AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

AIM413

Detect machine learning model drift in production

Aaron Sengstacken
Sr. Machine Learning Solutions Architect
AWS

Alan Tan
Sr. Product Manager, AI Platforms
AWS

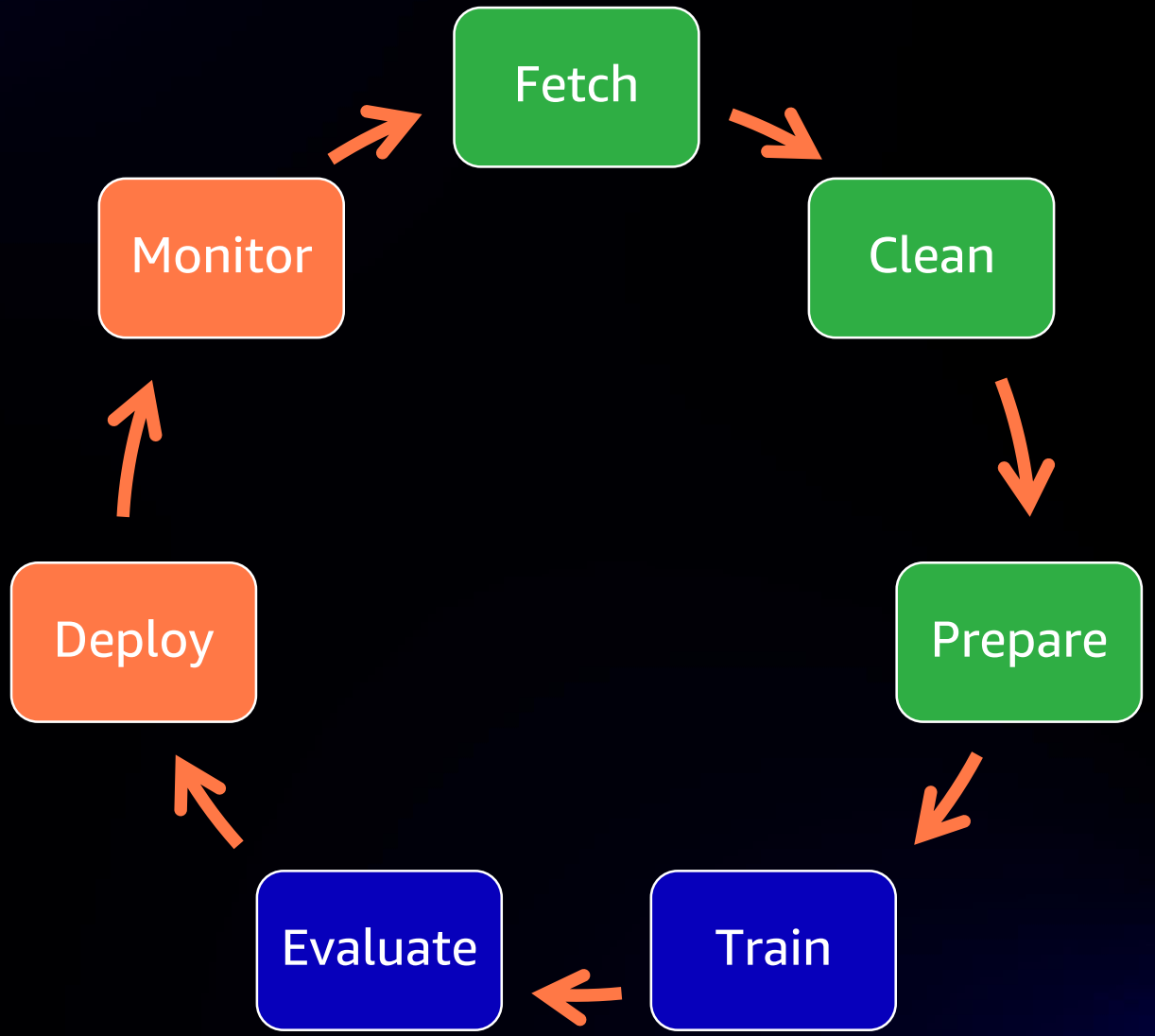
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

- Machine learning lifecycle
- Why monitor production models?
- Amazon SageMaker Model Monitor
- Strategies for retraining
- Demo
- Q&A

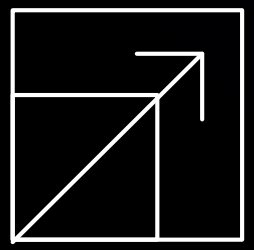
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Machine learning lifecycle

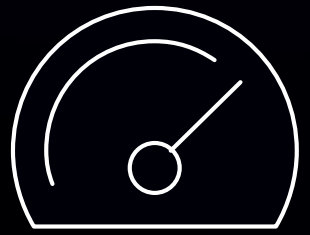


© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

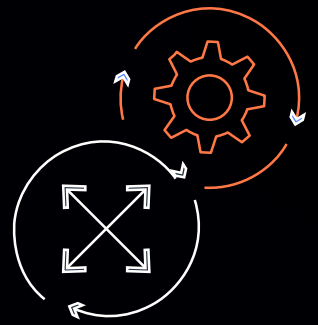
Serving system monitoring



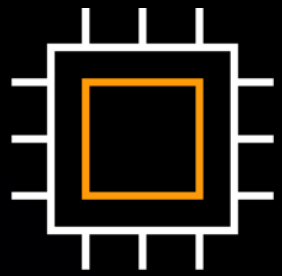
Prediction demand



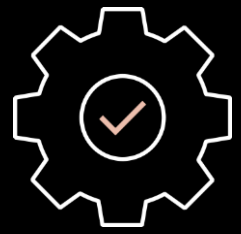
Prediction latency



Model age



Instance utilization



Number of errors

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Introducing Amazon SageMaker Model Monitor

Continuous monitoring of models in production



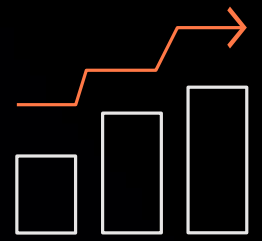
Easy to set up



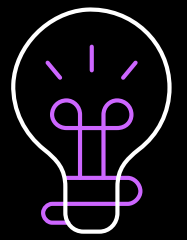
Continuous monitoring



Extensible



Visualize



Integrated

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Monitoring types supported

Amazon SageMaker Clarify



Detect divergence in data

Detect quality degradation over time

Track model balance

Feature attribution

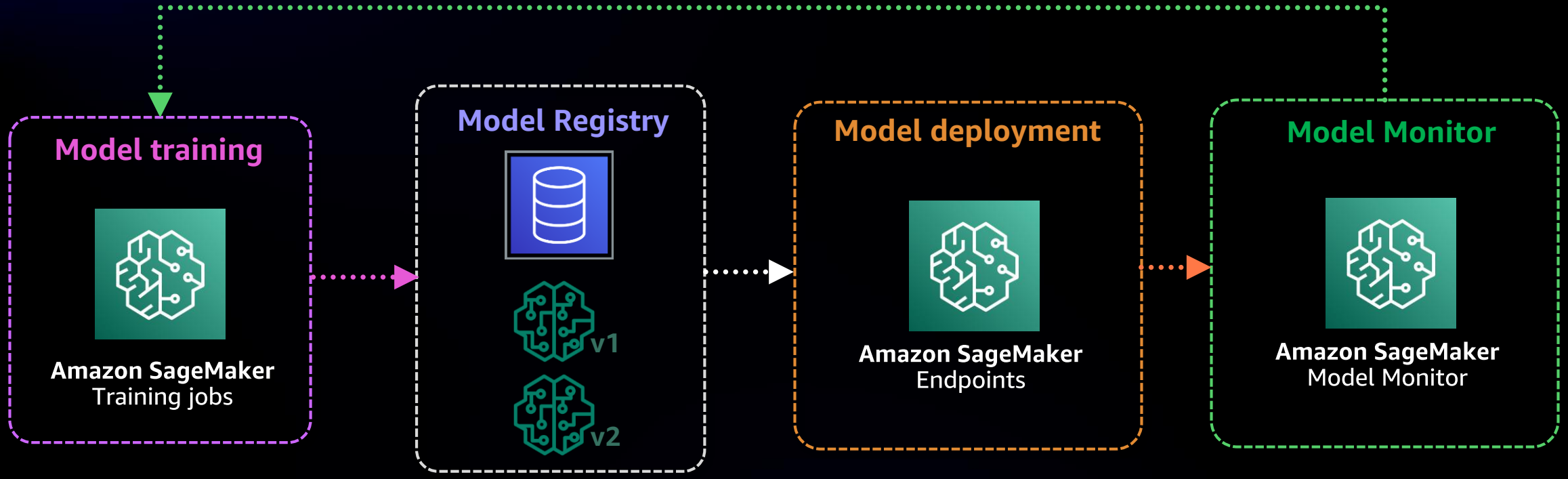
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Strategies for monitor alarms

- Retrain the model
- Label newly collected data
- Collect new training data
- Modify inference compute
- Adjust alarm thresholds

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Integration with Amazon SageMaker Pipelines



Model Registry:
Central store for versioned model packages including metadata associated with model packages such as model lineage, configuration, and metrics

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Deploying a model is not the end. You need to continuously monitor models in production and iterate.

The real-world data may not look like the training data.

Concept drift due to divergence of data

+

Model performance can change due to unknown factors

+

Continuous monitoring involves tooling and expense

=

Model monitoring is critical

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Further reading

- Amazon SageMaker Model Monitor documentation



- Amazon SageMaker Clarify documentation



- AI Fairness and Explainability whitepaper



- Amazon SageMaker Clarify whitepaper



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Aaron Sengstacken
awsaaron@amazon.com

Alan Tan
alantan@amazon.com

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.