4 Steps to Successful Model Operations
A ModelOps Guide
Introduction

Organizations are developing AI and ML models at an increasing rate to gain new insights, continue their digital transformation and reimagine their business. With multiple and often siloed teams developing AI models across the enterprise, a variety of tools and processes are more than likely employed.

Organizations have been using models to help with business decisions for decades. However, AI and machine learning models introduce new risks into model operationalization (post-development). Many model operations processes are manual or managed using home-grown solutions that constantly need to be updated as new technologies, tools and governance requirements are introduced.

As a result, over half of the models developed do not get deployed, and those that are take months to operationalize, often leading to suboptimal outcomes and delayed or diminished value.

Here are 4 steps that any organization can take to successfully operationalize AI/ML or any other type of model:

1. Define the end-to-end model operation process (referred to as the model life cycle)
2. Deploy models
3. Monitor the models in production
4. Govern model operations

“ModelOps lies at the center of any organizations’ enterprise strategy, it is an enabling technology that is key to converging various AI artifacts, platforms and solutions, while ensuring scalability and governance.”

Gartner, “Innovation Insight for ModelOps,” August 2020

Watch the Gartner webinar: “Use ModelOps for an Effective AI Strategy”
Define the Model Life Cycle

The first step of preparing a model for production use is establishing the end-to-end model operations process, referred to as the model life cycle (MLC).

A model life cycle (MLC) defines the requirements and processes for operationalizing a model. It includes detailed process flows with well-defined steps for operating, governing, monitoring and orchestrating the model throughout its life cycle, until it is retired. This includes steps for monitoring the model to ensure it continuously produces reliable results, and adheres to regulatory and compliance requirements. A model life cycle typically includes processes such as model registration, approvals, controls, retraining, testing, re-validation, and eventually retirement.

Every model has a unique life cycle that includes business and technical requirements. However, even though model life cycles are unique, there are MLC patterns found in models used across the enterprise. This allows for a “superset” model life cycle to be defined and used by many models, limiting the number of MLC processes that must be created and maintained.

The model life cycle establishes the technical and organizational scaffolding that unites data scientists, data engineers, developers, IT operations, model operations, risk managers and business unit leaders through clearly defined processes.

An Enterprise AI Architect typically has the responsibility for designing model life cycles.

How ModelOp Center Helps:

• Provides a library of pre-defined processes
• Creates a framework for governing, monitoring, and orchestrating models
Deploy Models

Deployment is the method by which you integrate a model into an existing production environment to make practical business decisions based on data.

There’s a growing awareness of the widening gap between the ability of data scientists to create models and the ability to deploy them in production. Deploying models consistently and efficiently to any consuming application is important for maximizing any model’s contribution to the business.

- Models that are not deployed in production have no value
- Models decline in value over time
- Some models can lose 50% or more of its value in weeks or even days

The first step of deployment is registering an abstraction of the model(s) in a central production model inventory. All the elements that compose the model—such as source code, tests, input and output schemas, training data, metadata, as well as outputs of training—should be included, along with all the elements required to execute it, including libraries.

By providing visibility into all models, the model inventory serves as a springboard for more efficient model deployment according to each model’s unique deployment path. The next step is to apply low-code techniques to deploy the model into the desired business application.

Models are deployed in a variety of environments and with a variety of tools. Regardless, the need for these tools to be integrated with a ModelOps solution cannot be understated. Without this integration, the ongoing operational efficiencies of management, monitoring and governance of the models is challenging at best and more than likely will diminish the long term value and revenue contribution of the model.

AI tools and IT and business systems integration is required for an enterprise AI orchestration platform

Typically the data scientist is responsible for deploying models.

How ModelOp Center Helps:

- Provides an execution environment
- Maintains a production model inventory
- Integrates with your existing tools, systems, and applications
- Ensures deployment steps are defined, executed and auditable
Monitor Models in Production

Monitoring begins when a model is first implemented in production systems for actual business use and continues until the model is retired, and sometimes even beyond as a historical archive. Monitoring should include verifying internal and external data inputs, tracking schema changes, statistical performance, data drift and ensuring the model performs within the control parameters set for it. Since each model is unique, monitoring frequency will most likely vary for each model.

Monitoring not only includes overseeing and tracking all model operational activities, it also includes remediation. For example, detecting model drift is not enough. Monitoring workflows needs to include executing, retesting or other corrective actions as required, initiating change requests, and gating activities that need approvals.

Models must be continuously monitored. Unlike software, models decay over time. Model performance needs to be tracked in three dimensions: statistically, technically, and from a business perspective:

- Statistically, is the use of input data and the output inferences performing as designed, or is there evidence of data “drift”? And how do the model’s inferences compare with one or more “better” models?
- Technically, is the model delivering inferences with the originally specified load and lag time and burden on operational systems?
- Finally, is the model continuing to provide useful business insights, or is the model at risk of exhibiting bias or violating business rules? And are new models being put into production quickly and efficiently?

If any of these metrics fall outside pre-set parameters, the best practice is to automate the updating of the model—including any necessary approvals—so that an optimized version can be quickly returned to production.

For monitoring to be most effective, it should include alerts and notifications of potential upcoming performance issues, and track and log the remediation steps until model health and performance is reinstated. With the speed at which AI and ML models perform, monitoring models has grown beyond human scale in most enterprises. Continuous monitoring and immediate remediation is essential for both reliability and governance.

A Model Operator typically has the responsibility for monitoring the health on models in production.

How ModelOp Center Helps:

- Immediately detects performance issues
- Initiates remediation steps
- Sends alerts and notifications
- Orchestrates monitoring steps and actions
Govern Model Operations

Models are a form of intellectual capital that should be governed as a corporate asset. They should be inventoried and assessed using tools and techniques that make auditing and reporting as efficient as possible.

The “black box” characteristics of AI and ML algorithms limit insight into the predictive factors, which is incompatible with model governance requirements that demand interpretability and explainability.

Some of the emerging challenges associated with governing AI/ML models are:

- Incomplete model inventories due to the increased use of AI models across business lines which creates silo’ed efforts
- Lack of insight into the operations of AI/ML models, hindering interpretability and explainability
- The increased frequency at which AI/ML models must be monitored due to the use of real time data and model decision making in high frequency digital channels

AI governance requires continuous compliance checking and ongoing enforcement of established regulatory and compliance controls and a central production model inventory that is routinely maintained with required documentation, metadata and assets throughout the life of each model. Compliance and auditability requires a systematic reproduction of training, evaluation and scoring of each model version and ultimately the transparency and auditability typically required for regulatory and business compliance.

How ModelOp Center Helps:

- Enforces conditional approvals and compliance controls
- Integrates with DevOps, MRM and ITSM applications
- Tracks all actions and changes for reproducibility and auditability
- Identifies feature importance for model inferences
Successful Model Operations

Automating and orchestrating all aspects of the model life cycle ensure reliable model operations and governance at scale. Each model in the enterprise can take a wide variety of paths to production, have different patterns for monitoring and various requirements for continuous improvement or retirement.

Automated management and orchestration provide the means to enforce every step in each model’s life cycle, providing the management oversight needed for ongoing operations and good governance and risk management.

A well-designed model life cycle leverages, not duplicates, the capabilities of the business and IT systems involved in developing models and maintaining model health and reliability. This includes integrating with model development platforms, change management systems, source code management systems, data management systems, infrastructure management systems and model risk management systems. This integration provides the connection points for orchestrating actions, streamlining the model operations processes and allowing for end-to-end management of the complete model lineage.

How ModelOp Center Helps:

• Accelerate model operationalization by up to 50%
• Double the productivity of the ModelOps team
• Increase model revenue contribution by up to 30%

Learn more at modelop.com/product
About ModelOp

ModelOp, the pioneer of ModelOps software, enables large enterprises to address the critical governance and scale challenges necessary to fully unlock the transformational value of enterprise AI and machine learning investments. Core to any AI orchestration platform, G2000 companies use ModelOp Center to govern, monitor and orchestrate models across the enterprise and deliver reliable, compliant and scalable AI initiatives.

For more information visit modelop.com.