



Dominando o Amazon SageMaker

**Construção, treinamento e otimização de
modelos de aprendizado de máquina**

Tiago Simão

Arquiteto de Soluções - Enterprise
AWS

Agenda

- Construção, Treinamento e Otimização de modelos de ML
- Sagemaker Studio
- Demonstrações



SageMaker: Construindo modelos de ML

Construindo modelos de ML

**Notebooks totalmente
gerenciados e
compartilháveis no
Amazon EC2**



Notebooks Jupyter, gerenciados e compartilháveis

Implemente notebooks em recursos de computação elástica



Algoritmos integrados

15 algoritmos integrados disponíveis como imagens de containers



Soluções integradas e modelos de código aberto

Mais de 150 modelos populares



AutoML

Crie modelos de ML automaticamente com total visibilidade



Suporte aos principais frameworks e ferramentas

Otimizado para estruturas populares de aprendizado profundo (DL), como TensorFlow, PyTorch, Apache MXNet e Hugging Face

Notebooks no SageMaker

O AMAZON SAGEMAKER OFERECE DOIS TIPOS DE NOTEBOOKS JUPYTER TOTALMENTE GERENCIADOS COM UM CLIQUE PARA EXPLORAÇÃO DE DADOS E CRIAÇÃO DE MODELOS DE ML

Studio Notebooks

Notebooks colaborativos de início rápido que se integram a ferramentas de ML criadas especificamente no SageMaker e em outros serviços da AWS

Notebook Instances

Instâncias autônomas e totalmente gerenciadas do Jupyter Notebook na console do Amazon SageMaker

Amazon SageMaker tem algoritmos integrados ou você pode trazer o seu

Classificação

Linear Learner | XGBoost | KNN

Visão Computacional

Image classification | Object detection |
Semantic segmentation

Modelagem de Tópicos

LDA | NTM

Trabalhando com Texto

BlazingText | Supervised | Unsupervised

Recomendação

Factorization machines

Previsão

DeepAR

Tradução de Sequências

Seq2Seq

Regressão

Linear Learner | XGBoost | KNN

Agrupamento

KMeans

Deteccão de Anomalia

Random cut forests | IP Insights

Redução de Recursos

PCA



SageMaker: Aprofundando no treinamento de modelos de ML

Treinamento de modelos de ML

Treinamento de modelos de ML de forma rápida e custo efetiva



Gerenciamento de experimentos e ajuste de modelos

Economize semanas de esforço rastreando automaticamente o treinamento e hiperparâmetros de ajuste



Defina o perfil e a depuração do treinamento

Use métricas em tempo real para corrigir problemas de desempenho



Treinamento distribuído

Complete distributed training up to 40% faster



Compilador de Treinamento

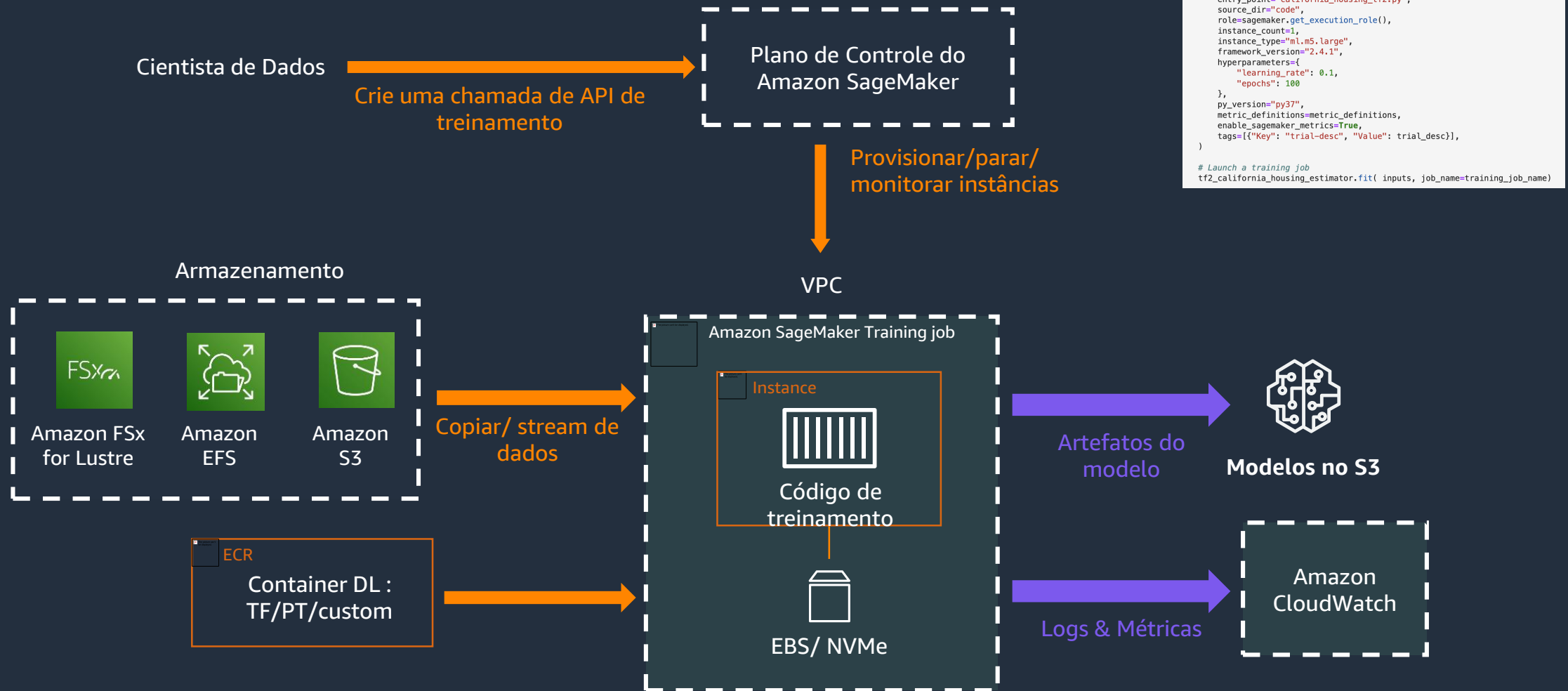
Acelere os tempos de treinamento em até 50% por meio do uso mais eficiente de GPUs



Treinamento gerenciado com Spot

Reduza os custos de treinamento em até 90%

Treinamento com o Amazon SageMaker



Estimador de Treinamento

```
# 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)
```

Amazon SageMaker Ajuste Automático de Modelos

**Ajuste automaticamente os
hiperparâmetros
em seus algoritmos**



Otimização em escala

Ajuste milhares de diferentes combinações de parâmetros do algoritmo



Automatizado

Usa ML para encontrar os melhores parâmetros



Mais rápido

Elimine dias ou semanas de trabalho manual



Árvores de decisão

Tree depth | Max leaf nodes | Gamma | Eta | Lambda | Alpha

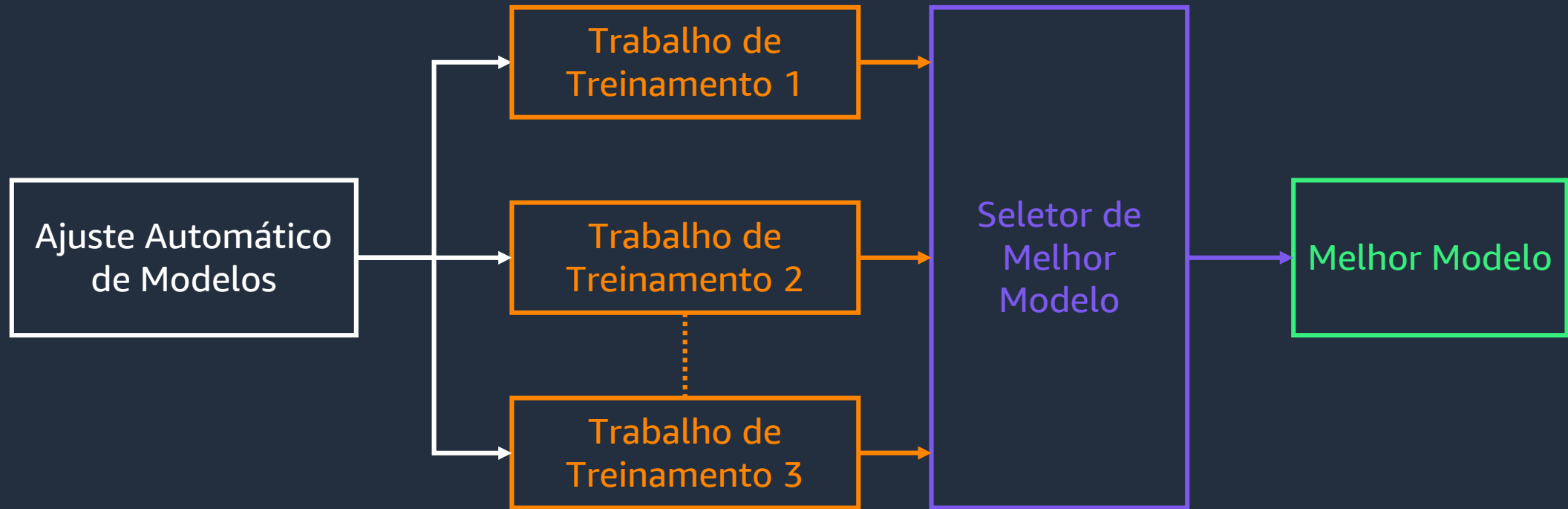


Redes Neurais

Number of layers | Hidden layer width | Learning rate | Embedding dimensions | Dropout

Amazon SageMaker Ajuste Automático de Modelos

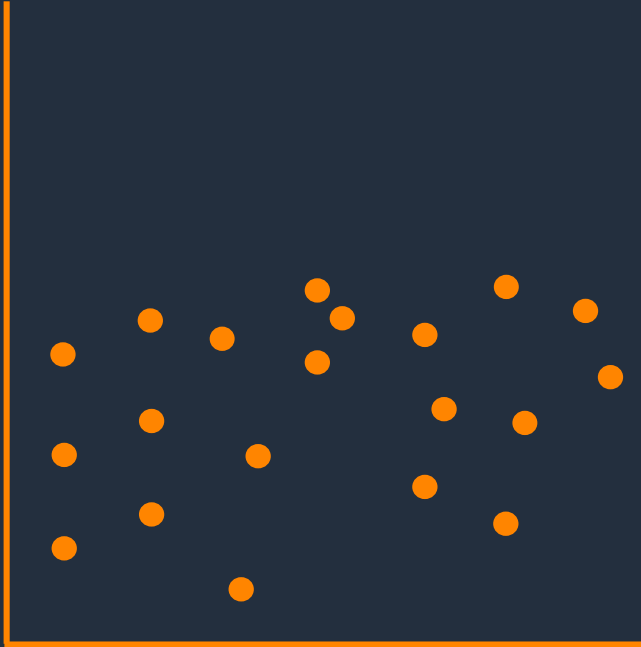
AJUSTE DE HIPERPARÂMETROS



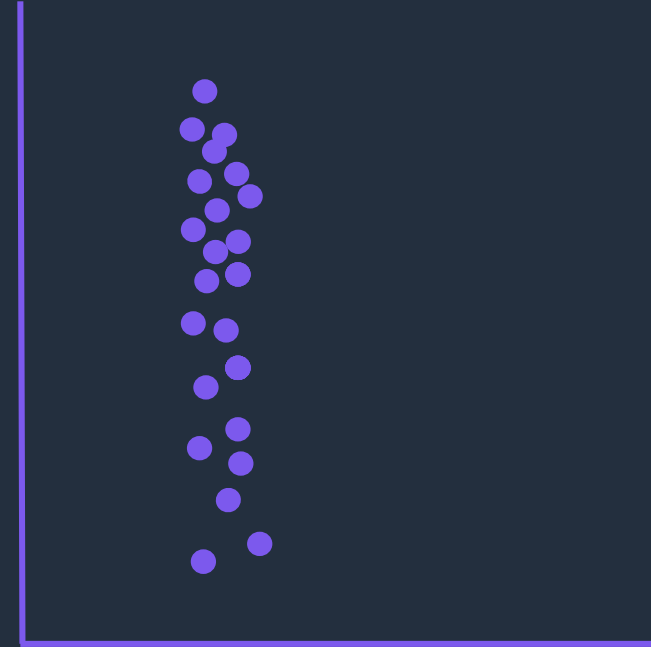
Amazon SageMaker Ajuste de Modelo Automático

E SE EU PRECISAR QUE TODOS OS MEUS TRABALHOS SEJAM AJUSTADOS AO MESMO TEMPO?

Pesquisa bayesiana



Pesquisa aleatória



Configurando a tarefa de ajuste de hiperparâmetros

1.

Escolha hiperparâmetros e intervalos

```
hyperparameter_ranges = {'eta': ContinuousParameter(0, 1),  
                          'min_child_weight': ContinuousParameter(1, 10),  
                          'alpha': ContinuousParameter(0, 2),  
                          'max_depth': IntegerParameter(1, 10)}
```

2.

Escolha a métrica objetiva

```
objective_metric_name = 'validation:auc'
```

3.

Escolha os parâmetros do trabalho

```
tuner = HyperparameterTuner(xgb,  
                             objective_metric_name,  
                             hyperparameter_ranges,  
                             max_jobs=20,  
                             max_parallel_jobs=3)
```

Amazon SageMaker Ajuste Automático de Modelos

POSSO USAR O AJUSTE DE HIPERPARÂMETROS COM MEU PRÓPRIO MODELO?

1

Algoritmos Integrados

2

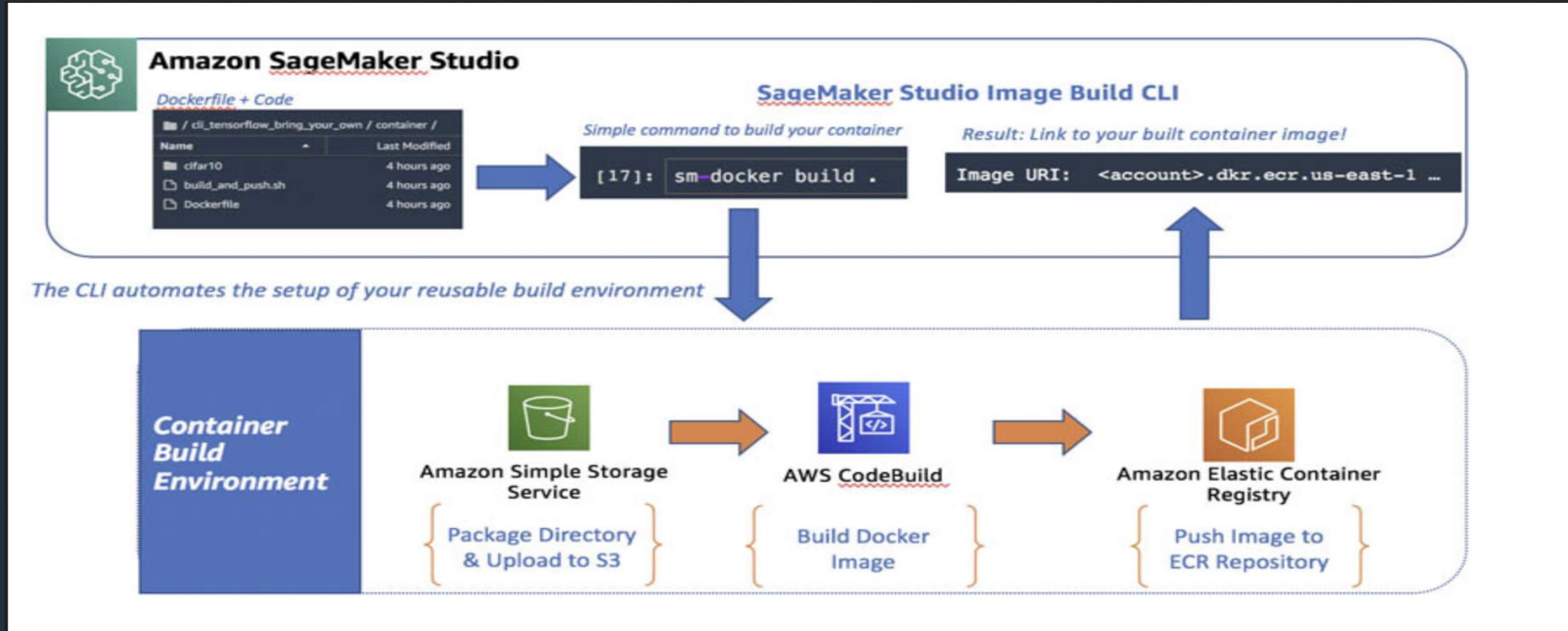
Docker
(Totalmente Customizável)

3

Script Mode
(Totalmente Customizável)

Amazon SageMaker Automatic Model Tuning

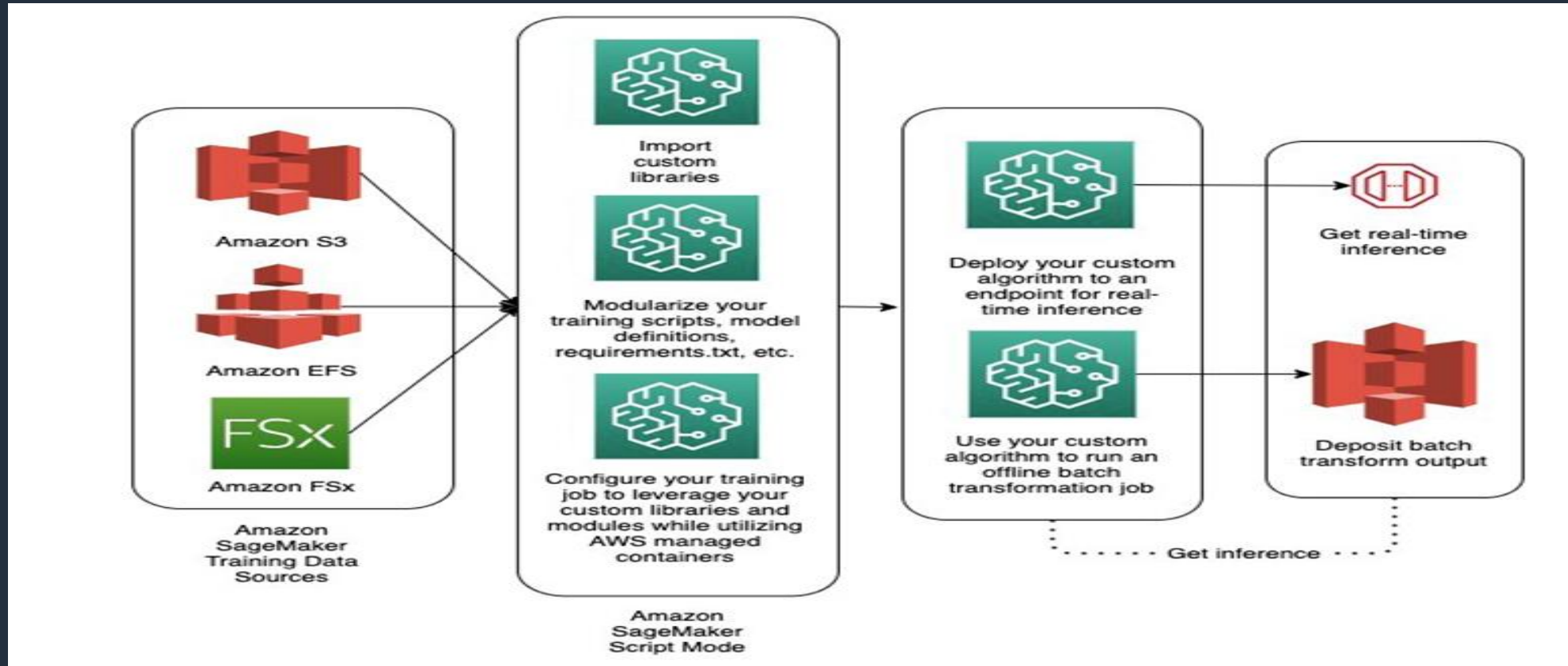
POSSO USAR O AJUSTE DE HIPERPARÂMETROS COM MEU PRÓPRIO MODELO?



✓ Docker

Amazon SageMaker Automatic Model Tuning

POSSO USAR O AJUSTE DE HIPERPARÂMETROS COM MEU PRÓPRIO MODELO?



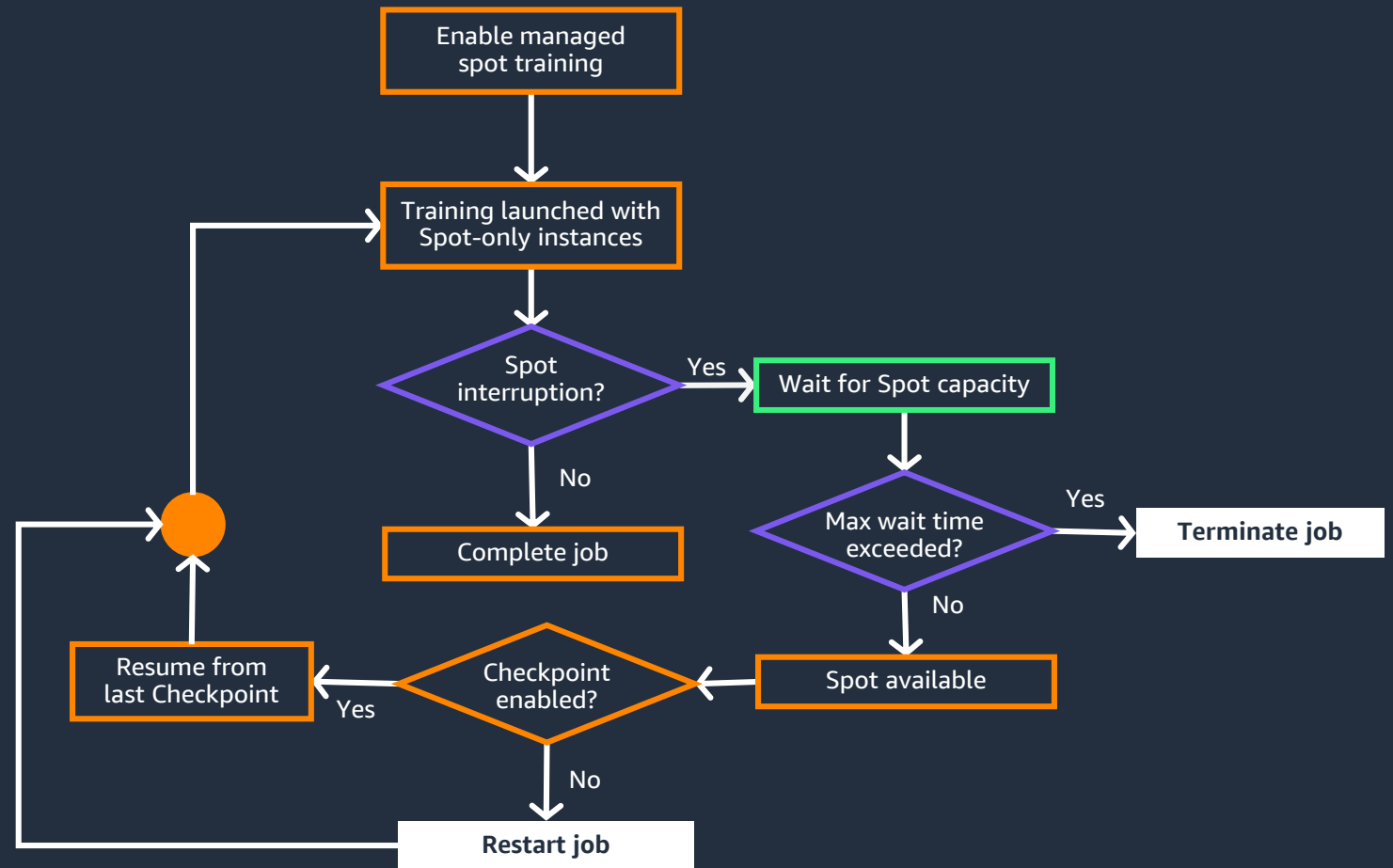
Script Mode



Treinamento gerenciado com Spot

Treinamento gerenciado com Spot

Economize até 90% em custos de treinamento de modelos



Principais considerações



Treinando apenas com Spot

Os trabalhos interrompidos são retomados se forem verificados e se as instâncias Spot estiverem disponíveis; os trabalhos serão reiniciados se não forem verificados*

Funciona com ajuste automatizado de modelos

Os trabalhos de treinamento só podem ser executados com uma única instância - digite uma única - AZ

Ainda não se integra com o Spot Fleet e Spot Block



Checkpoint

Algoritmos integrados checkpoints automatizados

Para modelos personalizados, o checkpoint deve ser ativado

Checkpoints são salvos no S3

Os modelos que não fazem o checkpoint estão sujeitos ao MaxWaitTime de 60 minutos



Preço

Veja as economias no console da AWS ou use a API DescribeTrainingJob

Cobrado pela duração do treinamento antes da conclusão ou encerramento; não é cobrado pelo tempo ocioso, o faturamento começa quando as instâncias estão prontas

Cobrado pelo tempo de download de dados apenas uma vez, mesmo que o trabalho seja interrompido várias vezes

* Checkpointing is a best practice and is highly recommended

Depurador



Gere modelos de ML com mais rapidez

Detecte gargalos e problemas durante o treinamento em tempo real e corrija problemas para implantar modelos mais rapidamente, com uma ferramenta unificada



Otimize recursos sem código adicional

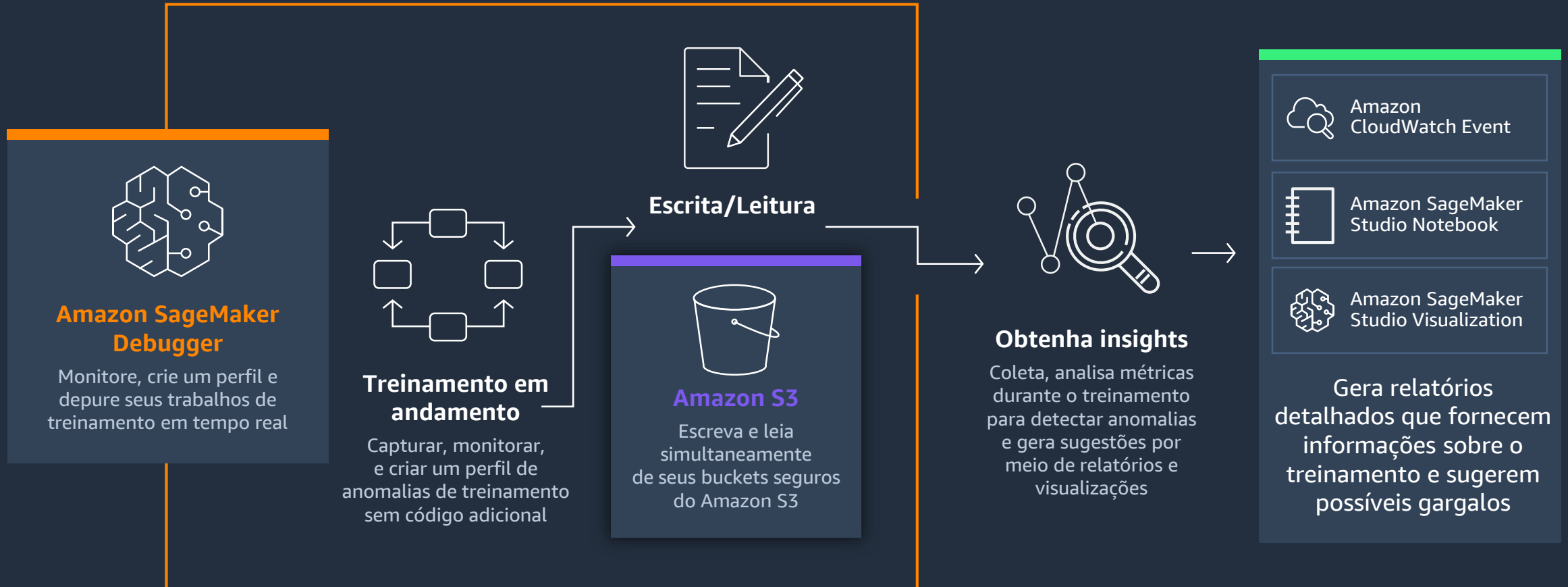
Monitore e crie o perfil dos recursos do sistema sem codificação e obtenha recomendações para otimizar recursos de forma eficaz



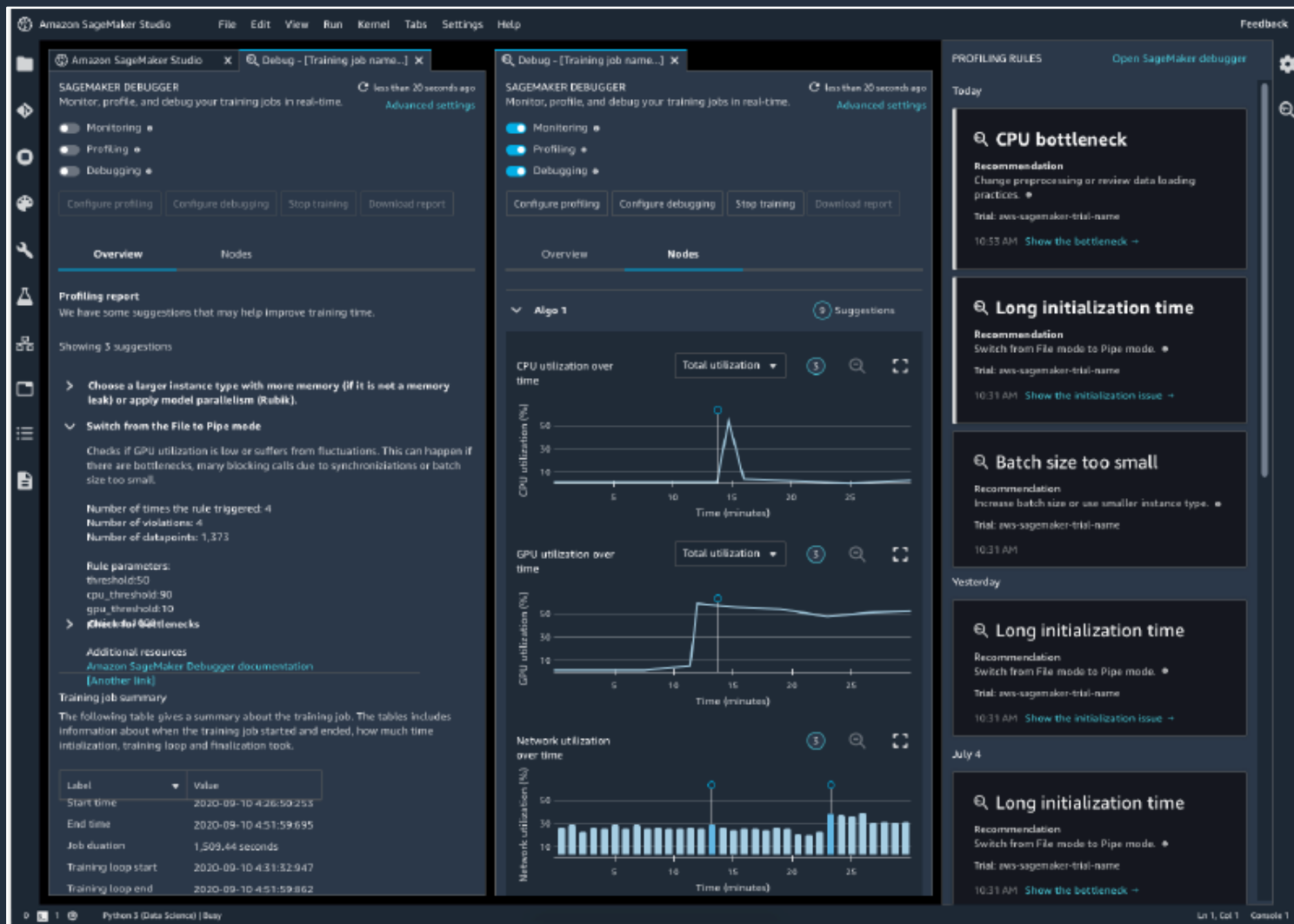
Torne o treinamento de ML transparente

Obtenha informações completas sobre o processo de treinamento de ML em tempo real e offline

Amazon SageMaker Debugger — Como Funciona



Monitore e crie o perfil de utilização dos recursos do sistema

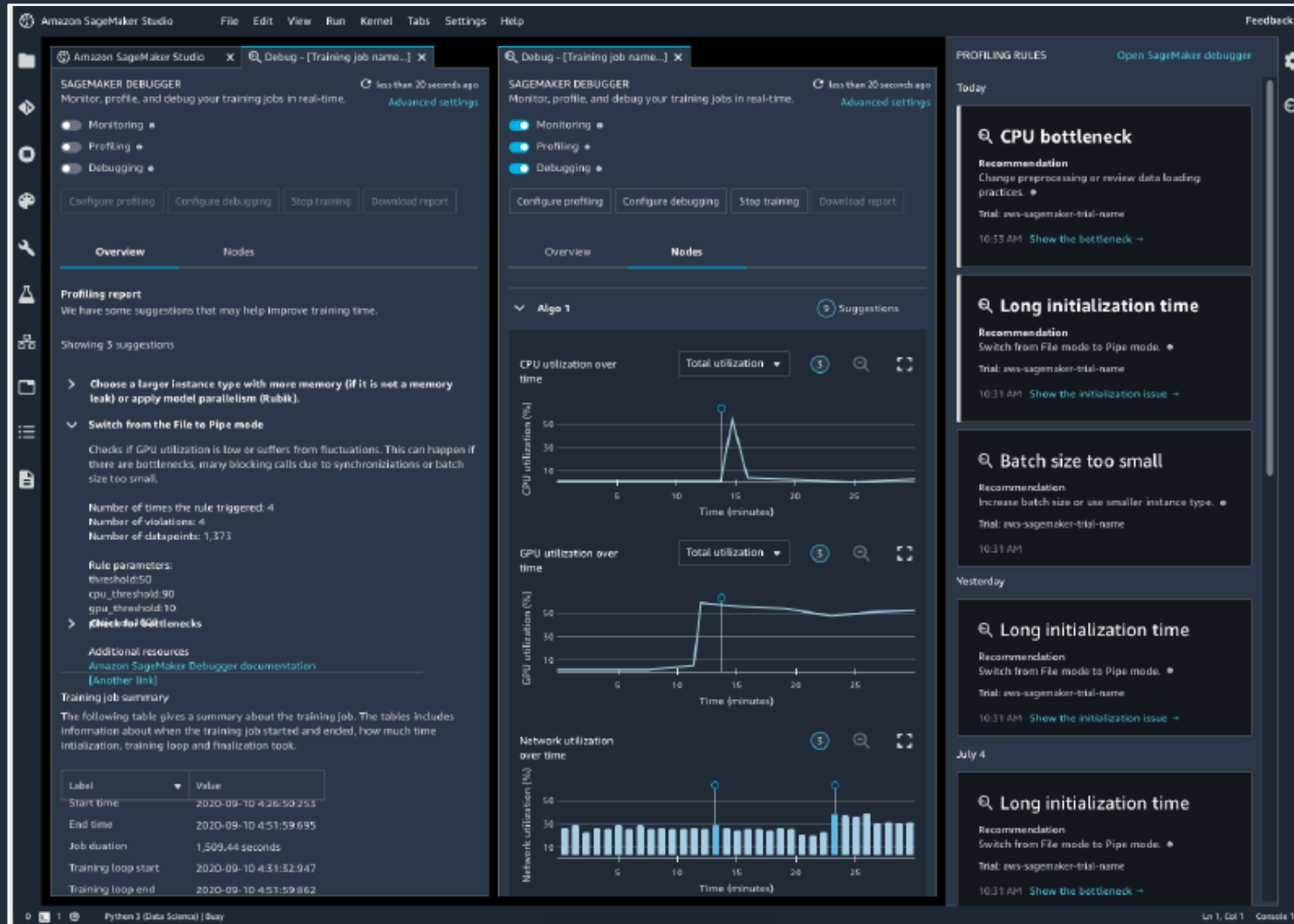


Monitore automaticamente a utilização dos recursos do sistema

Crie um perfil de trabalhos de treinamento para coletar métricas da estrutura de ML

Visualize a utilização de recursos do sistema para GPU, CPU, rede e memória no SageMaker Studio

Analyze errors and take action



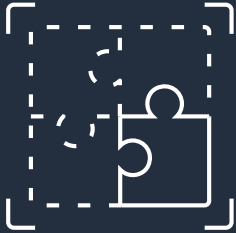
Built-in analysis in the form of rules

Automatically analyze training data including inputs, outputs, tensors

Detect if a model is overfitting or overtraining, or determine if gradient values are incorrect

Specify custom actions to stop training or send alerts

Amplo suporte a algoritmos e frameworks



1. Suporta

algoritmos de ML populares, como o XGBoost, e frameworks de aprendizado profundo, como TensorFlow, PyTorch, Apache MXNet e Keras, com contêineres integrados do SageMaker



2. Integra

com o AWS Lambda para agir com base nos resultados dos alertas



3. Invoque ações

para interromper automaticamente um trabalho de treinamento ao detectar uma ação não convergente, como perdas aumentando continuamente

Amazon SageMaker Studio

Traz ferramentas para cada etapa do ciclo de vida do ML em uma interface de usuário visual unificada





Obrigado!

Tiago Simão

simaot@amazon.com

www.linkedin.com/in/simaot