
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

AIM346

ML governance with Amazon SageMaker

Ankur Mehrotra

General Manager,
Amazon SageMaker
AWS

Sai Latha Rudrabhatla

Principal Product Manager,
Amazon SageMaker
AWS

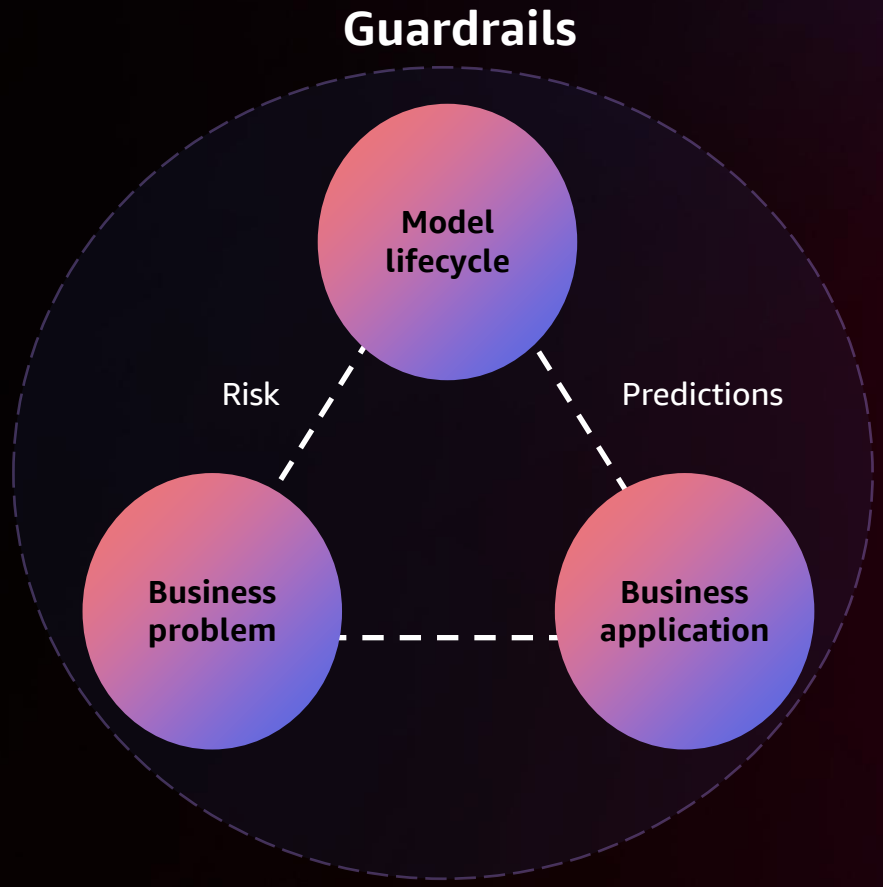
Domenic Centofanti

VP Group Squad Leader,
Data Science Platform
Fidelity Investments

aws

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

ML at an enterprise





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

Complexity with ML at an enterprise



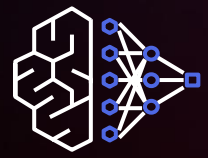
Types of activities

- Data preparation
- Model building
- Model training
- Hyperparameter tuning
- Deployment
- Building pipelines
- Monitoring



Types of users

- Business owner
- Compliance manager
- Risk officer
- Regulator
- Customer
- IT admin
- Data scientist
- Data engineer
- ML engineer



ML models



Regulations

aws

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

ML governance



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

Amazon SageMaker

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

PREPARE

Geospatial
Visualize geospatial data

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

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 set up 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

Shadow Testing
Validate model performance in production

Inference Recommender
Automatically select compute instance and configuration

Model Monitor
Maintain accuracy of deployed models

Kubernetes operators & components
Manage and monitor models on edge devices

Edge Manager
Manage and monitor models on edge devices

MLOps: Pipelines | Projects | Model Registry
Workflow automation, CI/CD for ML, central model catalog

Canvas
Generate accurate machine learning predictions – no code required

Studio | RStudio
Integrated development environment (IDE) for ML

Governance
Model Cards | Dashboard | Role Manager

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

ML governance with Amazon SageMaker

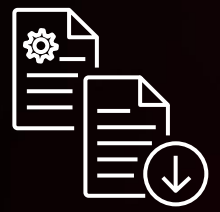
NEW!

GENERALLY AVAILABLE



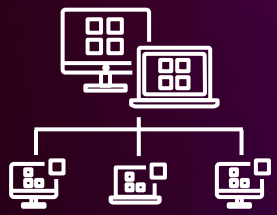
Amazon SageMaker Role Manager

Define custom permissions for SageMaker users in minutes



Amazon SageMaker Model Cards

Create a single source of truth for model information



Amazon SageMaker Model Dashboard

Audit performance and lineage of all your models, in one place

aws

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

Demo

Sai Latha Rudrabhatla

Principal Product Manager,
Amazon SageMaker
AWS

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

Demo scenario – SageMaker Role Manager

Alice is a data scientist at her company

Alice needs access to SageMaker Studio to build a new customer churn model for the Chief Marketing Officer

ML administrators want to ensure that Alice gets the permissions she needs without granting her access to resources and activities that aren’t necessary to do her job

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

Demo scenario – SageMaker Model Cards

The company requires various stakeholders, such as internal auditors, to approve the customer churn model that Alice and her team build before deploying it to production

To do so, during model development Alice needs to capture model information so that she can quickly generate a report once the model is ready for approvals

Alice wants to rely on the model information that is already available in SageMaker, such as training details and evaluation results, so that she can speed up the documentation process

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

Demo scenario – SageMaker Model Dashboard

Once approved, Alice deployed her customer churn model to an inference endpoint for use in marketing applications; there are other such models that marketing department uses for their daily business

The marketing department wants to stay informed on all the models that are being used for business applications and their performance

They also want to view alerts and lineage of a specific model and investigate if it is not performing as expected

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

Getting started

Control panel
Studio
Studio Lab
Canvas
RStudio

Admin console
SageMaker dashboard
Images
Lifecycle configurations
Search
Governance
Ground Truth
Notebook
Processing
Training
Inference
Edge Manager
Augmented AI
AWS Marketplace

Tutorials
Documentation

Jupyter Lab 3 will be the officially supported version in SageMaker Studio, while Jupyter Lab 1 will only receive security fixes from August 31, 2022. Learn more

Amazon SageMaker > Domains > Domain: default-1668112486067

Control Panel

Configure and manage SageMaker domain, users and apps.

Users

Add user

Search users

1

Name	Modified on	Created on

No users

To add a user, choose Add user and enter a user name.

Domain

Status
Ready

Authentication method
AWS Identity and Access Management (IAM)

Domain ID
d-tkhuwsahgsc1

Execution role
arn:aws:iam::364732211972:role/service-role/AmazonSageMaker-ExecutionRole-20221109T113409

Apps

Configure apps to manage your ML workflow using SageMaker.

Studio | Studio Lab | Canvas | RStudio

Studio

Amazon SageMaker Studio provides a single, web-based visual interface where you can perform all ML development steps, including building notebooks, experiment management, automatic model creation, debugging, and model and data drift detection.

Configure app | Learn more

Projects

Amazon SageMaker project templates enabled for this account

Launch constraint role: arn:aws:iam::364732211972:role/service-role/AmazonSageMakerServiceCatalogProductsLaunchRole
Product use role: arn:aws:iam::364732211972:role/service-role/AmazonSageMakerServiceCatalogProductsUseRole
API Gateway use role: arn:aws:iam::364732211972:role/service-role/AmazonSageMakerServiceCatalogProductsApiGatewayRole
CloudFormation use role: arn:aws:iam::364732211972:role/service-role/AmazonSageMakerServiceCatalogProductsCloudformationRole
CodeBuild use role: arn:aws:iam::364732211972:role/service-role/AmazonSageMakerServiceCatalogProductsCodeBuildRole
CodePipeline use role: arn:aws:iam::364732211972:role/service-role/AmazonSageMakerServiceCatalogProductsCodePipelineRole
Events use role: arn:aws:iam::364732211972:role/service-role/AmazonSageMakerServiceCatalogProductsEventsRole
Firehose use role: arn:aws:iam::364732211972:role/service-role/AmazonSageMakerServiceCatalogProductsFirehoseRole
Glue use role: arn:aws:iam::364732211972:role/service-role/AmazonSageMakerServiceCatalogProductsGlueRole
Lambda use role: arn:aws:iam::364732211972:role/service-role/AmazonSageMakerServiceCatalogProductsLambdaRole
SageMaker use role: arn:aws:iam::364732211972:role/service-role/AmazonSageMakerServiceCatalogProductsExecutionRole

Amazon SageMaker project templates enabled for Studio users

Execution role: arn:aws:iam::364732211972:role/service-role/AmazonSageMaker-ExecutionRole-20221109T113409

Images | Lifecycle configurations

Getting Started

Amazon SageMaker provides machine learning (ML) capabilities that are purpose-built for data scientists and developers to prepare, build, train, and deploy high-quality ML models efficiently.

Configure role

Establish roles and quickly configure permission policies for your users.

Create a role

Configure SageMaker Domain

Congratulations, your domain is set up.

Launch SageMaker

Learn by doing

Follow along the hands-on tutorials to learn how to use Amazon SageMaker to accomplish various machine learning lifecycle tasks, including data preparation, training, deployment, and MLOps.

Browse hands-on tutorials

Next step: Guides and resources

Get the most from SageMaker

After you've completed the setup steps, you can explore advanced features using the following resources.



Video tutorials

Watch instructional videos to learn more about SageMaker capabilities.

View tutorials



Developer resources

Watch instructional videos to learn more about SageMaker capabilities.

Explore resources



Full documentation

Get in-depth information about SageMaker features, capabilities, and integrations in our extensive documentation.

Read the docs

© 2022, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

customer-churn-....pdf customer-churn-....pdf Show All

Amazon SageMaker

Getting started

Control panel
- Studio
- Studio Lab
- Canvas
- RStudio

SageMaker dashboard
- Images
- Lifecycle configurations
- Search

Governance
- Model dashboard NEW
- Model cards NEW

Ground Truth

Notebook

Processing

Training

Inference

Edge Manager

Augmented AI

AWS Marketplace

Tutorials

Documentation

Getting Started

Amazon SageMaker provides machine learning (ML) capabilities that are purpose-built for data scientists and developers to prepare, build, train, and deploy high-quality ML models efficiently.

Configure role

Establish roles and quickly configure permission policies for your users.

Create a role

Configure SageMaker Domain

Congratulations, your domain is set up.

Launch SageMaker

Learn by doing

Follow along the hands-on tutorials to learn how to use Amazon SageMaker to accomplish various machine learning lifecycle tasks, including data preparation, training, deployment, and MLOps.

Browse hands-on tutorials

Next step: Guides and resources

Get the most from SageMaker

After you've completed the setup steps, you can explore advanced features using the following resources.



Video tutorials

Watch instructional videos to learn more about SageMaker capabilities.

View tutorials



Developer resources

Watch instructional videos to learn more about SageMaker capabilities.

Explore resources



Full documentation

Get in-depth information about SageMaker features, capabilities, and integrations in our extensive documentation.

Read the docs

© 2022, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences



UNITED

"With Amazon SageMaker's new governance capabilities, we have increased control and visibility over our machine learning models. We are saving significant amount of time and able to scale up."

Ashok Srinivas

Director of ML Engineering and Ops

aws

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



"Each of these contribute towards our model governance being sufficient to warrant the trust that our clients place in us as a financial service provider."

Dean Matter

ML Engineer

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

Spotlight: Fidelity Investments

Domenic Centofanti (he/him)

VP Group Squad Leader, Data Science Platform
Fidelity Investments

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

1946
Fidelity management & research company founded



1965
Fidelity purchases first computer



1995
Fidelity becomes an internet pioneer with first mutual fund to create a home page

2016
Fidelity's first product application deployed to the cloud

2019
Fidelity launches multi-cloud hybrid strategy

TODAY
4,300+ applications on public cloud

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

What are we sharing today?

How we use machine learning (ML) at Fidelity

Model governance

ML challenges

Addressing challenges

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

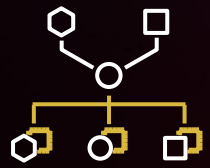
ML at Fidelity



100s of data scientists



100s of models in development/ training/production



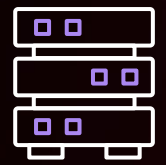
Multi-platform model development/training



Model governance automation



Algorithm review process spanning use case to deployment



Data across multiple AWS accounts with unique access



Numerous libraries and services

aws

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

My role at Fidelity

Scale and evolve machine learning to meet business use cases through SageMaker end-to-end adoption

Expand concept of “model governance” with accountability to wider set of stakeholders

- Business
- Architecture, data scientists, ML engineers
- Legal & compliance
- Algorithm review boards
- Consumers
- Model quality (entropy and bias)

Manage and enforce process from the model’s genesis to its ongoing sustainability

Evolve “paved paths” to reduce time to impact while scale ability to support data scientists

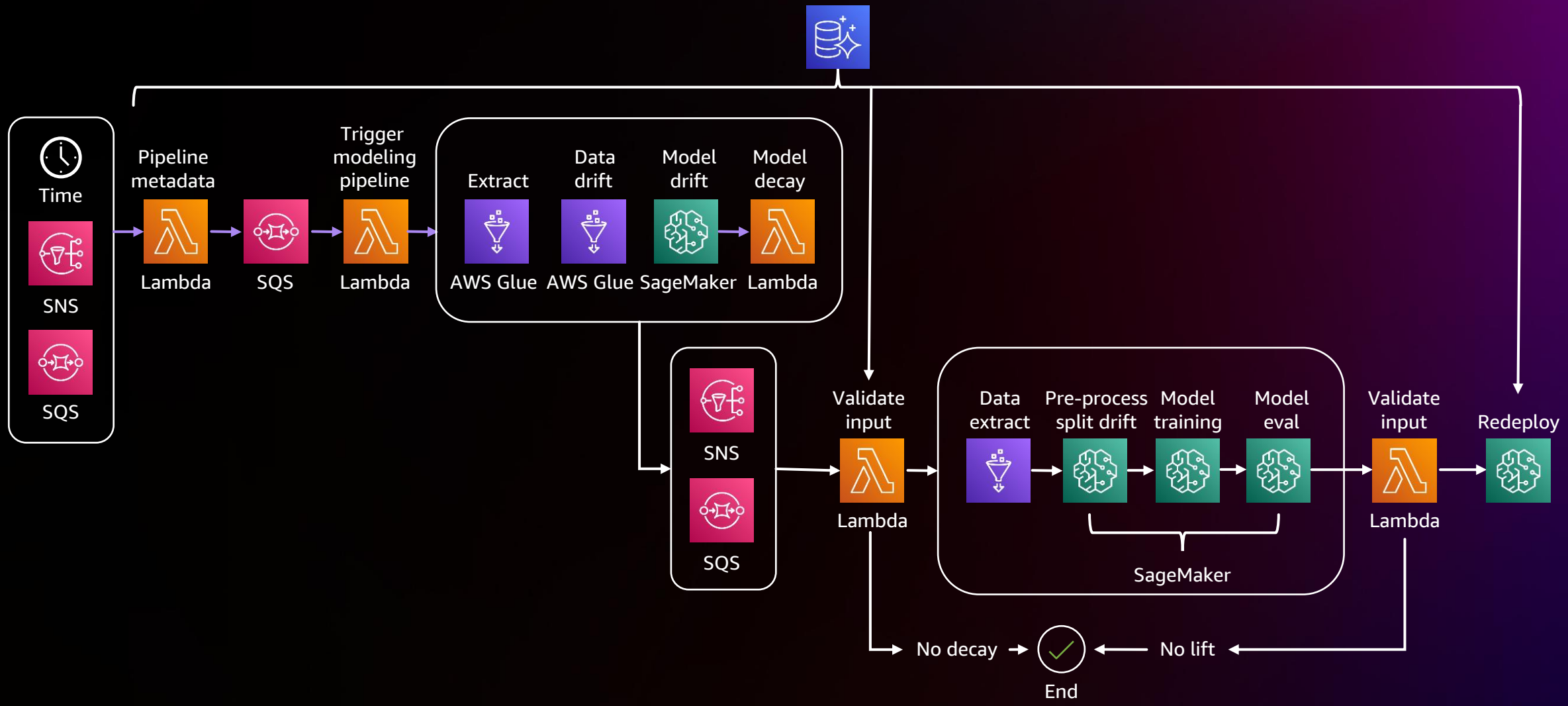
Develop installation ready AWS frameworks for other business units

Drive adoption through seamless and low resistance pipelines for development, training, and deployment

aws

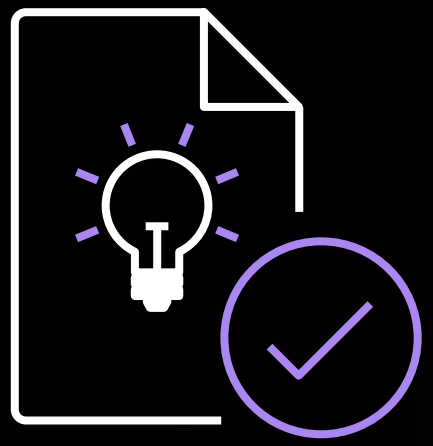
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Fully bespoke ML pipeline



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

ML use case



Large number of data scientists

Large number of models under development

Identify “paved paths” for particular types of models

- Onboarding
- NLP
- Video
- Online feature store integration
- Batch feature integration

Ability to monitor models by approach and components needed approach

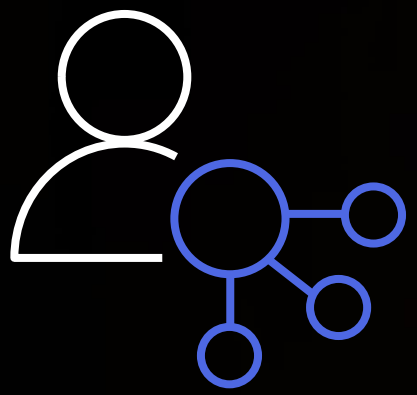
Build reusable frameworks and the ability to see models adhering to frameworks in SageMaker Model Dashboard

Govern and optimize SageMaker usage with Role Manager, Model Cards, Model Dashboard

aws

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

Addressing challenges: SageMaker Role Manager



Roles and policies

- Need for out-of-the box configuration capability customizable to data scientists' roles with least permission
- Intuitive workflows to assist data scientists to apply security constructs like AWS KMS, Amazon VPC, and session tags
- Tagging to apply like roles for data scientists with similar requirements
- Ability to manage user/role/custom policy to comply with security audits

aws

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

Addressing challenges: SageMaker Model Cards



Process flow

- With 100s of data scientists in various stages of deployment, ensuring the appropriate metadata is tracked throughout the development/training/process
- Ability to capture technical and business metadata (model name, location, business use case, algorithm, model development stage, etc.) is critical for review, support, and compliance
- Ability to add custom tags and links are critical
- Link to model and business artifacts providing single window to review status
- Post-production, provide contextual data

aws

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

Addressing challenges: SageMaker Model Dashboard



Ongoing support and quality

- With 100s of models in production and new use cases requesting new models, managing redundancy, quality and overlap of models are critical
- Ability to publish customizable model details from cards, for example lineage, is essential to governing the end-to-end process
- Identifying events that trigger a need for retraining
- History of alerts, execution and failures provides the basis for faster resolution and continuous improvement

aws

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

In summary

1. Evaluating governance capabilities provided by Amazon SageMaker to be the connective tissue across the model deployment pipeline to govern, support and monitor models for their performance and status

2. Evaluating integration with model governance capabilities

3. Evaluating ability to provide historical and current compliance information

4. Developing best practices for Role Manager, Model Cards, and Model Dashboard will take time and require evolution

aws

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

Get Started

Product webpage

Technical documentation

Upcoming Amazon SageMaker sessions

AIM341-R Transforming responsible AI from theory into practice (Dec 1, 2022 4:15 PM)

AIM401 Deploy deep learning models cost-effectively on GPUs with Amazon SageMaker (Dec 2, 2022 8:30 AM)

AIM326-R Prepare data and model features for ML with ease, speed, and accuracy (Dec 2, 2022 10:45 AM)

aws

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

Thank you!

Ankur Mehrotra

https://www.linkedin.com/in/mehrotraankur/

Sai Latha Rudrabhatla

https://www.linkedin.com/in/sailatharudrabhatla/

Domenic Centofanti

https://www.linkedin.com/in/domenico-centofanti-52a14b/

aws

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