CLOUDERA

WHITE PAPER

10 STEPS TO MAKING MACHINE LEARNING OPERATIONAL

A Guide for Business Leaders on Unlocking the Benefits of ML

EXPERIMENT

DEPLOY

SUSTAIN

Table of Contents

EXECUTIVE SUMMARY

Data science and machine learning are reshaping entire industries, making it possible to achieve previously impossible levels of scale through operational efficiencies and continuous learning and innovation. That's because data science and machine learning automate the extraction of useful insight from data, detecting patterns in a way that would take humans weeks, months, or years to complete—if at all.

You can use machine learning to automate business processes and enhance or invent new products and services. You can predict what a customer is likely to buy. You can automatically detect manufacturing inefficiencies or fraudulent behavior. And you can do all of this while harnessing the new data and insight your machine learning capabilities yield, allowing for even further optimization and innovation opportunities.

Modern data storage, processing, and software capabilities have progressed far enough to allow any organization to capture and use its wealth of diverse data to train, test, and validate even the most complex predictive machine learning models. Many companies have successfully embedded predictive models in their core business capabilities to develop game-changing products and services that would have otherwise been unachievable. And in doing so, they've proven that machine learning has already changed the business landscape forever.

Machine learning and AI can create **$2.6T** in new business value for marketing and sales by 2022, and up to **$2T** in new business value for manufacturing and supply chain planning.[1]

Over 75% of companies generating more than half a million dollars in revenue boosted sales through the use of machine learning to better understand customer behavior.[2]

84% of enterprises believe investing in AI will lead to greater competitive advantages.[3]

For many enterprises, however—and maybe you count your enterprise among them—machine learning remains elusive; while some businesses find at least a modicum of success in their machine learning and AI initiatives, others find themselves on the outside looking in, watching competitors gain the upper hand.

Global spending on AI systems will hit $77.6 billion by 2022.[4] Yet, while the most opportunity comes from adopting machine learning at enterprise scale, only 21% of enterprises embed AI across multiple business units.[5] That equates to widespread AI investments across enterprises and industries with very little internal proliferation. One explanation could be that many business leaders are exploring the novelty of AI and don't fully understand the numerous ways in which machine learning and AI can create value across their business. AI can solve problems that were once unsolvable, and it can provide answers to questions enterprises didn't even know to ask. Because of this, achieving success requires experimentation and incremental approaches to adoption.

Another possible reason why adopters are slow to embed AI across multiple business units is that enterprises that have a lot to gain from machine learning struggle to get it off the ground. Almost 80% of all AI and machine learning projects stall out due to problems with data quality, labeling, and building trusted models.[6] Businesses are coming up short in a big way because the obstacles to making machine learning operational are eventually perceived as simply too great to overcome.

Fortunately, you can use the 10 steps listed in this guide to find success in your own efforts in achieving the transformative business benefits of machine learning. As you read on, you'll also get a closer look at why many enterprises fail at deploying and adopting machine learning capabilities across the enterprise. So, while you discover the path toward operational machine learning, you'll also learn common mistakes and missteps to avoid.

STEP ONE:

Take a holistic approach to machine learning

If anyone offers you an out-of-the-box machine learning "solution," take your business elsewhere. You can't purchase truly effective machine learning off the shelf, tack it onto an existing application or process, and reap the rewards. That's because machine learning isn't a single tool or platform or solution; it's a capability, one that can never really be mastered over time by taking a software-only approach.

In truth, machine learning thrives best when it's supported by an organizational ecosystem. Before you can deliver machine learning capabilities, you must first have the right data governance and data engineering tools and standards in place to develop your machine learning models at their core. Likewise, ongoing data governance, model sustainability, and the integration capabilities of your architecture greatly impact a machine learning capability as it moves forward into production.

Machine learning must be viewed holistically as an integral part of your data strategy. By putting it in context alongside your existing IT environments, processes, applications, and workflows, you'll better support business processes and drive greater results.

It all comes back to data

How well you can continuously govern data across the entire organization will play a major role in the success and sustainability of your machine learning initiatives. While automation, business predictions, and product innovations are the goals of machine learning, those goals are achieved only by creating and maintaining algorithms—and an algorithm can only be as accurate as the data that shapes and feeds it.

If you were to try to boil the concept of machine learning down to its most basic definition possible, it may look something like this: Machine learning is the act of training machines to process information and make a recommendation of some kind with a high degree of accuracy, but doing so without human intervention (automation) and by analyzing often incredible amounts of data within a fraction of the time it would take a human to do (scale).

> "When we're talking about machine learning, we're talking about training machines to process information and make a recommendation in a similar way as a human would, either methodically or intuitively. And really we're talking about doing this in an automated way, at scale."
>
> Alex Bleakley, Manager, Machine Learning Solutions Architecture, Cloudera

So, if a human makes highly flawed recommendations because they relied on incomplete, inconsistent, unrefined, irrelevant, or biased data, then a machine learning algorithm will do the exact same—only much faster and on a much larger scale.

The fluid nature of models

Data fuels the model, which in turn drives a given set of machine learning capabilities. Your data changes over time. New data sources come in, patterns evolve, and how well you govern your data may fluctuate. All of these things impact a model.

Building enterprise capabilities with machine learning models at their core is different than traditional software application development. And unlike most apps or microservices, models have the potential to shift in real time, depending on their function and the data they interact with.

If you want to see immediate and long-term success of your machine learning initiatives, you have to understand the dynamic, fluid nature of the models that drive machine learning capabilities. Once models are deployed into production, they must be continuously monitored for updates because the data that feeds a model can change naturally over time. A model's data can also become corrupted and inaccurate. In either case, continuous updates to models ensure they continue to deliver the results and business outcomes they were designed for.

There also comes a time when a model reaches its natural end-of-life. This could happen for any number of reasons. Maybe the original business problem that the model helped to solve is no longer an issue for the organization. Or, perhaps a model existed only to push recommendations to a customer-facing service that's being removed. By continuously monitoring your models, you'll be able to retire ones that are no longer needed and free up your resources.

STEP TWO:

Be willing to experiment and, yes, fail

Machine learning brings the promise of automating business processes, of solving imminent and long-term problems that may not have been avoided otherwise, of optimizing and differentiating your products. But when it comes to achieving a machine learning capability itself, there really aren't any shortcuts. There are only right and wrong turns.

Before you map out your enterprise machine learning strategy, the very first question you should ask yourself is this: "What problem am I trying to solve?"

It's helpful to think on a granular level when you try to identify a problem or opportunity. What incremental, positive change can machine learning possibly make to a process or application? Let the problem you're trying to solve be your guide, and then let the solution you're after be the destination. Once you know what problem you want to fix, then ask yourself this follow-up question: "Can machine learning solve this problem?"

Think through both questions with an open mind and be open to the possibility that machine learning may not be the answer to the problem you're trying to solve.

> "The challenges that I see on a day-to-day basis are not actually tied to technology or infrastructure. The vast majority of the challenges that I run into with companies are just picking the right problems to solve to begin with."
>
> Michael Gregory, Director, Machine Learning Strategy and Customer Success, Cloudera

Walk before you run

Every step along the way, you have to treat making machine learning operational much like building a campfire: Be patient and set realistic expectations of yourself and your team. If you manage to make a spark but expose it to the wider elements before it's ready, it might extinguish.

Your first viable machine learning model is that spark—a small but bright proof of concept for what machine learning can do for your enterprise. When you have one, protect it and help it grow over time.

If at first you don't succeed...

Machine learning models and the algorithms behind them are by nature about science, not business results. Only their application can drive business results, and you should approach the business problem you're trying to solve with the application of machine learning as an experiment. When you're seeing whether or not machine learning can solve a problem, you're testing a hypothesis.

If your hypothesis proves correct, then you're ready to think in terms of how to apply that machine learning model for a business capability at large. If you find that machine learning can't solve the problem you tested, then take the lessons learned along the way and apply them to your next hypothesis. Regardless of what the outcome might be, you should be willing to take risks and understand that sometimes those risks will result in lessons learned. And even when faced with a series of consecutive failures, just remember that a single breakthrough can lead to monumental value and light the way forward to solving many other problems.

What matters most is perseverance and building momentum through continuous learning, even if you have to pivot along the way.

STEP THREE:

Build a multi-disciplined team—and don't box them in

Let's say you've identified a few problems and opportunities to enhance some existing processes. You want to see if machine learning can help, so you've gotten buy-in from your fellow executives and, more importantly, you've been able to bring some data scientists on board.

Before your small team of data scientists can experiment, you have to also loop in the people in charge of your data governance. These data custodians will help your data scientists get access to the data they need. They'll make sure the required data can be presented in such a way that it can be consumed by your data scientists. To do so, they'll build information models and the necessary data pipelines from data warehouses, data engineering, and other data services, for example.

Choosing the right platform and tools

Even if you're just beginning your machine learning journey, the platform and tools you select today will lay the groundwork for tomorrow, and they should have functionality that bridges two primary and continuous phases of machine learning:

PHASE ONE

Phase one covers holistic machine learning development and the building of the machine learning models. Does the platform you're considering give your team practical access to the data, compute resources, and libraries they need? Can your team collaborate efficiently across disciplines, and can the platform enable this access and collaboration in a way that's governed and secure?

PHASE TWO

Phase two focuses on production, scaling, and ongoing operations. Here, a platform with everything your team needs to put models into large-scale production is very important. If you find a platform that checks all the boxes for phase one, ask yourself if that platform works with continuous integration tools so you can deploy your models anywhere they're required to operate.

Ongoing monitoring, management, and scaling

Beyond experimentation and deployment, your platform of choice should also make monitoring your models running in production easy and intuitive, so you can keep pace with the ever-changing nature of those models and make adjustments when necessary. You'll also need a platform that enables continuous governance, security, and transparency via cataloging and lineage for your models in production.

And while your team may start small, your machine learning platform must be able to scale up to potentially hundreds or thousands of users. Other platform characteristics to look at and compare include costs that may be involved with moving data out of cloud environments, which could make experimentation expensive in some instances.

> "There's obviously a lot of tooling out there which is very low cost and easy to acquire but doesn't provide you a complete process. It doesn't bring in security and governance. It doesn't bring in the lineage and auditing you require, and actually to retrofit those to certain technologies is really hard."
>
> Chris Royles, Principal Systems Engineer, Cloudera

STEP FOUR:

Iterate quickly, optimize later

The experimentation in phase one requires flexibility. Let your data scientists select and use the tools and frameworks they want. They should have the freedom to iterate quickly and build models that can be optimized later. Don't worry about getting a model that's flawless the first time through; you may spend too much time trying to perfect a model only to learn that the solution or enhancement you were hoping for wasn't actually achievable through machine learning. Let your team experiment rapidly, fail early and often, continuously learn, and try new things.

A balancing act

One challenge you may encounter early on may be to find the right balance between giving your data scientists the freedom to experiment while simultaneously enforcing appropriate governance and security.

While standardization and control are important, your data scientists have to remain agile and nimble as they experiment. Try not to lock your team down and slow any progress when governance and security standards are enforced. Let them play in the sandbox—just keep that sandbox safe.

STEP FIVE:

Take your architecture into account

Your team of data scientists will initially work in a lab-like environment, but they should never work in a vacuum. Think about your wider data ecosystem from the very beginning. What if you do arrive at a model that's ready to be deployed as a predictive service? You need to know how that service should interface with whatever other services or apps will consume the predictions your model makes.

Your team needs to understand and consider architectural parameters and restraints, such as the number of predictions that can be served within a given timeframe based on the amount of data processing resources available. Otherwise, you may find yourself with a viable model full of potential that's still not in production 12 months after it was ready, merely because the architecture was an afterthought.

> "Start straight away… thinking production first and thinking about what your architecture looks like."
>
> Alex Bleakley

Your enterprise undoubtedly has standards in place for code source control and integration. Have your team adopt and leverage the tools and best practices that already exist within your enterprise. Doing so will make integration much easier down the road.

STEP SIX:

Embrace machine learning by evolving your organization

Your team of data scientists have iterated on a machine learning model that shows promise. Maybe they discovered it is in fact possible to predict a target variable with a high degree of accuracy using a large data set. The breakthrough provides a morale boost, and other viable machine learning algorithms follow suit. You started your machine learning journey with a list of problems and opportunities. For at least a few of those items on your list, you're able to say, "yes, machine learning can help." Eureka!

Now it's time to take the leap, and this is where many organizations can unfortunately fall short.

There's a wall that seems to exist between experimentation and large-scale production. Many organizations hit this wall because they don't know how to weave machine learning development, production, and maintenance into their existing processes, workflows, architecture, and culture.

Some organizations try to force machine learning into a rigid structure where it doesn't fit. Others completely isolate it where it can't benefit anyone or anything. But solving the problem isn't really about trying to fit machine learning into your existing organizational scheme; it's about making the structure of your organization more flexible so that machine learning can be embraced.

> "Let's say you wanted to try skiing. You went skiing a couple of times and then you told everybody that you ski. 3 years later you haven't skied again, but you still own skis and you're still telling people that you ski. That's exactly what's happening with AI. People are dabbling in it. They put it off to the side, they don't integrate it into the day-to-day life of their organization, and then 3 years later, they're doing artificial intelligence over here in the corner where it's not benefiting anybody."
>
> Santiago Giraldo, Senior Manager Product Marketing, Cloudera

Integrating machine learning across teams, departments, and stakeholders

From an organizational perspective, there are any number of ways to structure the business for optimized machine learning. It's all about identifying what works best for your company.

One way is a centralized team strictly focused on data science. This team primarily builds machine learning models and then puts them into production across other parts of the enterprise where these capabilities can be used. This approach revolves around the idea of building a center of excellence.

Another organizational approach embeds one or more data scientists into business product teams across the enterprise. This lets your data scientists get close to a business problem so they can better understand it. However, they still need to collaborate together as a team of data scientists, which again underlines the importance of choosing a machine learning platform that lets them efficiently collaborate and share knowledge across departments.

The best organizational approach for you depends on your particular business needs. Whatever structure you feel is best for your enterprise, what's most important is maintaining the ability for your data science teams to collaborate and share ideas and best practices with each other.

STEP SEVEN:

Maintain the integrity of your models

It's easy to forget that a model is never the end point, especially during the early iterative stages. Moving models into production is one thing, but maintaining those models is something else entirely. Some of the most successful machine learning adopters have hundreds—if not thousands—of models in production, and every last one of those models undergoes a continuous lifecycle of improvement. Models have to be validated on an ongoing basis.

As your underlying data changes and shifts, and as your models themselves have an impact on the data, the models using that data have to be updated and improved upon. Maintaining the integrity of your models demands vigilance. Otherwise, your models may drift and become inaccurate—directly impacting the quality of your predictions and, ultimately, your business.

STEP EIGHT:

Close the skills gap

Building the right team structure up front is important, but the very nature of machine learning blurs organizational lines and breaks down the barriers between traditional roles.

If your team consists solely of data scientists, your models may grow more sophisticated with every iteration and experiment, but then you might hit a wall when it's time to try take a model into production. Traditional data scientists have trouble making the leap from building models to putting those models into production and integrating them with other systems of services and applications, which is why cross-functionality is important.

Try to build a team whose experience, talents, and capabilities—including data engineering, data science, software development, DevOps, product development, and domain expertise—overlap. Look for candidates with the core skills that are necessary to accomplish your most important tasks, and then get them together and let them learn from one another.

Make sure the individuals that make up your team have the aptitude and willingness to expand their skill sets and knowledge base over time. For example, look for data scientists who want to invest in their data engineering skills.

> "I'll go into an organization, they'll have a lot of SQL skills, database management skills... they might have a lot of analytical skills. You need to be able to transition those skills across."
>
> Chris Royles

The "do-it-all" data scientist is a rarity. That's because new roles around data science and machine learning are still being defined, and unique skill sets will continue to coalesce into new job titles as time goes on. For now, however, it's not about finding the person who can tackle every aspect of machine learning; it's often about cultivating the people you already have.

> "I typically advise businesses to stop trying to hire these unicorn data scientists that can do it all as one person, and just start investing in what you might call T-shaped skills. And the goal is to have a team of people whose skills overlap."
>
> Chris Wallace, Research Engineering Lead, Cloudera Fast Forward Labs

STEP NINE:

Treat models in production like living software

You assembled a cross-functional team. You iterated and experimented on a series of machine learning models. Some models proved they can benefit the business; other models proved a particular problem couldn't be solved by machine learning.

You gave your dedicated team of data scientists and data engineers access to the platform and tools of their choice while still managing to enforce a strong set of data governance and security standards. Early on, you considered how models would be integrated with your existing architecture and applications, and you invested in the ongoing education of your team members. So far, so good.

Fast forward, well beyond your initial experimentation days. Your machine learning team has been federated across business units, but they continue to collaborate with one another on a unified machine learning platform, sharing workflows and helping one another.

Your enterprise now has hundreds of machine learning models in production. What started as a tiny spark from a bit of wood and a lot of patience has grown into a healthy fire.

But remember that your models, like that flame, are alive. Like any living thing, those models that power your machine learning capabilities must be fed, sustained, and controlled.

Protecting your models

You need clear visibility over your models at all times, which means you also have to understand and monitor the entire lifecycle of each one—from the data feeding it to how it's operating in production. Sometimes data naturally progresses out of range and a model becomes confused. Other times, bad actors or simple human error can taint data or production environments and steer a model off course.

To protect your models in production, it's important that you have the ability to keep them secure. This means having the visibility into model lineage and the control over who can access and make changes to your models. Additionally, it's essential that you're able to monitor the ongoing technical, mathematical, and business performance of each and every model you have in production—and that you're able to take corrective action as soon as you see that a model is drifting.

While you should be able to trust your models, never follow them blindly. This all goes back to strong data governance and data integrity.

Understanding how your models work and change over time

Compliance has historically centered around the storage, movement, and application of data. Another reason that clear, real-time visibility into your models is important is that regulatory demands will eventually catch up with the dynamic nature of machine learning. Some industries, like the financial sector, already have compliance requirements pertaining to the use of machine learning models.

That's why it's important to gain a thorough understanding of how a model arrives at its output. This necessitates not only ongoing understanding and evaluation of the underlying data, but also an ability to "unpack" an outcome driven by a model and provide some explanation for how the resulting decisions were made. Or, at the very least, it's important to demonstrate sufficient consideration of potential harms, risks, and steps taken to mitigate them. For example, it's worth exploring the use of interpretability frameworks to help demonstrate —internally as well as to a group of regulators—what features affect your model's output. This kind of capability helps you to detect, investigate, and reduce the risk of your ML-driven application or service amplifying some hidden bias in the underlying data set—and that means a better outcome for everyone.

> "It's not enough that you have data scientists and data engineers who are doing a good job. You also have to think, 'How are we going to explain this and explain this well?' The business teams that are working with data scientists—do they understand that at some point they are going to have to be speaking with customers who will want to know how this works, where the data came from, what training data was used?"
>
> Ade Adewumi, Manager, Strategy and Advising, Cloudera Fast Forward Labs

STEP TEN:

Understand—and abide by—your ethical obligations

There is certainly no shortage of ethical considerations in machine learning. And like many elements of machine learning, maintaining and adhering to a rigorous set of ethical standards is an ongoing process.

Far too many companies have found themselves in the news for their mishandling of data and breaching customer trust. The same can occur with machine learning models; therefore, the onus is on each individual enterprise to be responsible data and machine learning model custodians. You must protect the integrity of your models (protecting them from bad actors) and understand how they think and what they're doing (monitoring them for drifting, biased outcomes, or abuse of data).

Some ethical considerations

Make sure you truly have consent from customers and other stakeholders to apply the necessary data against a machine learning model. If a person is contributing data to a model, do they have a clear understanding of how their data is being used? Also, always be open to the possibility that a model has a bias. Ask yourself if you're applying a data set in a way that crosses a line and might be objectionable: Are you exposing any sensitive customer information? Can other actors use these predictions in nefarious ways? Does your model make predictions based on unethical attributes such as race or gender?

You also have an ethical responsibility and expectation to safeguard the data you collect and feed to machine learning models.

> "We're building new systems that interact with people that are unlike systems we've had in the past. When you're building systems, dealing with people, you need to think about those people's rights and what is fair."
>
> Doug Cutting, Chief Architect, Cloudera

A few last ethical considerations: Eventually, your burgeoning machine learning capabilities will need to be integrated into your existing corporate sustainability framework. What are the downstream effects of your machine learning initiatives? As you begin to move hundreds of machine learning models into production, what economic and ecological impact will doing so have? Think about these questions before you begin any machine learning initiative and re-evaluate your answers often.

CONCLUSION

Machine learning holds great promise. But like most great things, it asks for a bit of patience, an open mind, and a willingness to persevere when small wins prove elusive.

If you aspire to bring machine learning into full-scale operational use, then start small and scale your efforts with a well thought-out and informed strategy. Remember, effective and innovative machine learning capabilities are cultivated over time.

Between you and the benefits of machine learning is a path filled with common pitfalls and mistakes. Now that you're guided by these 10 steps, your journey to making machine learning operational in your enterprise can become much smoother.

Sources

1 McKinsey
2 MIT Sloan Management Review
3 Statista
4 IDC
5 McKinsey
6 Dimensional Research

Cloudera, Inc. 395 Page Mill Road Palo Alto, CA 94306 USA cloudera.com