AIM321

Productionize ML workloads using Amazon SageMaker MLOps, feat. NatWest

Usman Anwer (he/him)
Principal Product Manager
Amazon SageMaker

Shelbee Eigenbrode (she/her)
Principal ML Specialist
Solutions Architect, AWS

Dr. Greig Cowan (he/him)
Head of Data Science
NatWest Group

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AIM321

Productionize ML workloads using Amazon Sagemaker MLOps featuring NatWest

Speakers

- Shelbee Eigenbrode, Principal AI/ML Specialist, WW MLOps Lead, AWS
- Greig Cowan, Head of Data Science, NatWest Group
- Usman Anwer, Principal Product Manager, AWSRelevant

Announcements

- n/a – although new improvements are shown.

Takeaways

- NatWest Case Study is best section of this session. How to scale your MLOps w/Vision, Metrics and Outcomes.

Azure

- Azure ML standardizes on the Mlflow open-source framework for MLOps.

< Section Header
< Table of Contents

What is MLOps

- The process of continuously delivering high-performance ML Models at scale. DevOps for Machine Learning.
- People, Process, Tools
- Overall, a weak definition of MLOps. A better overview can be found on Nvidia's site here.

How does Sagemaker Support MLOps

- This session focuses on the Model Lifecycle
  - Project Environment Creation
  - Model Training
  - Model Registry
  - Model CI/CD Pipelines
  - Model Deployment
  - Model Monitoring

What's new in SageMaker for MLOps

- Pipeline (CI/CD) Improvements
  - Local Mode: Helps to run, test pipelines locally, to reduce cycle time and costs during pipeline creation. Deploy to cloud when ready.
  - AutoML Support: Allows the pipelining of AutoML jobs in workflow.
  - Cross-Account Sharing: Coordinate pipelines across enterprise accounts
  - SDK Simplifications: Improved Python SDK Capabilities
- Model Monitor
  - Batch Inference: New capability to Monitor Batch Inference Models
  - Demo of Batch Inference Model Monitoring

Scaling MLOps

- Initial, Repeatable, Reliable, Scalable

NatWest Case Study

- Challenges around People, Process, Data, Technology.
- Established Metrics to solve these for Faster Time to Value
  - Faster delivery of end-to-end solutions (> 12 mos) => (< 3 mos)
  - Simplified discovery and access to data (> 5 Days) => (< 1 day)
  - Simplified ML Model Route to Production (> 6 mos) => (< 2 weeks)
  - End User self-service environment creation (> 50 days) => (< 2 days)

MLOps use cases

CONTINUOUSLY DELIVER HIGH-PERFORMANCE ML MODELS AT SCALE, FASTER

Develop

Provision environments

Perform experiments

Deploy

Automate model training

Package and test models

Deploy and monitor models

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Press Esc to exit full screen

Scaling MLOps: Maturity framework

KPI

Time to POC	Model deployment time	ML defect rate	MLDC cycle time

Initial → Repeatable → Reliable → Scalable

Capabilities

Initial
- Access to resources
- Standardize experimentation environments & processes
- Track experiments

Repeatable
- Automate governed access to resources
- Automate model-build and deployment tasks
- Standardize source-code repositories
- Centralize management of models

Reliable
- Integrate full CI/CD practices into ML pipelines for traceability, improved quality, and lineage tracking

Scalable
- MLDC is reproducible through standardized templates shared across teams

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NatWest Group challenges before MLOps

NatWest Group

People	Process	Data	Technology
Lack of correct training, capabilities	Manual handovers, governance, and long time to value	Data discovery and access difficult	Fragmented and out-of-date environments

Our MLOps vision

1. Automation & standardized patterns
2. Optimized governance
3. Simplified data access
4. Modern tech stack

5. Faster time to value

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Adoption and engagement

BUILDING A LONG-TERM CAPABILITY ACROSS THE ENTERPRISE

NatWest Group

Training

Trained **100s** of data scientists, engineers and leaders across >700 AWS courses and Immersion Days

Dedicated support

Established well-resourced (12 FTE) and capable **enablement** team…

…gathering continuous feedback

Visible metrics

Continually **track and report** on KPIs related to speed of setup, training completion, use-case success and data access

User champions

Early adopters identified and given priority support to validate and test the capability

Reach far and wide

Project teams and use cases from **across the organization**, e.g.,

Fraud
Financial crime
Finance
Audit
Credit
Climate
Pricing
…

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key takeaways

NatWest Group

Hearts & minds

Create a **compelling** and business-led vision for MLOps

Adoption & engagement support for the user community

Build for complexity

Data scientists & engineers are a smart bunch!

Consider **orchestration** beyond SageMaker

Bring your own containers & templates

Build for flexibility

MLOps is a **journey** that you won't take in one step

Deliver value quickly to users and be able to respond promptly to new requests

Operating model

Consider carefully how your users will build ML-driven products

Centralized vs. federated

Standardization frees devs to focus on the value-add

Legacy tech

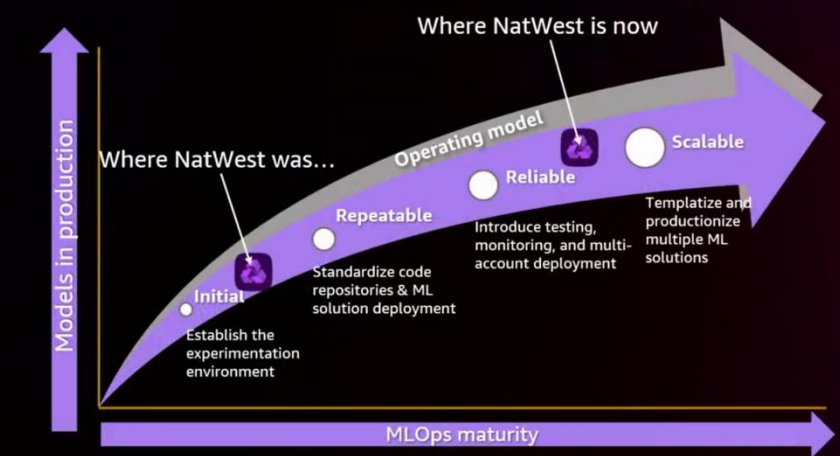
Build your system to work with legacy estate

Working between **cloud and on-premises** environments

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

A metric-based approach focused on business value

NatWest Group



Metric	Where NatWest was…	Where NatWest is now
Faster delivery of end-to-end solutions	~12 months	< 3 months
Simplified discovery and access to data	~5 days	< 1 day
Simplified ML-model route to live	3-6 months	< 2 weeks
End user self-service environment creation	> 30 days	< 2 hours

* Other metrics also used during project delivery

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.