MLOps: Monitoring All of Your Models with DataRobot Agents
Today’s Presenters:

Gautam Rewati  Rajiv Shah  Tim Whittaker
1. What is MLOps
2. MLOps Agents Framework
3. MLOps Agents Libraries
4. Collab and AWS Demo
The Emergence of MLOps

Machine Learning Operations (MLOps) technology and practices provide a scalable and governed means to rapidly deploy and manage ML applications in production environments.
<table>
<thead>
<tr>
<th>Experience challenges with model development</th>
<th>Cite obstacles to moving AI models into production</th>
<th>Monitor their AI and ML assets once in production</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>96%</strong></td>
<td><strong>90%</strong></td>
<td><strong>1%</strong></td>
</tr>
</tbody>
</table>

Successful AI Adoption is Yet to be Figured Out
The Fundamentals of AI Adoption at Scale

**Production Model Deployment**
Variety of ML tools, languages and platforms makes it difficult to collaborate between Data Science and IT teams.

**Production Model Monitoring**
Traditional software tools not designed to track ML model behaviour.

**Model Lifecycle Management**
Models have complex lifecycle and become stale requiring frequent updates.

**Production Model Governance**
Governance is required to minimize risk and ensure compliance with regulations.
MLOps Overview

Monitor Everything, Everywhere

- Any type of model, deployed anywhere, monitored in the same way in the same real-time command center.
- Monitor everything between your model and business value.
- Automation enables scale to tens of thousands of items monitored.

Superior ML Health

- Improve the performance of models in production by automatic challengers, retraining, and rollouts.
- Free up data scientists to tackle more problems instead of maintaining and debugging.

Automation-Accelerated Governance

- Move faster while increasing scrutiny, compliance, and transparency.
- Design and execute an enterprise-wide AI governance strategy.

Any type of model, deployed anywhere, monitored in the same way in the same real-time command center.
Monitor everything between your model and business value.
Automation enables scale to tens of thousands of items monitored.
Improve the performance of models in production by automatic challengers, retraining, and rollouts.
Free up data scientists to tackle more problems instead of maintaining and debugging.
Move faster while increasing scrutiny, compliance, and transparency.
Design and execute an enterprise-wide AI governance strategy.
DataRobot MLOps UI
GO TO AN ACTUAL EXTERNAL DEPLOYMENT
DataRobot MLOps Dashboard
DataRobot MLOps Service Health
DataRobot MLOps Data Drift
DataRobot Predictions Over Time
DataRobot Accuracy Over Time
Why MLOps Monitoring Agents?

Data scientists build great models on their laptops

Engineers deploy those models to org infrastructure.

But where is the monitoring? MLOps agents allows customers to monitor deployments running on their own infrastructure as if it were running inside of DataRobot MLOps
Use Cases:

- Stored procedures running in a modern database.
- Python (scikit-learn and other) models developed and deployed outside of DataRobot (AWS, Azure, GCP, Bare Metal)
- DataRobot models downloaded and deployed on org infrastructure
- H2O models deployed as POJO (or MOJO) on org infrastructure
- Spark ML Models running on org infrastructure
MLOps Agents Framework

The framework is composed of 3 parts

1. A remote / external model
2. The Agent Service
3. The MLOps libraries
MLOps Agents Service
MLOps Agents Service

The Agents Service is a Java service which runs on org infrastructure and is configured to monitor deployments. Once new information is detected, the agents process the information and forwards it to DataRobot MLOps. The agent is designed to handle multiple deployments.

Agents Service requires configuration, specifically

1. Valid MLOps URL
2. Valid Authorization Token
3. Channel Configuration: Buffer to disk, message broker, etc.
MLOps Agents Library
MLOps Library

The MLOps Library provides a way for you to get the same monitoring features with models deployed in your own environment, as you can with models deployed within the DataRobot environment.

The MLOps library is available in Java, Python and R.
MLOps Library - Simple Example

Code must be added to:

- Initialize MLOps Agent
- Report Predictions and Features
- Report timing
- Shutdown Agent
Google Colab + MLOps Agents
Google Colab + MLOps Agents Demo Overview

1. Download MLOps Agents jar
2. Configure and setup DataRobot MLOps Agents on Google Colab
3. Create a binary classification model
4. Create deployment placeholder in DataRobot
5. Make Predictions
   - `predictions = clf.predict_proba(test_data).tolist()`
6. Reporting Metrics/Stats:
   - `mlops.report_deployment_stats(len(predictions_array), execution_time)`
   - `mlops.report_predictions_data(features_df, predictions_array, association_ids)`
7. Upload Actuals
8. Check reported metrics on DataRobot Deployment dashboard
SageMaker + MLOps Agents
Architecture (some more details)

MLOps Agent

- Report To MMM
- Channel API
  - Spool File
  - SQS
  - Pub/Sub
  - RabbitMQ

MLOps Library

- Channel API
  - Spool File
  - SQS
  - Pub/Sub
  - RabbitMQ

Buffer Mechanism
Sagemaker + MLOps Agents Demo Overview

1. High level Overview of Sagemaker Example Notebook
2. Create Sagemaker Endpoint
3. MLOps Agents config for AWS SQS
4. Running Tracking Agents as Fargate container
5. Instrument Inference Code with MLOps Library

- runtime_client.invoke_endpoint(EndpointName=endpoint_name,Body=payload,Accept='Accept')
- mlops.report_deployment_stats(len(predictions_array), execution_time)
- mlops.report_predictions_data(features_df, predictions_array, association_ids)

6. Check reported metrics on DataRobot Deployment dashboard
Questions & Answers
**DataRobot Community**

- **Engage, learn, and accelerate** your AI/ML journey
- **Connect** with peers to find solutions to AI challenges
- **Explore** helpful content to take your AI to the next level
- **Build** your brand as an AI expert & thought leader
- **Join** your peers today at [community.datarobot.com](http://community.datarobot.com)
- **Questions:** aisuccess-webinars@datarobot.com

**Latest Topics**

- **Challenge of Predicting Time**
  - 2 posts, 1 reply, 3 hours ago
- **Unsupervised Learning with no Target**
  - 1 post, 0 replies, 4 hours ago
- **Multisolvability dataset in Regression**
  - 0 posts, 1 reply, 6 hours ago
- **US Cloud—Planned DB Maintenance (March 14th—March 15th)**
  - 1 post, 0 replies, 5 hours ago
- **Can DR Indicate Trends in my Data**
  - 4 posts, 0 replies, 3 days ago
- **Academic License**
  - 3 posts, 0 replies, 6 days ago
- **Predictive Maintenance (of NASA turbines) using DataRobot**
  - 2 posts, 0 replies, 11 days ago
- **Europe User's Guide**
  - 1 post, 1 reply, 12 days ago
- **Student/Academic License For Europe**
  - 3 posts, 0 replies, 16 days ago
- **Webinar March 10: Model Building—Automate Done, Now What?**
  - 0 posts, 3 replies, 18 days ago

**Announcements**

- **DataRobot Platform**
  - 143 posts, 4 hours ago
- **Learning Center**
  - 64 posts, Monday
- **AI & ML General**
  - 199 posts, 3 hours ago
- **About Community**
  - 18 posts, a week ago

**Questions:**

- **Join your peers today at [community.datarobot.com](http://community.datarobot.com)**
- **Connect** with peers to find solutions to AI challenges
- **Explore** helpful content to take your AI to the next level
- **Build** your brand as an AI expert & thought leader
- **Questions:** aisuccess-webinars@datarobot.com

**Featured Posts**

- **AI & ML General Blog**
  - 3 hours ago
- **Automated Machine Learning Discussions**
  - 6 hours ago
- **Automated Machine Learning Discussions**
  - 6 hours ago
- **Automated Machine Learning Discussions**
  - 6 hours ago
- **Automated Machine Learning Discussions**
  - 6 hours ago
- **Automated Machine Learning Discussions**
  - 6 hours ago
- **Automated Machine Learning Discussions**
  - 6 hours ago
- **Automated Machine Learning Discussions**
  - 6 hours ago
- **Automated Machine Learning Discussions**
  - 6 hours ago
- **Automated Machine Learning Discussions**
  - 6 hours ago

**Top Solution Authors**

- **Emily**
  - 2
- **Uncertaintyflow**
  - 1
- **opemak**
  - 1
- **Carlos**
  - 1

**Category Kudos Leaderboard**

- **Uncertaintyflow**
  - 14