WHITE PAPER

Get the most from your AI investment by operationalizing analytics

Contents

Organizations realize they need analytics to run, grow, and differentiate their products and services. More than just driving strategic decision making, they need to incorporate analytics into their high-volume operational and transactional decisions every minute of every day. This means that analytics is no longer just the responsibility of the data science team. Organizations have to continuously deliver analytics into operational systems in a systemic, high-quality and dependable way.

But that's the challenge. Organizations haven't been very successful in crossing the last mile of analytics to deploy models into operational systems. They have given serious attention over the last decade to accessing data and building models using the latest techniques. But putting the models into production has been an afterthought. And the research shows this.

According to a recent SAS survey of IT and analytics teams, organizations deploy less than 50% of the best models, and take more than three months to deploy 90% of them.

Most models are sitting on the shelf not adding business value. Since it takes so long to deploy models, they may be out of sync with business needs. They may also be out of date because of changing market conditions and data. This can result in less than optimal results. Imagine a credit card company using a risk model that wasn't up to date to reflect the latest fraud scam.

The ability to operationalize analytics has never been more critical. As organizations explore the potential of AI to solve complex business problems, they have to find a sustainable way to put those models into production in a governed, trustworthy and automated way.

Fortunately, there is a better way. Analytics needs to take a page from application development (DevOps) and embrace ModelOps – a practice that puts in place the culture, process and technology to operationalize analytics faster and more efficiently. ModelOps ensures maximum business impact from analytics, automates repeatable tasks, builds collaboration between stakeholders and streamlines the flow of analytics into decision making processes.

A model's value lies in how quickly it's deployed

Consider the role that analytical models play in these industries:

- **Retail banking:** In the case of credit card products, models enable all of the interactions along the cardholder's journey, from initial underwriting and credit-line assignment to real-time card authorizations and collections.
- **Health care:** Providers look to analytics to reduce readmission rates, improve health outcomes and increase patient safety. Health insurers rely on analytics to reduce eligibility fraud.
- **Retail:** Analytics drives relevant customer interactions through timely offers and optimized prices in an omnichannel world.
- **Government:** Agencies look to analytics to administer their benefit programs more efficiently. Transportation agencies rely on analytics to predict traffic and road conditions.

In every case, models drive actions that generate revenue, minimize costs, or improve customer and citizen experiences. But this can only happen when those models are efficiently integrated into the production systems that drive automated decisions. A model offers no value when it's not in production.

Regardless of industry, the ability to produce good operational models quickly while incorporating increasing complexity of data, customer behaviors and market conditions can mean the difference between success and failure. With so much at stake, why does it take so long to get a model into production?

Model deployment: What gets in the way?

The main reason models take so long to implement is because the groups that produce and deploy models are from separate teams – analytics and IT.

Data scientists and analysts build models using a range of analytics languages depending on their skill set, preferences and the task at hand. They are not typically focusing on the operational implications of a model but rather on how well the algorithm performs. They work in environments that are outside production systems.

Then it's up to IT to take the champion model and deploy it. This group focuses on how to incorporate a model into an operational system, so it runs efficiently. But IT may not understand the data transformations or the statistical methods used in the model. Both groups rely on a different set of processes, tools and programming languages.

And the handoffs from the modeling team to the operational team can be challenging, with many poorly defined steps. There can be a lot of iteration and manual work that makes it hard to react rapidly to changing business conditions and create a consistent pipeline of models.

These development and deployment problems become more evident at large organizations where hundreds of data scientists and analysts are producing many models to address specific challenges.

Some of the most common issues include:

Operational and development environments are different

Organizations rarely deploy models in their development language. This is typically because the model code may not be executable in the organization's operational systems and may need to be converted to another language. For example, porting or recoding a Python model into a production language like Java can be difficult. In addition, there may be performance issues due to its complexity of calculations.

Recoding a single model to integrate it into an operational system can extend the project timeline by weeks or months due to manual effort, rework and revalidation.

Each time the organization develops or updates a model, it must repeat the cycle. Think about how the workload escalates once an organization builds tens or hundreds of models using different languages, tools or platforms.

Lack of transparency in model construction

IT may not have a clear understanding of the data transformation, feature engineering and algorithms used to build a model. Model features used during the training phase may need to be manually replicated or refactored in at the inference (production) phase. As a result, when it is introduced into an operational system, the model may behave differently than intended because some of the elements were either lost in translation or removed due to complexity. This can reduce the value of the model significantly – or potentially eliminate its value altogether.

No common model repository

Without a common model repository, it is impossible to keep track of an organization's model inventory. It is impossible to know how many models (champion and challenger) an organization has; who is modifying, validating and approving the models; assess model "freshness," etc.

In addition, it is easy for organizations to lose track of which models are the "right" models to move into production. This can lead to organizations using out-of-date models or prevent the deployment of models that are production-ready.

Lack of established processes

Many organizations lack a repeatable ModelOps process. They don't have an established process for verifying that a deployed model is working as designed in the real world, integrating the model within the existing operational processes, and monitoring and retraining the model on an ongoing basis.

No performance monitoring

Lack of continuous monitoring of model performance means degradation and lost business opportunities. It's critical to identify when a model's predictive capabilities are declining as new data is generated, or when the organization needs to incorporate new market, customer, business or competitive scenarios. If models underperform or degrade, organizations are leaving money on the table, downgrading their service levels or losing customers.

Deployment architecture scalability

Typically, there are many models running in a production system (and there could be multiple versions of the same model). Analytical processes are scheduled to run at specific intervals or with constrained service-level agreements. This means that the load on real-time systems will vary by time of day, season or external events, which could have a negative effect on model performance. That's why you need to closely look at optimizing the deployment architecture based on business needs.

Inability to comply with regulatory requirements

In many industries, there is increasing regulatory pressure to better and more transparently explain analytic models in production and the controls around those models. Organizations must document and explain the model inputs, the model itself, how it works and the decisions it affects. With more frequent use of AI, these explanations become even more challenging.

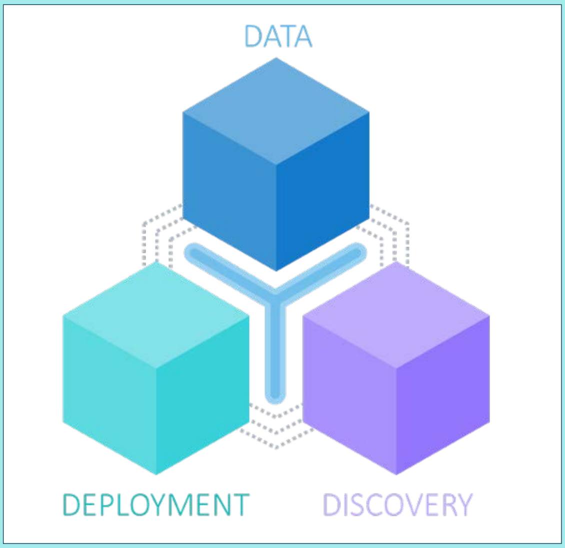
ModelOps: Accelerating the analytics life cycle

In order to actually operationalize analytics, you have to develop, validate and activate your analytical assets within existing systems or processes on a continuous, accelerated basis. You should have an organizational philosophy and practice that supports this continual flow of analytics. That's what ModelOps provides.

The Analytics Life Cycle Today

Data

Data scientists and analysts need common access to data from all relevant sources – data that is continually refreshed to provide the best environment to rapidly build models. They also need to have the ability to perform additional data wrangling within the discovery environment to get data into the right form for a particular analysis.



Discovery

Data scientists should be able to use a range of coding languages (e.g., Python, R and SAS®) for a wide range of algorithms and techniques. Business analysts may prefer a visual point-and-click interface for data exploration and model building. Others may want to use code but manage the code within a visual interface showing the analytic pipeline. An organization's analytic ecosystem should accommodate every user need. And setting up your analytics to avoid silos and local model processing will streamline discovery and improve efficiency.

Deployment

Organizations used to bring all the data into a central analytics platform in order to run analytic production processes, typically in batch jobs. This may still be appropriate in some cases. But moving data does have costs in terms of storage, processing and latency. And big data amplifies the problems. There is a trend to push analytics to where the data resides, such as a data platform (e.g., Hadoop), data streams (e.g., video) or web services (e.g., recommendation engines). You can deploy lightweight analytical applications as an API or in a container. The goal is to deliver analytics to wherever it's needed.

ModelOps brings model development and operations into harmony. The goal of ModelOps is to create a shared approach to the creation and deployment of models so models are no longer "tossed over the wall" to operations for deployment.

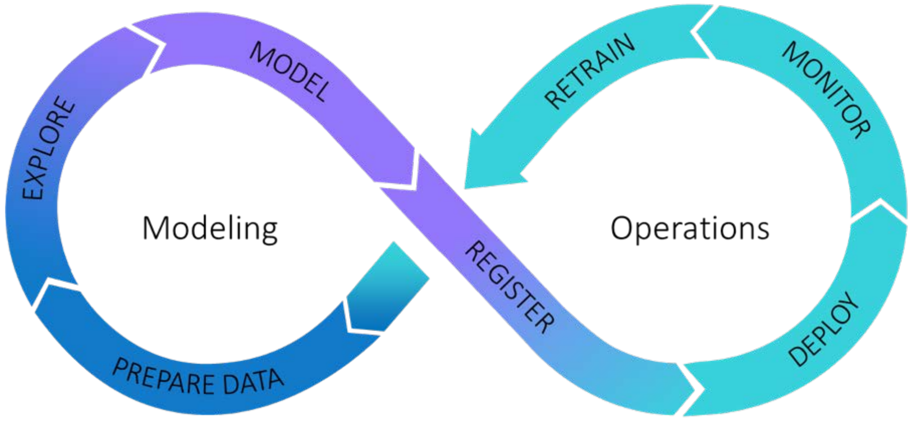


Figure 1: ModelOps life cycle.

There is a reason why analytics needs its own approach to operations. Analytical models are different from traditional IT applications and have different deployment requirements. There are three critical differences.

First, a deployed application does the same thing over and over – and only software changes or updates can affect its behavior. An analytical model, though, can change behavior due to changes to the code itself, the model produced or the associated data. As a result, the behavior of analytic models is much more complex and difficult to predict than application behavior. They are also more difficult to explain, test and improve. Also, models naturally degrade over time – especially AI models. So you must monitor them regularly, retrain or update them using new data, and then test them for accuracy before redeploying them.

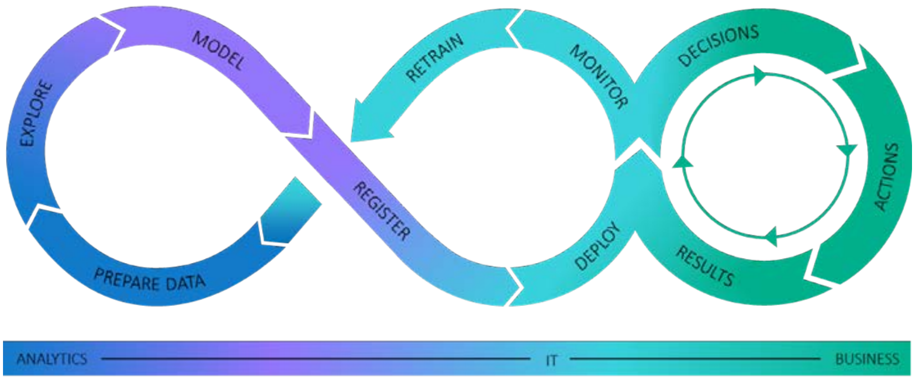
Second, analytics employs a wide variety of open source and commercial software packages. These analytics packages do not use the same coding as operational systems. This requires translating or readying the model for deployment.

Finally, models may use data from multiple sources, which may or may not be available in the operational systems. You will need to create (or make available) any data features to deploy a model for production. Organizations must understand and govern that data.

These differences point to the need for a specific model operationalizing process that accelerates deployment through closer integration of the two teams and their systems so everyone is collaborating and working toward the same goals.

Operationalizing analytics

It is not enough to put models into production. If you want your models to drive business results, you must embed them into operational decision-making processes. The critical steps to this process start with the handoff of a champion model and end when you monitor and improve model performance in a decision flow. An important aspect of ModelOps is how a model is affecting business decisions and results. In the figure below, an additional loop has been added to illustrate how important the business dimension is to a successful modeling program. A model can be performing as designed but if it is not having a positive business impact, then it becomes irrelevant.



The 7 key steps for operationalizing analytics

Register

A centralized model repository, life cycle templates and version control capabilities provide visibility into commercial and open source analytical models, ensuring complete traceability and governance. It will also promote collaboration among stakeholders and manage the analytics workflow effectively. Letting organizations store data, code, properties and metadata associated with models enables transparency and assessment of the real value of your analytical assets. It also ensures that intellectual property related to your analytics assets is retained despite staffing changes.

Deploy

The deployment phase is about integrating analytical models into a production environment and using them to make predictions. It is often the most cumbersome step for IT (and a DevOps process) to handle, but essential in delivering value. Ideally, you should be able to combine commercial and open source models in the same project to compare and select the champion model to deploy. Depending on the use case, you can publish models to batch operational systems (e.g., in database, in Hadoop, Spark), on-demand systems (e.g., web application), cloud or a real-time system using an event-stream processing engine.

Decide

While models can be put into production using the methods defined in the deploy step, incorporating them into a decisioning engine provides a powerful set of capabilities for operationalizing analytics. Capabilities to look for include the ability to create decision processes using a GUI that allows business users and data scientists to incorporate analytical models, business rules and custom code (such as Python) into decision flows used in the automation of operational decisions across the enterprise. This makes complex decision flows and instructions such as nested flows, conditional branching and the ability to process multiple records from a data table easier.

The GUI should allow you to construct decisions from scratch or create them by selecting from prebuilt sets of rules, variables, registered models and custom code. It should also incorporate governance capabilities, including versioning, predeployment testing and lineage. Using decision flows designed within the decisioning engine improves the traceability, replicability, relevance and trustworthiness of analytics.

Act

Actions are a critical step in operationalizing analytics because this is where the decisions are published to the operational business processes. To ensure the most accurate decision outcomes, decision flows should be able to access the most current customer information as the decision is executing.

Embedding SQL query statements that are carried out as the decision is executed is one way to accomplish this. Users should also be able to publish decision flows across a variety of channels and processes – digital and traditional. You must be able to execute decision flows in batch or in real time with publishing options such as REST APIs or real-time options such as in stream or at the edge. In the case of real-time decisioning services, speed and the ability to scale are important.

Because decision volumes can be high – particularly when you deploy decisions in a real-time environment – data throughput is critical. And the number of transactions per second and response times are important, as is the ability to execute decision flows asynchronously in high-volume environments.

Measure

Understanding how well the decision flows (and the related models) are performing is a critical but sometimes overlooked step in operationalizing analytics. For proper implementation of the analytics life cycle, you should record and track information about a decision in real time, during execution of the decision flow. This should include when you present an action or contact to customers and how they responded.

You can collect these contact-and-response results from the channel or business process, aggregated and accessed by analytics and reporting tools via an API. You can feed the results into other decision flows, fine-tune them and feed them back into the model monitoring process to improve model performance.

Monitor

Once organizations start seeing the value from analytics, the next phase of operationalizing analytics begins. You need to analyze and monitor scores for ongoing performance. And you should evaluate whether models are still behaving as expected based on changing market conditions, business requirements and new data. You can produce performance reports for champion and challenger models using various fit statistics.

Retrain

If model performance degrades, organizations should take one of three approaches:

- Retrain the existing model using new data.
- Revise the model with new techniques, feature engineering, or perhaps new data elements, etc.
- Replace the model with a better model.

This requires collaboration and communication between data scientists and business stakeholders to secure commitment on indicators, metrics or KPIs for success. And you must continuously measure these indicators for business impact.

ModelOps Roles and Responsibilities

- Data Scientists
  - Perform a key role in the model development stage of the analytics life cycle and support operationalizing analytics efforts.
  - Monitor model performance and retrain or create new models using new data, new algorithms or feature engineering.
  - Work with data engineers to ensure alignment between training data, validation data and production data.
- IT/Ops
  - Collaborate with data scientists to understand programming, coding or API needs for integrating analytical models into IT and operational systems.
  - Understand operational data flows, source data and data preparation required for model publishing and scoring.
  - Design, tune and manage IT infrastructure and operations for analytical models.
  - Integrate or incorporate analytics results into existing or new applications.
- Business Stakeholders
  - Clearly define the benefits they are expecting from the analytical models being developed to solve a specific business challenge.
  - Collaborate with data scientists and business experts to review the results once models are put into production and monitor the results on a continuous basis.
  - Understand how decisions influenced by the analytical models affect business processes and personnel and take steps toward building a stronger analytics culture.

The value of operationalizing analytics

Operationalizing analytics is a critical step for organizations to align their analytics investments with business objectives (i.e., KPIs). This phase is the last mile they must cross before realizing business value from analytics and AI. Operationalizing analytics will create a clear path for your organization to quickly generate value from the insights it generates from analytics.

When you operationalize analytics, you're more likely to find opportunities and generate revenue. This is because healthy businesses quickly get models into production and monitor their effectiveness.

If your strategy changes, the market changes, your competitive positions evolve, customer behavior changes or new data emerges, then monitoring your models and regularly refining them will help deliver consistent decisions and expected value.

A comprehensive approach to operationalizing analytics that incorporates ModelOps allows you to simplify the management of analytics assets (open source and commercial models). Automated model management enables more effective collaboration by letting users track progress through each step of the modeling process and providing all of the stakeholders with a unified view of each model's freshness, definition and value. Automated steps let users operationalize models in batch and in real time in different operational environments, including streaming, in database and at the edge.

Automation also allows users to focus on reducing the time it takes to operationalize analytics so models remain at peak performance. You can gain complete control of your model collections – that includes model properties, input/output data and variables, model version history and business rules – to ensure transparency and understanding.

4 things you can do to help operationalize analytics in your organization

As leaders, you can have direct and positive influence on the direction of your organization's analytics program. You can:

- Involve IT and DevOps teams at the outset of a data science project so that operationalizing analytics is not an afterthought. You should develop analytical models with operationalizing in mind.
- Agree on the quantifiable outcomes before preparing data and building analytical models. Do not underestimate the need to monitor and validate the business results of analytical models.
- Ensure the data science, business and IT and DevOps teams have a clear understanding of the roles, processes and required handoffs involved, from data preparation and model development to operationalizing analytics through adoption of ModelOps.
- Evaluate how to reduce technical debt associated with approving a data science project, especially one involving multiple tools, toolkits, languages, data sets, pipelines, models and systems.

Learn more

Learn more at: sas.com/operationalize-analytics.

To contact your local SAS office, please visit: sas.com/offices