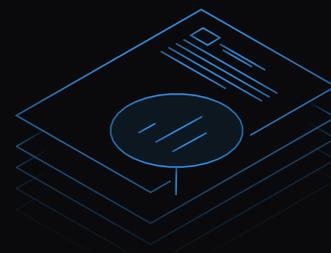




Democratizing Data Science at DemystData



For New York-based software company DemystData, it's no surprise to learn that data is a core part of their business; after all, it's right there in their name. The company's goal is to "demystify" data, with a platform that helps their clients discover, explore, and access the exploding world of data.

Financial institutions, especially legacy big banks are using more data than they were in decades past but, as a whole, data is still being heavily underutilized in this space. That leads to a lot of business decisions being made on suboptimal or incomplete pictures of people, businesses, and properties.

DemystData aims to close that gap and help solve the problem by opening up their clients' access to new and more data. But as datasets get bigger and data sources more varied, that also means increased complexity and more time-consuming work for an already limited pool of data science resources at the New York-based company.

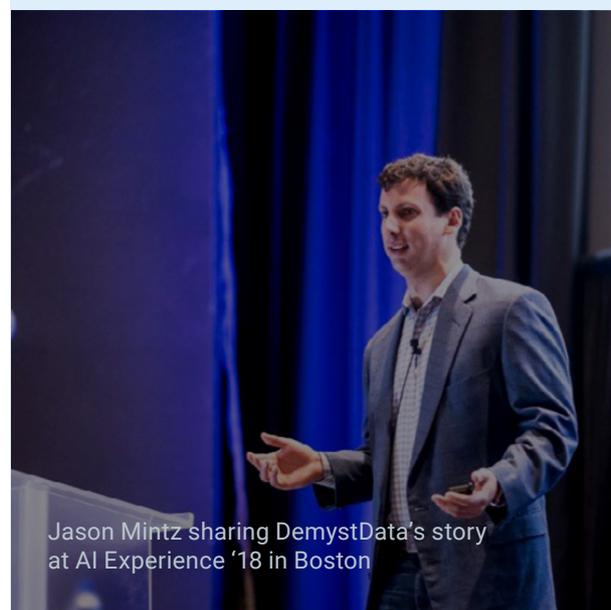
That's when DataRobot came in to meet the challenge.

To really crystallize the impact that DataRobot's automated machine learning platform has made at DemystData, VP of Product Jason Mintz explained the difference between deductive and inductive problem solving:

- **Deductive problem solving** is the typical scientific method approach. You start with a general principle or idea, look for evidence to prove or disprove it, adjust your hypothesis accordingly, iterate, and repeat. It's a perfectly valid way to solve problems, one undertaken at most financial services companies.

If a legacy financial institution wants to build a new business credit model, the data science team might sit down and ask themselves, "What will be predictive of a business' credit? Is it their online rating? Their revenue? The number of people who visit their website?" The team would discuss potential fields that could be relevant to solving the issue, find sources of that data, build models, test them, and see what's predictive. It works, but they would need to know what they're looking for, and are ultimately limited by their imagination; they have to hope that they find what they need to confirm or disprove their hypothesis.

DemystData



Jason Mintz sharing DemystData's story at AI Experience '18 in Boston

The reason we can focus all our time on the data is because we have DataRobot as our modeling superpower.

— Jason Mintz
VP of Product at DemystData

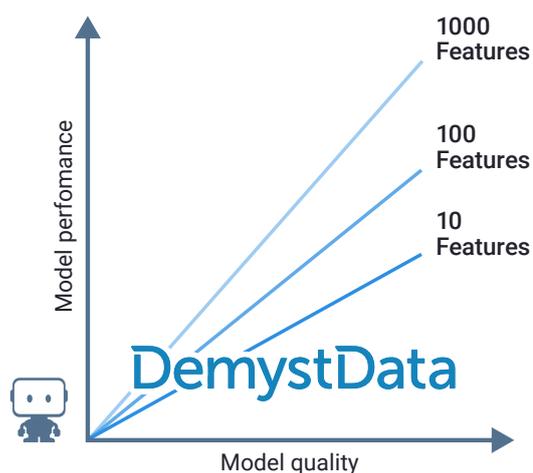


- **Inductive problem solving** meanwhile takes a bottom-up approach, one that focuses more on pattern recognition. Instead of starting with your hypothesis or a general idea you're trying to solve for, you start with the data and look for answers within the data, to see what pops.

"DataRobot allowed us to take an inductive approach to problem solving; a throw-spaghetti-at-the-wall approach to see what sticks," said Jason. "We could take a step back and from a blank slate, find all the data we could and throw it into DataRobot and see what pops, what's correlated, what's truly predictive."

DemystData works with their clients to take them through the entire machine learning lifecycle — from data discovery, through initial modeling, optimization, tuning and fine-tuning some more, finding value in the data, and productionalizing solutions — taking those data and models and actually implementing it so that it solves business problems. Both the data and the models are equally important to the success of the final predictive solution, something Jason compared to a racecar.

"You can build the best racecar possible — optimize the tires for road conditions, train a great driver, make it as aerodynamic as possible, and all those things will make your car go faster, but if you don't put in high-quality gasoline, it won't perform as well," said Jason. "The car is like our models, and the gasoline is our data. We need to put good data into our modeling systems if we're going to get high performance out of them."



With DataRobot, instead of spending all our time building models, iterating through five or six combinations, we now had this awesome tool that, in minutes, could iterate through 100s of models far better and faster than we ever could.

— Jason Mintz

From Jason's perspective, there are two ways of improving your predictive models. The first is to improve your model quality: with better machine learning techniques or more skilled data scientists working on your models. DataRobot's automated machine learning platform — with its ability to generate dozens of algorithmic-agnostic models in minutes — was the answer to improving DemystData's model quality.

The second way is to add more features, more signals, and more data. You don't want to just pour in meaningless data and drown out any signal, but in general the more data you have, the more you can learn from it and pick out the value in it. DemystData focuses mostly on the data piece, helping clients find more data and identify more signal, and DataRobot plays a huge role in allowing DemystData to focus on their core competency.

"The reason we can focus all our time on the data is because we have DataRobot as our modeling superpower," said Jason. "Before DataRobot, all our data scientists needed to spend more of their time cleaning the same set of variables and trying and testing different modeling combinations than actually building models.



With DataRobot, instead of spending all our time building models, iterating through five or six combinations, we now had this awesome tool that, in minutes, could iterate through 100s of models far better and faster than we ever could.”

“Life before DataRobot was long, slow, and painful. DataRobot has really been able to revolutionize how we approach problem solving, both internally and with our clients.”

Democratizing Data Science Across the Organization

By automating many of the previously manual and time-consuming steps of the machine learning lifecycle, DataRobot was able to help DemystData improve not only the quality of their models, but also their overall data science productivity. But, according to Jason, the most significant impact that DataRobot had on DemystData was in democratizing data science across the entire organization.



What DataRobot really did was open up data science to everybody at Demyst. We now have salespeople competing in Kaggle competitions to get a better feel for the software. They were able to speak the language, and upload the dataset, press one button, and get a result. That really got them excited about data science.

— Jason Mintz

“Our incumbent data scientists had a fear that DataRobot coming in would take away their jobs,” said Jason. “Instead, they became the biggest proponents of DataRobot because it allowed them to do work more efficiently on boring stuff and spend more of their time on the more interesting parts of data science.”

DemystData used to struggle to get by with 2 or 3 data scientists, working at one tenth the efficiency and 10 times the cost of DataRobot. By using DataRobot to automate and streamline the more basic and mundane parts of data science, those data scientists saw their overall productivity skyrocket.

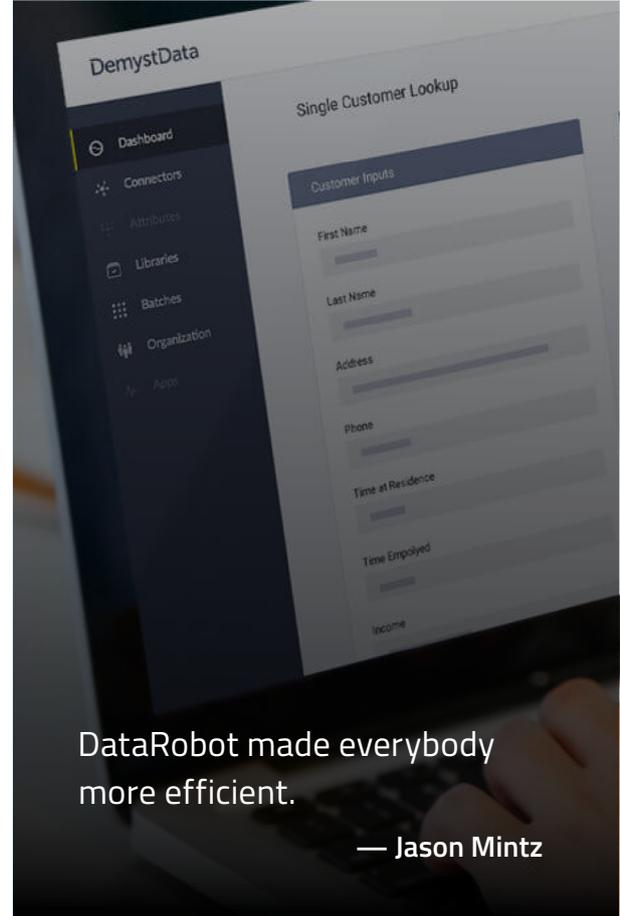
But DataRobot didn’t just impact the data scientists at DemystData; because of its simplicity and ease-of-use, even DemystData workers who didn’t have a background in mathematics or data science were now able to contribute to machine learning projects. According to Jason, DataRobot played a significant role in bringing the whole organization together and getting non-technical employees informed and aligned with the company’s data science mission and technology.

“What DataRobot really did was open up data science to everybody at Demyst”, said Jason. “We now have salespeople competing in Kaggle competitions to get a better feel for the software. They were able to speak the language and upload the dataset, press one button, and get a result. That really got them excited about data science.”

DataRobot also fundamentally changed how DemystData approached hiring and growing their team. Previously, when Jason and his team were looking to build out their data science team, they were trying to hire generalist data scientists, who can do a little bit of everything across all data science functions. With DataRobot delivering a baseline and wide range of data science skills on its platform to users at DemystData, Jason could now focus on hiring for very specific capabilities. While the challenges of hiring for this scarce resource still exist, Jason is optimistic that DataRobot is addressing that gulf and altering the data science landscape.

“DataRobot made everybody more efficient,” said Jason. “We were able to do a lot of the work with fewer people, and we were able to grow the team based on different skill sets to solve the problems that our customers have, rather than needing to fill this base need [for machine learning] that DataRobot was able to fill for us.”

“I think it’s a really exciting time to be in data science”, continued Jason. “It’s being applied to more industries and more companies. There’s a really long tail of companies and opportunities where data science has traditionally not worked. And what Datarobot and automated machine learning enables is for those companies to start using data science and get the same benefits that more tech-focused companies had to themselves for years.”



DataRobot made everybody more efficient.

— Jason Mintz

Contact Us

DataRobot
One International Place, 5th Floor
Boston, MA 02110
USA

www.datarobot.com

info@datarobot.com



© 2019 DataRobot, Inc. All rights reserved. DataRobot and the DataRobot logo are trademarks of DataRobot, Inc. All other marks are trademarks or registered trademarks of their respective holders.