

# AWS AI/ML

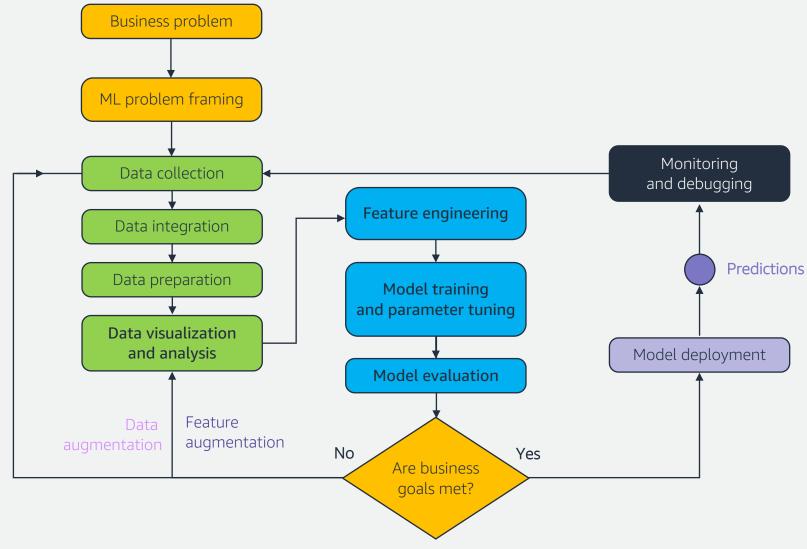
Gabriel Bella Martini Arquiteto de Soluções AWS

## **Agenda**

- 1- Introdução aos serviços de AI da AWS
- 2- Introdução ao SageMaker
- 3- Treinando modelos no SageMaker
- 4- Customizando o SageMaker



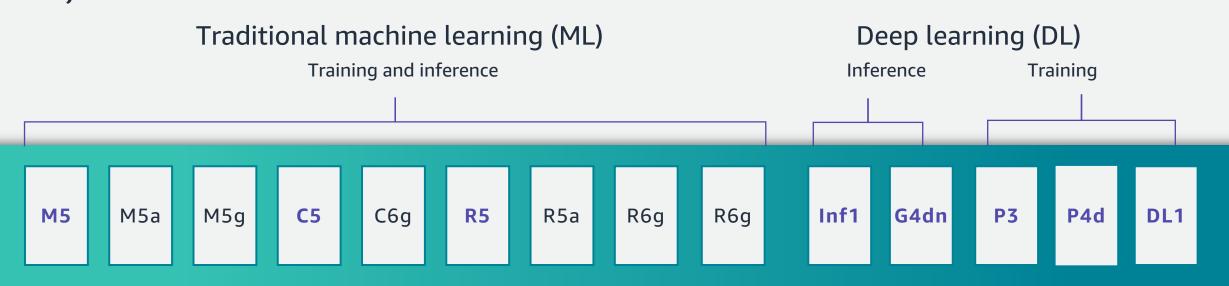
## O processo de um projeto de ML





## Infraestrutura de computação

Escolha de CPUs, GPUs e aceleradores para suas necessidades de desempenho e orçamento





Habana Gaudi accelerators



**EPYC CPU** 



Graviton CPU Inferentia chip

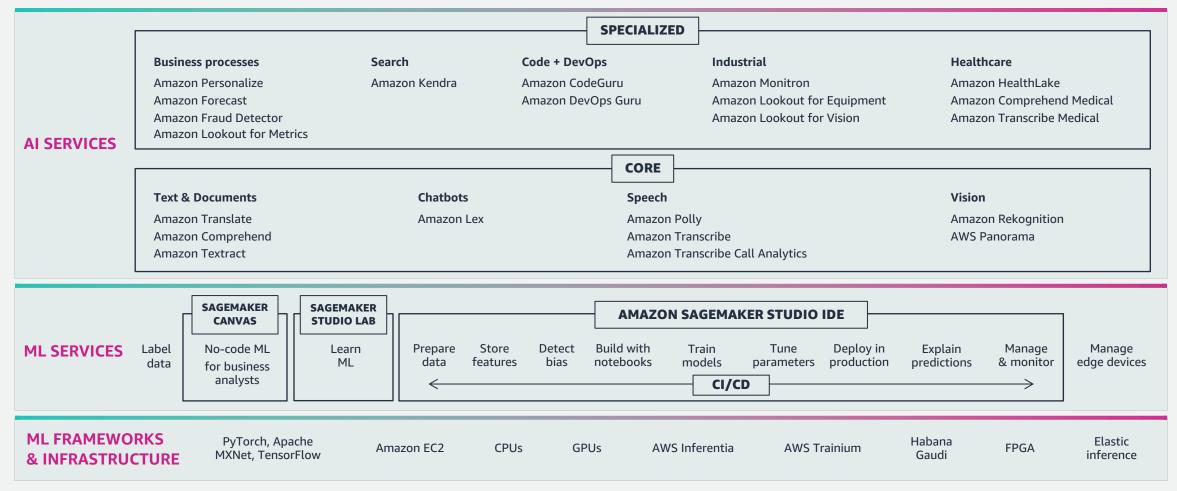


A100, V100, T4 GPUs



## Serviços AI/ML na AWS

### Conjunto mais amplo e completo de recursos de aprendizado de máquina





# Serviços de IA

Labs: <a href="https://github.com/gabrielmartinigit/aiml-notebooks">https://github.com/gabrielmartinigit/aiml-notebooks</a>



# SageMaker

Labs: <a href="https://github.com/gabrielmartinigit/aiml-notebooks">https://github.com/gabrielmartinigit/aiml-notebooks</a>



### **Amazon SageMaker**

#### **PREPARE**

#### SageMaker Ground Truth

Label training data for machine learning

#### SageMaker Data Wrangler NEW

Aggregate and prepare data for machine learning

#### SageMaker Processing

Built-in Python, BYO R/Spark

#### SageMaker Feature Store NEW

Store, update, retrieve, and share features

#### SageMaker Clarify NEW

Detect bias and understand model predictions

#### **BUILD**

#### SageMaker Studio Notebooks

Jupyter notebooks with elastic compute and sharing

### Built-in and Bring your-own Algorithms

Dozens of optimized algorithms or bring your own

#### **Local Mode**

Test and prototype on your local machine

#### SageMaker Autopilot

Automatically create machine learning models with full visibility

#### SageMaker JumpStart NEW

Pre-built solutions for common use cases

#### **TRAIN & TUNE**

#### **Managed Training**

Distributed infrastructure management

#### SageMaker Experiments

Capture, organize, and compare every step

#### Automatic Model Tuning

Hyperparameter optimization

#### **Distributed Training NEW**

Training for large datasets and models

#### SageMaker Debugger NEW

Debug and profile training runs

#### **Managed Spot Training**

Reduce training cost by 90%

#### **DEPLOY & MANAGE**

#### **Managed Deployment**

Fully managed, ultra low latency, high throughput

### Kubernetes & Kubeflow Integration

Simplify Kubernetes-based machine learning

#### **Multi-Model Endpoints**

Reduce cost by hosting multiple models per instance

#### SageMaker Model Monitor

Maintain accuracy of deployed models

#### SageMaker Edge Manager NEW

Manage and monitor models on edge devices

#### SageMaker Pipelines NEW

Workflow orchestration and automation

#### SageMaker Studio

Integrated development environment (IDE) for ML



## **Tipos de Treinamento**



ML Marketplace



Algoritmos Built-in

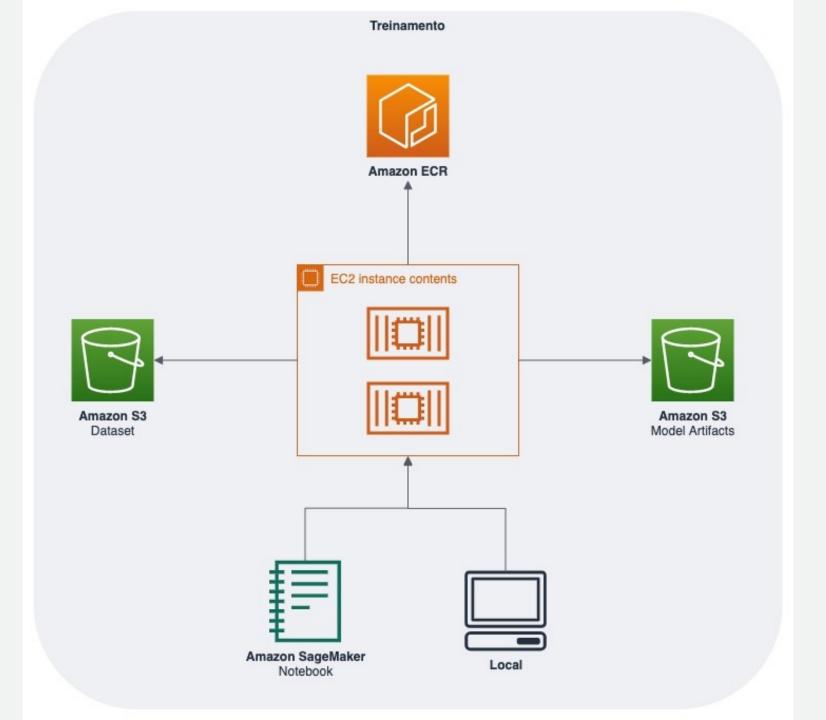


Script Mode

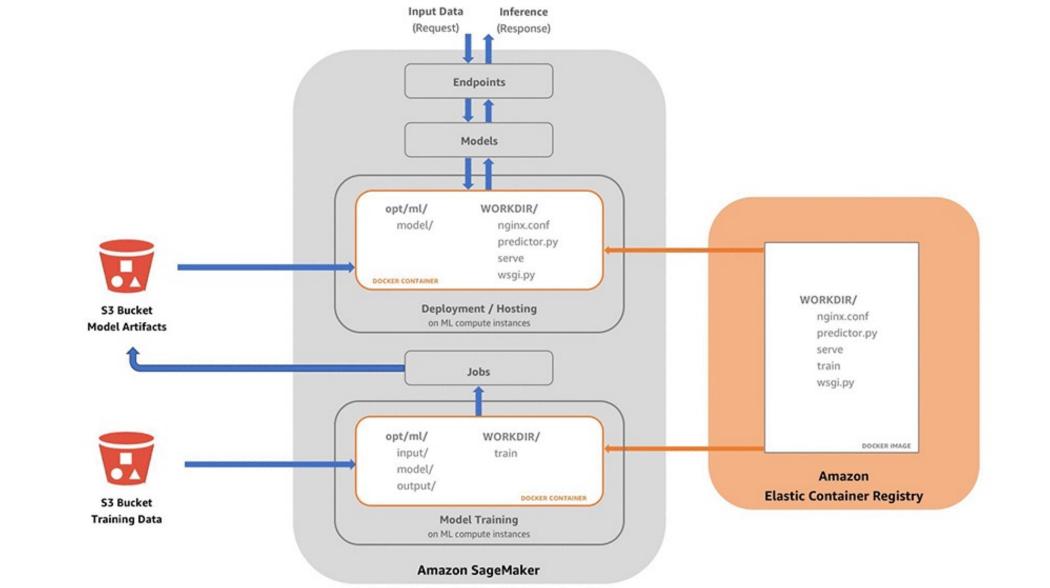


Custom Container











## **MLOps**

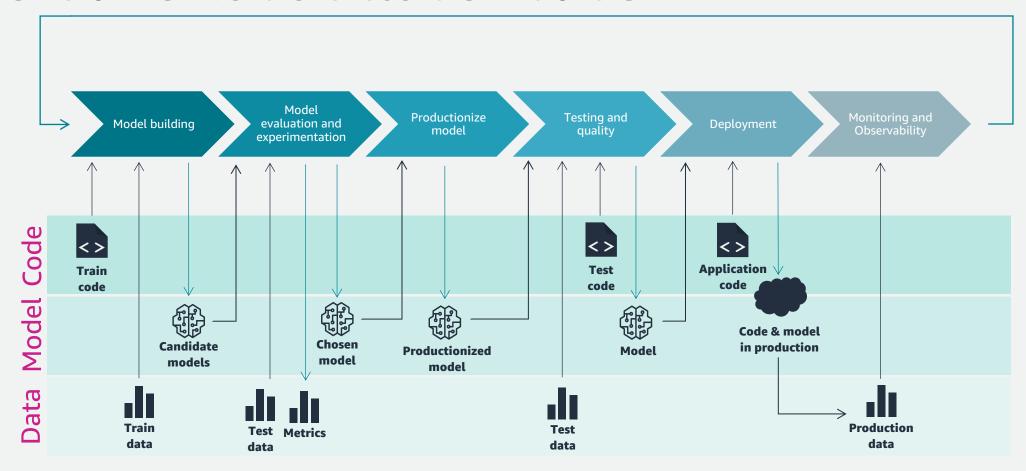
AWS MLOPS whitepaper: <a href="https://d1.awsstatic.com/whitepapers/mlops-continuous-delivery-machine-learning-on-aws.pdf">https://d1.awsstatic.com/whitepapers/mlops-continuous-delivery-machine-learning-on-aws.pdf</a>



	ROLE	PRIORITIES	NEEDS
	Data Scientist	Makes sense of data, generates and communicates insights to improve or create business processes, creates predictive ML models to support them	<ul><li>Data access</li><li>ML compute environments</li><li>Robust ML tools</li></ul>
	Data Engineer	Builds scalable pipelines, transforms and loads data into structures complete with metadata that can be readily consumed by the Data Scientist	<ul><li>Ad hoc querying</li><li>Quick visualization</li></ul>
	Security	Risk and Compliance across the enterprise. Prevent data leakages. Audit user activity. Detect model bias.	<ul> <li>Alerts – data leakages, breaches</li> <li>Reports – data, user activities</li> </ul>
\ /	MLOps / ML Engineer	Monitoring for reliability, quickly diagnose deployment or availability issues	<ul><li>Data drift, Model drift</li><li>Dashboards</li></ul>
\ /	SysOps Engineer	Provision the right infrastructure for the right team without incurring idle resources expenses	<ul><li>Capacity planning</li><li>Resource usage</li><li>Governance at scale</li></ul>
	Business Sponsor	Vetting the priortization and ROI, funding projects, providing ongoing feedback	<ul><li>Reporting</li><li>Results</li><li>Dashboards</li></ul>



### Gerenciamento do ciclo de vida de ML





### Estudar AI/ML na AWS

- Machine Learning University: <a href="https://www.youtube.com/channel/UC12LqyqTQYbXatYS9AA7">https://www.youtube.com/channel/UC12LqyqTQYbXatYS9AA7</a>
   <a href="https://www.youtube.com/channel/UC12LqyqTQYbXatYS9AA7">Nuw</a>
   <a href="https://www.youtube.com/channel/UC12LqyqTQYbXatYS9AA7">Nuw</a>
   <a href="https://www.youtube.com/channel/UC12LqyqTQYbXatYS9AA7">https://www.youtube.com/channel/UC12LqyqTQYbXatYS9AA7</a>
   <a href="https://www.youtube.com/channel/UC12LqyqTQYbXatYS9AA7">https://www.you
- AWS Skill Builder: <a href="https://explore.skillbuilder.aws/lms/">https://explore.skillbuilder.aws/lms/</a>
- SageMaker Studio Lab: <a href="https://studiolab.sagemaker.aws/">https://studiolab.sagemaker.aws/</a>





# **Obrigado!**

Gabriel Bella Martini Arquiteto de Soluções AWS