AWS re:Invent

WIN407-R

Performance tune SQL Server using AWS Machine Learning

Jugal Shah
Partner Solutions Architect
Amazon Web Services

AWS re:Invent

aws

Agenda

- Machine learning (ML) introduction
- ML stack
- ML process
- Demo
- Lab

What is machine learning?

Machine learning: AI machines that improve their predictions by learning from large amounts of input data

Deep learning: ML AIs that use complex “artificial neural networks” to greatly improve their predictions and decisions

Artificial intelligence (AI): Machines that are programmed to perform the tasks that normally require humans

The Amazon Machine Learning stack

AI SERVICES

(App developers with little knowledge of ML)

Vision

- Amazon Rekognition Image
- Amazon Rekognition Video

Speech

- Amazon Textract — NEW
- Amazon Polly
- Amazon Transcribe

Language

- Amazon Translate
- Amazon Comprehend

Chatbots

- Amazon Lex

Forecasting

- Amazon Forecast — NEW

Recommendations

- Amazon Personalize — NEW

ML SERVICES

(ML developers and data scientists)

Amazon SageMaker

- Notebooks
- Ground Truth — NEW
- Algorithms
- Marketplace — NEW
- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning — NEW
- Training
- Optimization (Neo) — NEW
- Deployment
- Hosting

ML FRAMEWORKS & INFRASTRUCTURE

(ML researchers and academics)

Frameworks

- TensorFlow
- mxnet
- PYTORCH
- Chainer
- HOROVOD

Interfaces

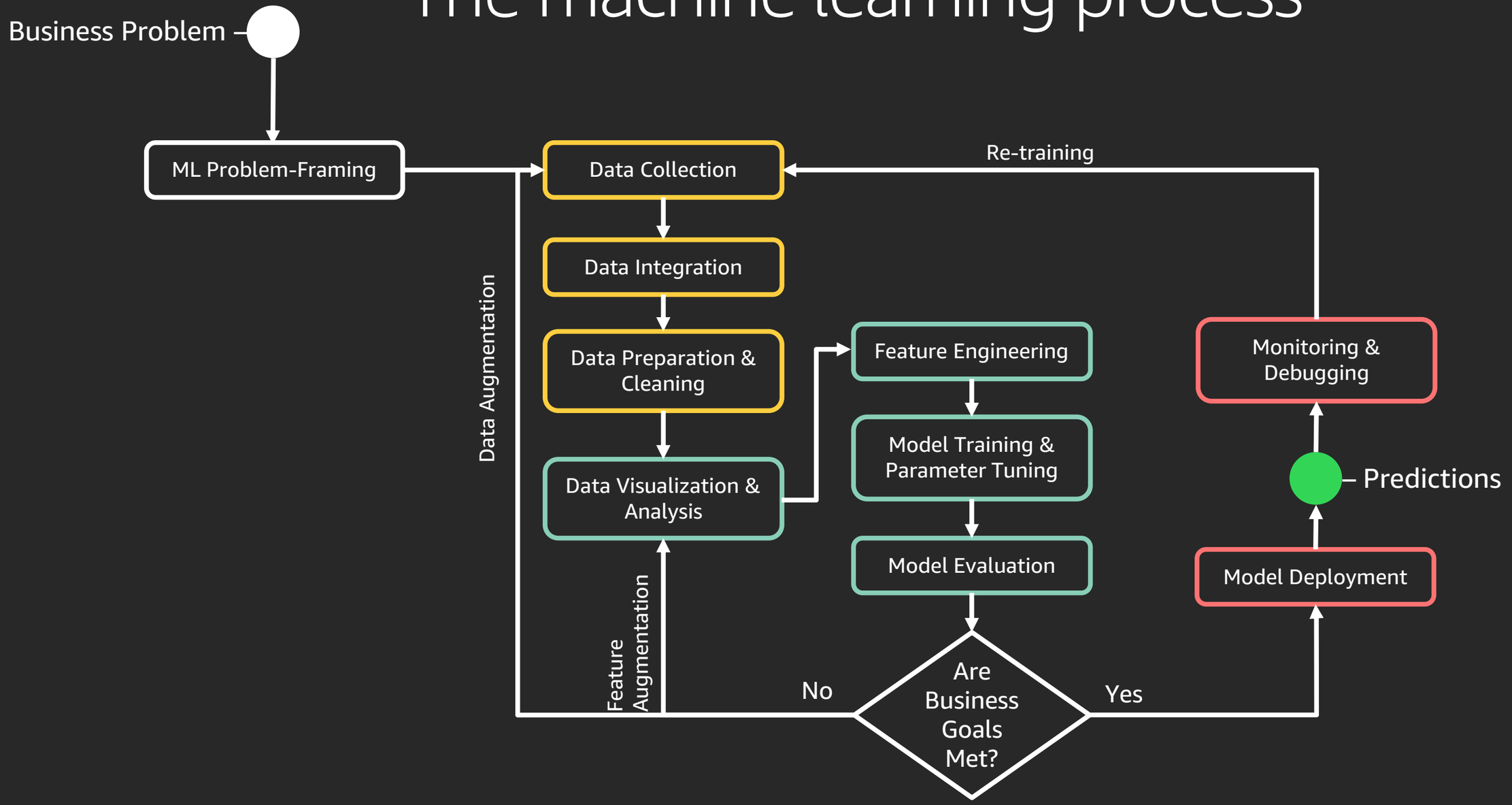
- GLUON
- Keras

Infrastructure

- Amazon EC2 P3 & P3DN — NEW
- Amazon EC2 C5
- FPGAs
- AWS IoT Greengrass
- Amazon Elastic Inference — NEW
- AWS Inferentia — NEW

The machine learning process

The machine learning process



Demo

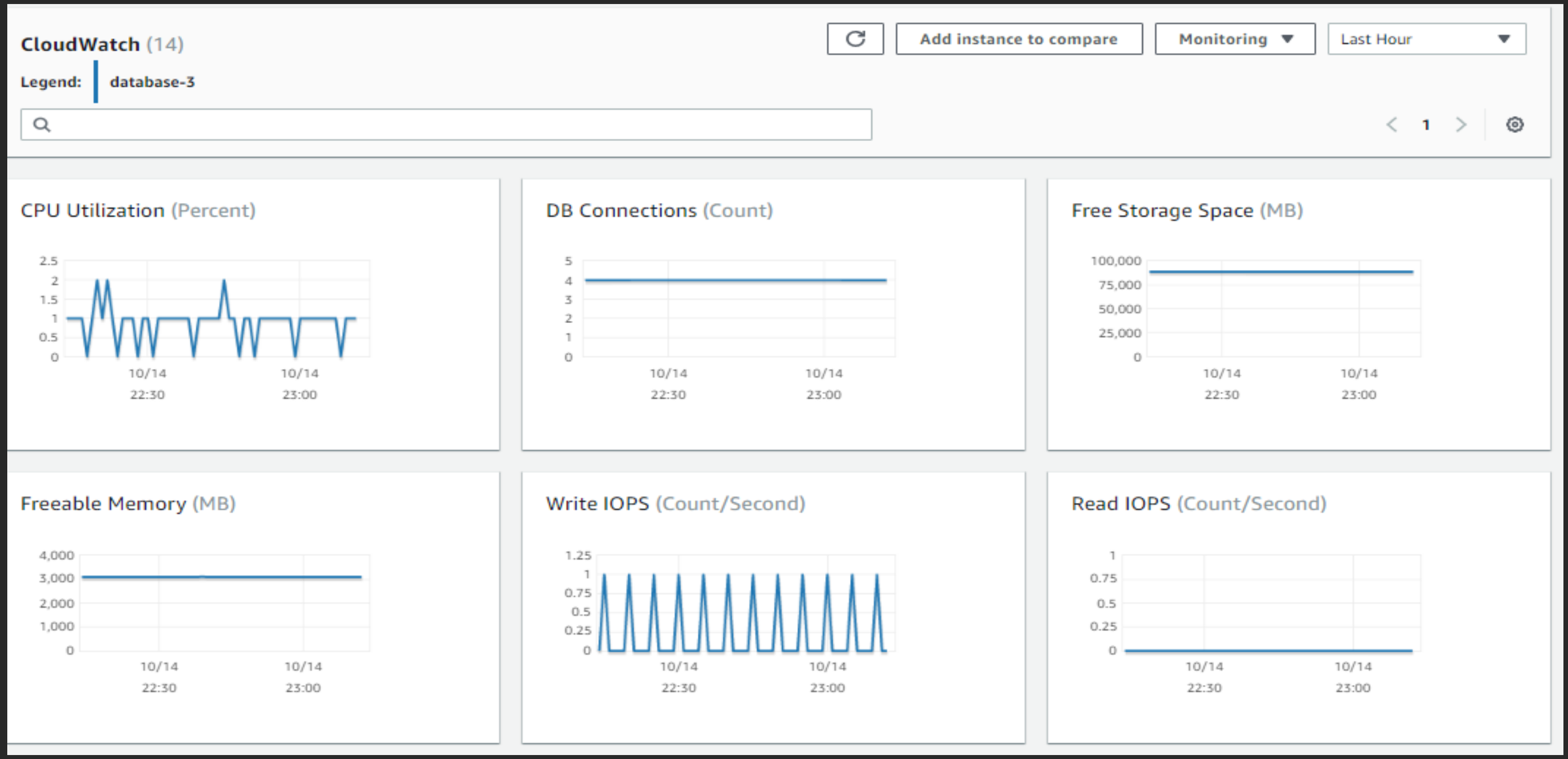
Problem:

Monitoring the multiple RDS metrics and identifying the problematic instance configuration is a time-consuming task. Manual tasks can also lead to human errors.

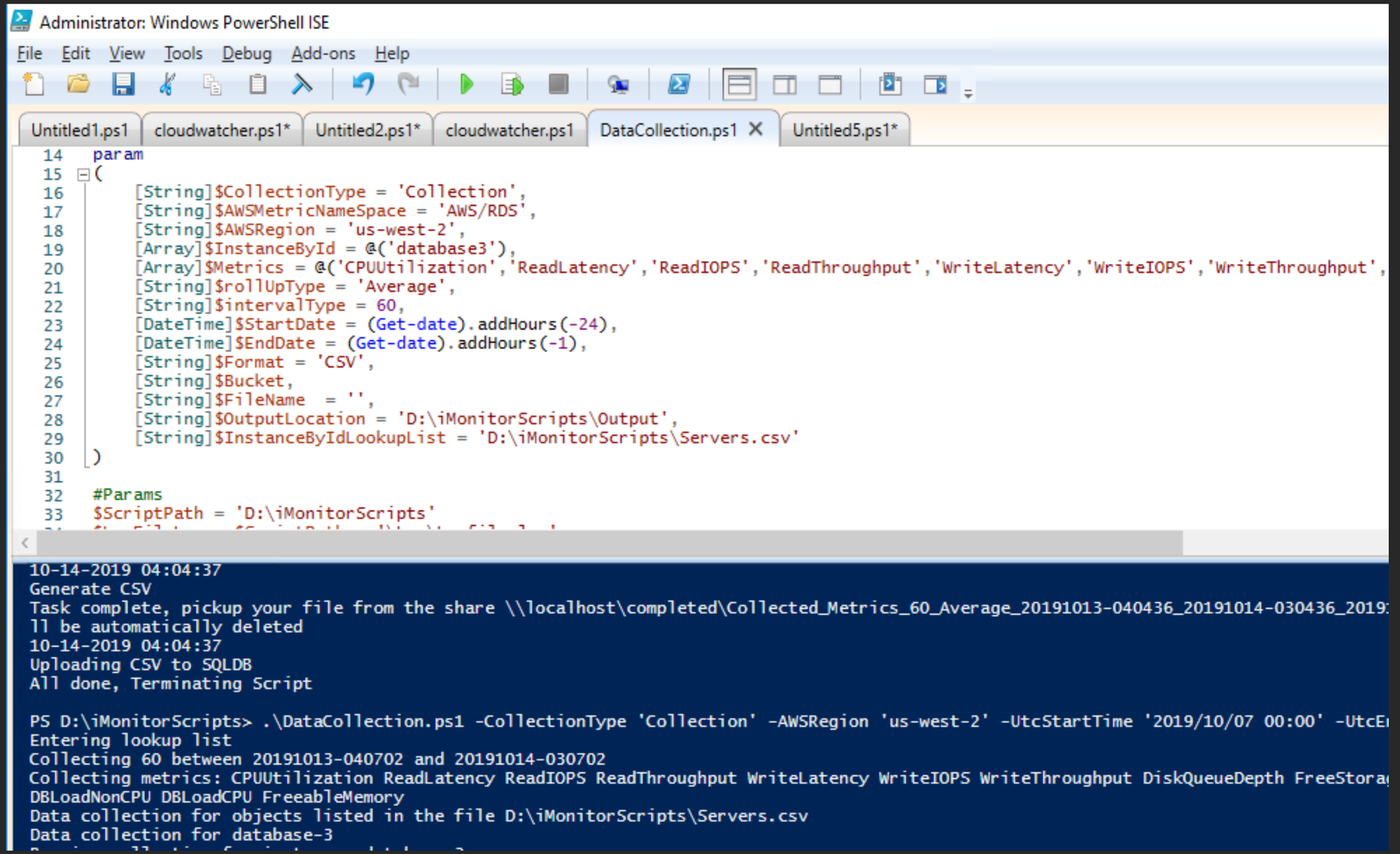
Solution:

In this session, we create the solution using machine learning to analyze the workload.

RDS: Amazon CloudWatch metrics



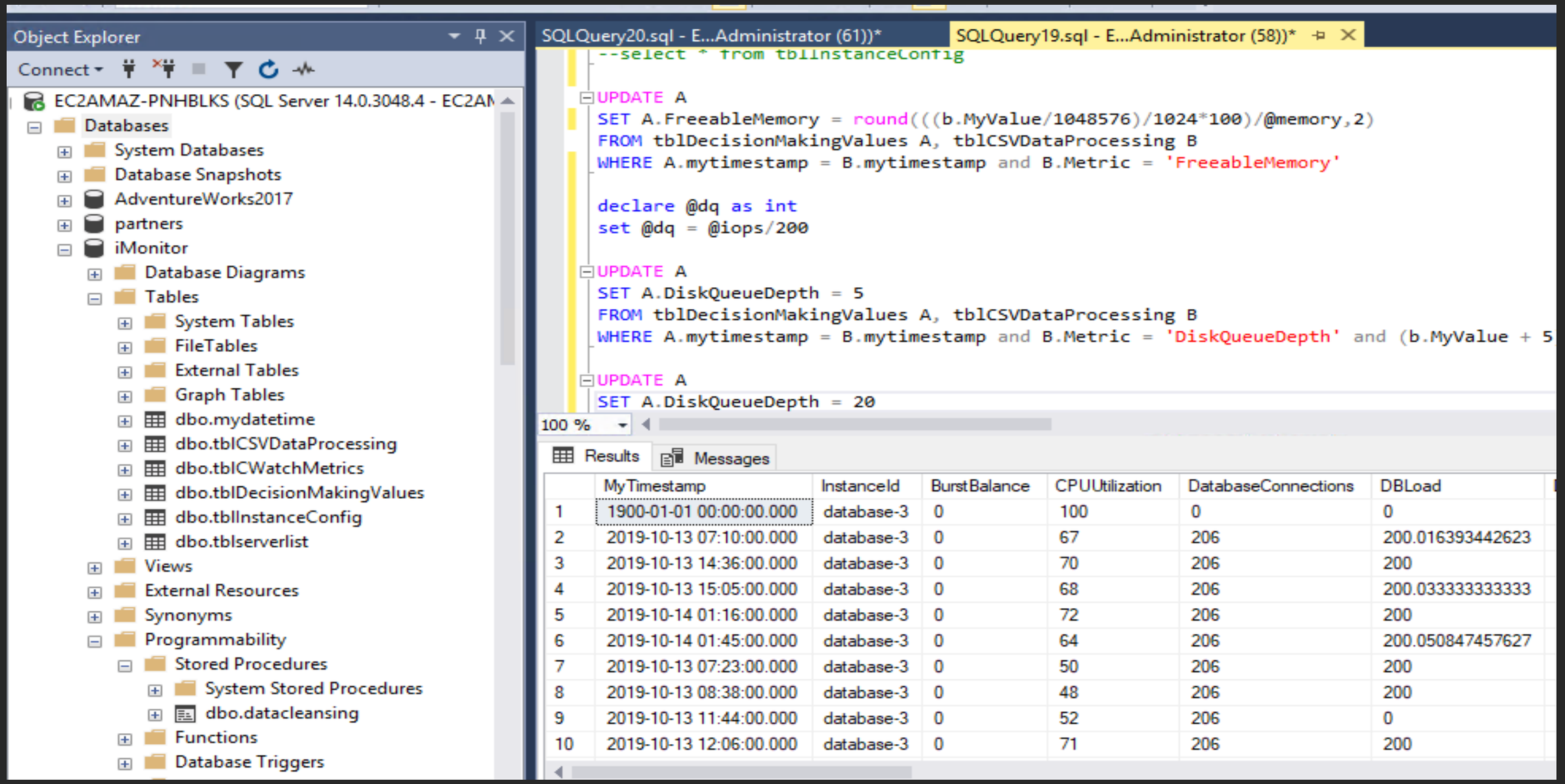
Data collection



CloudWatch metrics: Collected by PowerShell in CSV format

	A	B	C	D
1	Timestamp	InstanceId	Metric	value
2	10/13/2019 4:07	database-3	ReadThroughput	0.00
3	10/13/2019 4:07	database-3	CPUUtilization	1.00
4	10/13/2019 4:07	database-3	FreeableMemory	13927755776.00
5	10/13/2019 4:07	database-3	WriteThroughput	2303.00
6	10/13/2019 4:07	database-3	DiskQueueDepth	0.00
7	10/13/2019 4:07	database-3	WriteLatency	0.00
8	10/13/2019 4:07	database-3	FreeStorageSpace	107032346624.00
9	10/13/2019 4:07	database-3	ReadLatency	0.00
10	10/13/2019 4:07	database-3	ReadIOPS	0.00
11	10/13/2019 4:07	database-3	WriteIOPS	0.00
12	10/13/2019 4:07	database-3	DatabaseConnections	0.00
13	10/13/2019 4:08	database-3	ReadLatency	0.00
14	10/13/2019 4:08	database-3	FreeStorageSpace	107032346624.00
15	10/13/2019 4:08	database-3	ReadThroughput	0.00
16	10/13/2019 4:08	database-3	FreeableMemory	13927292928.00
17	10/13/2019 4:08	database-3	ReadIOPS	0.00
18	10/13/2019 4:08	database-3	WriteLatency	0.00
19	10/13/2019 4:08	database-3	DatabaseConnections	0.00
20	10/13/2019 4:08	database-3	DiskQueueDepth	0.00
21	10/13/2019 4:08	database-3	WriteThroughput	2756.00
22	10/13/2019 4:08	database-3	CPUUtilization	1.00

Data preparation & cleaning



	MyTimestamp	InstanceId	BurstBalance	CPUUtilization	DatabaseConnections	DBLoad
1	1900-01-01 00:00:00.000	database-3	0	100	0	0
2	2019-10-13 07:10:00.000	database-3	0	67	206	200.016393442623
3	2019-10-13 14:36:00.000	database-3	0	70	206	200
4	2019-10-13 15:05:00.000	database-3	0	68	206	200.033333333333
5	2019-10-14 01:16:00.000	database-3	0	72	206	200
6	2019-10-14 01:45:00.000	database-3	0	64	206	200.050847457627
7	2019-10-13 07:23:00.000	database-3	0	50	206	200
8	2019-10-13 08:38:00.000	database-3	0	48	206	200
9	2019-10-13 11:44:00.000	database-3	0	52	206	0
10	2019-10-13 12:06:00.000	database-3	0	71	206	200

Output file to train the ML model

	A	B	C	D	E	F	G	H
1	CPUUtilization	DiskQueueDepth	FreeableMemory	FreeStorageSpace	Latency	IOPSPercent	ThroughputPercent	Decision
2	67	5	74.19	20	0.7340323	1	32.99	Disk,Low Storage
3	70	6	19.34	20	1.33792073	1	33.78	Memory,Disk,Low Storage,High IO Latency
4	68	7	19.28	86.83	0.92694825	1	32.65	Memory,Disk
5	72	8	18.81	86.81	2	1	32.71	Memory,Disk,High IO Latency
6	64	8	18.77	86.8	0.95284423	1	33.12	Memory,Disk
7	50	9	72.14	95.23	3	1	35.61	Disk,High IO Latency
8	48	10	58.63	92.97	4	1	27.88	Disk,High IO Latency
9	52	11	23.8	87.53	0.50071423	1	29.15	Disk
10	71	12	19.43	86.84	5	1	34.61	Memory,Disk,High IO Latency
11	41	13	19.22	86.83	6	1	31.45	Memory,Disk,High IO Latency
12	71	15	19.03	86.82	7	1	32.32	Memory,Disk,High IO Latency
13	68	15	18.97	86.82	8	1	32.38	Memory,Disk,High IO Latency
14	67	16	18.96	86.82	0.88515428	1	32.54	Memory,Disk
15	72	18	18.81	86.81	9	1	32.51	Memory,Disk,High IO Latency
16	58	19	18.97	86.82	10	1	31.77	Memory,Disk,High IO Latency
17	42	5	18.97	86.81	11	1	33.7	Memory,Disk,High IO Latency
18	84	5	50.91	91.79	0.68257465	1	20.42	Disk
19	53	5	45.63	90.91	12	1	33.88	Disk,High IO Latency
20	66	5	45.38	90.91	0.67843437	1	20.63	Disk

Jupyter Notebook: Build, train & deploy the model

jupyter My-Dry-Run-Demo1-Copy1 (autosaved)

File Edit View Insert Cell Kernel Widgets Help Trusted conda_python3

Run Code nbdiff

```
import pandas as pd
import numpy as np
from datetime import datetime
import io
import sagemaker.amazon.common as smac

import boto3
from sagemaker import get_execution_role
import sagemaker

import matplotlib.pyplot as plt
import seaborn as sns
```

In [18]:

```
role = get_execution_role()
bucket='dryrun-2'
sub_folder = 'Input'
data_key = 'jd-cwdata.csv'
data_location = 's3://{}/{}/{}'.format(bucket, sub_folder, data_key)
```

Amazon SageMaker > Models

Models

Create endpoint | Create endpoint configuration | Actions | Create model

Search models

1

Name	ARN	Creation time
cwm991-linear-learner-job-20191014212441	arn:aws:sagemaker:us-west-2:712820468924:model/cwm991-linear-learner-job-20191014212441	Oct 14, 2019 21:29 UTC

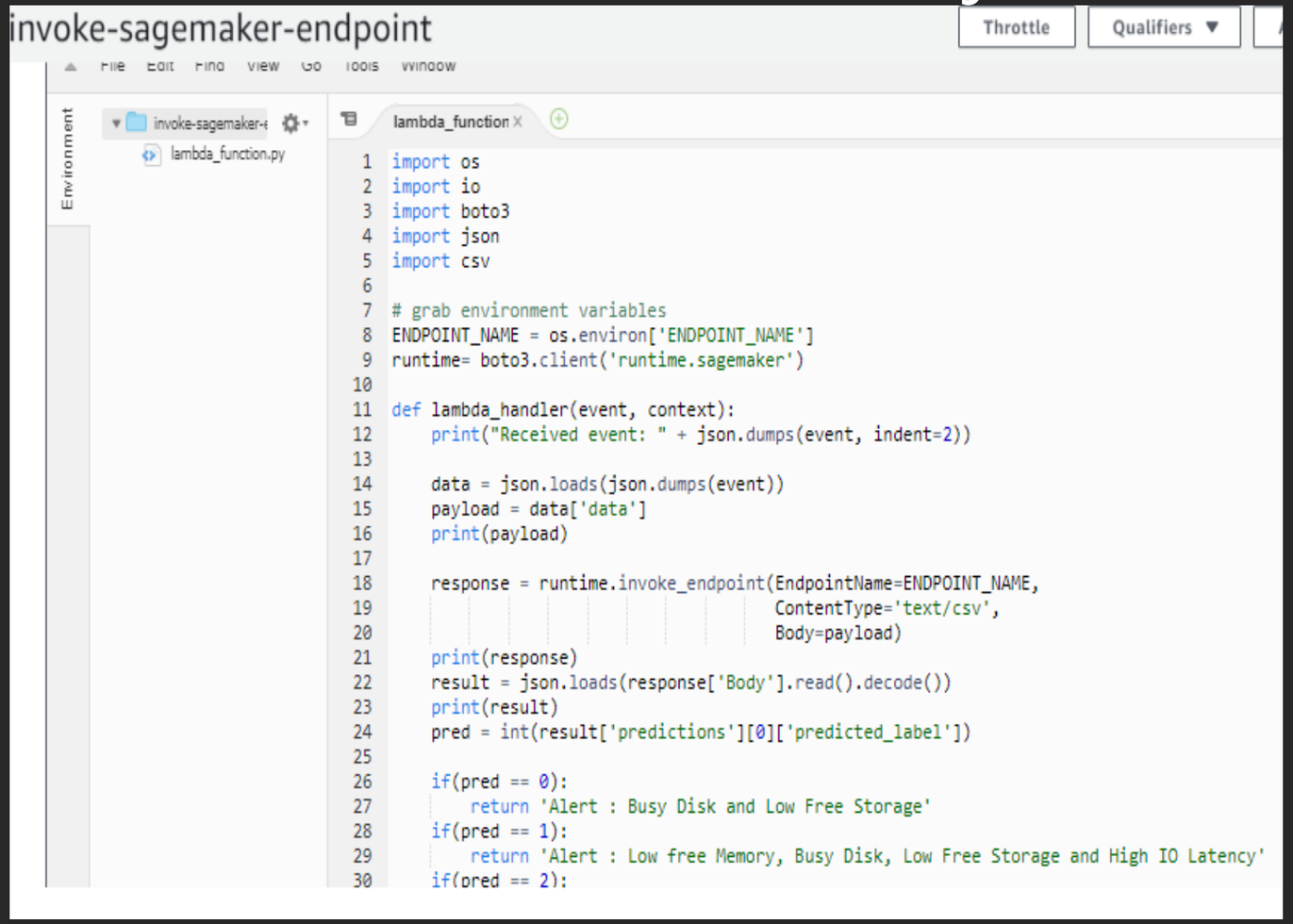
Amazon SageMaker > Endpoints

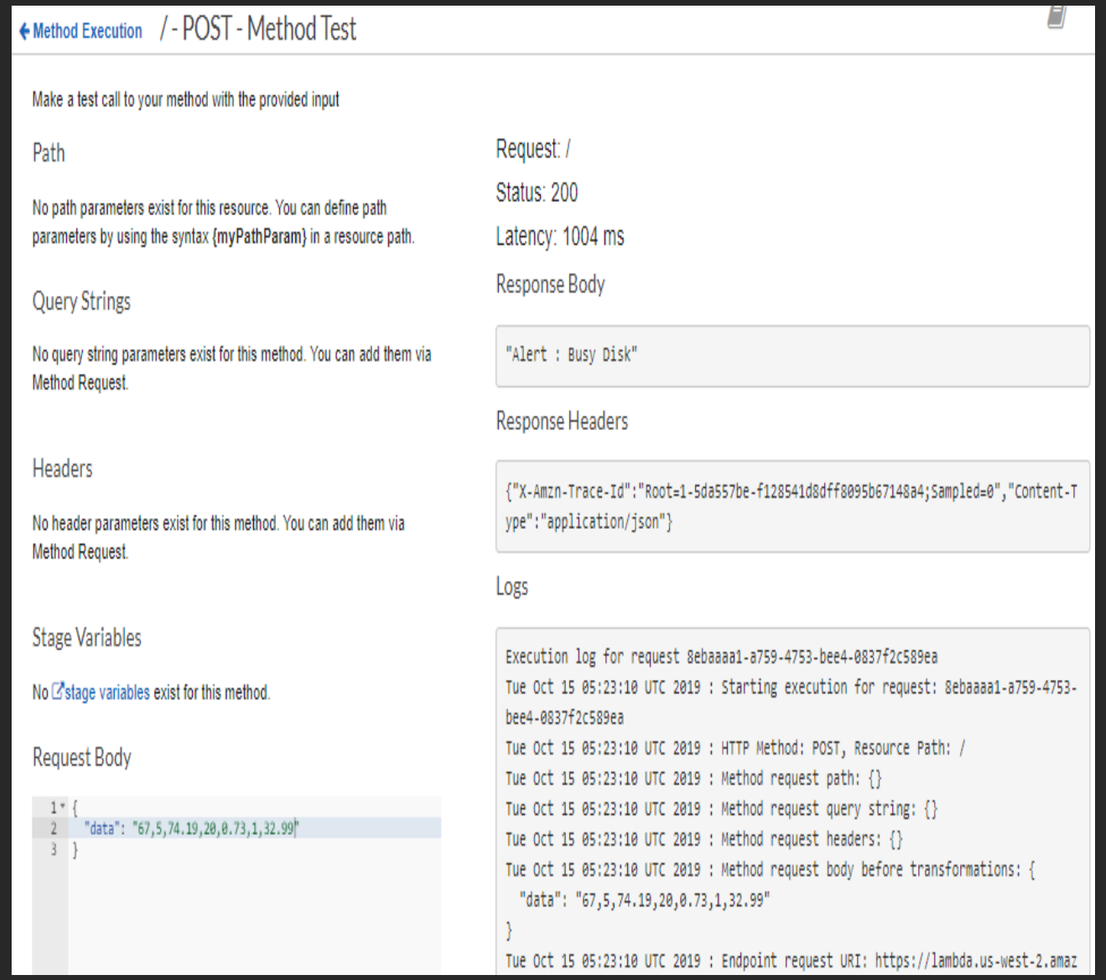
Endpoints

Search endpoints

Name	ARN
cwm991-linear-learner-job-20191014212441	arn:aws:sagemaker:us-west-2:712820468924:endpoint/cwm991-linear-learner-job-20191014212441

Lambda & API Gateway





Lab

Thank you!

Jugal Shah
jugashah@amazon.com

AWS re:Invent

aws

Please complete the session survey in the mobile app.