business insights for every organization. DataRobot is trusted by global customers across industries, including a third of the Fortune 50, delivering over a trillion predictions for leading companies around the world. Learn more at [datarobot.com](https://www.datarobot.com).

Sign up for a free trial today to find out how DataRobot can help your organization at [datarobot.com](https://www.datarobot.com)