aws
INNOVATE
DATA AND AI/ML EDITION
22 February 2023

aws INNOVATE
DATA AND AI/ML EDITION

Beyond model development, training and deployment - Deep Dive on Amazon SageMaker Model Monitoring

Sahil Verma
Solutions Architect
AWS India

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

- What is Amazon SageMaker?
- Why do we need model monitoring? Why is it important?
- What is Amazon SageMaker Model Monitor and how it works?
- Types of model monitoring
- Deep Dive on Data Quality Monitoring
- Demo - Data Quality Monitoring
- How to get started with Amazon SageMaker Model Monitoring?

aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What is Amazon SageMaker?

Amazon SageMaker Studio

Access ML data
Connect to many data sources such as Amazon S3, Apache Spark, Amazon Redshift, CSV files, and more

Prepare data
Transform data to browse data sources, explore metadata, schemas, and write queries in popular languages

Build ML models
Optimized with 150+ popular open-source models and frameworks such as TensorFlow and PyTorch

Train and tune ML model
Correct performance problems in real time

Deploy and monitor results
Create, automate, and manage end-to-end ML workflows to improve model quality

aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What is Amazon SageMaker?

Amazon SageMaker Studio

Access ML data
Connect to many data sources such as Amazon S3, Apache Spark, Amazon Redshift, CSV files, and more

Prepare data
Transform data to browse data sources, explore metadata, schemas, and write queries in popular languages

Build ML models
Optimized with 150+ popular open-source models and frameworks such as TensorFlow and PyTorch

Train and tune ML model
Correct performance problems in real time

Deploy and monitor results
Create, automate, and manage end-to-end ML workflows to improve model quality

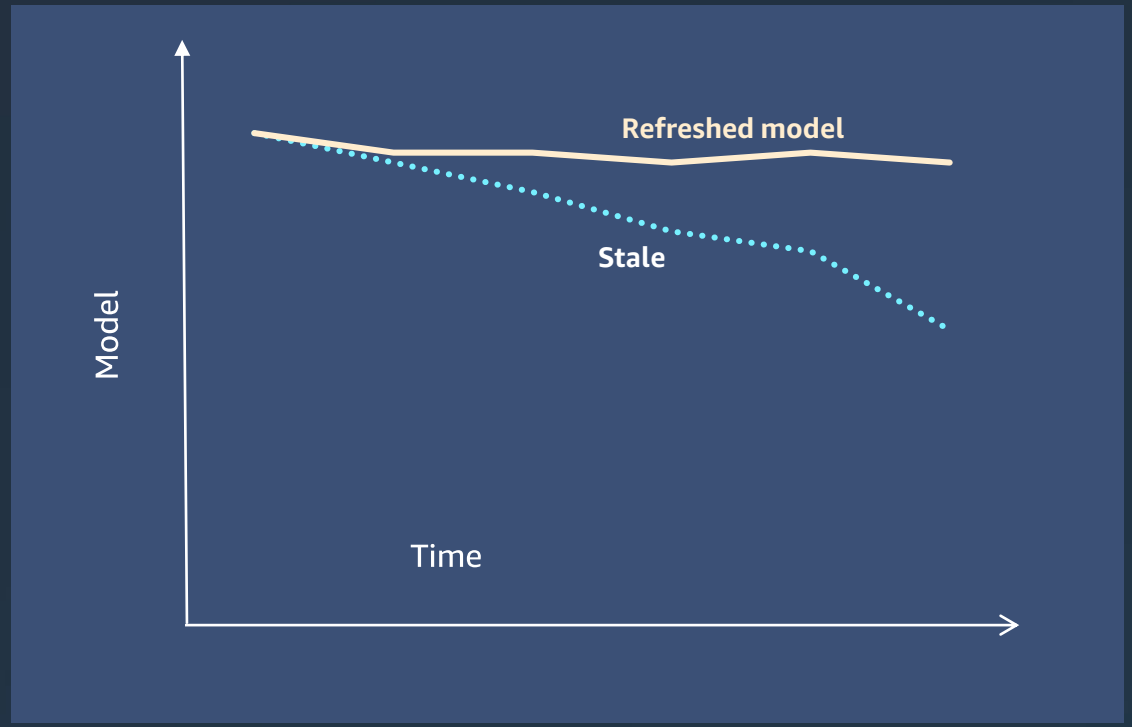
© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Model Development Life Cycle

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What happens after the model deployment ?

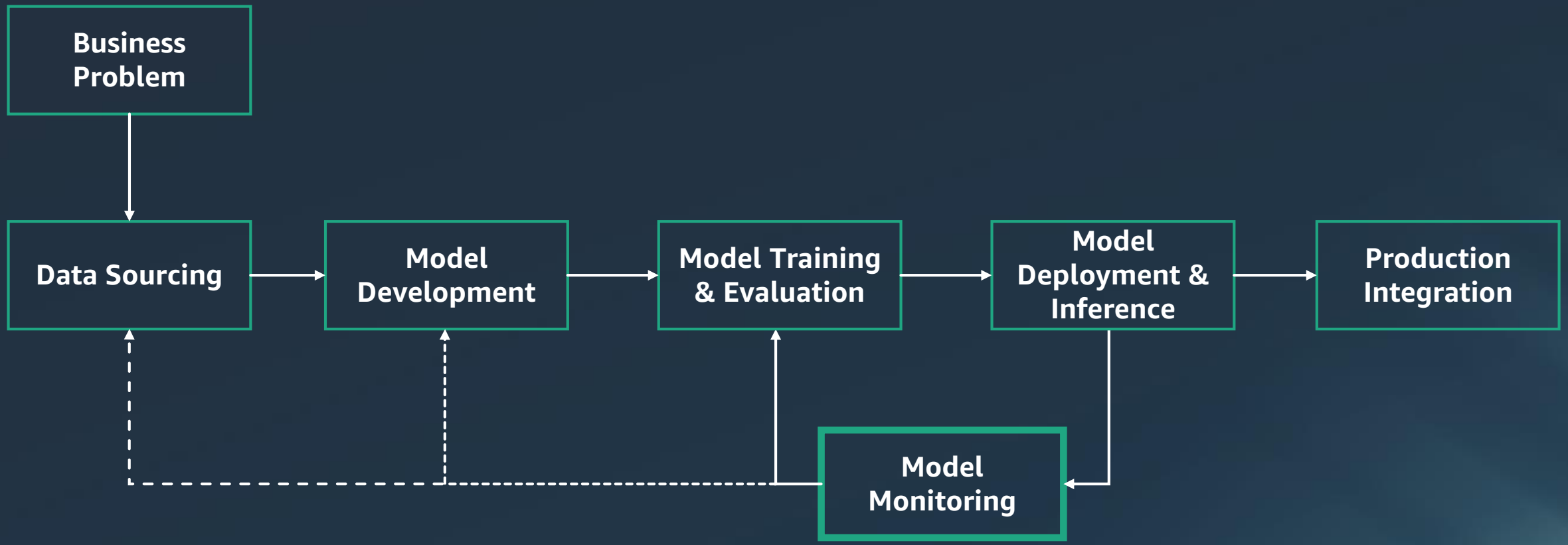
Deploying a model is not the end. You need to continuously monitor models in production and iterate.



- Model accuracy degrade over time
- Bias/change in feature attributions
- Concept drift due to divergence of data

aws
© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Model Development Life Cycle

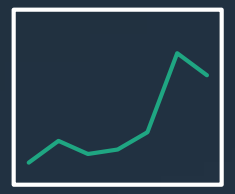


© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon SageMaker Model Monitor

Automatic data collection

Data collected from endpoints is stored in Amazon S3 Buckets



Continuous monitoring

Define a monitoring schedule and detect changes in quality against a pre-defined baseline

Flexible monitoring Rules

Use built-in rules to detect drift or write your own rules for custom analysis

Visual data analysis

See monitoring results, data statistics, and violation reports in Amazon SageMaker Studio; Analyze in Notebooks

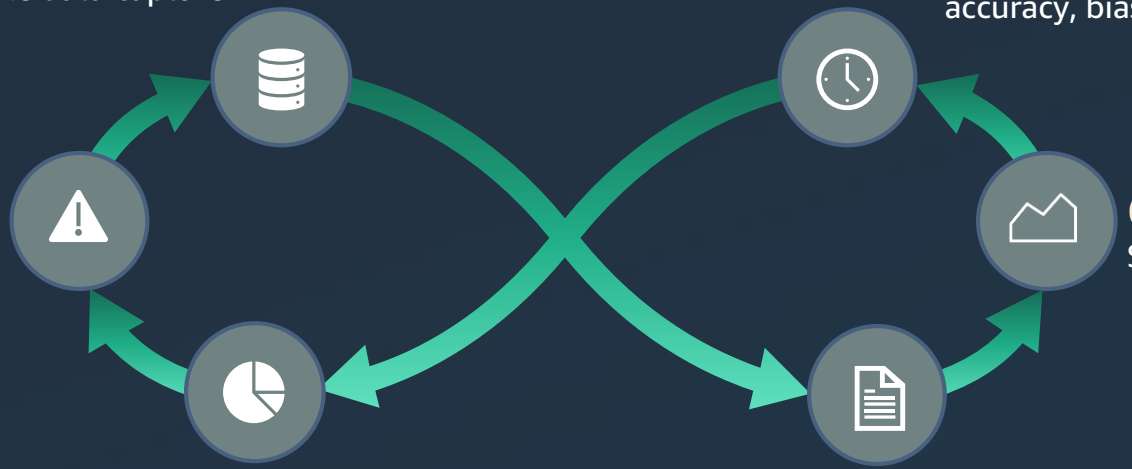
Amazon CloudWatch integration

Metrics emitted to Amazon CloudWatch make it easy to alarm and automate corrective actions

aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How does Amazon SageMaker Model Monitor work?



Deploy trained model
Enable data capture

Schedule monitoring jobs
Detect and monitor data quality, model accuracy, bias, and explainability drift

CloudWatch Alerts
Retrain models / Update data

Generate baseline
Statistics and constraints

View drift monitoring results
Reports in Amazon S3, visualize metrics in Amazon SageMaker Studio, analyze in Amazon Notebooks

Collect ground truth
Merge with predictions

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Monitoring types supported

Model Monitor supports monitoring and detection of following types of drift

Data Quality

Detect divergence in data

- Real time data capture from endpoints
- Define Baseline
- Pre-built container and custom analysis
- Type, Num Present, Num Missing
- Mean, Sum, Std Dev

Model Quality

Detect quality degradation over time

- Merge predictions with GroundTruth
- Compare predictions to GroundTruth
- Generate reports and violations
- MAE, RMSE, F1

Model Bias

Track model balance

- Overfitting
- Underfitting
- Class Imbalance
- DPL
- KL Divergence
- LP-Norm

Model Explainability

Feature attribution

- How much each feature contributed to predictions
- Shapley Values

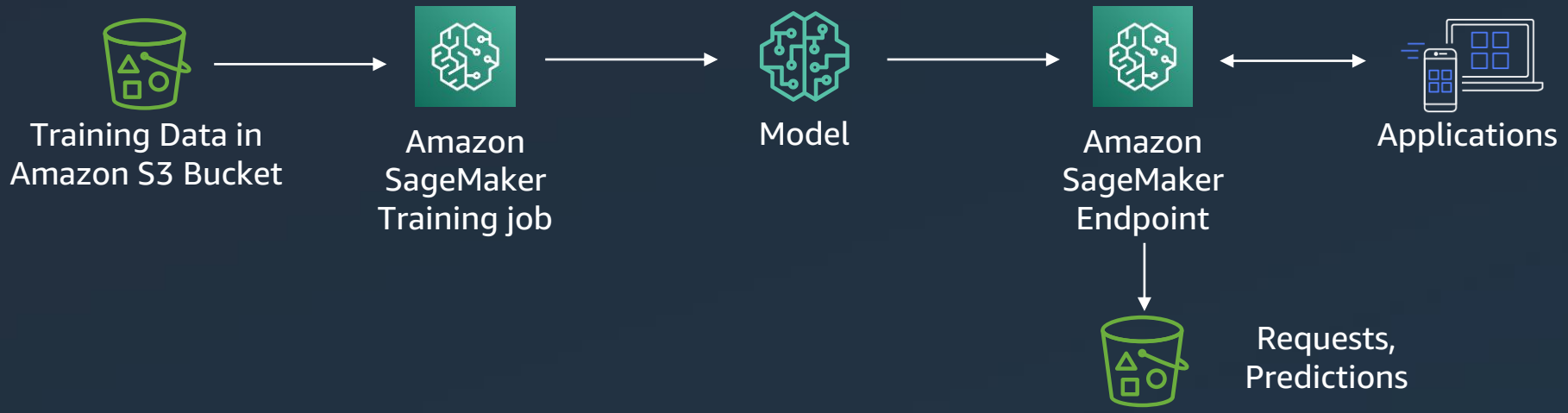
aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data Quality Monitoring

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Step 1 - Deploy Model and Configure Data Capture



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Step 1 - Deploy Model and Configure Data Capture

```
from sagemaker.model_monitor import DataCaptureConfig

endpoint_name = 'DEMO-xgb-churn-pred-model-monitor-' + strftime("%Y-%m-%d-%H-%M-
%S", gmtime())

data_capture_config = DataCaptureConfig(
                        enable_capture=True,
                        sampling_percentage=100,
                        destination_s3_uri=s3_capture_upload_path)

predictor = model.deploy(initial_instance_count=1,
                instance_type='ml.m4.xlarge',
                endpoint_name=endpoint_name,
                data_capture_config=data_capture_config)
```

aws

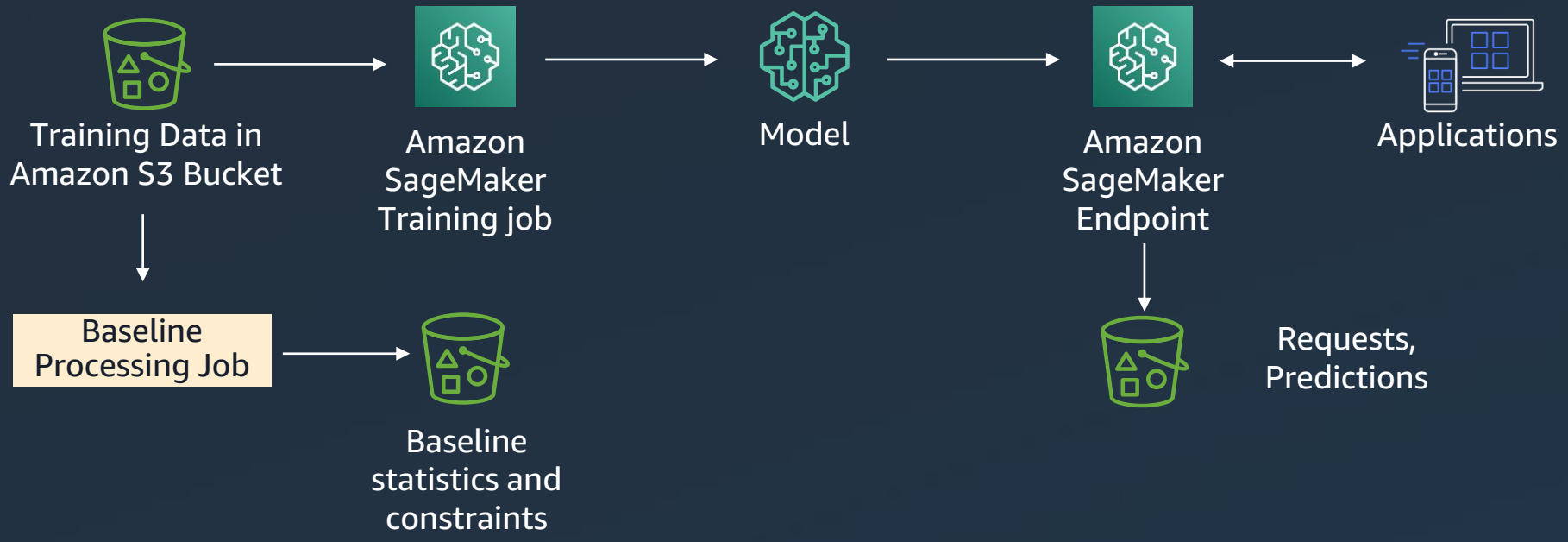
© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Step 1 - Data Capture – Input and Output


```
{
  "captureData": {
    "endpointInput": {
      "observedContentType": "text/csv",
      "mode": "INPUT",
      "data":
"92,0,176.3,85,93.4,125,207.2,107,9.6,1,2,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,1,0",
      "encoding": "CSV"
    },
    "endpointOutput": {
      "observedContentType": "text/csv; charset=utf-8",
      "mode": "OUTPUT",
      "data": "0.039806101471185684",
      "encoding": "CSV"
    }
  },
  "eventMetadata": {
    "eventId": "81f77a8a-314e-4c95-8908-189d5c9656e0",
    "inferenceTime": "2022-08-11T13:18:40Z"
  },
  "eventVersion": "0"
}
```


© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Step 2 – Create Baseline



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Step 2 – Create Baseline

A *constraints.json* file is used to express the constraints that a dataset must satisfy.

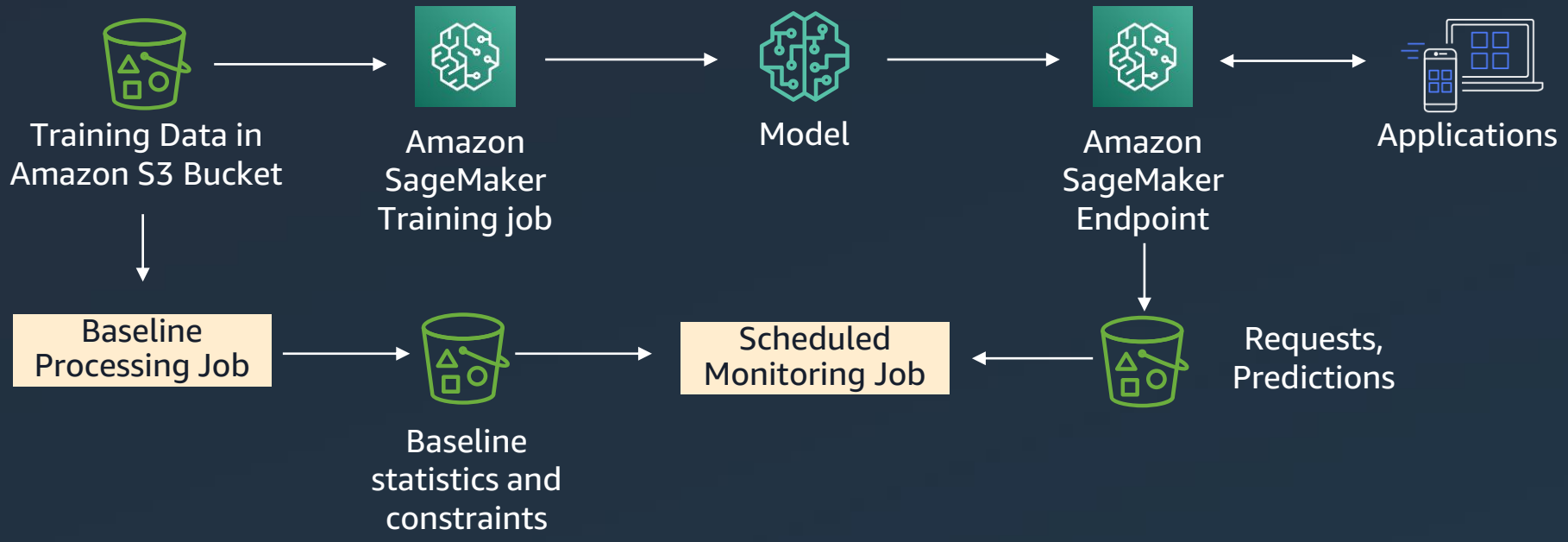

```
{
    "name" : "Night Calls",
    "inferred_type" : "Integral",
    "completeness" : 1.0,
    "num_constraints" : {
      "is_non_negative" : true
    }
  }
```


The schema defined in the *statistics.json* file specifies the statistical parameters to be calculated for the baseline and data that is captured.


```
{
    "name" : "Night Calls",
    "inferred_type" : "Integral",
    "numerical_statistics" : {
      "common" : {
        "num_present" : 2333,
        "num_missing" : 0
      },
      "mean" : 100.22717531075868,
      "sum" : 233830.0,
      "std_dev" : 19.28202933863294,
      "min" : 42.0,
      "max" : 175.0,
      "distribution" : { …
```


© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Step 3 – Schedule Monitoring Job



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

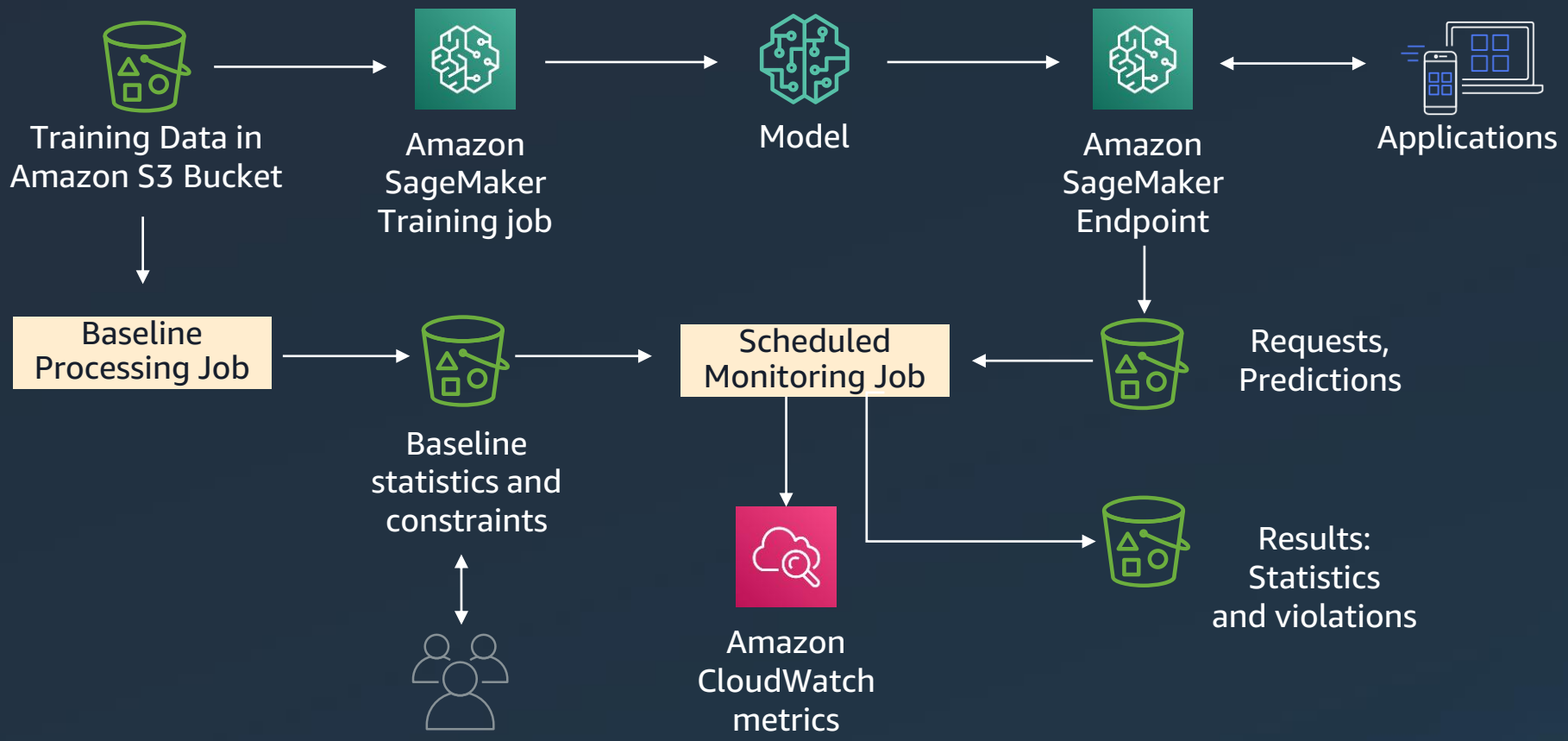
Step 3 – Schedule Monitoring Job

1. Amazon Model Monitor runs a **ProcessingJob** on your behalf at the schedule you select –i.e. Monitoring Jobs
2. It analyzes the data collected using your choice of analysis container (pre-built or custom)
3. It compares results against the **baseline**
4. You can generate results for each **Monitoring Job**
    - **Violations report** for each job in Amazon S3
    - Statistics report for data collected during the run
    - Emits summary metrics and statistics to Amazon CloudWatch

aws

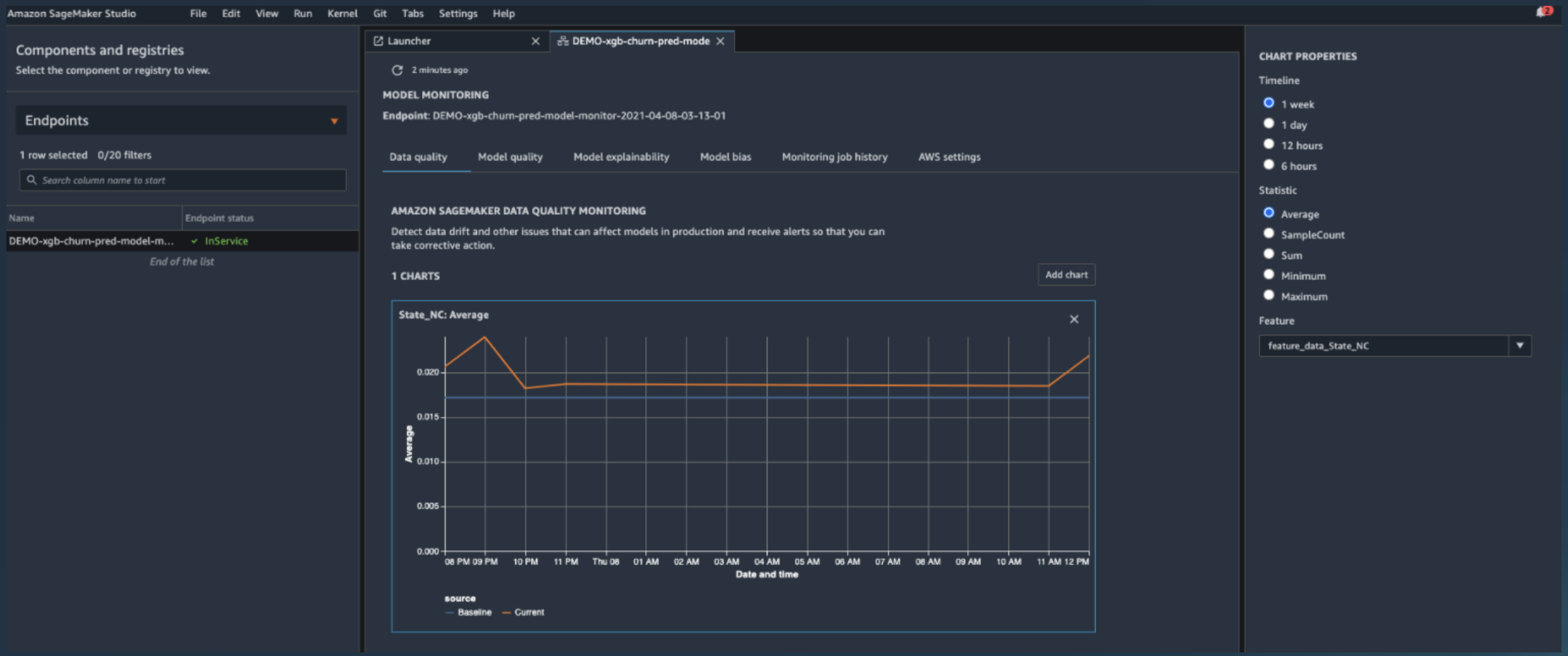
© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Step 4 – View Monitoring Results



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Step 4 – View Monitoring Results



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

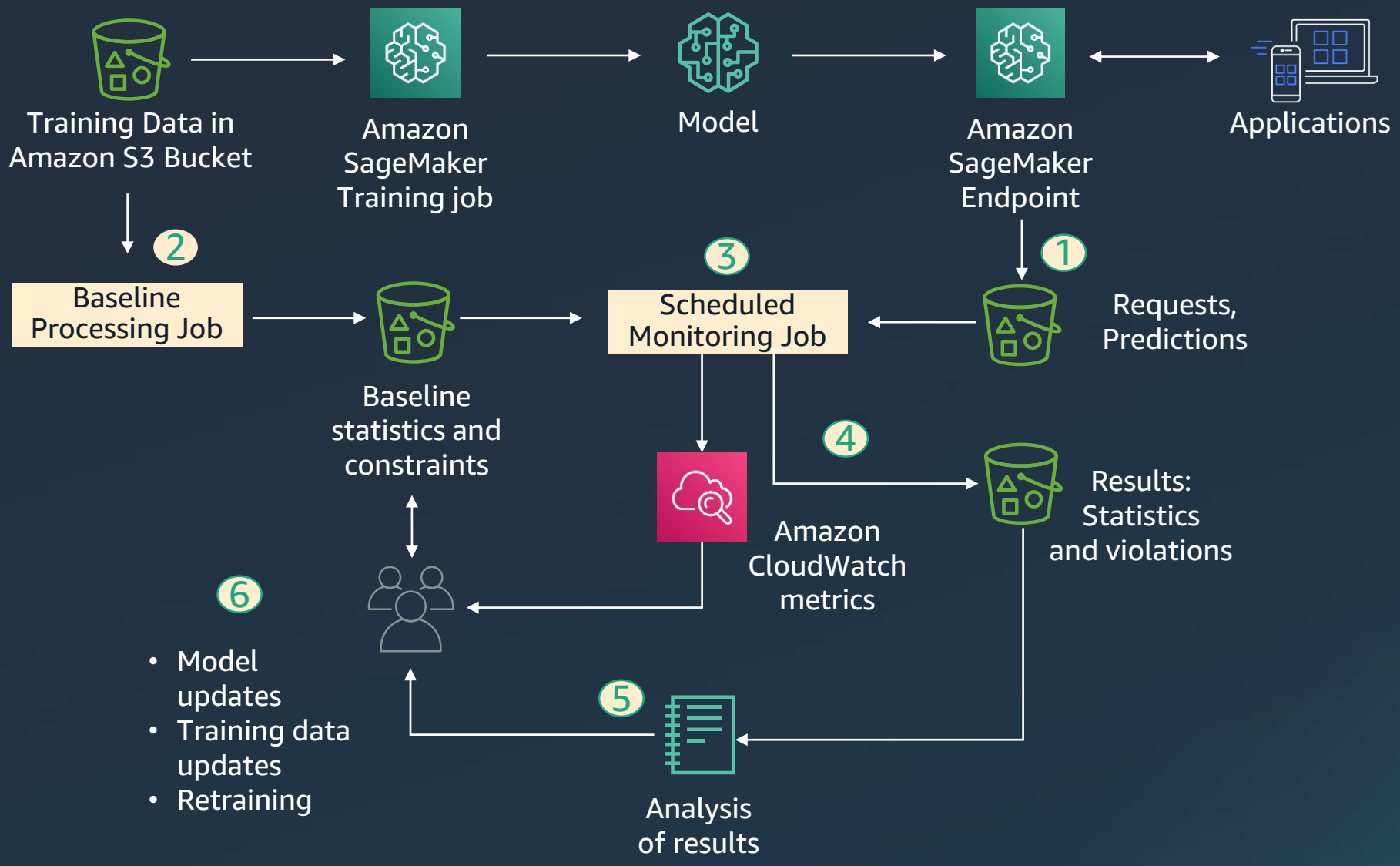
Step 4 – View Violations

	feature_name	constraint_check_type	description
0	State_MA	data_type_check	Data type match requirement is not met. Expected data type: Integral, Expected match: 100.0%. Observed: Only 99.6236559139785% of data is Integral.
1	State_AL	data_type_check	Data type match requirement is not met. Expected data type: Integral, Expected match: 100.0%. Observed: Only 99.6236559139785% of data is Integral.
2	State_AZ	data_type_check	Data type match requirement is not met. Expected data type: Integral, Expected match: 100.0%. Observed: Only 99.6236559139785% of data is Integral.
3	State_WY	data_type_check	Data type match requirement is not met. Expected data type: Integral, Expected match: 100.0%. Observed: Only 99.6236559139785% of data is Integral.
4	State_NE	data_type_check	Data type match requirement is not met. Expected data type: Integral, Expected match: 100.0%. Observed: Only 99.6236559139785% of data is Integral.
5	State_GA	data_type_check	Data type match requirement is not met. Expected data type: Integral, Expected match: 100.0%. Observed: Only 99.6236559139785% of data is Integral.
6	State_OR	data_type_check	Data type match requirement is not met. Expected data type: Integral, Expected match: 100.0%. Observed: Only 99.6236559139785% of data is Integral.
7	VMail Plan_no	data_type_check	Data type match requirement is not met. Expected data type: Integral, Expected match: 100.0%. Observed: Only 99.6236559139785% of data is Integral.
8	VMail Message	data_type_check	Data type match requirement is not met. Expected data type: Integral, Expected match: 100.0%. Observed: Only 99.6236559139785% of data is Integral.
9	State_DE	data_type_check	Data type match requirement is not met. Expected data type: Integral, Expected match: 100.0%. Observed: Only 99.6236559139785% of data is Integral.

aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Step 5 – Take corrective actions

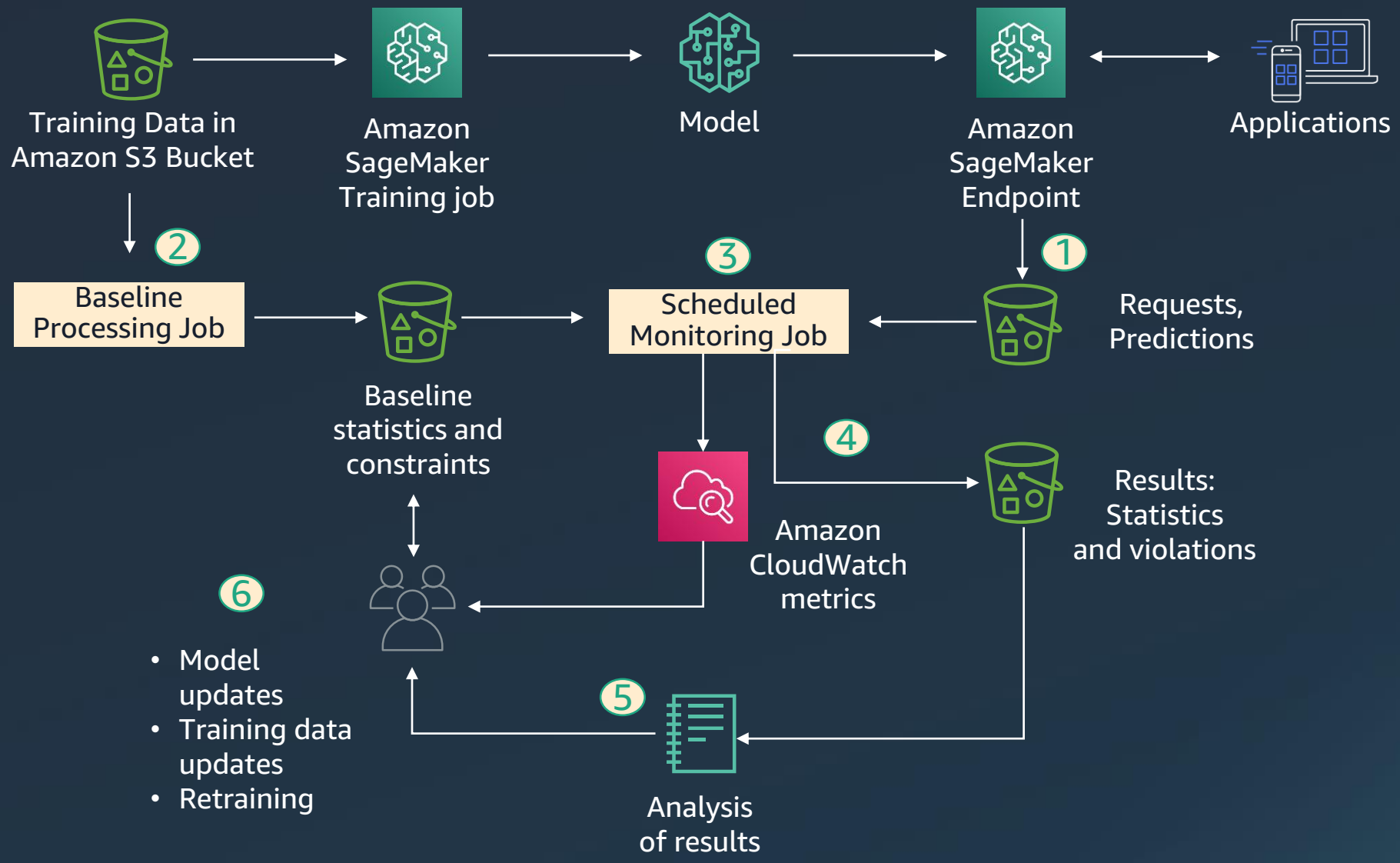


© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Demo – Data Quality Monitoring

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Demo – Data Quality Monitoring



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Recap

- What is Amazon SageMaker?
- Why do we need model monitoring? Why is it important?
- What is Amazon SageMaker Model Monitor and how it works?
- Types of model monitoring
- Deep Dive on Data Quality Monitoring
- Demo - Data Quality Monitoring
- How to get started with Amazon SageMaker Model Monitoring?

aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How can you get started?

- Introduction to Amazon SageMaker
- Amazon SageMaker Model Monitor details
- Amazon SageMaker examples (GitHub)
- Amazon SageMaker Model Monitor examples (GitHub)
- Amazon SageMaker Machine Learning tutorials (with hands-on labs)
- Amazon SageMaker pricing
- AWS Free Tier

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Visit the Data & AI/ML resource hub

Dive deeper into these resources, get inspired and learn how you can use AI and machine learning to accelerate your business outcomes.

- 6 steps to machine learning success e-book
- 7 leading machine learning use cases e-book
- Machine learning at scale e-book
- Achieving transformative business results with machine learning e-book
- Tackling our world's hardest problems with machine learning e-book
- Accelerating machine learning innovation through security e-book
- … and more!

https://bityl.co/FqdC

Visit resource hub

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Training and Certification

Access the AI & ML learning plan courses built by AWS experts on AWS Skill Builder

- Get started with digital self-paced, on-demand training and ramp-up guides to help you grow your technical skills
- Learn how to apply machine learning, artificial intelligence, and deep learning to unlock new insights and value in your role
- Take the steps today, towards validating your expertise with an AWS Certified Machine Learning – Specialty Certification

https://bit.ly/3FnxDH7

Learn your way explore.skillbuilder.aws »

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you for attending AWS Innovate – Data & AI/ML Edition

We hope you found it interesting! A kind reminder to **complete the survey.**
Let us know what you thought of today's event and how we can improve the event experience for you in the future.

aws-apj-marketing@amazon.com

twitter.com/AWSCloud

facebook.com/AmazonWebServices

youtube.com/user/AmazonWebServices

slideshare.net/AmazonWebServices

twitch.tv/aws

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.