AWS re:Invent

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

AIM402

Develop your ML project with Amazon SageMaker

Aashmeet Kalra (she/her)
Principal Solutions Architect
AWS

Raju Penmatcha, PhD (he/him)
Sr. Solutions Architect (AI/ML)
AWS

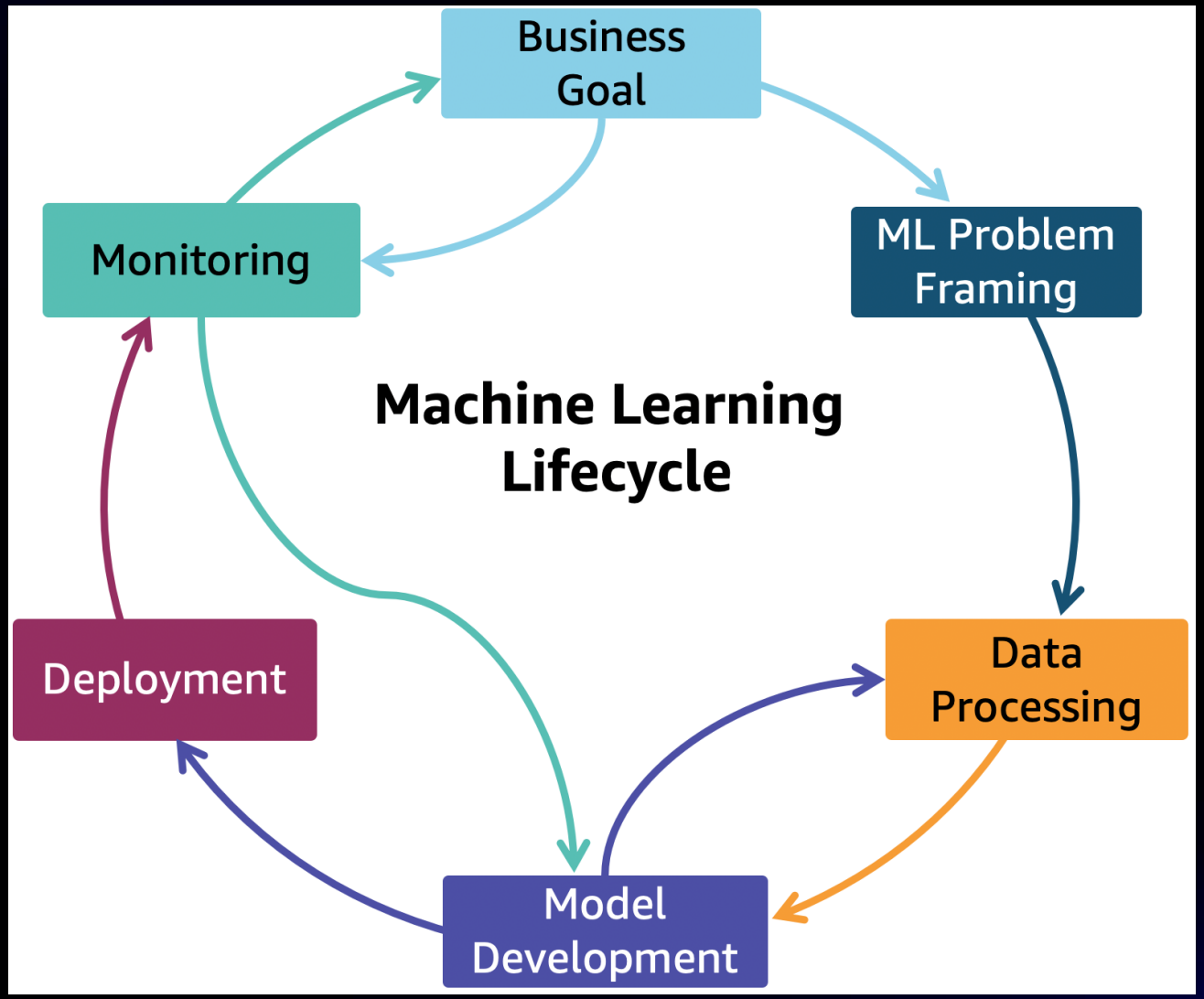
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

- Machine learning lifecycle
- Workshop use case and dataset
- Workshop logistics
- 4 hands-on labs in Amazon SageMaker
- Q&A

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

Machine learning (ML) lifecycle

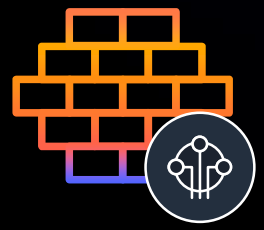


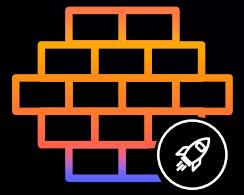
Reference: Machine Learning Lens, AWS

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

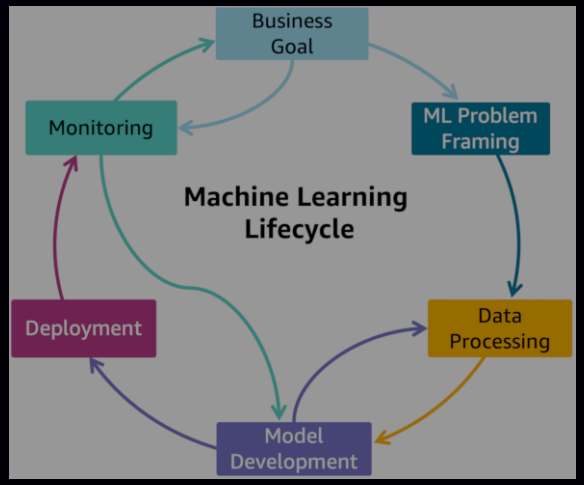
Workshop summary

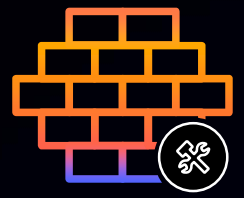
SageMaker Pipelines



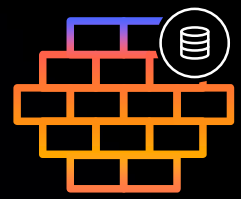


SageMaker Deploy



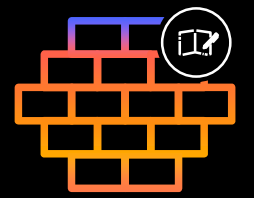


SageMaker train and tune



Data Wrangler and Feature Store

Sagemaker Studio



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

Workshop use case and dataset

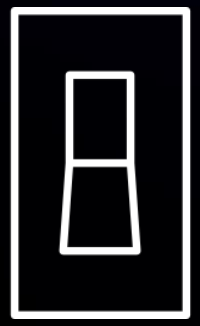
Objective:

Develop a machine learning model to predict whether a customer will enroll for a certificate of deposit (CD) at a bank

age	job	marital	education	default	housing	loan	contact	month	day_of_week	duration	campaign	pdays	previous	poutcome	deposit
56	housemaid	married	basic.4y	no	no	no	telephone	may	mon	261	1	999	0	nonexistent	no
57	services	married	high.school	unknown	no	no	telephone	may	mon	149	1	999	0	nonexistent	no
37	services	married	high.school	no	yes	no	telephone	may	mon	226	1	999	0	nonexistent	no
40	admin.	married	basic.6y	no	no	no	telephone	may	mon	151	1	999	0	nonexistent	no
56	services	married	high.school	no	no	yes	telephone	may	mon	307	1	999	0	nonexistent	no
45	services	married	basic.9y	unknown	no	no	telephone	may	mon	198	1	999	0	nonexistent	no
59	admin.	married	professional.course	no	no	no	telephone	may	mon	139	1	999	0	nonexistent	no

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

Workshop logistics



Event engine



Labs

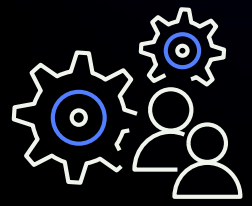


Moderators

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

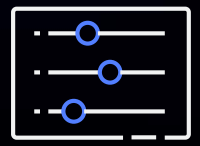
Amazon SageMaker Studio

FULLY INTEGRATED DEVELOPMENT ENVIRONMENT (IDE) FOR MACHINE LEARNING



Collaboration at scale

Share notebooks without tracking code dependencies



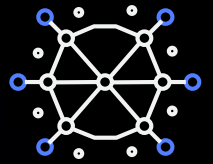
Easy experiment management

Organize, track, and compare thousands of experiments



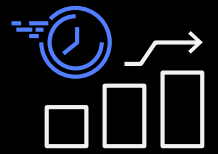
Automatic model generation

Get accurate models with full visibility and control without writing code



Higher quality ML models

Automatically debug errors, monitor models, and maintain high quality



Increased productivity

Code, build, train, deploy, and monitor in a unified visual interface

aws

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

Workshop Access

https://dashboard.eventengine.run

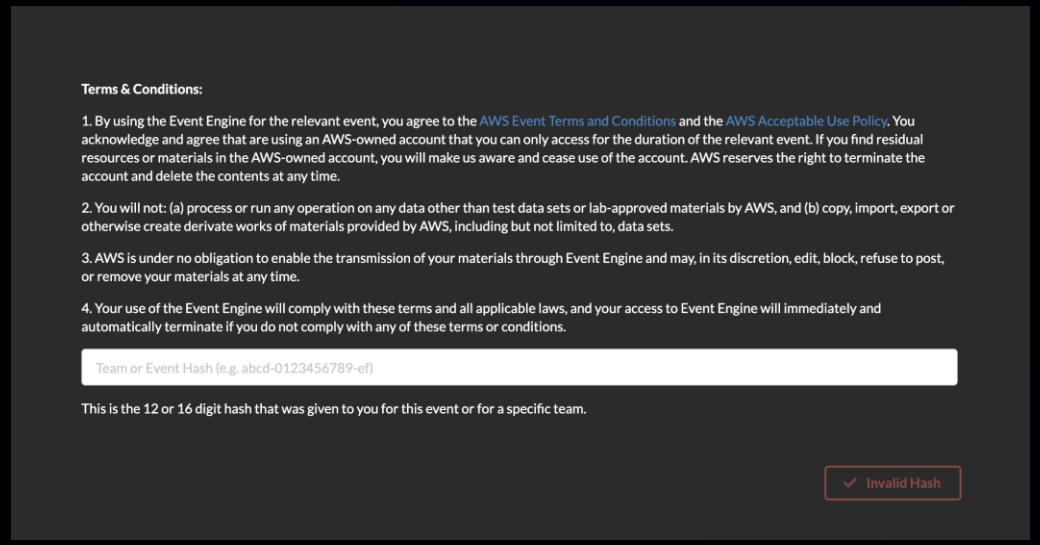
Event Hash # 6c00-13603dbc34-27

Workshop URL: https://bit.ly/3xB9vLn

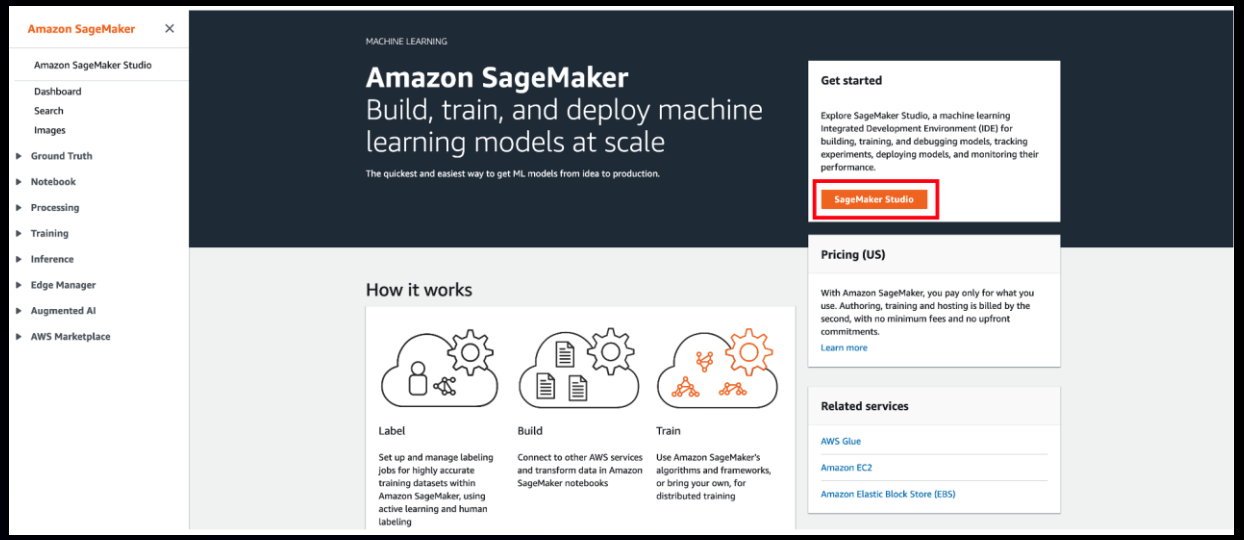
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SageMaker Studio setup

Event engine



SageMaker Studio



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

Prerequisites

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

How SageMaker Data Wrangler works



Amazon SageMaker Data Wrangler

A faster, visual way to aggregate and prepare data for machine learning



Select and query

Select and query data from a variety of data sources such as Amazon Athena, Amazon Redshift, AWS Lake Formation, and Amazon S3



Cleanse and enrich

Cleanse and explore data, perform feature engineering with built-in data transforms, and detect statistical bias with SageMaker Clarify



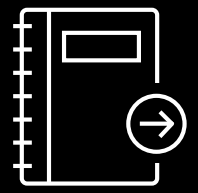
Visualize

Graphically understand data and detect outliers with preconfigured visualization templates



Understand

Use a sample dataset to quickly estimate model performance, accuracy, and diagnose potential issues



Operationalize

Export data preparation workflow to a notebook or code script to bring the workflow into production

Import data from a feature store such as Amazon SageMaker Feature Store

Integrate data preparation workflow with Amazon SageMaker Pipelines

Export prepared data into Amazon SageMaker Feature Store

Data Wrangler	Feature Store	Train/tune/deploy	Pipelines

aws

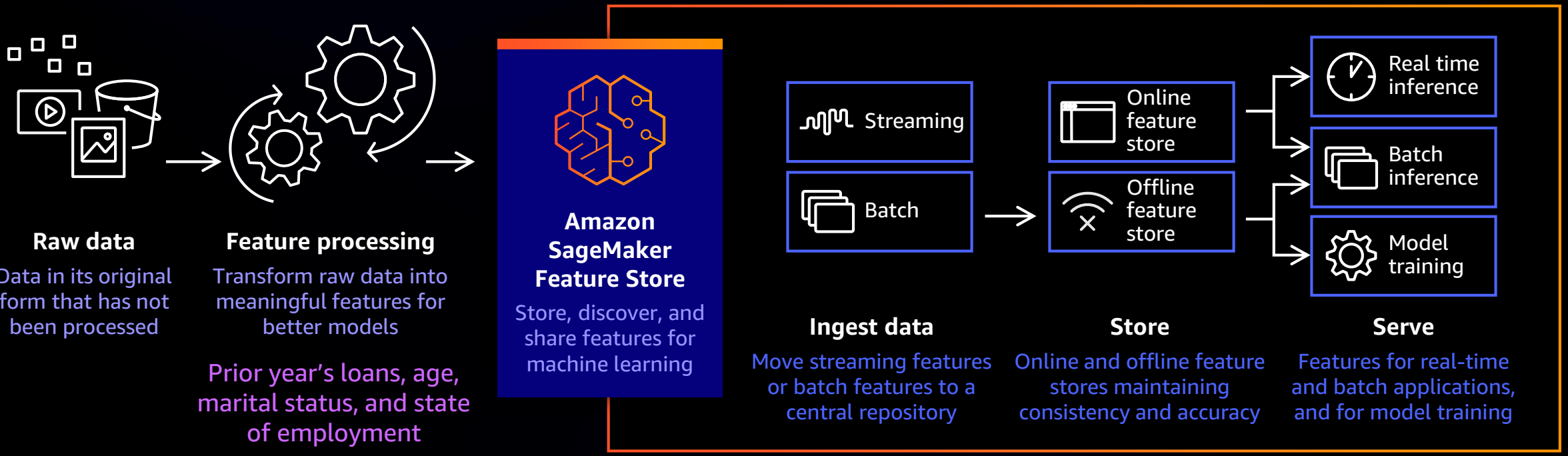
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Lab 1: Data Wrangler

Data Wrangler	Feature Store	Train/tune/deploy	Pipelines

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

Amazon SageMaker Feature Store – Features are the foundation of high-quality models



Data Wrangler	**Feature Store**	Train/tune/deploy	Pipelines

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

Lab 2: Feature Store

Data Wrangler	**Feature Store**	Train/tune/deploy	Pipelines

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

Training with SageMaker



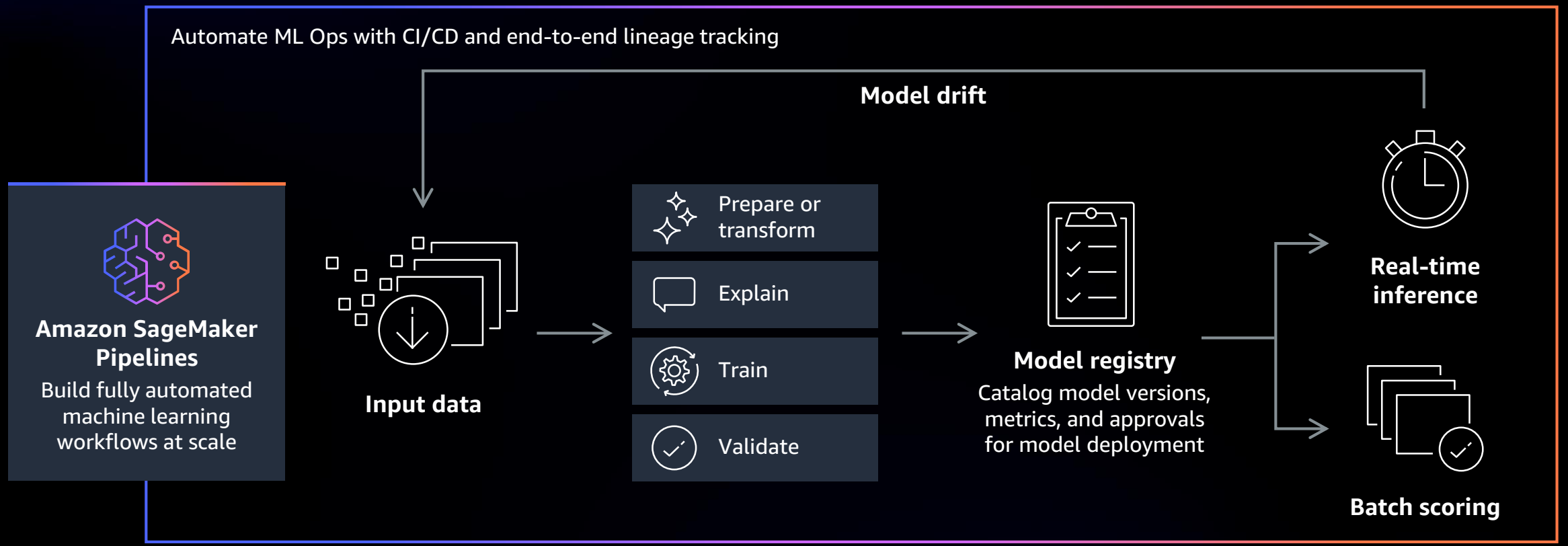
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Lab 3: Train, tune, and deploy a model

Data Wrangler	Feature Store	**Train/tune/deploy**	Pipelines

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

Amazon SageMaker Pipelines overview



Data Wrangler	Feature Store	**Train/tune/deploy**	Pipelines

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

Lab 4: SageMaker Pipelines

Data Wrangler	Feature Store	Train/tune/deploy	**Pipelines**

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

Summary

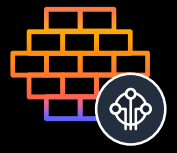
WORKSHOP KEY LEARNINGS



Prepared data and store features for reuse using Data Wrangler and Feature Store



Trained, tuned, and deployed an ML model on SageMaker



SageMaker Pipelines helped organize ans manage the entire ML workflow



SageMaker Studio gave an integrated development experience (IDE)

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

Resources

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

Additional Resources

- SageMaker – https://aws.amazon.com/sagemaker/
- Data Wrangler – https://aws.amazon.com/sagemaker/data-wrangler/
- Feature Store – https://aws.amazon.com/sagemaker/feature-store/
- SageMaker Pipelines – https://aws.amazon.com/sagemaker/pipelines/
- Machine Learning (Well-Architected Lens) – https://docs.aws.amazon.com/wellarchitected/latest/machine-learning-lens/welcome.html

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

Thank you!

Aashmeet Kalra
@aashmee

Raju Penmatcha, PhD
@awsraju

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