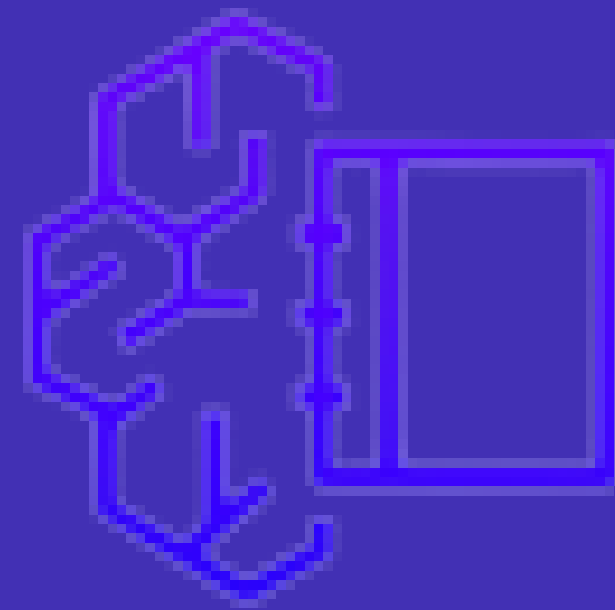# MLOPS WITH AWS SAGEMAKER

Session 5

Agenda

> Understand ML Governance

> Learn Sagemaker Model Monitoring

> How to Use Sagemaker Dashboard and Model Cards

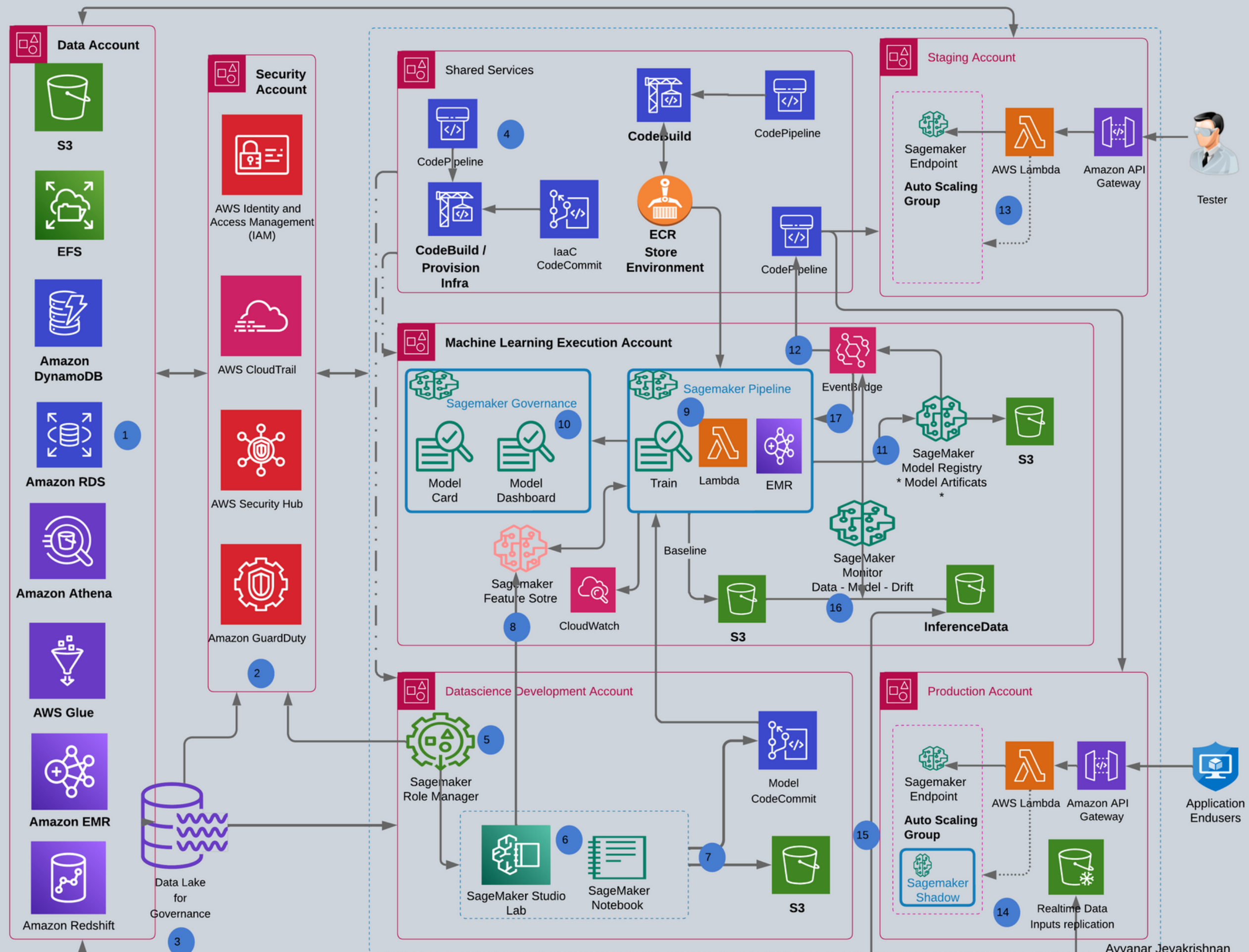> Demo - Model Monitoring, Model Dashboard and Cards

RECAP

- In previous sessions, We discuss about what is MLOps, MLOps Challenges, MLOps Benefits. and highLevel Overview of AWS Sagemaker.
- Demo -Use the terraform code to Deploy the Sagemaker Studio with User profiles and Spaces.
- Demo - Sagemaker Pipeline
- Demo - Sagemaker Studio - Projects, MLOps Templates
- Datawrangler, Dataflow, Feature store Overview
- Demo - Use DataWrangler transform Data to publish to Feature Store for ML Pipeline, Create a Model using AutoPilot and Deploy.
- Sagemaker Canvas and How to Create a Model using No-Code, Demo: Predict ITC Stock Price

- You can find the All Session Presentation and Recordings in GitHub
https://github.com/aws-data-usergroup-bangalore/sagemaker-mlops/tree/main

LET US RECALL OUR ROADMAP

Ayyanar Jeyakrishnan

ML GOVERNANCE

ACCOUNTABILITY - ML governance is the set of policies, processes, and procedures that ensure the responsible development, deployment, and use of machine learning systems.

ETHICS - It encompasses the ethical, legal, and technical aspects of machine learning

RISK MITIGATION - ML governance aims to mitigate risks associated with machine learning, such as algorithmic bias, privacy violations, and safety concerns.

FAIRNESS- Fairness in ML governance is critical to building trust in machine learning systems and ensuring that they operate in a way that aligns with ethical and legal standards,

TRANSPARENCY - Transparency in ML governance involves making the inputs, outputs, and algorithms of ML systems transparent and accessible, as well as explaining how decisions are made and how errors are detected and corrected. This allows stakeholders to identify biases, understand how decisions are made, and hold the system accountable for its actions.

AWS ML SAGEMAKER - GOVERNANCE TOOLS

Sagemaker Role Manager

AWS Sagemaker Role Manager is a feature that simplifies the management of AWS IAM roles for SageMaker resources.

Clarify

AWS Sagemaker Clarify is a tool that helps identify and mitigate bias in machine learning models.

Model Dashboard

Amazon SageMaker Model Dashboard is a portal accessible from the SageMaker console that allows you to track and explore all models in your account, including deployed models and their real-time performance when used for inference using Model Monitoring.

Model Cards

Use Amazon SageMaker Model Cards to document critical details about your machine learning (ML) models in a single place for streamlined governance and reporting

Presented by Ayyanar Jeyakrishnan - DataOps Labs User Group

SAGEMAKER ROLE MANAGER - MLOPS

ML activities (5 activities selected)

	Name	Description
☐	Access Required AWS Services	Permissions to access S3, ECR, Cloudwatch and EC2. Required for execution roles for jobs and endpoints.
☑	Run Studio Applications	Permissions to operate within a Studio environment. Required for domain and user-profile execution roles.
☐	Manage ML Jobs	Permissions to manage SageMaker jobs across their lifecycles.
☑	Manage Models	Permissions to manage SageMaker models and Model Registry.
☑	Manage Endpoints	Permissions to manage SageMaker Endpoint deployments and updates.
☑	Manage Pipelines	Permissions to manage SageMaker Pipelines and pipeline executions.
☐	Manage Experiments	Permissions to manage experiments and trials.
☑	Search and visualize experiments	Permissions to audit, query lineage and visualize experiments.
☐	Manage Model Monitoring	Permissions to manage monitoring schedules for SageMaker Model Monitor.
☐	S3 Full Access	Permissions to perform all S3 operations
☐	S3 Bucket Access	Permissions to perform operations on specified buckets.
☐	Query Athena Workgroups	Permissions to execute and manage Amazon Athena queries.
☐	Manage Glue Tables	Permissions to create and manage Glue tables for SageMaker Feature Store and Data Wrangler.

Persona

Data Scientist

Data Engineer*

MLOps

Sagemaker Compute

SAGEMAKER CLARIFY

Detect bias in your data and model

Explain model behavior

Identify imbalances in data

Understand your model

Check your trained model for bias

Monitor your model for changes in behavior

Monitor your model for bias

Explain individual model predictions

SAGEMAKER MODEL MONITORING

DATA QUALITY
- **Model Input and Output lie in appropriate range - Eg) Average Age of Human cannot be 200**
- **Values should not be Negative - E.g) Cricket Score**

MODEL QUALITY
- **Evaluated based on Problem and type of Model Used - Eg) Regression - RMSE, For Binary Classification - Confusion Matrix, Precision, Recall, F1**
- **Based on Problem statement, We use the Metrics - Eg) Dog in the Photo (Its not a High risk), At a same time Credit Card Fraud Detection. We need to monitor the Recall and using Model Monitor you can Configure alert if certain Threshold exceeds**

MODEL BIAS
- **A Model with Biased Predictions will lead to bad customer experience e.g) Loan Processing.**
- **We need to ensure that we have inclusivity in Data to train the model and not excluding any Candidates based on race, origin, etc.**
- **Model Monitor Alert when Bias Beyoid certain threashold**

FEATURE DRIFT
- **Different Features contribute to different extend to the Models output, We have seen this DataWrangler session which show the feature weightage on Model pridiction.**
- **E.g) House Prediction, Size and Location Near to Beach have high features.**
- **In case if there is commercial and sport complex added to some location, the price of the property will go up. If the attribution values goes beyond certails threshold, It need to be Alerted**

SAGEMAKER MODEL DASHBOARD

- ▶ Amazon SageMaker Model Dashboard is a centralized portal that allows users to view, search, and explore all the models in their account.

- ▶ The dashboard allows users to track which models are deployed for inference, if they are used in batch transform jobs or hosted on endpoints, and monitor their performance as they make real-time predictions on live data.

- ▶ The dashboard provides a comprehensive presentation of all monitor results, helping users quickly identify models that don't meet set metrics for data quality, model quality, bias, and explainability

- ▶ The Model Dashboard aggregates model-related information from several SageMaker features, such as model cards, workflow lineage, and endpoint performance, providing an out-of-the-box model governance solution for ensuring quality coverage across all models

- ▶ To use the Model Dashboard, users should have one or more models in their account. While it's not mandatory, they can gain the most value out of the dashboard if they set up model monitoring jobs using SageMaker Model Monitor for models deployed to endpoints.

Presented by Ayyanar Jeyakrishnan - DataOps Labs User Group

SAGEMAKER MODEL CARDS

Intended uses of a model
Specifying a model's intended uses helps ensure responsible model development and usage by describing appropriate and inappropriate scenarios, including use cases, assumptions, and considerations beyond technical details.

Risk Ratings
Model cards include a risk rating field to categorize a model's risk level as unknown, low, medium, or high to comply with rules and regulations when deploying the models with varying levels of risk.

Model card JSON schema
Evaluation details for a model card must be provided in JSON format, including metrics generated by SageMaker Clarify or SageMaker Model Monitor, and model content must be in the model card JSON schema as a string.

How to Integrate Sagemaker Model Monitoring, Model Dashboard and Cards inside the Sagemaker Pipeline

DEMO

Understand the Flow of Model Monitor

CHRUN PREDICATION USING XGBOOST MODEL → Deploy a Model Using Sagemaker Pipeline with Endpoint with Data Capture → Invoke the Model using the test data. It will capture all the input payload and store in S3 → Display the Capture Data in S3

MODEL MONITORING DATA QUALITY → Create a Baseline job with the training dataset. Create a Contraints and statistics of the features. → Create a Schedule (Hourly to check the statistics and Contraints. → Generate a Traffic, Compare the traffic data with the baseline data created → Generate and Validate the Violated Report