

AWS re:Invent

NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

Deploying a complete machine learning fraud detection solution

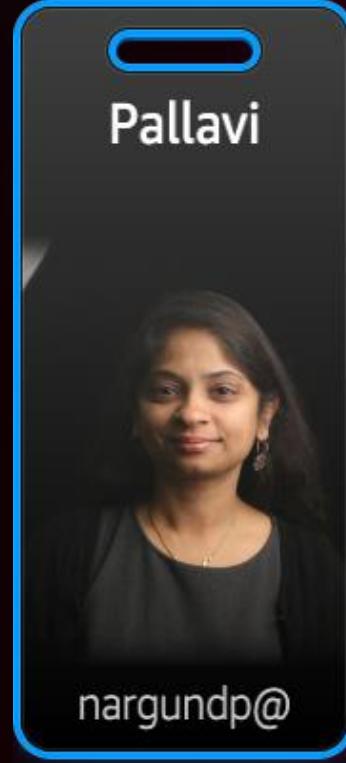
Manish Chugh (he/him)

Principal Solutions Architect
Amazon Web Services

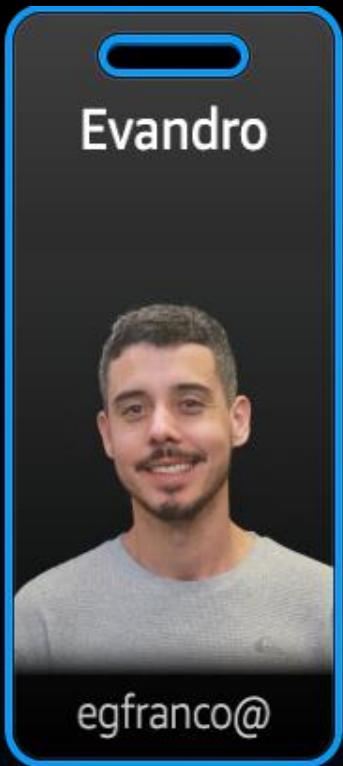
Pallavi Nargund (she/her)

Senior Solutions Architect
Amazon Web Services

Speakers



Workshop Core Team



Agenda

Auto insurance fraud

Machine learning with Amazon SageMaker

Hands-on labs

Problem statement

31%¹

\$40B
\$400 - \$700²

72%³

¹ Personal auto research: Application integrity down, fraud up - <https://vrsk.co/3UBQRhm>

² Federal Bureau of Investigation (FBI) Publication - <https://bit.ly/3hyHj80>

³ Coalition Against Insurance Fraud Stats - <https://bit.ly/3g41LNU>

Challenges



Traditional business policies, procedures and rules



Lack of meaningful insights in data collected



Slow and time consuming, impacts customer experience

Use of technology to combat fraud



Predictive modeling and
reduce false positives



Advance analytics
and Machine
learning to rescue



Cloud computing
accelerate innovation
and reduces cost

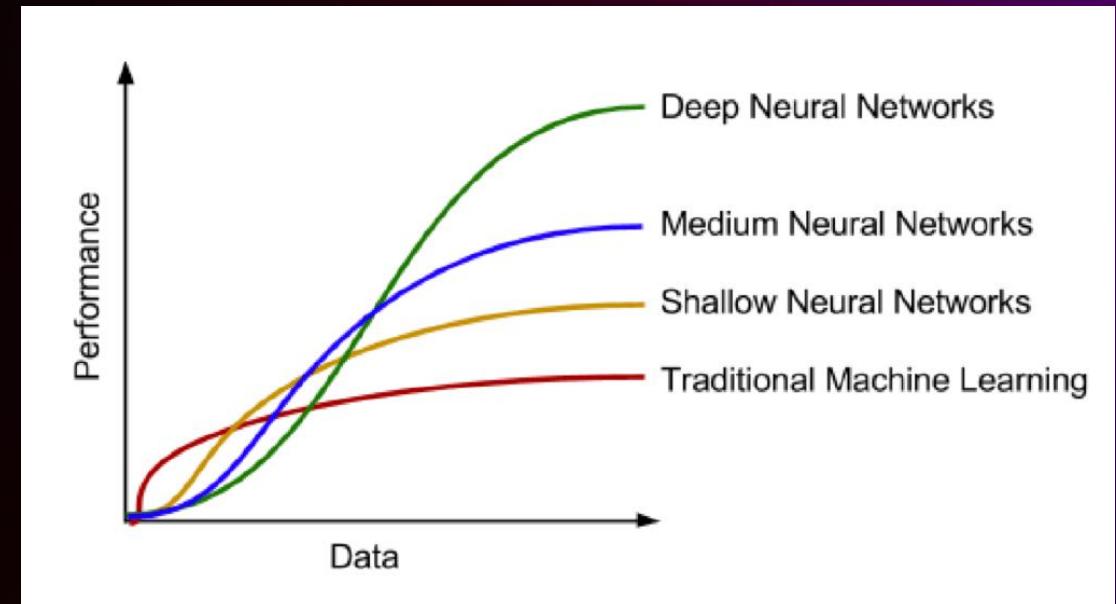
Machine learning challenges

- Extremely unbalanced data (aka “one-class” dataset)
- Often unlabeled data
- Requires low-latency inference; models can’t be too complex



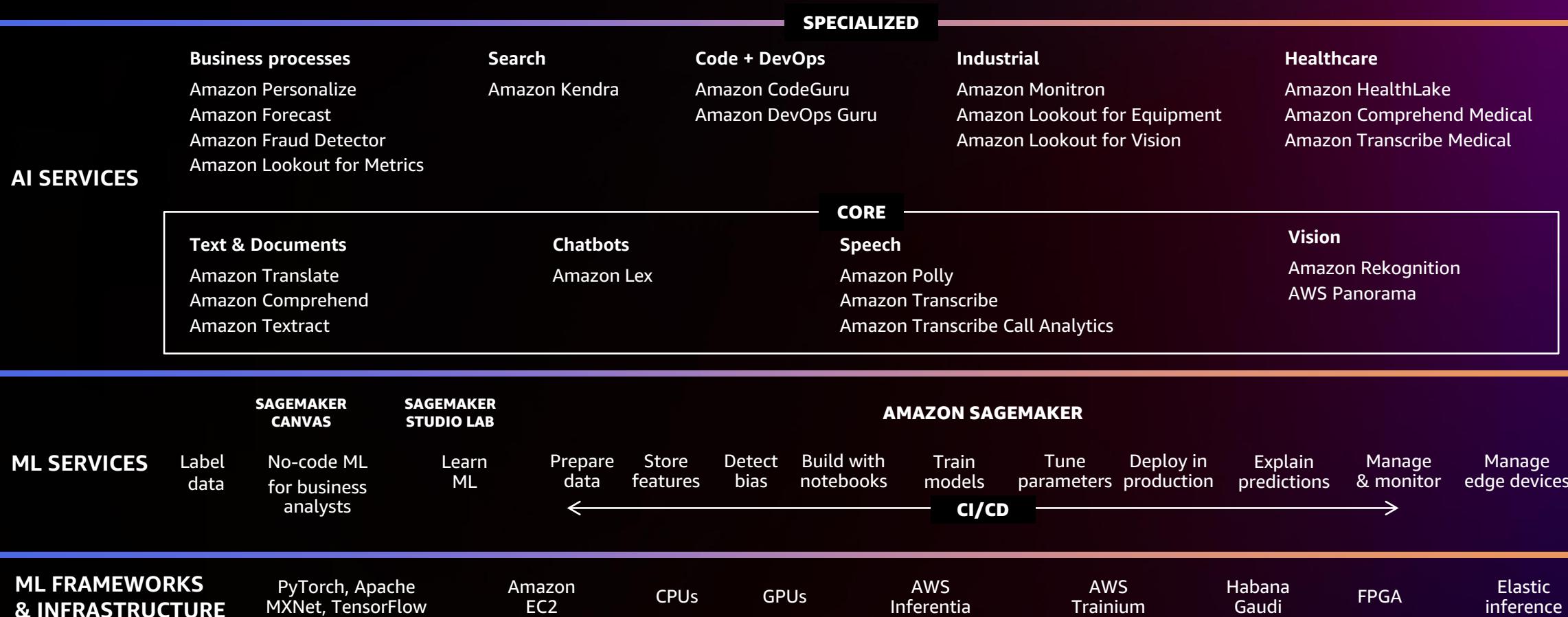
Machine learning solution option

- Labeled data/supervised learning
 - Build a Classifier
- Unlabeled data/unsupervised learning
 - Build an anomaly detector
 - Isolation Forest, Random Cut Forest
 - Build deep learning autoencoders for semi-supervised learning



The AWS AI/ML stack

Broadest and most complete set of machine learning capabilities



Amazon SageMaker

PREPARE DATA AND BUILD, TRAIN, AND DEPLOY ML MODEL FOR ANY USE CASE

PREPARE

Ground Truth

Create high quality datasets for ML

Data Wrangler

Aggregate and prepare data for ML

Processing

Built-in Python, BYO R/Spark

Feature Store

Store, catalog, search, and reuse features

Clarify

Detect bias and understand model predictions



Data Scientist



ML Engineer



Business Analyst

BUILD

Studio Notebooks & Notebook Instances

Fully managed Jupyter notebooks with elastic compute

Studio Lab

Free ML development environment

Built-in Algorithms

Integrated tabular, NLP, and vision algorithms

JumpStart

UI based discovery, training, and deployment of models, solutions, and examples

Autopilot

Automatically create ML models with full visibility

Bring Your Own

Bring your own container and algorithms

Local Mode

Test and prototype on your local machine

TRAIN & TUNE

Fully Managed Training

Broad hardware options, easy to setup and scale

Distributed Training Libraries

High performance training for large datasets and models

Training Compiler

Faster deep learning model training

Automatic Model Tuning

Hyperparameter optimization

Managed Spot Training

Reduce training cost by up to 90%

Debugger and Profiler

Debug and profile training runs

Experiments

Track, visualize, and share model artifacts across teams

Customization Support

Integrate with popular open-source frameworks and libraries

DEPLOY & MANAGE

Fully Managed Deployment

Ultra low latency, high throughput inference

Real-Time Inference

For steady traffic patterns

Serverless Inference

For intermittent traffic patterns

Asynchronous Inference

For large payloads or long processing times

Batch Transform

For offline inference on batches of large datasets

Multi-Model Endpoints

Reduce cost by hosting multiple models per instance

Multi-Container Endpoints

Reduce cost by hosting multiple containers per instance

Inference Recommender

Automatically select compute instance and configuration

Model Monitor

Maintain accuracy of deployed models

Kubernetes & Kubeflow Integration

Simplify Kubernetes-based ML

Edge Manager

Manage and monitor models on edge devices

Studio | RStudio

Integrated development environment (IDE) for ML

MLOps: Pipelines | Projects | Model Registry

Workflow automation, CI/CD for ML, central model catalog

Canvas

Generate accurate machine learning predictions—no code required

Workshop use case – Labeled data set



Binary classification problem



Improve predictive accuracy using boosting



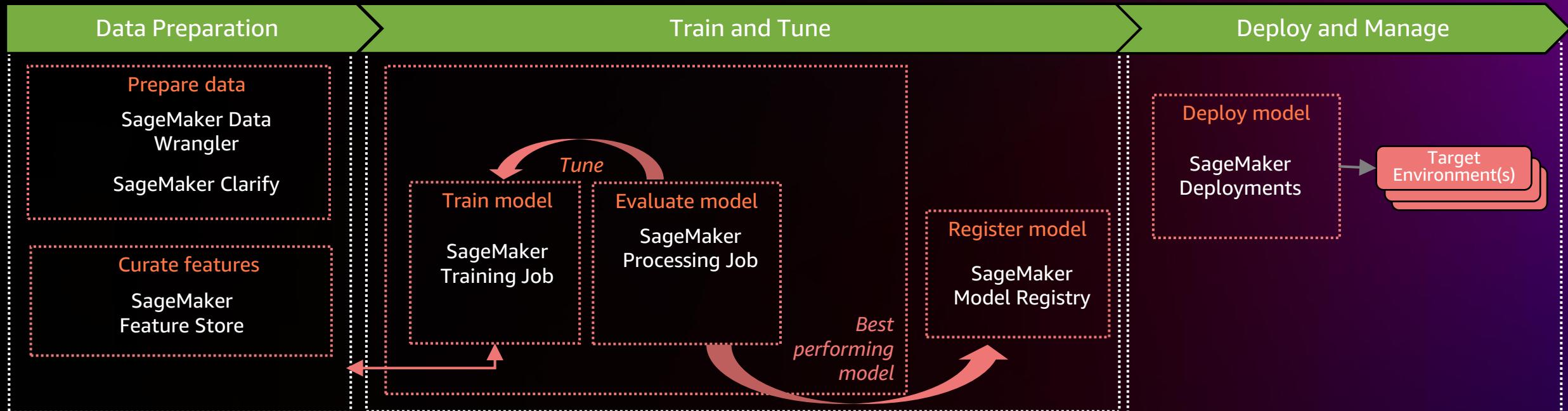
Gradient boosting



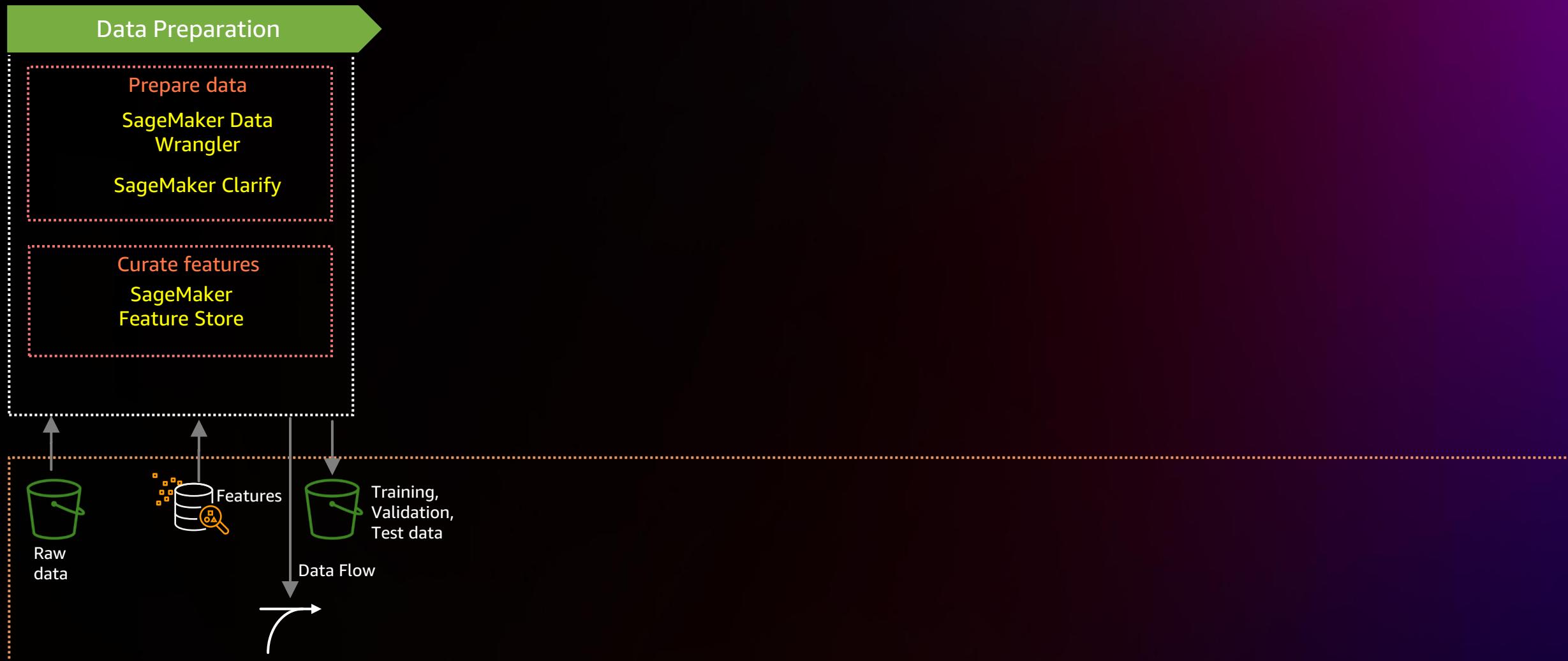
Extreme gradient boosting

Building, training, and deploying models

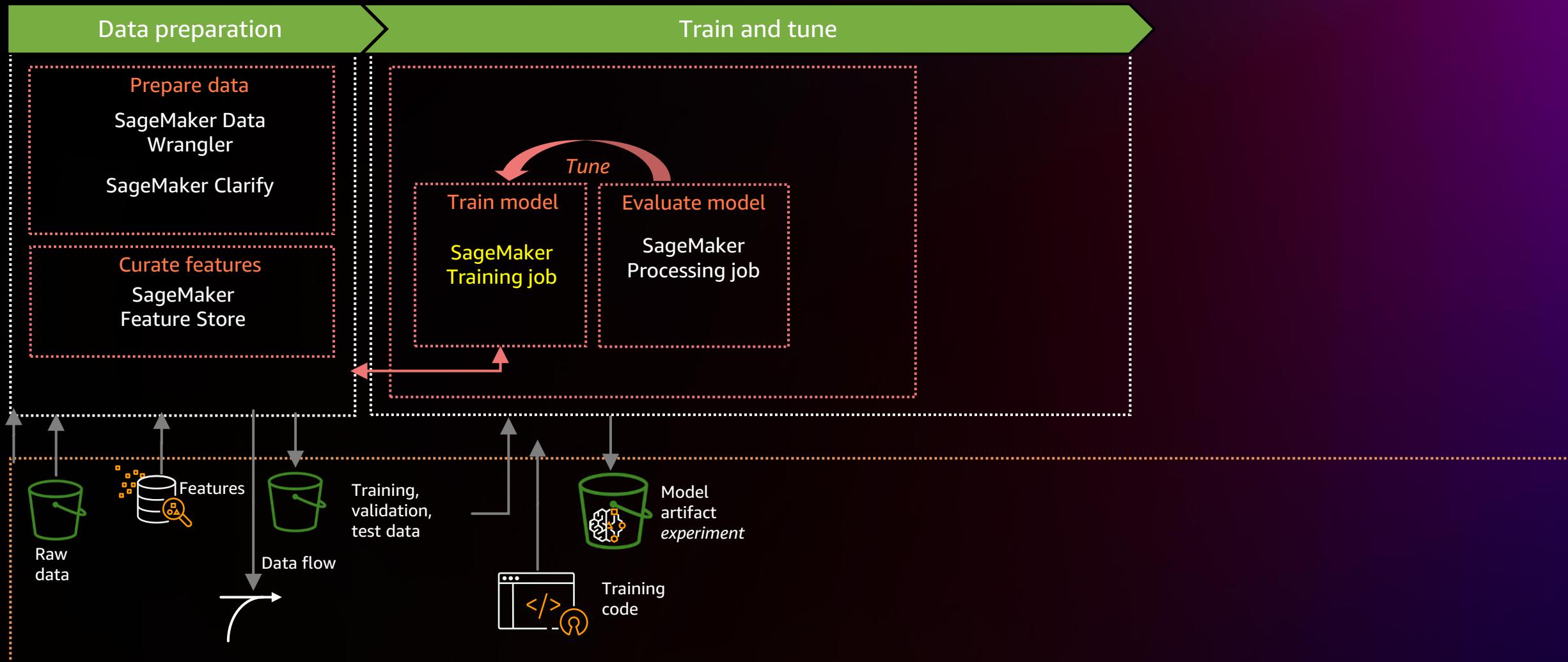
Experimentation to Production



Data preparation



Training and tuning



Model options

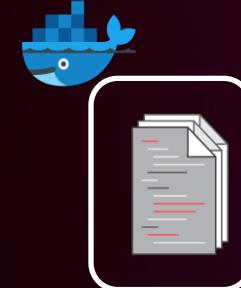
Built-in algorithms (17)
No ML coding required

- XGBoost - Gradient boosted trees
- Matrix factorization
- Regression
- Principal component analysis
- K-means clustering
- And more!

Bring your own script
Amazon SageMaker builds the container
Open-source containers



Bring your own container
Full control, you build the container
R, C++, etc.



AWS Marketplace
for
machine learning



Fully managed, distributed, auto scaled, secured

Training on SageMaker

```
# Input data from s3
inputs = {"train": s3_inputs_train, "test": s3_inputs_test}

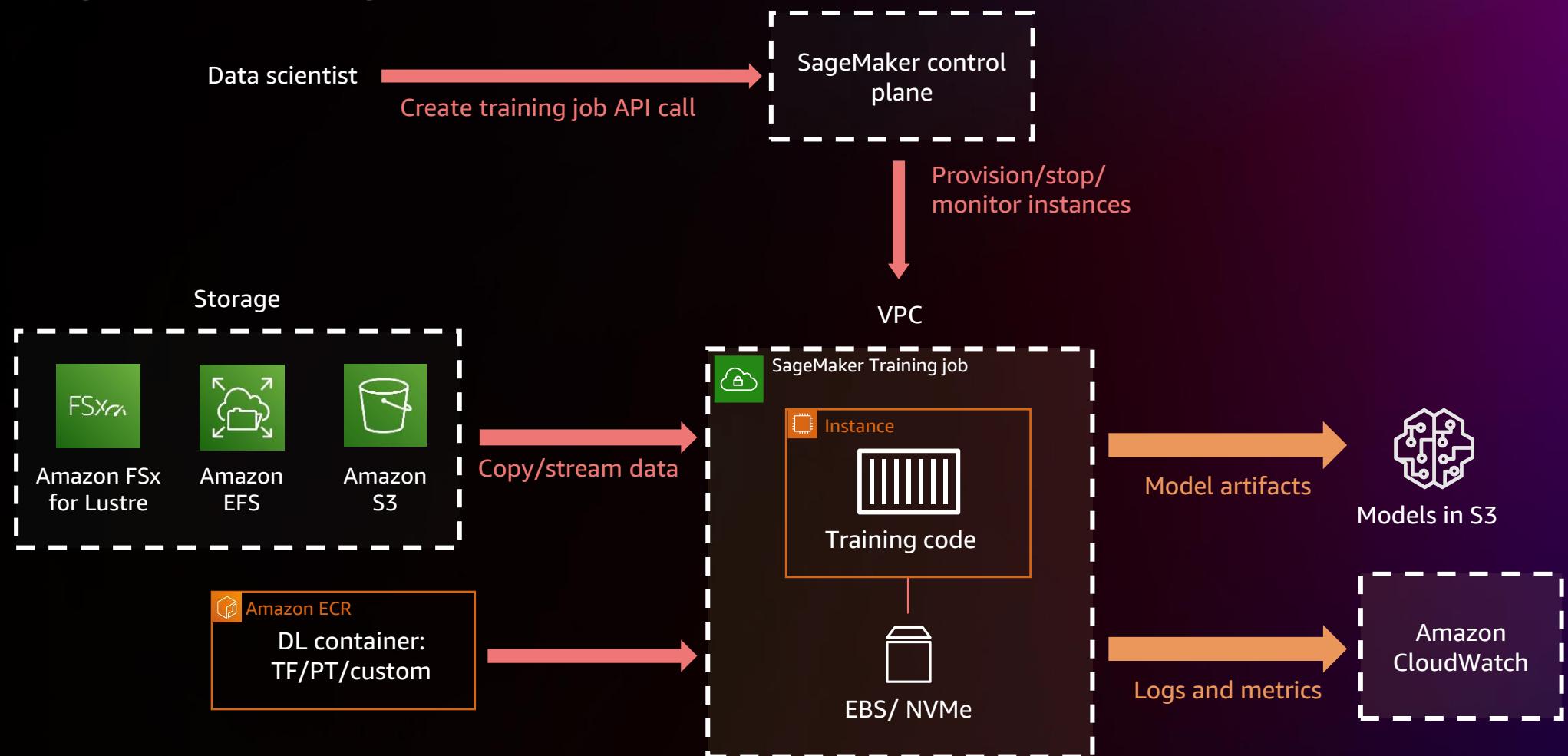
metric_definitions = [
    {"Name": "loss", "Regex": "loss: ([0-9\\.]+)"}, 
    {"Name": "accuracy", "Regex": "accuracy: ([0-9\\.]+)"}, 
    {"Name": "val_loss", "Regex": "val_loss: ([0-9\\.]+)"}, 
    {"Name": "val_accuracy", "Regex": "val_accuracy: ([0-9\\.]+)"}, 
]

# Create a TensorFlow Estimator
tf2_california_housing_estimator = TensorFlow(
    entry_point="california_housing_tf2.py",
    source_dir="code",
    role=sagemaker.get_execution_role(),
    instance_count=1,
    instance_type="ml.m5.large",
    framework_version="2.4.1",
    hyperparameters={
        "learning_rate": 0.1,
        "epochs": 100
    },
    py_version="py37",
    metric_definitions=metric_definitions,
    enable_sagemaker_metrics=True,
    tags=[{"Key": "trial-desc", "Value": trial_desc}],
)

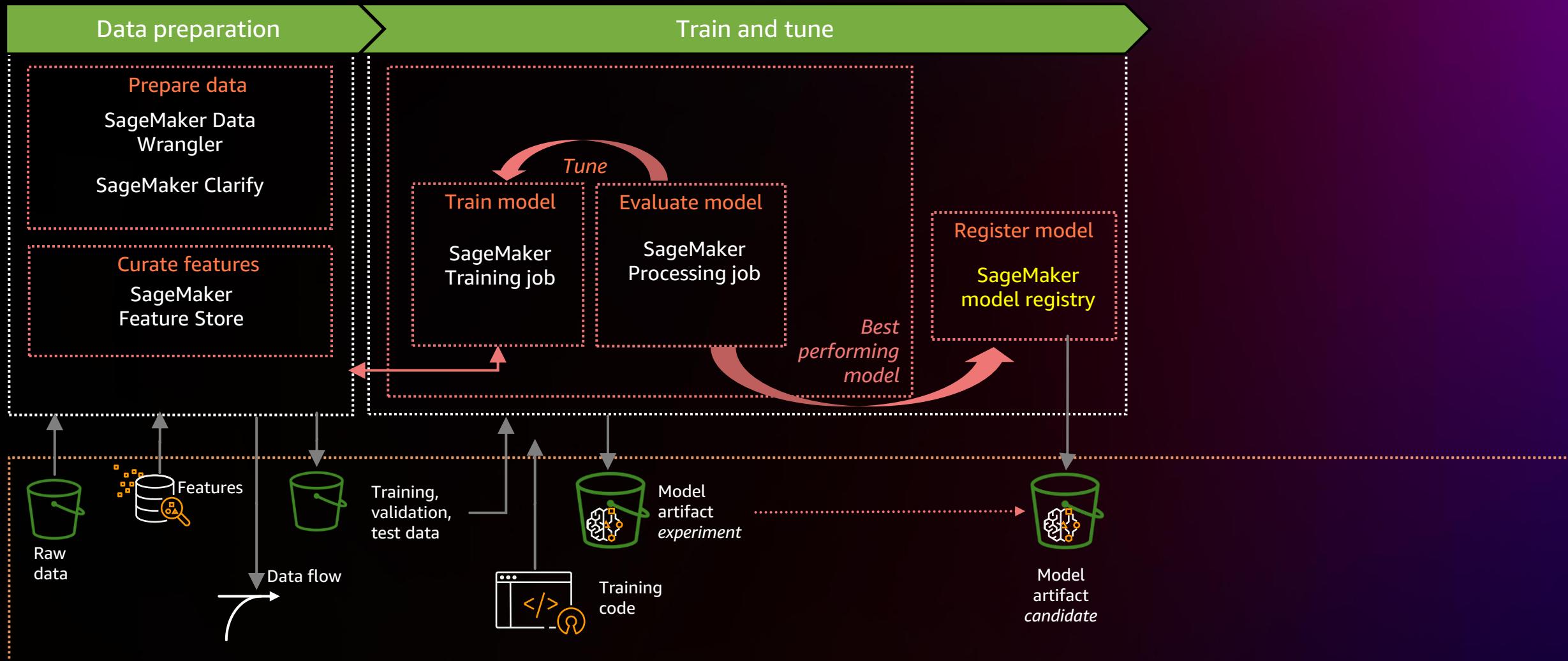
# Launch a training job
tf2_california_housing_estimator.fit( inputs, job_name=training_job_name)
```



Training on SageMaker



Register model



Amazon SageMaker Model Registry

fraud-detect-demo

Insurance claim fraud detection

Turbine

Versions **VERSION 13**

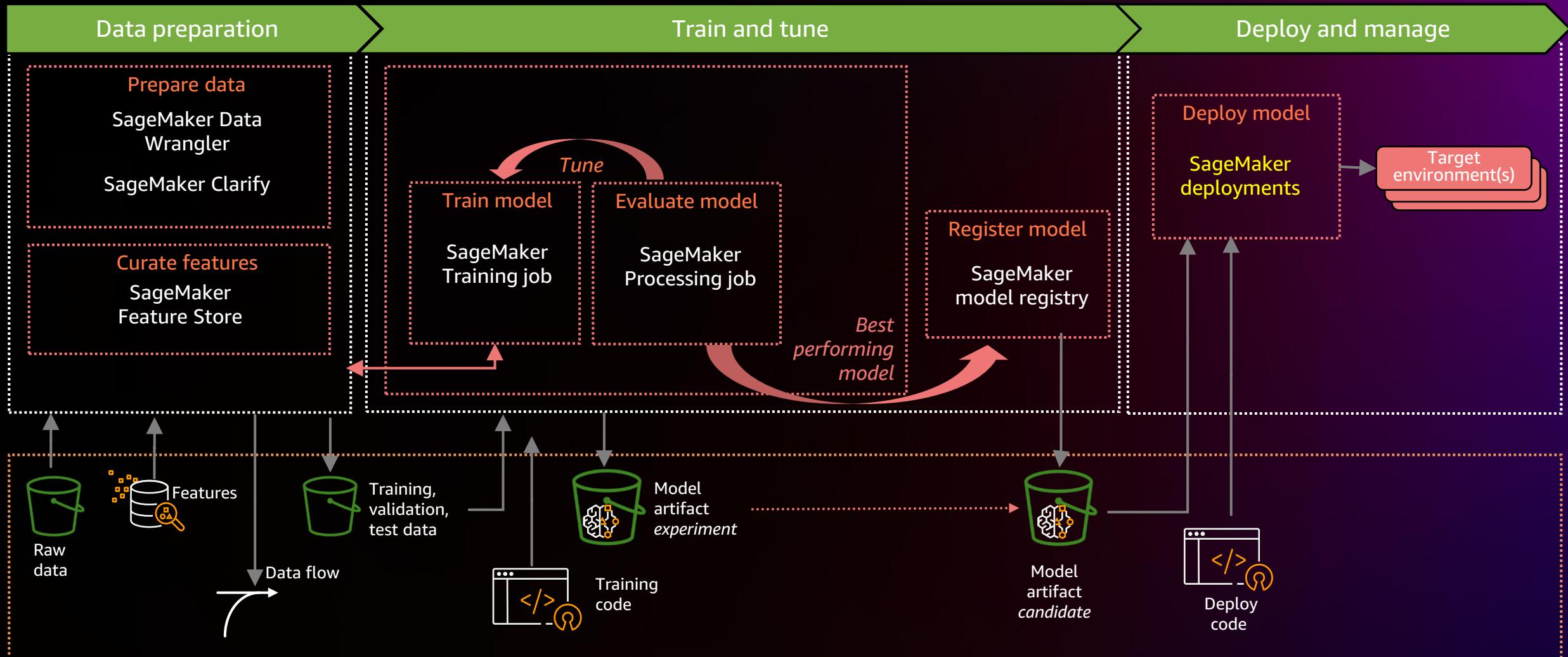
Search column

XGBoost classifier to detect insurance fraud.

Status **Approved** **Model group** **fraud-detect-demo** **Update status**

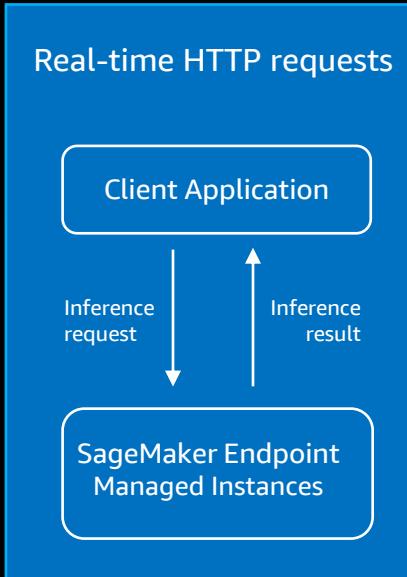
Version	Activity	Model quality	Explainability	Bias report	Inference recommender	Load test	Settings
13							
12							
11	Event type	Event	Comment	Modified by	Last modified		
10	ModelDeployment	Endpoint: fraud-...			7 days ago		
9	Approval	Status updated ...		pallavi	7 days ago		
8	Approval	Status updated ...			7 days ago		
7							
6	None	Approved	XGBoost classifier to ...	pallavi	1 month ago		
5	None	Approved	XGBoost classifier to ...	pallavi	1 month ago		
4	None	Approved	XGBoost classifier to ...	pallavi	1 month ago		

Deploy and manage

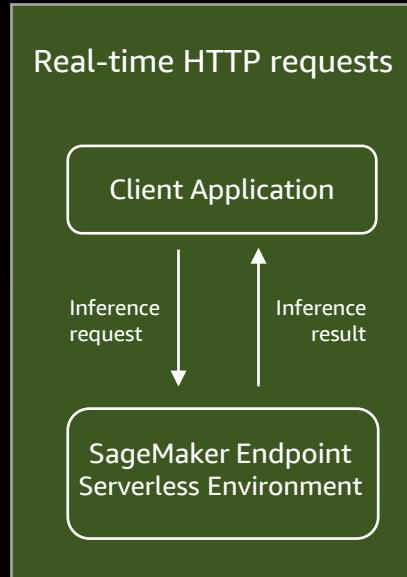


Amazon SageMaker model deployment modes

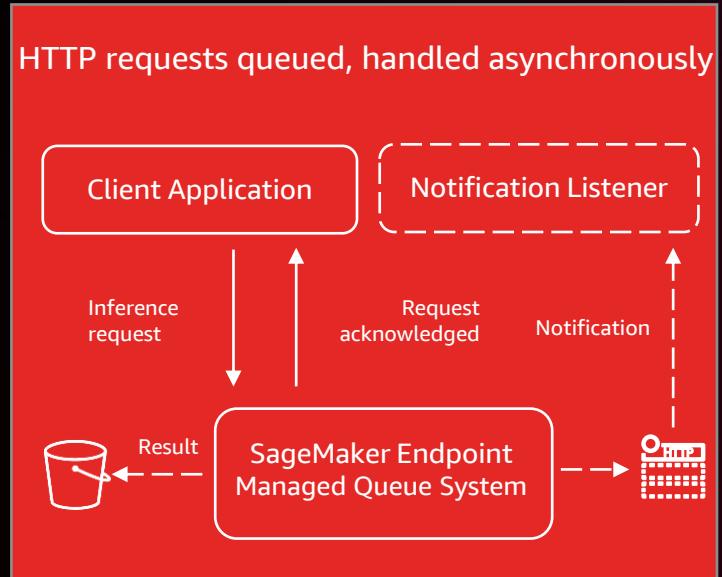
Real-time Inference



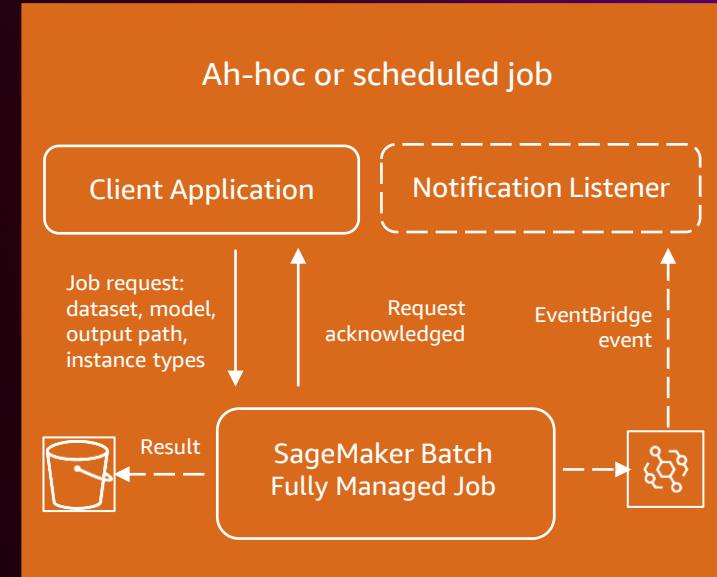
Serverless Inference



Asynchronous Inference

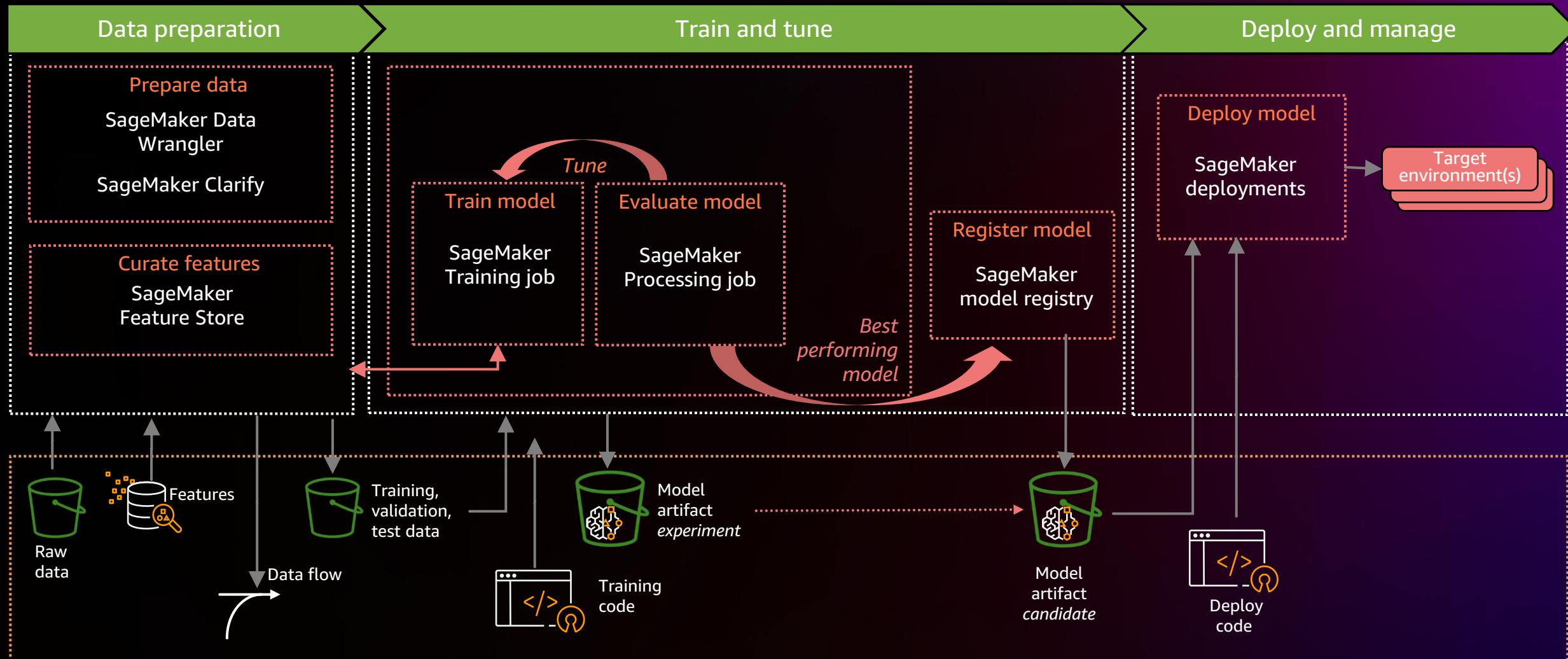


Batch Transform

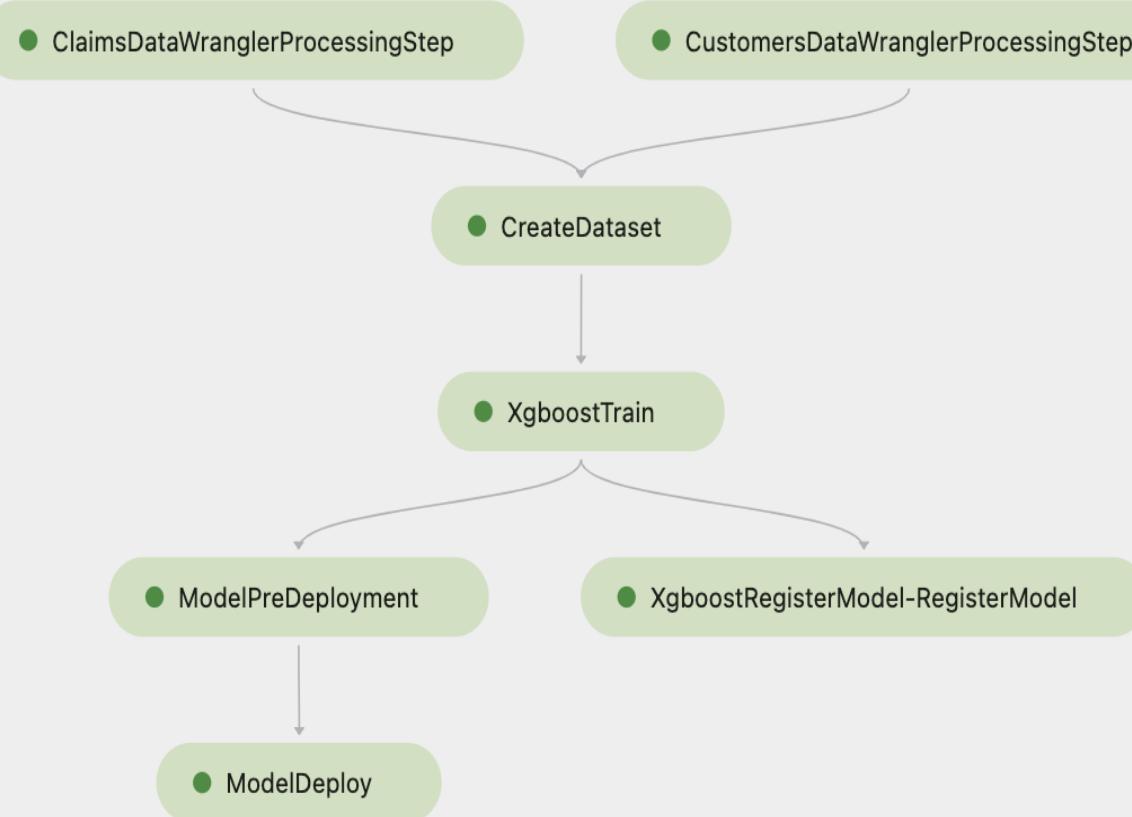


MLOps

MLOps - SageMaker Pipeline



SageMaker pipelines



Supported steps:

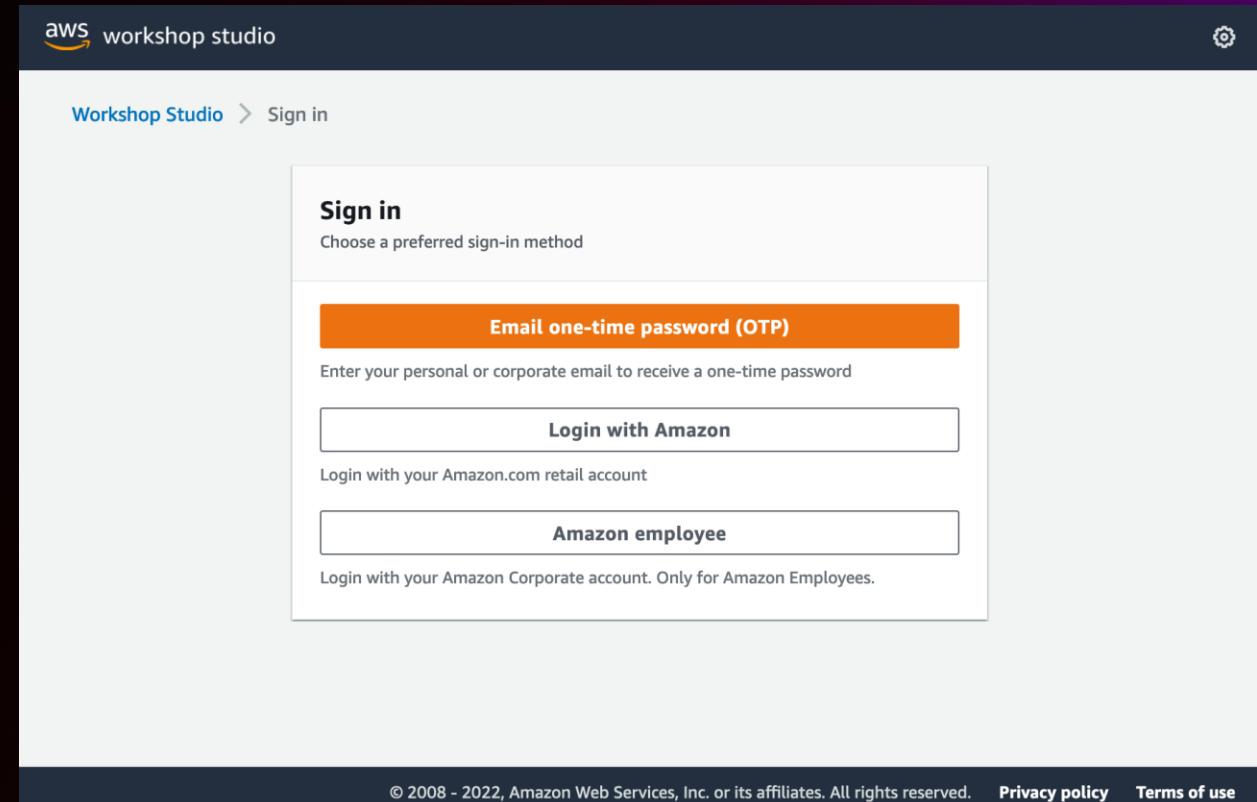
- **Processing:** Data processing and model evaluation
- **Training:** Model training using SageMaker Training jobs
- **Tuning:** Hyperparameter tuning jobs
- **Condition:** Conditional step execution
- **Transform:** Batch predictions
- **RegisterModel:** Create model package resource
- **Model/CreateModel:** Package model for deployment
- **Callback:** Incorporate additional tasks or AWS services into your workflow
- **AWS Lambda:** Run a Lambda function
- **ClarifyCheck:** Run baseline drift checks against previous baselines
- **QualityCheck:** Run baseline suggestions and drift checks against a previous baseline
- **Amazon EMR:** Run tasks on a running cluster

Getting started with this workshop

- As a participant, you will have access to an AWS account with any optional pre-provisioned infrastructure and IAM policies needed to complete this workshop.
- The AWS account will only be available for the duration of this workshop. You will lose access to the account thereafter.
- The optional pre-provisioned infrastructure will be deployed to a specific region. Check your workshop content to determine whether other regions will be used.
- Be sure to review the terms and conditions of the event. Do not upload any personal or confidential information in the account.

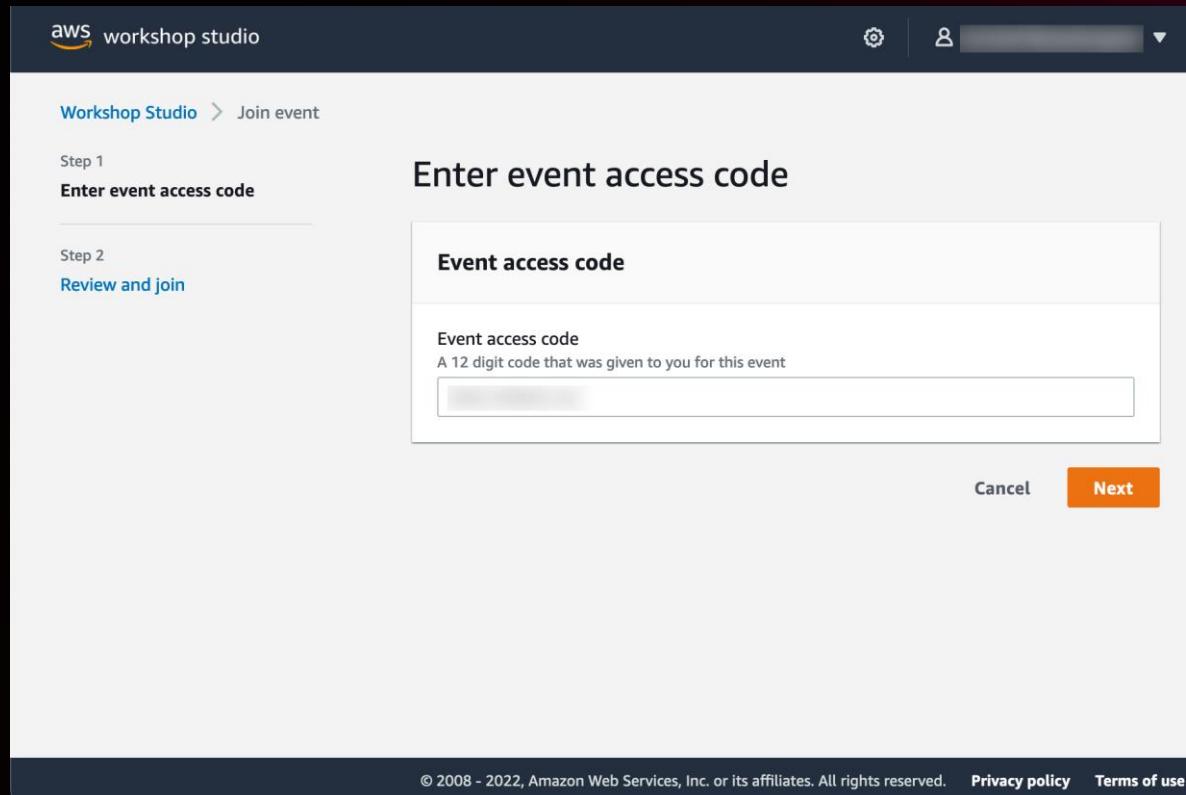
Step 1: Sign-In to AWS Account

bit.ly/STP303-2022

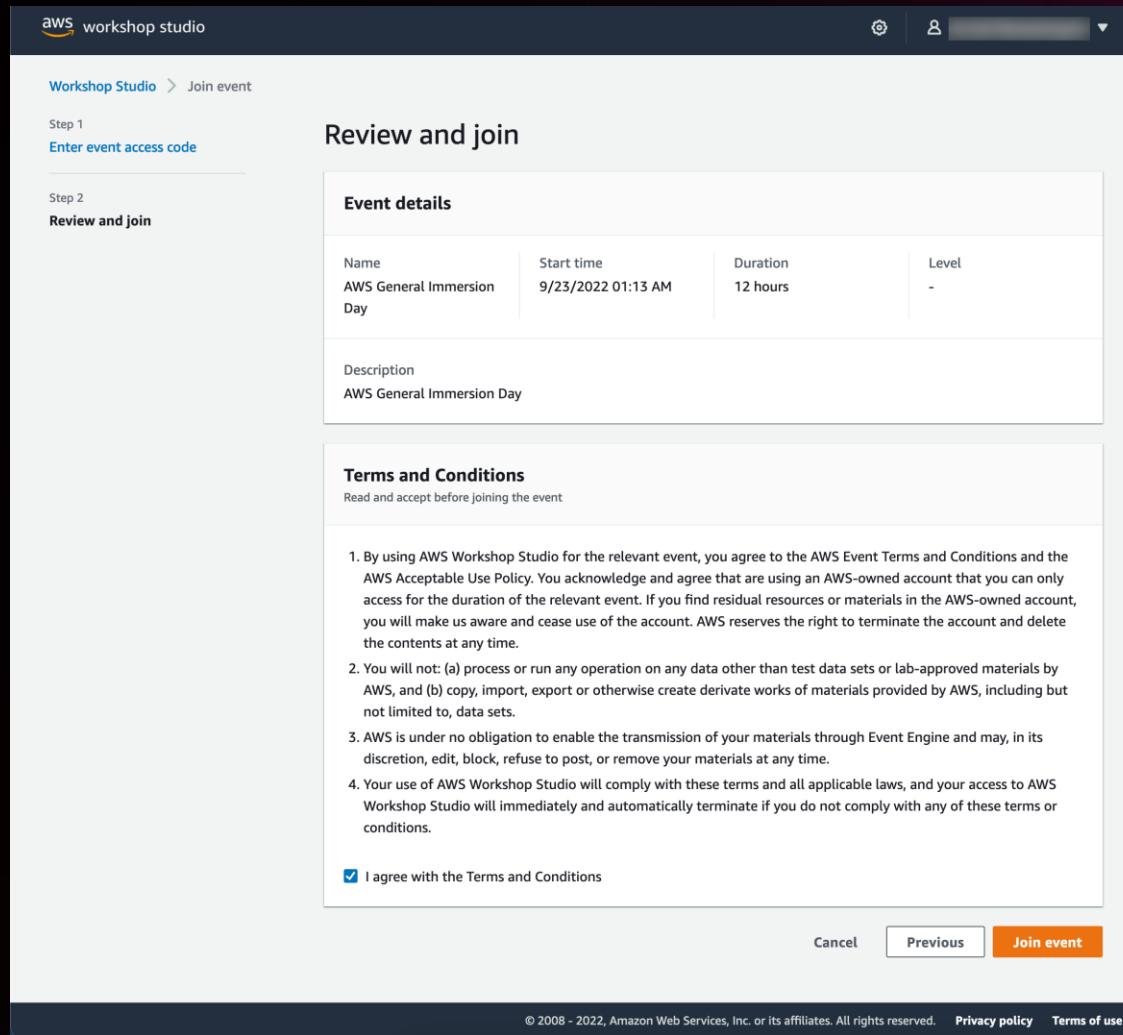


Step 2: Enter event access code

Enter 12 digit event access code.



Step 3: Review terms and join event



The screenshot shows the 'Review and join' step of the AWS Workshop Studio 'Join event' process. The interface is divided into two main sections: 'Event details' and 'Terms and Conditions'.

Event details

Name	Start time	Duration	Level
AWS General Immersion Day	9/23/2022 01:13 AM	12 hours	-

Description
AWS General Immersion Day

Terms and Conditions
Read and accept before joining the event

1. By using AWS Workshop Studio for the relevant event, you agree to the AWS Event Terms and Conditions and the AWS Acceptable Use Policy. You acknowledge and agree that are using an AWS-owned account that you can only access for the duration of the relevant event. If you find residual resources or materials in the AWS-owned account, you will make us aware and cease use of the account. AWS reserves the right to terminate the account and delete the contents at any time.
2. You will not: (a) process or run any operation on any data other than test data sets or lab-approved materials by AWS, and (b) copy, import, export or otherwise create derivative works of materials provided by AWS, including but not limited to, data sets.
3. AWS is under no obligation to enable the transmission of your materials through Event Engine and may, in its discretion, edit, block, refuse to post, or remove your materials at any time.
4. Your use of AWS Workshop Studio will comply with these terms and all applicable laws, and your access to AWS Workshop Studio will immediately and automatically terminate if you do not comply with any of these terms or conditions.

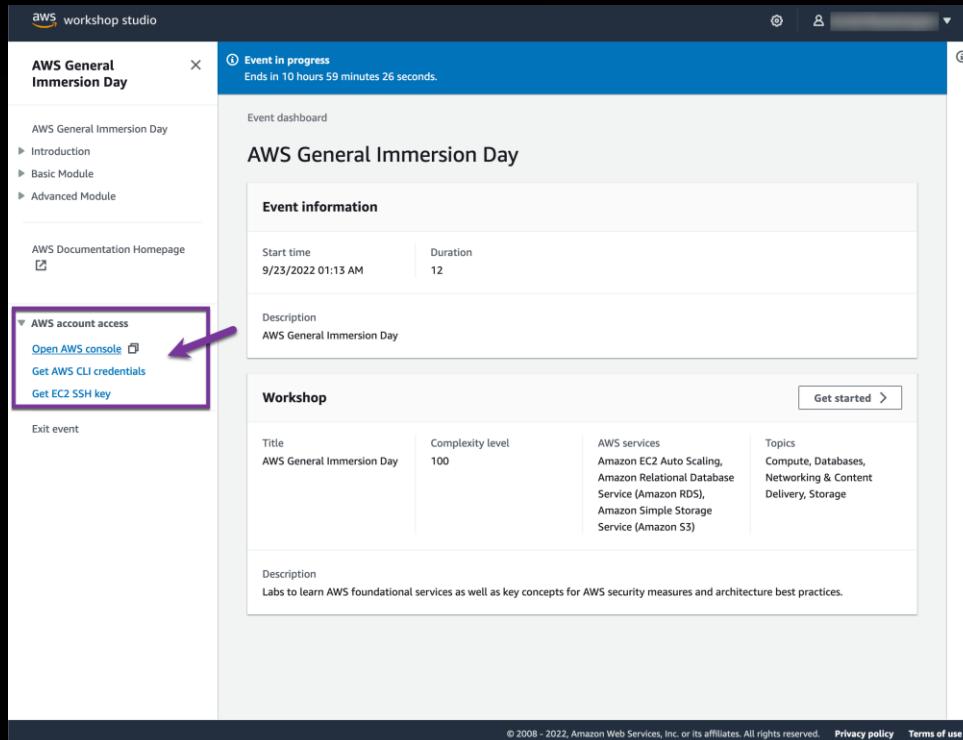
I agree with the Terms and Conditions

Cancel Previous **Join event**

© 2008 - 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy policy](#) [Terms of use](#)

Step 4: Access AWS account

Access the AWS Console, or generate AWS CLI credentials as needed



AWS General Immersion Day

Event in progress
Ends in 10 hours 59 minutes 26 seconds.

Event dashboard

AWS General Immersion Day

Event information

Start time	9/23/2022 01:13 AM	Duration	12
------------	--------------------	----------	----

Description
AWS General Immersion Day

Workshop

Title	AWS General Immersion Day	Complexity level	100
AWS services	Amazon EC2 Auto Scaling, Amazon Relational Database Service (Amazon RDS), Amazon Simple Storage Service (Amazon S3)	Topics	Compute, Databases, Networking & Content Delivery, Storage

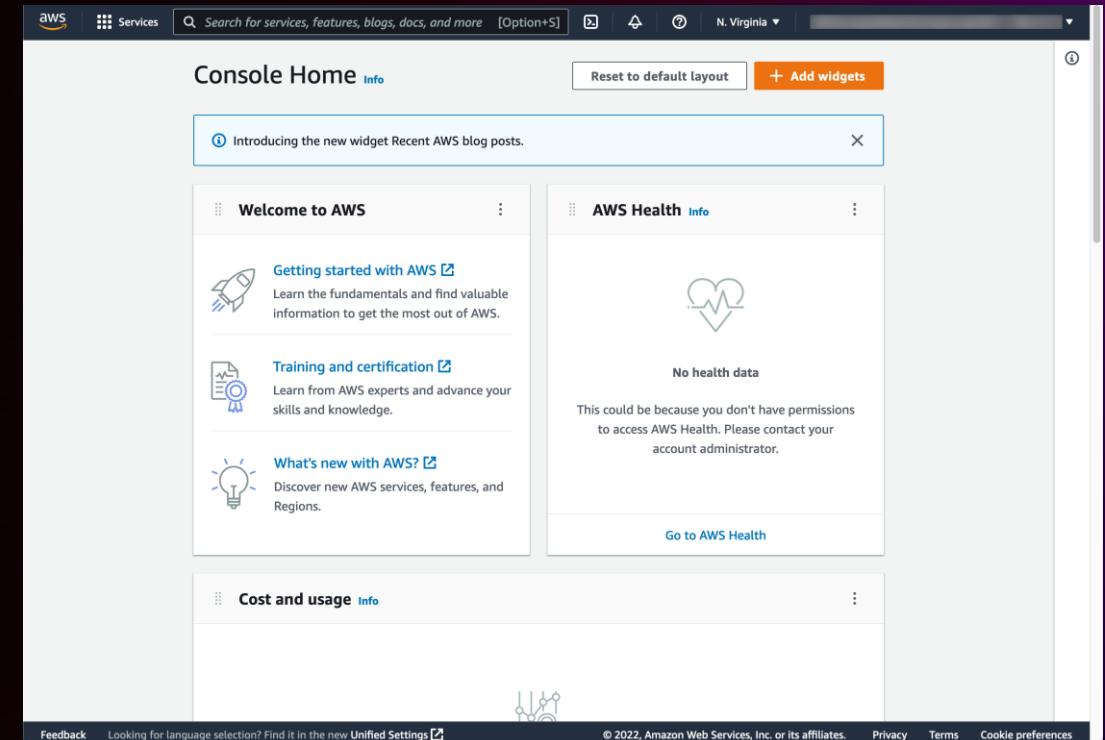
Get started >

AWS account access

- [Open AWS console](#)
- [Get AWS CLI credentials](#)
- [Get EC2 SSH key](#)

Exit event

© 2008 - 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy policy](#) [Terms of use](#)



Console Home

Search for services, features, blogs, docs, and more [Option+S]

Reset to default layout + Add widgets

Welcome to AWS

- Getting started with AWS** Learn the fundamentals and find valuable information to get the most out of AWS.
- Training and certification** Learn from AWS experts and advance your skills and knowledge.
- What's new with AWS?** Discover new AWS services, features, and Regions.

AWS Health No health data

This could be because you don't have permissions to access AWS Health. Please contact your account administrator.

Cost and usage

Feedback Looking for language selection? Find it in the new Unified Settings

© 2022, Amazon Web Services, Inc. or its affiliates. [Privacy](#) [Terms](#) [Cookie preferences](#)



Step 5: Get started with the workshop

AWS workshop studio

AWS General Immersion Day

Event in progress
Ends in 10 hours 59 minutes 26 seconds.

Event dashboard

AWS General Immersion Day

Event information

Start time	Duration
9/23/2022 01:13 AM	12

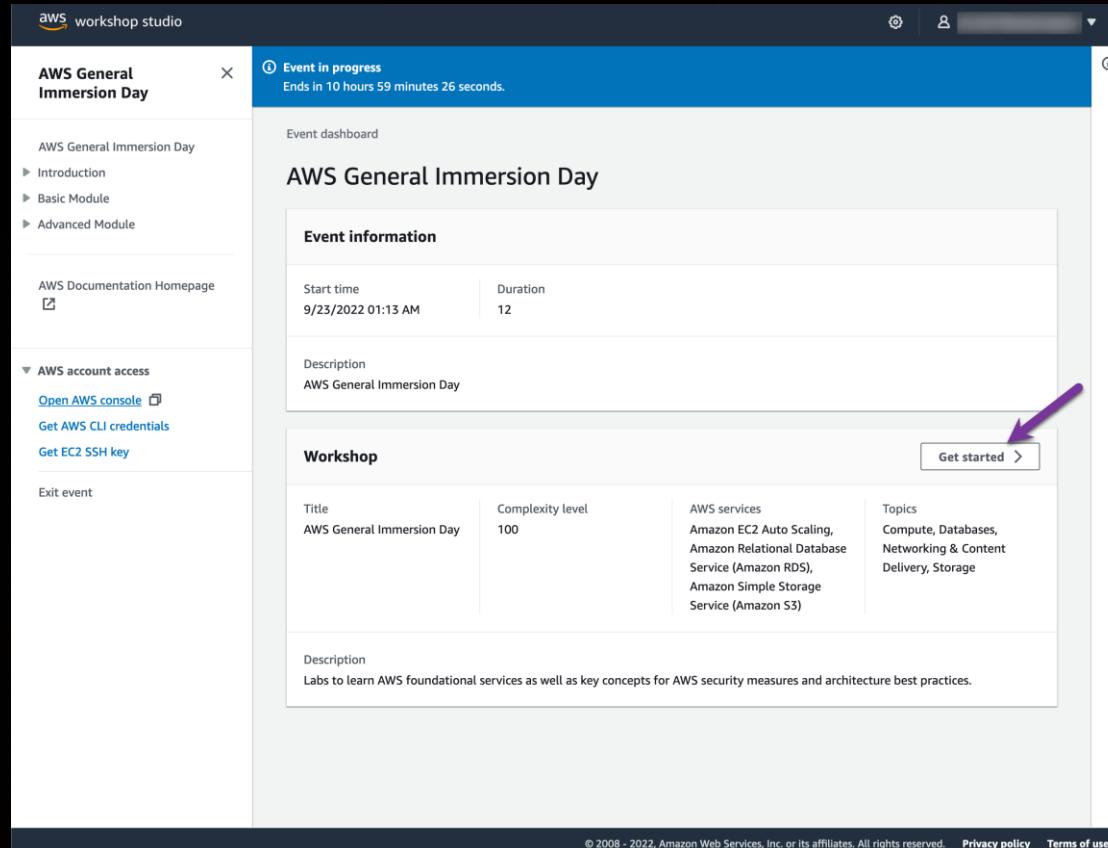
Description
AWS General Immersion Day

Workshop

Title	Complexity level	AWS services	Topics
AWS General Immersion Day	100	Amazon EC2 Auto Scaling, Amazon Relational Database Service (Amazon RDS), Amazon Simple Storage Service (Amazon S3)	Compute, Databases, Networking & Content Delivery, Storage

Description
Labs to learn AWS foundational services as well as key concepts for AWS security measures and architecture best practices.

Get started >



© 2008 - 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy policy](#) [Terms of use](#)



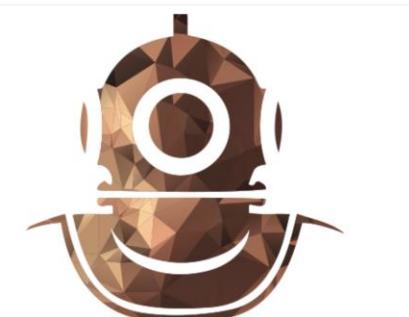
AWS workshop studio

AWS General Immersion Day

Event in progress
Ends in 10 hours 57 minutes 12 seconds.

Event dashboard > AWS General Immersion Day

AWS General Immersion Day



IMMERSION DAYS

In this General Immersion Day workshop, through a mix of service explanation and hands-on labs led by AWS, you will learn about AWS foundational services as well as key concepts for AWS security measures and architecture best practices.

The hands-on labs are largely divided into **basic** and **advanced** modules.

In basic modules, you can learn various features of each AWS foundational service. In advanced modules, you can learn how to connect each service organically to create architecture like 3-tier web application.

[Previous](#) [Next](#)

© 2008 - 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy policy](#) [Terms of use](#)



Thank you!

Manish Chugh

manchugh@amazon.com

Pallavi Nargund

nargundp@amazon.com



Please complete the session
survey in the **mobile app**