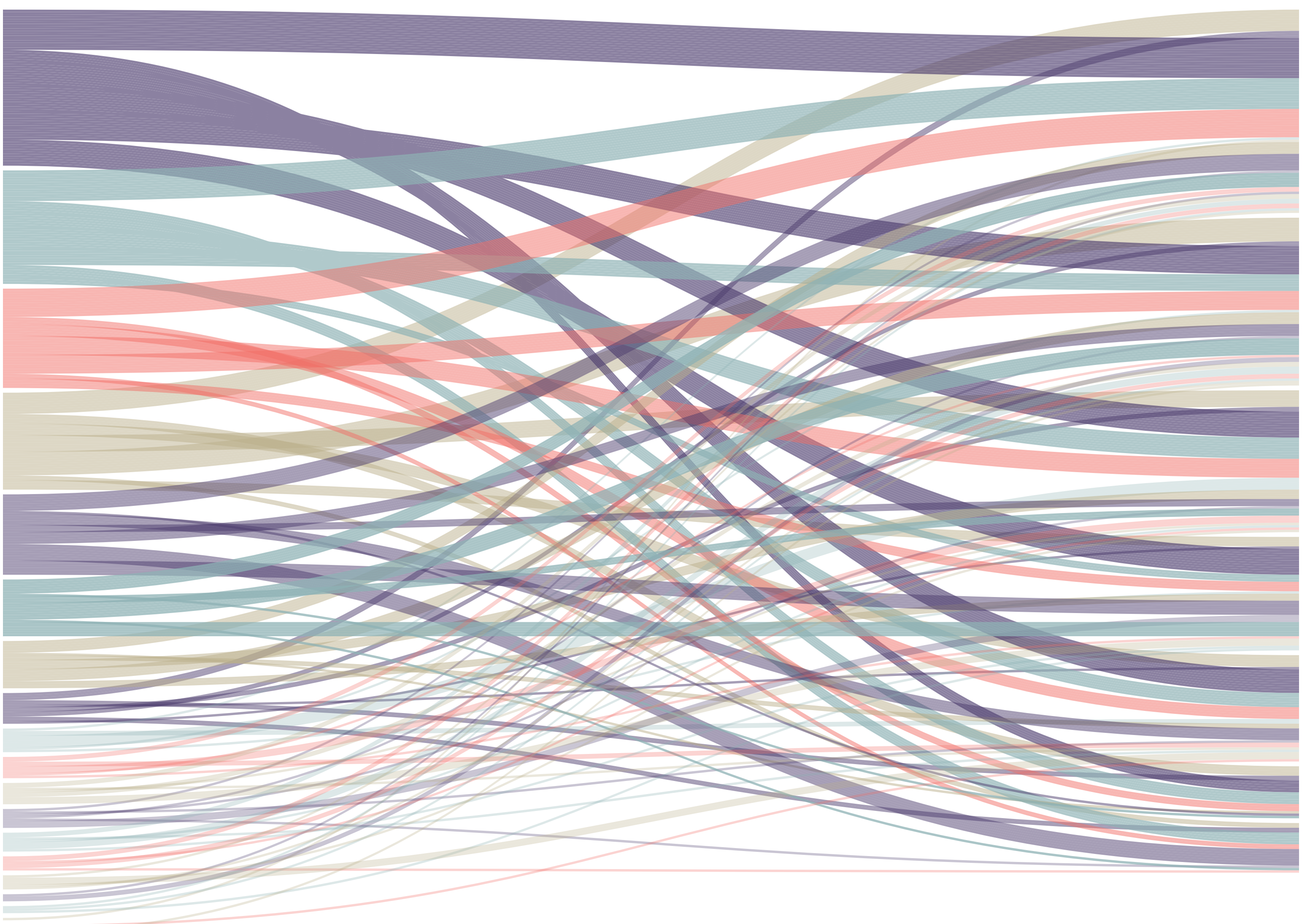
research

ALGORITHMIA

2020 state of enterprise machine learning



Contents

Introduction

In the last 12 months, there have been myriad developments in machine learning (ML) tools and applications, and hardware for AI and ML is also progressing. Google's TPUs are in their third generation, the AWS Inferentia chip is a year old, Intel's Nervana Neural Network Processors are enabling deep learning, and Microsoft is reportedly developing its own custom AI hardware.

This year, Algorithmia has had conversations with thousands of companies in various stages of machine learning maturity. From them we developed hypotheses about the state of machine learning in the enterprise, and in October, we decided to test those hypotheses. Building on the State of Enterprise Machine Learning report we published in 2018, we conducted a new two-prong survey this year, polling nearly 750 business decision makers across all industries from companies actively developing machine learning lifecycles or just beginning their machine learning journey.

One set of respondents was administered a blind version of our survey by a third-party (we refer to this group in the report as Group A); the other set was sent a survey by Algorithmia and was aware of the author (referred to herein as Group B). Group A contained 303 respondents and Group B contained 442.

We analyzed the responses from both groups for insight into their work, their companies' machine learning roadmaps, and the changes they've seen in recent months with regard to ML development. Where applicable, we state when only one group is being cited in a given statistic. The Methodology section provides further detail on the specifics of the survey prongs and how we processed the data.

The following are the findings of that effort, presented with our original hypotheses, as well as our analysis of the results. Where possible, we have provided a year-on-year comparison with data from 2018 and included predictions about what is likely to manifest in the ML space in the near term. We will soon make our survey data available on an interactive webpage to foster greater understanding of the ML landscape, and we are committed to being good stewards of this technology.

Algorithmia seeks to empower every organization to achieve its full potential through the use of artificial intelligence and machine learning by delivering the last-mile solution for model deployment at scale.

Survey at a glance

The main takeaway from the 2020 State of Enterprise Machine Learning survey is that a growing number of companies are entering the early stages of ML development, but challenges in deployment, scaling, versioning, and other sophistication efforts still hinder teams from extracting value from their ML investments. As a result, we will likely see a boom in the number of ML companies providing services to overcome these obstacles in the near term.

In this report, we focus on seven key survey findings and what they say about the machine learning landscape. Those key findings are as follows:

1. The number of data scientist roles at companies is often less than 10, but is growing rapidly across all industries.
2. Business use cases for machine learning are becoming more varied but currently, customer-centric applications are the most common.
3. Machine learning operationalization (having a deployed ML lifecycle) is fledgling but maturing across all industries with software and IT firms leading the charge.
4. The main challenges people face when developing ML capabilities are scale, version control, model reproducibility, and aligning stakeholders.
5. The time it takes to deploy a model is stuck somewhere between 31 and 90 days for most companies.
6. Budgets for ML programs are growing most often by 25 percent, and the banking, manufacturing, and IT industries have seen the largest budget growth this year.
7. Organizations are determining ML success by both business unit and statistical metrics with a significant divide by job level.

The report will go into each finding in detail and provide analysis and our outlook.

Key finding 1: The rise of the data science arsenal for machine learning

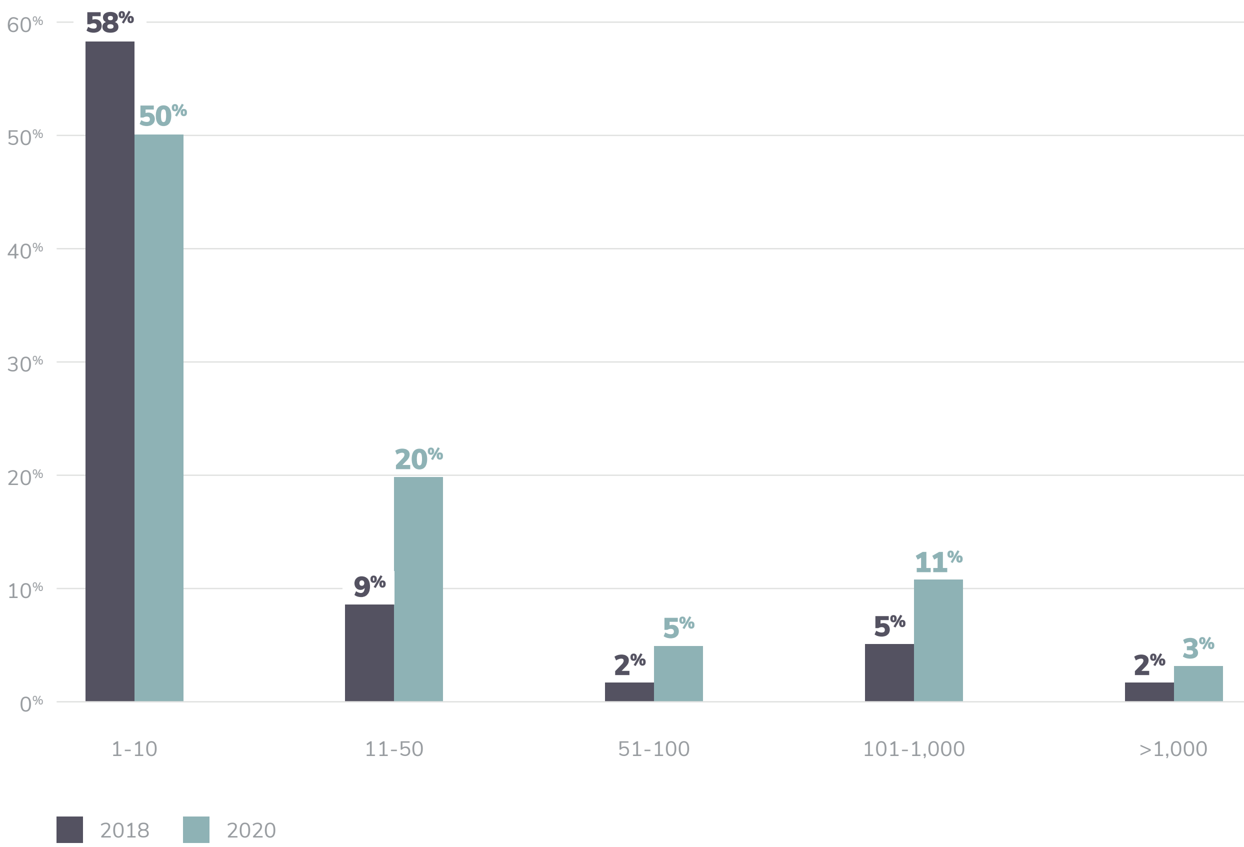
One of the pieces of data we collected this year was the number of data scientists employed at the respondent's place of work. In conversations we regularly have with companies, we repeatedly hear that management is prioritizing hiring for the data science role above many others, including engineering, IT, and software. Here is what the survey results showed.

Half of people polled (across both survey groups) said their companies employ between one and 10 data scientists. This is actually down from 2018 wherein 58 percent of companies claimed to employ between one and 10 data scientists.

We would have expected the number to increase over time because investment in AI and ML is known to be growing (Gartner). When assessed in the context of the full data, however, a likely reason for the downward trend presents itself. In 2018, 18 percent of companies employed 11 or more data scientists. This year, however, that number soared to 39 percent, suggesting that across all industries, organizations are ramping up their hiring efforts to build larger data science arsenals, with some of them starting from close to 10 data scientists already.

Another observation is that in 2018, barely 2 percent of companies had more than 1,000 data scientists; today that number is just over 3 percent, indicating small but significant growth. These companies include the big FAANG tech giants—Facebook, Apple, Amazon, Netflix, and Google (Yahoo); their large data science teams are working to maintain competitiveness as more third-party solutions crop up.

Data scientists employed, a year-on-year comparison



Reflects data from both survey groups. Note that respondents who did not know or were unsure are not depicted in this graph.

Demand for data scientists

In 2016, Deloitte predicted a shortage of 180,000 data scientists by 2018, and between 2012 and 2017, the number of data scientist jobs on LinkedIn increased by more than 650 percent (KDnuggets). The talent deficit and high demand means that hiring and maintaining data science roles will only become more difficult for small and mid-sized companies that cannot offer the same salary and benefits packages as the FAANG companies.

As demand for data scientists grows, we may see a trend of junior-level hires having less opportunity to structure data science and machine learning efforts within their teams, as much of the structuring and program scoping may have already been done by predecessors who overcame the initial hurdles. It could also mean, however, that leadership alignment has already been attained so ML teams will have more ownership and leeway in project execution.

New roles, the same data science

Finally, we may also see the merging of traditional business intelligence and data science in order to fill immediate requirements in the latter talent pool since both domains use data modeling (BI work uses statistical methods to analyze past performance, and data science makes predictions about future events or performance).

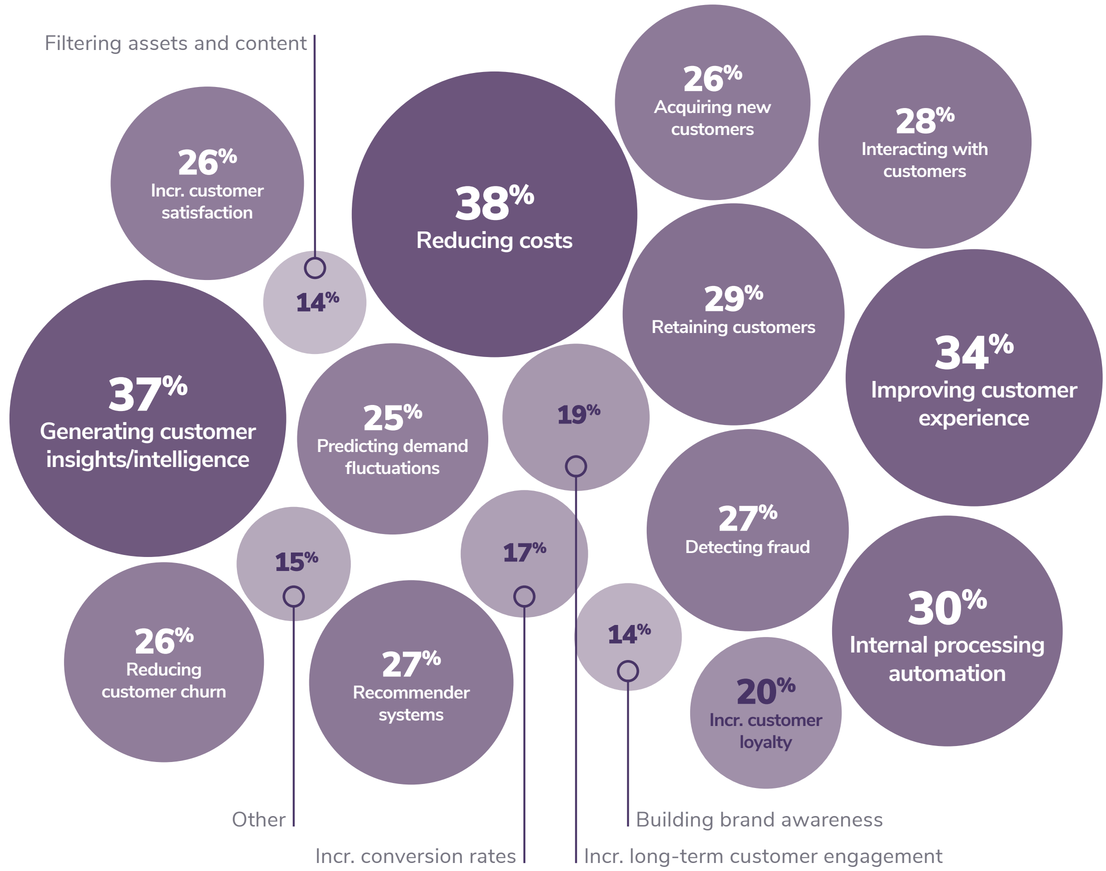
Gartner predicts that the overall lack of data science resources will result in an increasing number of developers becoming involved in creating and managing machine learning models (Gartner CIO survey). This blending of roles, will likely lead to another phenomenon related to this finding: more role names and job titles for the same sorts of work. To that end, we are seeing an influx of new job titles in data science such as Machine Learning Engineer, ML Developer, ML Architect, Data Engineer, Machine Learning Operations (ML Ops), and AI Ops as the industry expands and companies attempt to distinguish themselves and their talent from the pack.

Key finding 2: Cutting costs takes center stage as companies grow

As a company, we are interested in machine learning applications in the enterprise and we strive to keep a pulse on how industries are using emerging ML tech to automate workflows. There are countless ways to apply ML to a particular business problem, such as using prediction modeling to make assessments about customer churn or applying natural language processing to millions of tweets to analyze the percentage of negative sentiments.

In this year’s survey, we polled respondents about the ways their companies are using machine learning to ensure our understanding of the landscape is accurate or that we aren’t missing a key use case entering the enterprise. We anticipated a trend toward using ML to automate time-consuming processes and cut down on the number of human resources needed to do a given task. The results are depicted below.

Machine learning use case frequency



Reflects data only from survey Group B. Note that respondents were allowed to choose more than one answer.

In this year's survey, we provided a wide-ranging list of possible use cases and a write-in option. Respondents were encouraged to select all answers that applied to how their companies use AI and ML models today. The top three machine learning use cases across the board (for companies of all sizes) were as follows:

1. Reducing company costs
2. Generating customer insights and intelligence
3. Improving customer experience

When we break down the data by company size, we start to see some differentiation in priorities.

The top five ML use cases for companies with 10,000 employees or more:

1. Reducing company costs
2. Process automation for internal organization
3. Improving customer experience
4. Generating customer insights and intelligence
5. Detecting fraud

The top five ML uses cases for companies with 1,001-5,000 employees:

1. Reducing company costs
2. Retaining customers
3. Process automation for internal organization
4. Recommender systems
5. Increasing customer satisfaction

The top five ML use cases for companies with fewer than 100 employees:

1. Generating customer insights and intelligence
2. Improving customer experience
3. Reducing company costs
4. Increasing customer satisfaction
5. Retaining customers

Smaller companies focus on customers

The survey data showed that large companies are using ML primarily for internal applications (reducing company spend and automating internal processes), and smaller companies are primarily focused on customer-centric functions (increasing customer satisfaction, improving customer experience, and gathering insights). This suggests that as companies grow, they prioritize customer service less than cost-saving measures and applications that improve their product lines. Doing so comes at a price, however, as one-third of Americans consider switching companies after just one instance of poor customer service (Qualtrics). Conversely, an increase in customer retention rate of just 5 percent can produce more than a 25-percent increase in profits (Bain & Company).

Fortunately, machine learning is a solution for both types of business problems—cutting costs and customer satisfaction—and will likely shift business priorities in the near term as workflows are drastically augmented by new tech. For comparison, in our 2018 survey, 48 percent of respondents from companies with 10,000 or more employees said cost savings was a major ML priority, and 59 percent said increasing customer loyalty was the top ML use case, depicting a notable shift away from customers this year. It will be important to monitor this metric in future years to see if this is the beginning of a trend or an anomaly.

Before conducting this year's survey, we anticipated a more even spread of use cases across companies of all sizes independent of industry because of the number of companies and applications in development in the AI/ML space (Forbes). The percentages for cost reduction, robotic process automation, and customer service applications may be an indicator of ML's general newness and immaturity, which our next key finding discusses, or it may be demonstrative of the fact that those types of repetitive applications lend themselves more readily to automation. As machine learning becomes more sophisticated with time, we are likely to see a wider pool of use cases designed for specific organizational initiatives.

Breakdown of use cases by industry

Understandably, industries with customer-facing products or services (retail, manufacturing, healthcare, etc.) prioritize ML applications that improve customer service, and industries involved with security, compliance laws, and proprietary data (financial institutions, government agencies, insurers, etc.) focus more so on ML use cases that help solve those challenges. The following are a few noteworthy examples:

- Respondents in both survey groups who work in consulting and professional services industries said that reducing customer churn was their top ML priority.
- The education/edtech sector's top ML use case was interacting with customers, which is reasonable considering that students and instructors are likely a primary customer set in those industries.
- For the healthcare, pharmaceutical, and biotech industries, increasing customer satisfaction was the leading use case, suggesting that customer dissatisfaction or churn may be a continual challenge in those fields.
- IT companies use ML primarily to acquire new customers, and software development organizations prioritize ML recommender systems to guide users toward viewing new products or features to buy.
- Banks and financial services firms are focusing their ML efforts on retaining customers and detecting fraud—keeping customers happy and mitigating vulnerabilities to the company.
- Finally, the energy sector, including utility companies, are focusing on forecasting demand fluctuations using ML, likely to prevent power outages, reduce response times during disruptions of service, and plan for power consumption for coming years (NeuralDesigner).

Key finding 3: Overcrowding at early maturity levels and AI for AI's sake

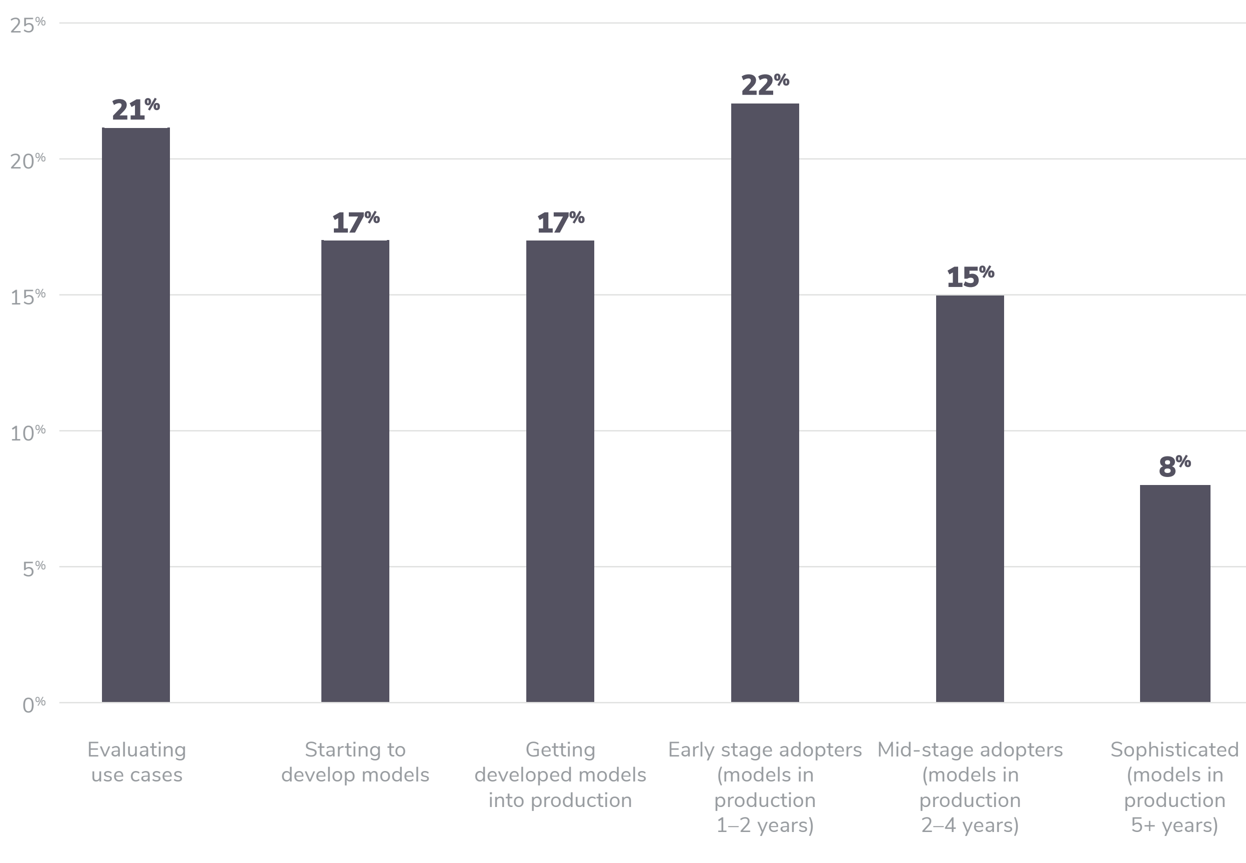
Understanding how companies view their own machine learning maturity provides insight into future developments in the ML space. For this survey, we asked respondents to gauge where they think their companies are located currently on the machine learning roadmap. That is to say, we sought to determine if they are just starting to consider machine learning applications for business problems or if they are operating a fully developed machine learning program, or somewhere in the middle of that spectrum, and whether their positioning has changed in the previous 12 months.

In 2018's survey report, nearly 40 percent of respondents said they were just beginning to develop ML plans (ie. evaluating use cases, starting to build models). Moreover, in 2018 fewer than 10 percent of respondents considered themselves at a sophisticated ML maturity level.

This year, we asked respondents to select one of the following options to gauge ML maturity levels:

- Not actively considering ML as a business solution
- Evaluating ML use cases
- Just starting to develop/build models
- Developed models; working toward production
- Early stage adoption (models in production for 1-2 years)
- Mid-stage adoption (models in production for 2-4 years)
- Sophisticated (models in production for 5+ years)

2020 machine learning maturity levels



Reflects data from both survey groups.

55%

of companies surveyed have not deployed a machine learning model

Of the respondents who are actively engaging in ML (removing the first category of those who are not evaluating ML as a business solution), about one-fifth said they are evaluating use cases, based on an average of both survey groups. Those just starting to develop and build models numbered 17 percent, and a separate 17 percent of companies have developed models but are still working toward production. This means that 55 percent of companies surveyed have not deployed a machine learning model.

ML in early stages of development

The number of companies with undeployed models is up 4 percent from last year, likely because there are more companies across the board beginning ML journeys, inflating the category of newcomers. It is important to note as well that our survey sample increased by more than 200 people from last year.

9%

more companies have gotten models into production since 2018

Just over 22 percent of companies have had models in production for 1-2 years; last year, 13 percent of respondents claimed this, demonstrating a fairly significant migration toward productionization even if it is still early days for most companies

Moreover, one-fifth of companies said they plan on getting models into production within the next year, suggesting that we may see a noticeable portion of companies moving into the next maturity category (mid-stage) in the near term.

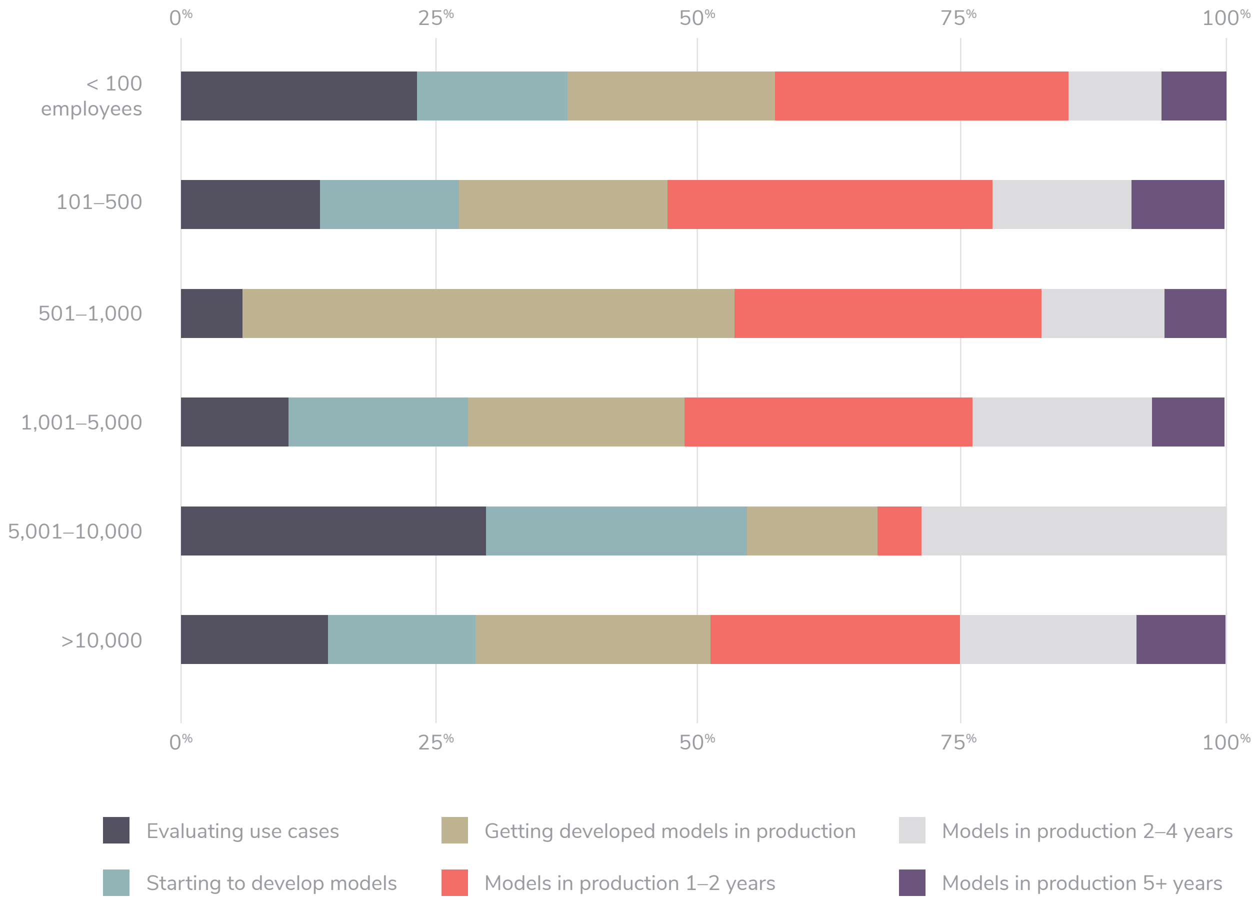
Year and company size comparison

In 2018, only 6 percent of respondents considered their companies to have sophisticated ML programs. This year, 8 percent do, and the majority of companies in the sophisticated category either have fewer than 500 employees or more than 10,000. In 2018, 39 percent of sophisticated companies had fewer than 100 employees and 29 percent had more than 10,000 employees.

There are several ways to read this maturity breakdown. First, large companies typically have more budget for innovation hubs and emerging technology, thus streamlining the development of sophisticated ML initiatives. Smaller companies, however, can be quite agile technologically, able to build, buy, and iterate quickly.

They can also be highly motivated to build reputation, profit, brand loyalty, and a competitive edge right out of the gate—machine learning can be an effective and efficient tool to reach all those goals. That the largest and smallest companies are leading in ML maturity is significant and may speak to and encourage a more equal tech landscape wherein the largest tech voices are not the only voices at play.

Machine learning maturity and company size



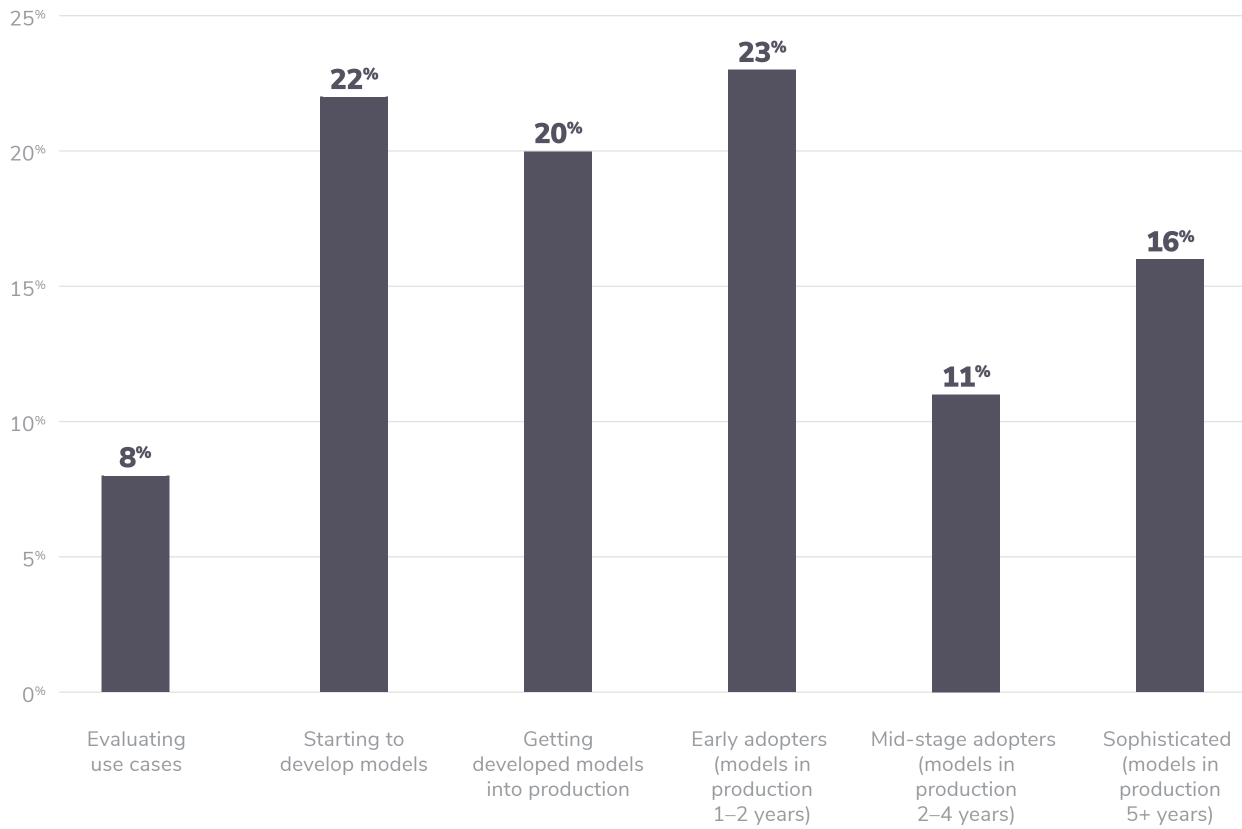
Reflects data only from survey Group B.

Mid-sized companies span all maturity levels with the highest concentration in the early-to-mid-stage levels of maturity, suggesting that they may have a bit of both worlds—the agility of smaller companies to tackle new projects quickly and growing budgets dedicated to emerging tech ([Digitalist Magazine](#)).

Gauging maturity in the year ahead

In the next 12 months, we expect the number of companies in the earliest machine learning stages (evaluating use cases and starting to develop models) to expand and then decline as ML becomes ever more ubiquitous in the enterprise. Eventually the early stages will decline as companies proceed through the machine learning lifecycle.

Anticipated maturity stage in the next 12 months



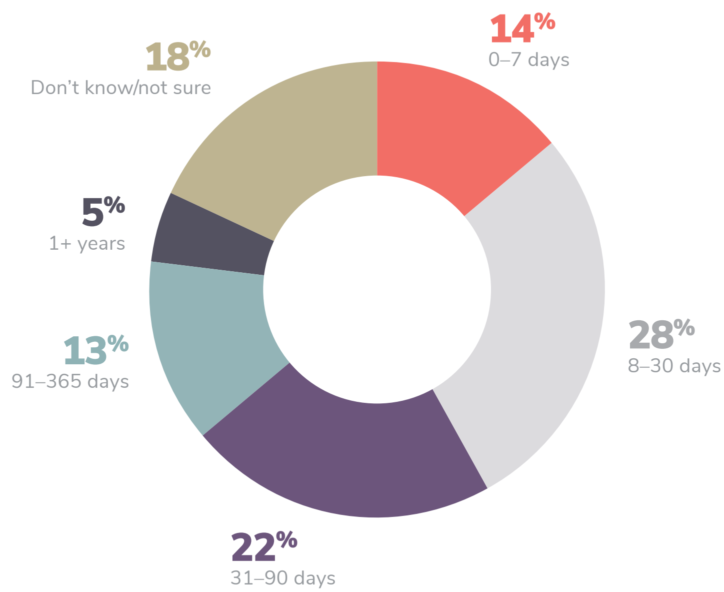
Reflects data only from survey Group A.

The bottom line is that we do see a shift toward greater ML maturity in all companies surveyed, however, those in mid-to-late stages of maturity are still quite low in number. We expect to see that group grow over the course of the next 12 months as companies overcome last-mile ML problems and align stakeholders toward building sophistication into their ML programs.

Key finding 4: An unreasonably long road to deployment

A new metric we are beginning to track this year is the time it takes an organization to deploy a single ML model. Of companies surveyed, just about half say they spend between 8 and 90 days deploying one model. And 18 percent of companies are taking longer than 90 days—some spending more than a year productionizing!

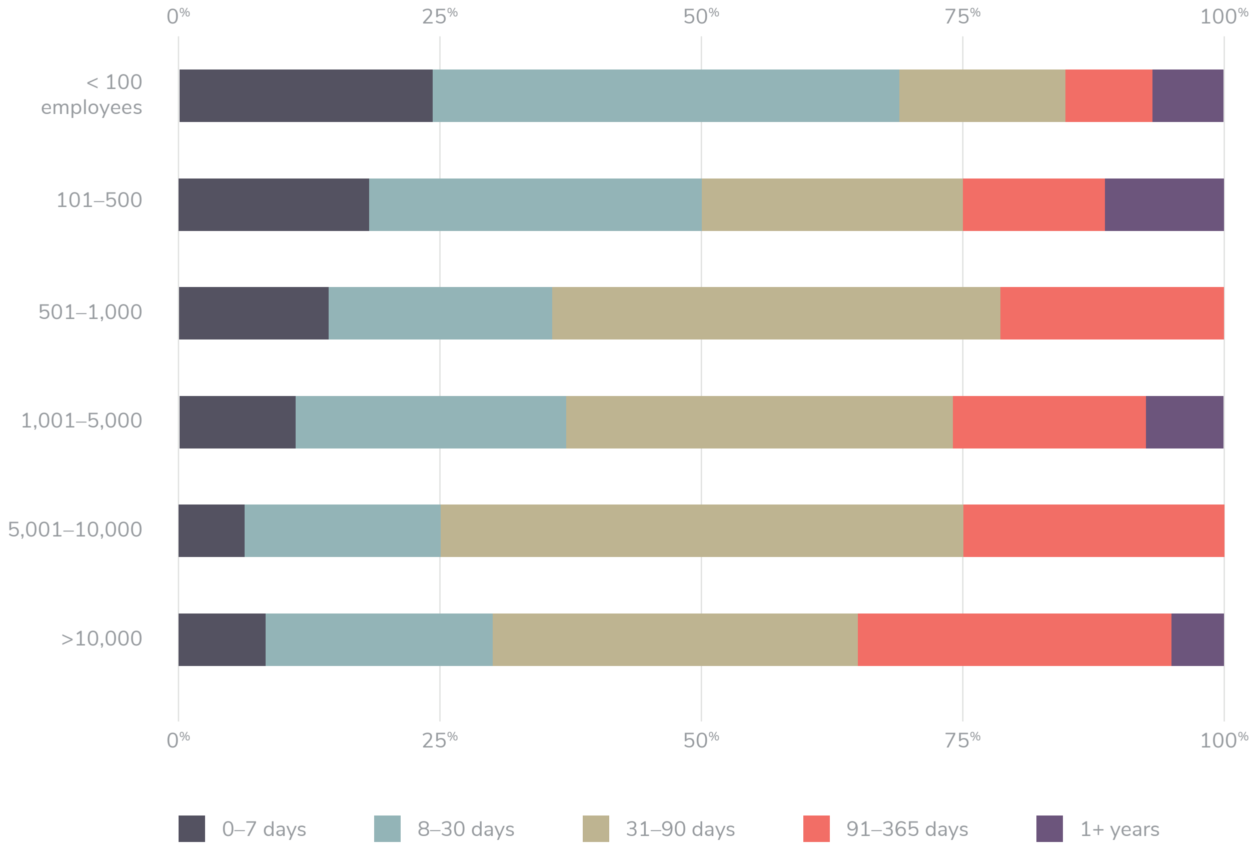
Model deployment timeline



Reflects data only from survey Group B.

We thoroughly understand that there are many challenges to overcome when building a robust ML lifecycle, deployment being a large one (Medium). That being said, we would still have expected the percentage of companies who deploy models in less than a week to be significantly larger than 14 percent, based on the number of companies in the early stage maturity level (models in production for 1-2 years). Company size and maturity level provide some context to explain this relatively low number.

Model deployment timeline and company size

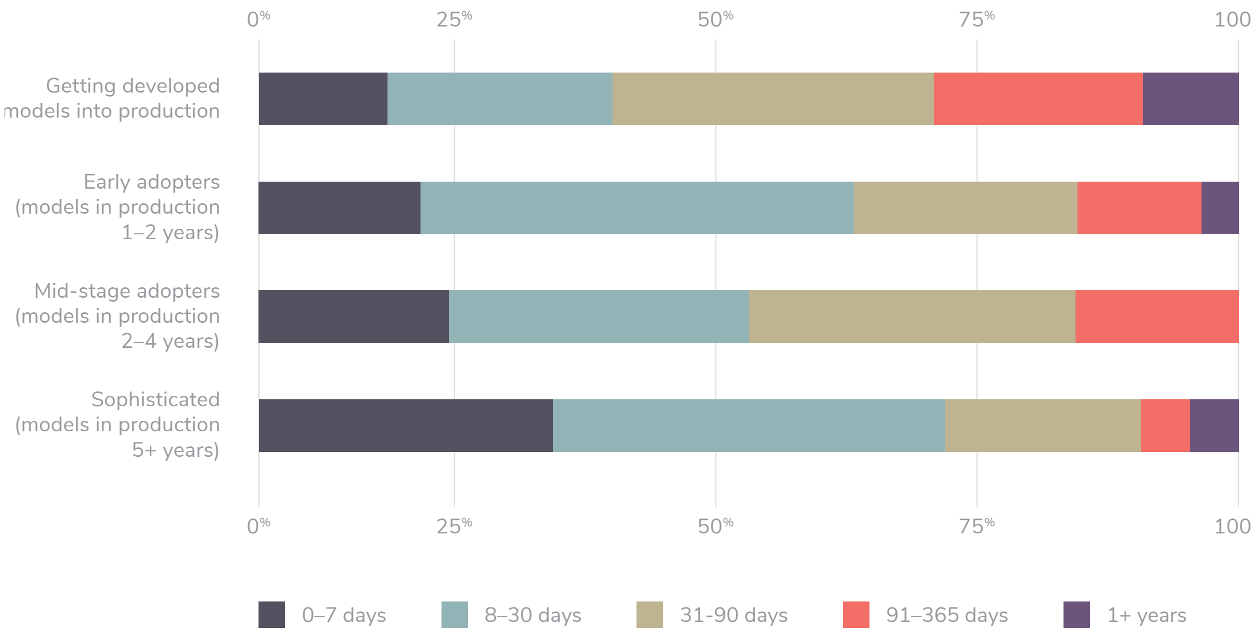


Reflects data only from survey Group B.

Companies of all sizes typically spend between 8 and 90 days deploying one model, with a few notable exceptions. A small (and we expect decreasing) number of companies is spending more than a year deploying models, and of those, mostly small-to-midsize organizations. Moreover, a fairly significant portion of companies with 100 employees or fewer is spending somewhere between 8 and 30 days deploying a single model.

Moreover, there is a slight decrease in the ideal 0-7 day range as company size increases, and on the other side, there is a somewhat uniform indication that the larger the company, the more likely it will spend between 4 and 12 months deploying a model. We assess, however, that the time to deployment phenomenon is less dependent on company size alone and more so on maturity level.

Model deployment timeline and ML maturity



Reflects data only from survey Group B. Note that the earliest two maturity levels (evaluating use cases and just starting to develop models) are not depicted here as those respondents have not deployed a model.

Most noticeable—and understandable—is the increase in the 0-7 day range as companies' machine learning programs mature. It follows that the more sophisticated a company's ML efforts are, the more likely it is to deploy a model quickly.

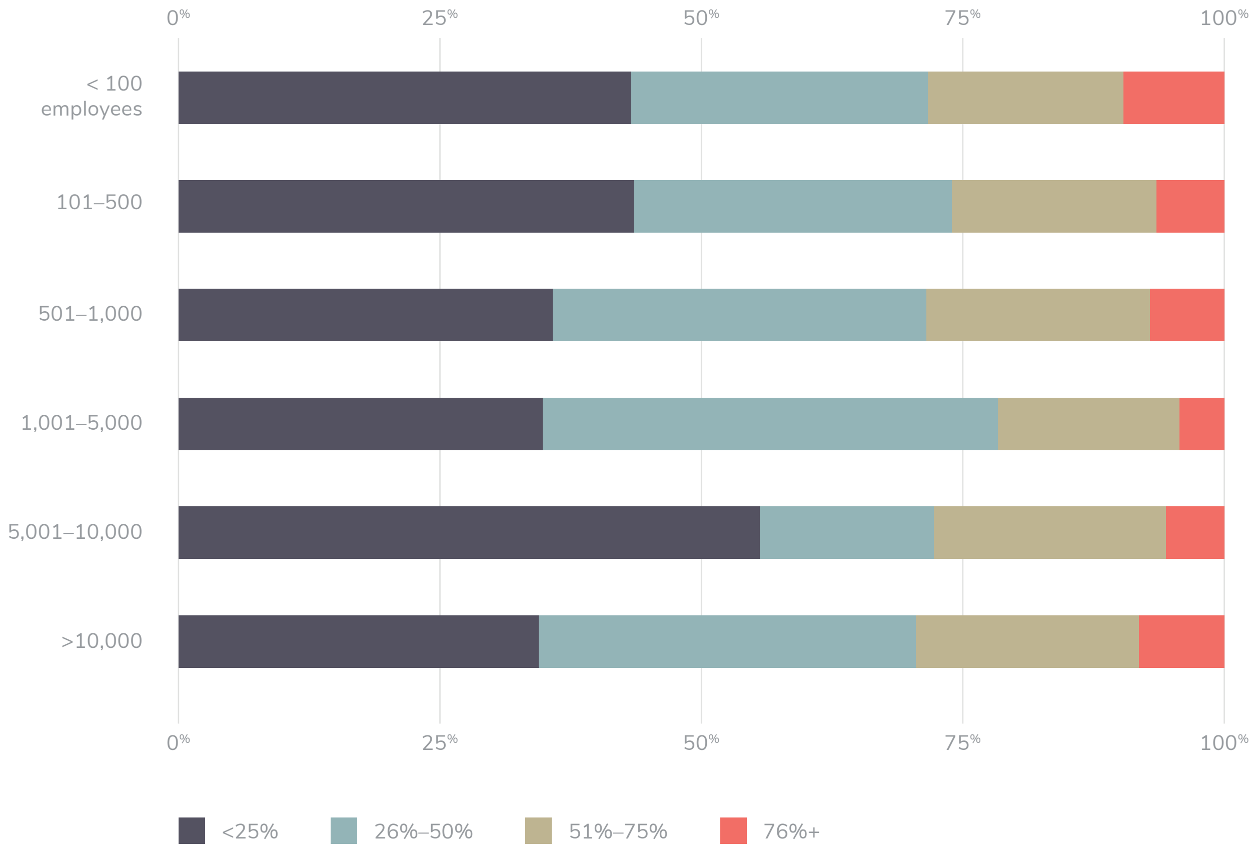
Also noteworthy is that the more sophisticated a company becomes, the more time it spends in the 8-30 day range for model deployment. Our best guess as to why is discussed in Key finding 5, the challenges associated with machine learning. In short, struggles with scale and aligning all stakeholders can add to timelines.

Data science workload and the last mile to deployment

When we look at the actual time spent deploying models, we see that at companies of all sizes, at least 25 percent of data scientist time is spent on deployment efforts. Put simply, a quarter of data science capability is lost to infrastructure tasks. In 2018, closer to 70 percent of data science capability was spent lost to deploying models, which at face value, appears to imply drastic improvement. However, the data cannot tell us definitely why this large decrease occurred. Ideally, it's due to data scientists having the tools they need to deploy with ease, but based on the low number of companies in the deployed category, we are not confident in that assessment. It is more likely that data science teams are handing off more of their models to a DevOps or IT team to deploy, if it's happening at all.

"I've heard many variants of this story: they all capture a misaligned pace of work between product and machine learning teams. Ultimately, this leads to machine learning research never making it out of the lab. And yet, the best measure of *impact* for machine learning, if you work in a non-research institution, is whether you can use it to help your customers—and that means getting it out of the door" (Medium).

Time data scientists spend deploying models by company size



Reflects data only from survey Group B.

Data science teams need to be able to deploy their work as quickly as possible to prevent their insights from being overcome by events (OBE); models and data change quickly as do market opportunities. As such, an insight that comes 10 days too late is OBE and no longer useful. To that end, much of the potential of ML may yet to be seen. "This is why AI has yet to reshape most businesses: For many companies, deploying AI is slower and more expensive than it might seem" (MIT Technology Review).

In November 2019, Gartner said that the “increased use of commercial AI and ML will help to accelerate the deployment of models in production, which will drive business value from these investments” (Gartner). It went on to assess that the majority of teams developing ML capabilities are doing so using open-source tooling because of the dearth of viable commercial options. We assess that that gap is soon to be filled with companies offering a full suite of ML tooling as companies seek to become more mature in their ML lifecycles and look for third-party solutions rather than spending valuable time building ML infrastructure.

It is worth sharing here a word of warning from Ryan Calo, an associate law professor at the University of Washington, who is also a co-founder of the Tech Policy Lab and a leading voice on law and emerging tech issues in the media. At a recent Seattle Times–sponsored panel discussion on AI and the future of work, Calo cautioned attendees about snake oil AI companies. He described a plausible future scenario in which countless third-party AI solutions flood the market, creating a cacophony of messaging about AI necessities.

The resulting confusion might allow AI firms to take advantage of non-technical customers who hope to stay competitive in their spaces. They may pay for services that are inappropriate or unnecessary for their business in order to mature their ML programs quickly. It is important at a time of rapid technological innovation, such as now, to tread intelligently and not fall victim to the “AI for AI’s sake” adage.

Key finding 5: Innovation hubs and the trouble with scale

A crucial component of realizing ML’s full potential is scale. Can it scale? Scaling models was the biggest overall challenge cited by respondents this year (43 percent). For comparison, that percentage is up 13 percent from last year. But multiple requirements factor into scaling—hardware, modularity, data sourcing, etc.—and optimizing for it can lead to cumbersome team cross-cutting (Arc.dev).

Of respondents from companies of more than 10,000 employees, 58 percent said scaling up was their top ML challenge. This may be demonstrative of decentralized organizational structures—data science teams siloed throughout company org charts—which can cause tooling, framework, and even programming language friction when scaled.

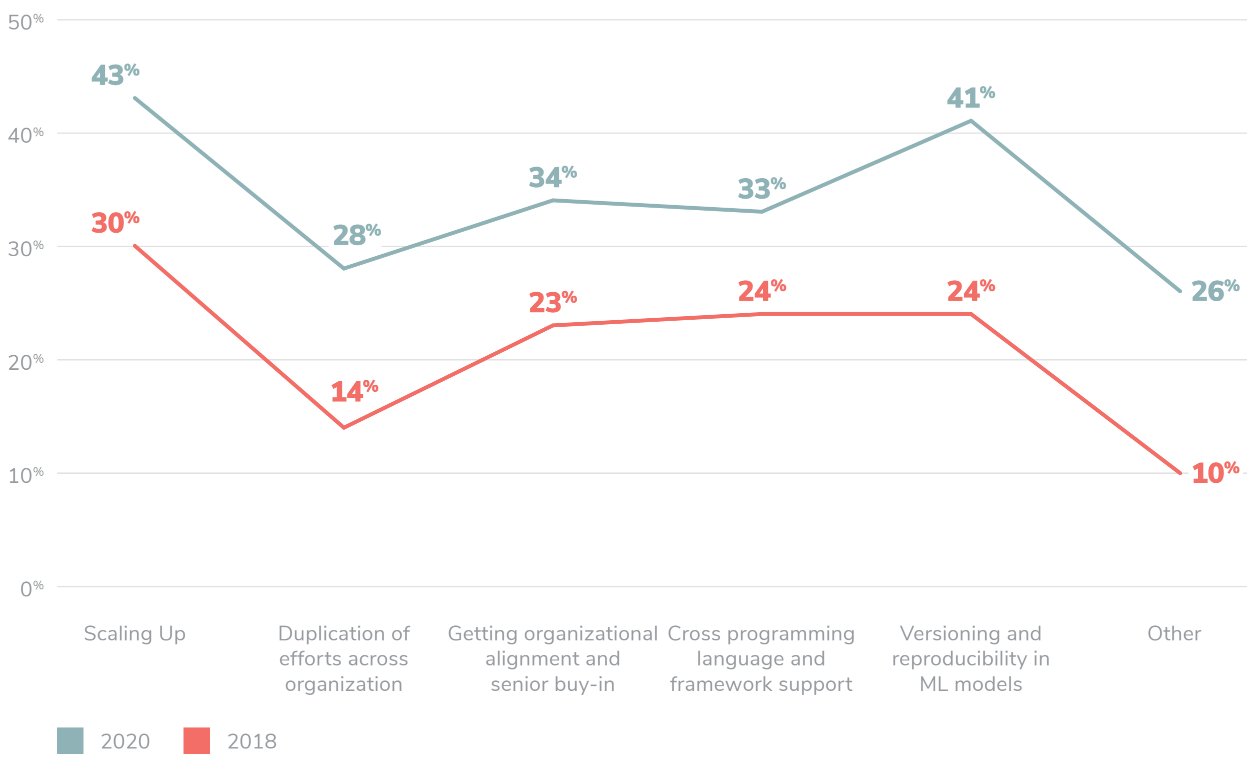
Earlier this year, Gartner predicted that “through 2020, 80 percent of AI projects will remain alchemy, run by wizards whose talents will not scale in the organization” (Gartner). This outlook may prove true, but, we are skeptical based on our observations of a continuous increase in centralized innovation hubs and emerging tech centers (see Ericsson, IBM, Pfizer, etc.). We assess these hubs will be more efficient at maturing ML for their companies than the decentralized alternative (ie. data science components siloed throughout organizations working on one-off projects and models). An innovation hub can iterate quickly, work with agility across an organization, and standardize ML efforts. They can often vet new technologies quickly, ensuring their companies keep at the bleeding edge of technological development. We anticipate this kind of centralized focus on ML and AI technologies may just turn lead into gold, so to speak.

Model reproducibility impedes ML maturity

The second most cited ML challenge was versioning and reproducibility of models (41 percent of respondents reported this). This number is much higher than the 24 percent of respondents who cited this challenge in 2018. Machine learning requires faster iteration than the traditional software development lifecycle, and ironclad version-control is paramount for pipelining, retraining, and evaluating models for accuracy, speed, and drift.

Versioning is one of the hurdles that data science and ML teams must overcome to reach more sophisticated levels of ML maturity, so in future surveys, we will be monitoring this metric closely. We expect the number of times this is cited as a challenge to decrease in the coming year.

Year comparison of machine learning challenges



Reflects data only from survey Group B. Note that respondents were allowed to select more than one challenge.

Organizational misalignment and ML progress

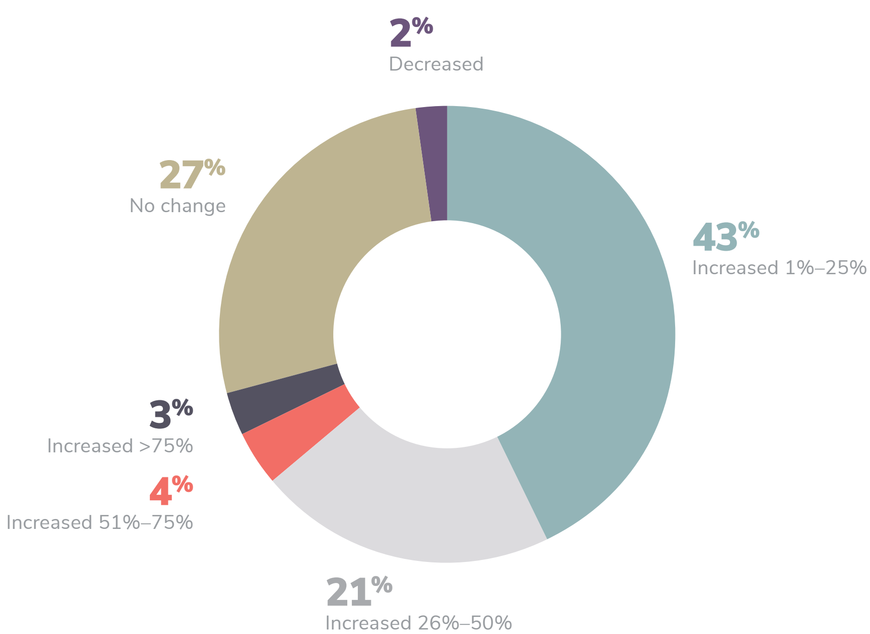
The third most cited ML challenge was getting organizational alignment and senior buy-in for ML initiatives (34 percent). Notably of the respondents who cited this challenge, 47 percent are from companies with more than 10,000 employees. Especially for decentralized organizations (no central innovation hub), trying to obtain multiple team and stakeholder concurrence may take a lot of time.

In 2018, 23 percent of respondents noted stakeholder alignment as a challenge. The 24-percent increase this year might contribute to other metrics, such as number of models deployed, time to deployment, and scaling. We expect this problem to decline in coming years as ML becomes more routine, reliable, and measurable.

Key finding 6: Budget and machine learning maturity, priorities and industry

This year's survey shows that ML budgets vary across industry and stage of maturity, but on the whole, are growing at companies of all sizes. This is in line with estimates that in 2018, the compound annual growth rate (CAGR) of AI was $23.94 billion and is expected to reach $208.49 billion by 2025 (MarketWatch).

AI/ML budgets FY18 to FY19



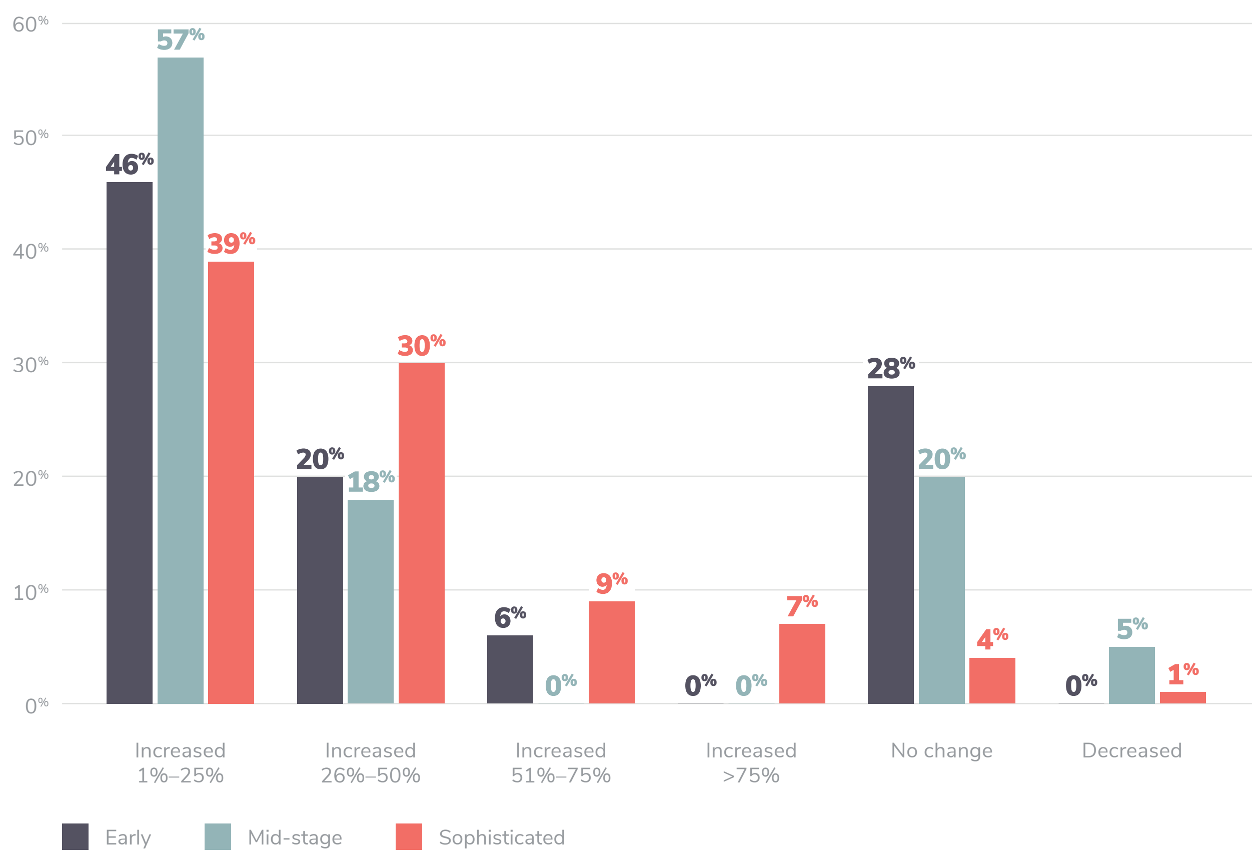
Reflects data only from survey Group A.

Twenty-one percent of respondents said budgets for AI/ML programs were growing between 26-50 percent. Forty-three percent of companies have increased their AI/ML budgets between 1 and 25 percent in the last year. And just under one-third (27 percent) of respondents noted that their budgets have not changed. This may be a reason why the majority of companies are still at early-stage maturity levels.

Budgets and ML maturity

Fifty-seven percent of companies in the mid-level maturity range (ML models in production for 2-4 years) increased their budgets between 1 and 25 percent. Close to 50 percent of companies at early stage maturity levels also increased their budgets between 1 and 25 percent. And nearly 40 percent of companies already at sophisticated ML maturity levels increased their budgets by as much as 25 percent. Finally, 30 percent of sophisticated maturity respondents said they increased their AI/ML budgets by 26-50 percent. We expect to see mid-stage and sophisticated companies increase their AI/ML budgets by more than a quarter in the very near term once ML has proven itself.

FY19 AI/ML budgets and ML maturity level

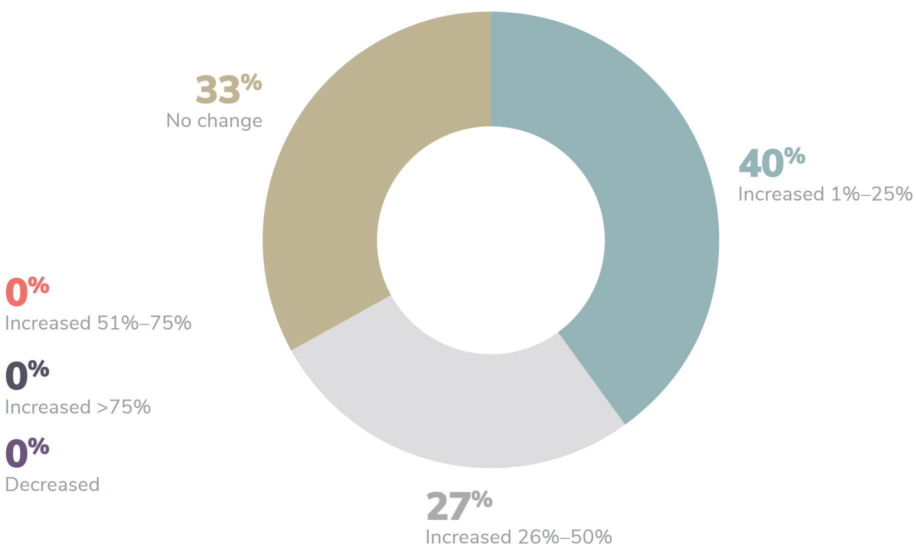


Reflects data only from survey Group A.

We assess that this upward budgetary trend is due to the fact that companies already at a mid or sophisticated level of ML maturity (models built and deployed for 2-5 years), are doubling down on their tech investment efforts. This means that companies in very early deployment stages (just starting to develop ML models) will have to triple their efforts to stay competitive in their industry. Now is the time to start planning a 2020 (and beyond) ML strategy.

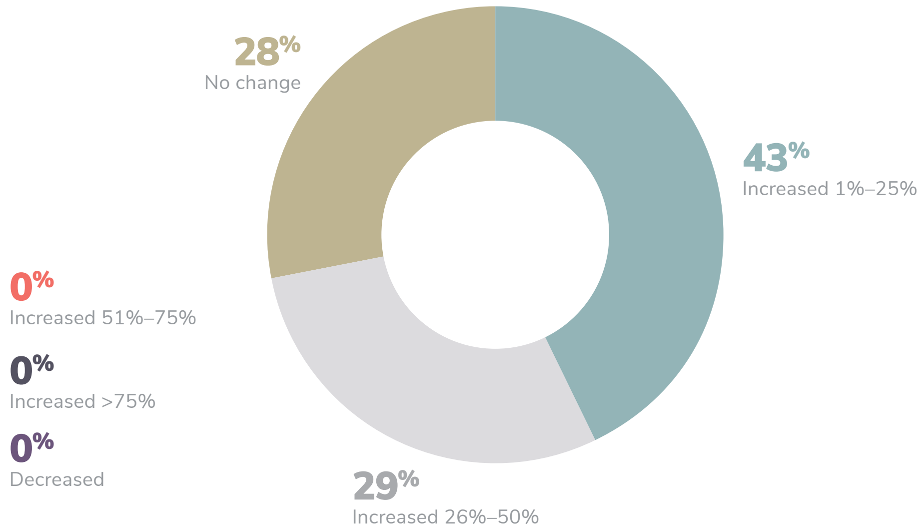
From an industry perspective, there was budgetary growth in specific industries, suggesting some jockeying for ML prowess in the nascent space.

AI/ML budgets for banking and financial services



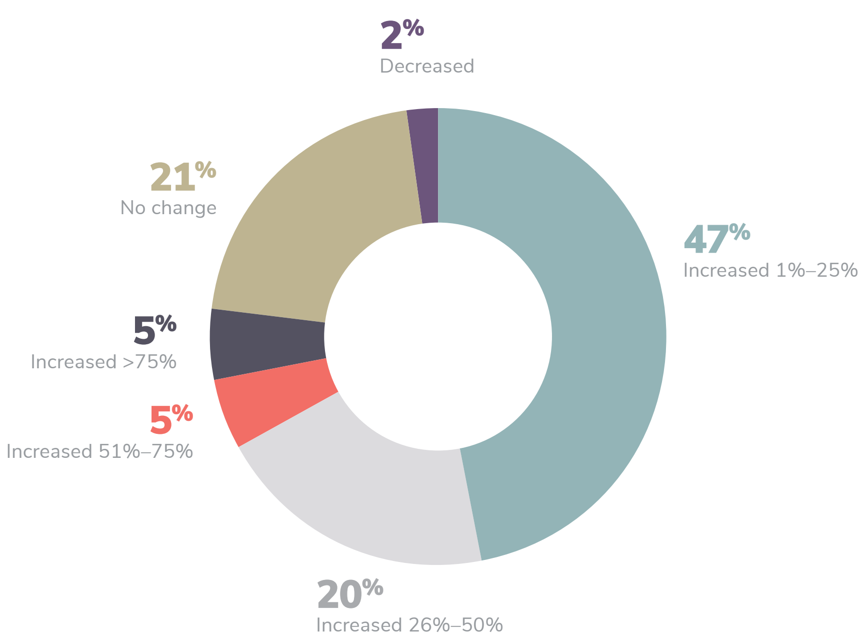
Reflects data only from survey Group A.

AI/ML budgets for manufacturing



Reflects data only from survey Group A.

AI/ML budgets for information technology



Reflects data only from survey Group A.

Key finding 7: Determining machine learning success across the org chart

While it is still early days for ML in the enterprise, our seventh key finding (how companies are determining what success means for their ML efforts) is likely an indicator of the path that machine learning will take as it develops throughout the enterprise (Emerj). Our hypothesis was that if ML success is primarily measured by dollars saved, then ML models designed to reduce costs are likely to be developed in droves, more so than any of countless other ML applications.

The top two metrics for discerning ML success as noted by respondents this year were tied for first place: business metrics, such as guaranteed ROI, and a more technical evaluation of ML model performance.

Across all industries and company size, 58 percent of respondents said ML efforts are successful if they produce ROI, reduce customer churn, aid in product adoption, and/or promote brand fidelity. And another 58 percent of respondents said ML efforts are successful when model accuracy, precision, speed, and drift meet threshold. (Note that respondents were encouraged to select more than one answer option, accounting for the more than 100 percent total.)

When those percentages are broken down by role, an interesting separation occurs. The individual contributor level (data scientist, software developer) values technical measures of ML success more so than the business metrics, and C-level executives and VPs generally place more value on the opposite—measuring ML success by how it ultimately benefits the company at a strategic level.

The director level is in the middle, valuing both the business unit impact (ROI, budgetary, strategic planning metrics) as well as the more technical metrics surrounding model performance. We assess that the director level will prove to be the crux of ML decisions made within organizations in the coming years as they seek to demonstrate their teams' capabilities but also prove to senior management that ML is a worthwhile investment to make.

To that end, we expect to see an increase in the number of proof of concepts demanded of ML tooling companies by organizations looking to build ML programs or mature their current efforts. This expectation works into our assessment that aligning stakeholders and obtaining senior buy-in will also become less of a challenge in the near term.

> "
> Machine learning will be the biggest technological shift of our generation, enabling businesses to achieve their full potential."
> Diego Oppenheimer, CEO Algorithmia

The future of machine learning

In our survey report from last year, we concluded that ML was very much in pioneering days, with most companies only just beginning to develop use cases, build models, and align teams. Twelve months later, we see the ML landscape already changing as early efforts to build healthy ML lifecycles become more streamlined.

Our hypotheses for the near-term future include the following:

- A growing number of data scientists employed at mid-sized companies to help gain industry edge using ML
- Lower levels of customer satisfaction at large corporations as they prioritize cutting costs
- The advent of more innovation hubs to drive ML adoption within organizations
- An increase in director-level roles stewarding ML progress across all industries

We are convinced that company size is not a determinant of ultimate ML maturity level, and we look forward to the future where companies of all sizes in all industries can implement machine learning to automate and augment their business goals.

We are particularly curious about what the near-term future holds for machine learning use cases. The trend of using ML to automate fairly formulaic tasks will soon give rise to more complex and pipelined ML workflows. As that happens, the infrastructure needed to compute those more compound applications will also change, requiring practitioners to make choices about tech tooling that may affect infrastructural performance or flexibility down the road.

This year's survey report should confirm for readers that machine learning in the enterprise is progressing in haste. Though the majority of companies are still in the early stages of ML maturity, it is incorrect to think there is time to delay your ML efforts. If your company is not currently ML–minded, rest assured your competitors are, and the rate of AI's development is bound to increase exponentially. Now is the time to future-proof your organization with AI/ML.

Join the 2020 state of enterprise machine learning conversation @algorithmia #2020StateOfML

Methodology

The purpose of the 2020 State of Enterprise Machine Learning report is to examine the progression of ML across the business landscape and compare the current state with that from 12 months ago to begin to identify trends, anomalies, or patterns of behavior. This report is based on data Algorithmia collected in the fall of 2019 in a two-prong survey effort that returned 745 respondents.

The first prong (referred to herein as Group A) comprised a set of 20 questions pertaining to machine learning efforts, capabilities, and company demographics, and was disseminated by an independent third-party company on Algorithmia's behalf. This was done to ensure survey attribution anonymity and remove bias for or against Algorithmia on the part of the respondents. The third party sourced a random sample panel of business leaders and ML practitioners (individual contributor, manager, director, and executives) at companies using data science for machine learning.

Group A respondents voluntarily participated in the survey and were offered a small compensation by the third party for doing so. Algorithmia received the raw data following the third party's survey completion after all identifiable respondent demographic information was removed by the third party.

The third party screened respondents using the following questions:

- Does your company employ data scientists?
- Which role best describes your title/role within your organization?
- Which industry do you currently work in?
- Which stage of ML maturity is your company in?

If respondents gave specific "I do not know or I am unsure" or null answers, they were removed from the respondent pool. In this way, Algorithmia amassed a group of 303 individuals with a level of insight into the machine learning efforts of their companies across a random sampling of industries, company sizes, and machine learning maturity levels.

The second prong (referred to herein as Group B) consisted of 21 questions pertaining to machine learning efforts, capabilities, and company demographics and was disseminated internally by Algorithmia with the company name and logo on it. Most questions overlapped with those in the third party's survey, but there were several exceptions, one of which was email address, which was collected in order to fulfill 10 random $50 gift card incentives. The survey explained that a respondent's email would be used solely for the gift card purpose, and Algorithmia maintained survey integrity by ensuring the respondents' answers were not connected with their email addresses in any way.

Group B was sent to individuals who have engaged with Algorithmia in the past in various capacities (ie. attended a company webinar, read an internal whitepaper, met with our team at an industry trade show, etc.). The 442 respondents in this group voluntarily participated in the survey and represented a diverse sampling of industries, company sizes, ML maturity levels, and organizational structures and roles.

The survey was conducted in two prongs to account for unintentional bias in either group. Where possible and appropriate, the researchers averaged Groups A and B for the most accurate readings and specified when one or both groups were represented in the text or in the graphs. All percentages were rounded to the nearest whole number.

The 2020 State of Enterprise Machine Learning questionnaires (Groups A and B) were developed collaboratively by the Product and Marketing teams at Algorithmia. The teams identified the key issues to measure, determined critical survey questions, and provided feedback on the draft questionnaire. The survey was designed to be self-administered and completed online in an average of six minutes.

We will continue to conduct this annual survey to increase the breadth of our understanding of machine learning technology in the enterprise and share with the broader industry how ML is evolving. In doing so, we can track trends in ML development across industries over time, ideally making more informed predictions with higher degrees of confidence.

About Algorithmia

Algorithmia is a leader in the machine learning space. We aim to empower every organization to achieve its full potential through the use of artificial intelligence and machine learning by delivering the last-mile solution for model deployment at scale.

Our technology is trusted by more than 100,000 developers, Fortune 100 financial institutions, government intelligence agencies, and private companies. Algorithmia enables customers to:

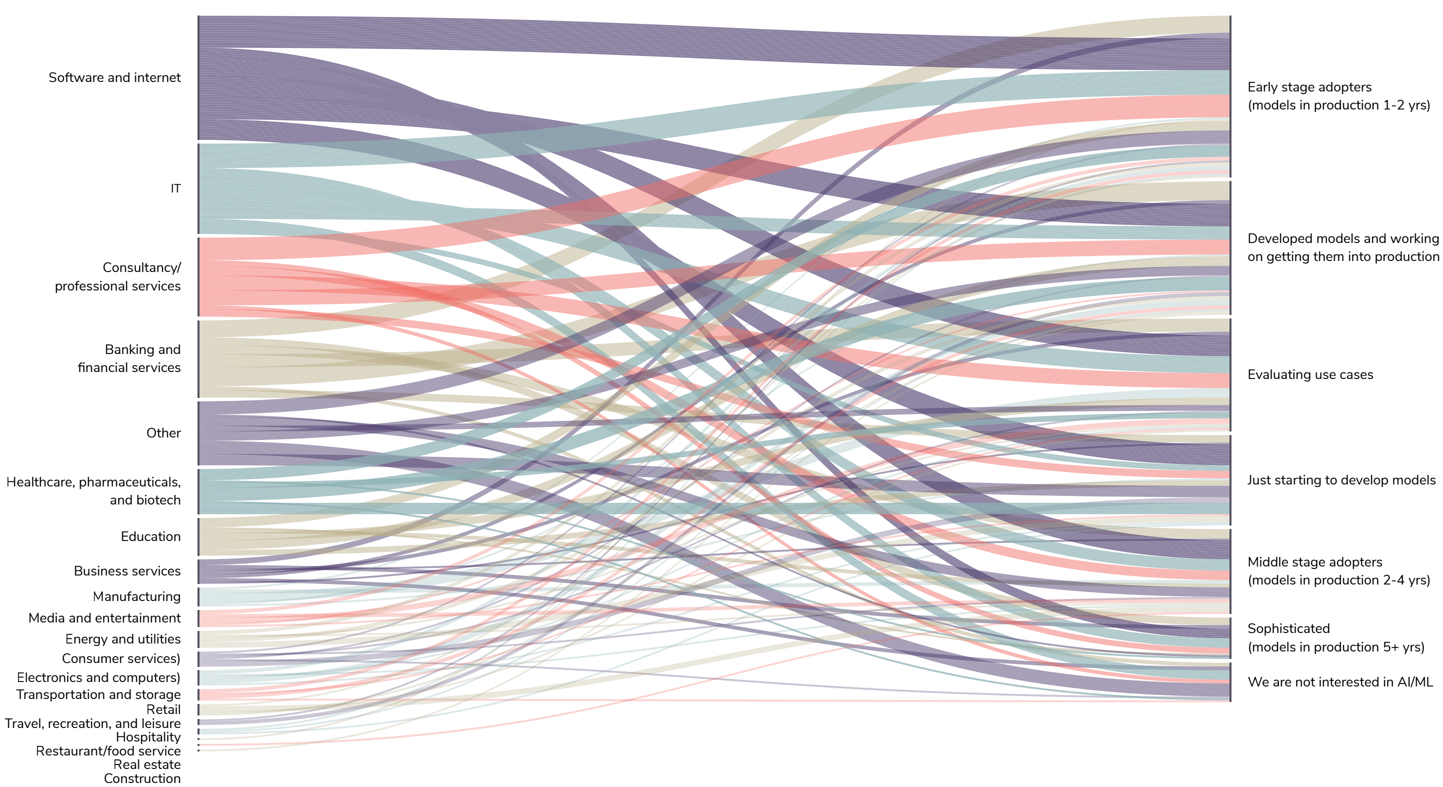
- Deploy models from a variety of frameworks, languages, and platforms
- Connect popular data sources, orchestration engines, and step functions
- Scale model inference on multiple infrastructure providers
- Manage the ML lifecycle with tools to iterate, audit, secure, and govern

To learn more about how Algorithmia can help your company accelerate its ML journey, visit our website at algorithmia.com.

About the cover

The cover image is a parallel set chart—similar to a Sankey diagram. Each line-set represents a specific data category. The width of each line-set's path is determined by the proportional amount of the category total.

The line-set on the left depicts the different industries that our survey participants come from. The line-set on the right of the diagram displays the machine learning maturity level of our respondents' companies. Reflects data only from survey Group B.



ALGORITHMIA