Public Preview

AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

Public Preview

AIM234

Introducing Amazon SageMaker Studio Lab

Michele Monclova
Principal SageMaker PMT
AWS

Emily Webber
Sr. ML Specialist SA
AWS

Our vision

Make machine learning and data science accessible to all customers (learners, developers, data scientists)

Lack of data science and machine learning talent slows down innovation

250K

Quanthub-estimated data science skills gap in 2020

122K

Open data science jobs on LinkedIn in November 2021

13%

VentureBeat percentage of DSML projects that make it into production

How do most people learn about data science?

MOOCs and universities

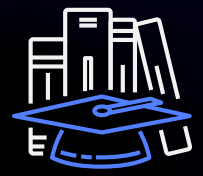
Blog posts

GitHub

On-the-job training and hackathons

Jupyter IDEs

What do our customers want?



Academics
I want the right skills for a great career

Basic theory and learn Python/R

Developers
I want to expand my technical skills with data science

Learn Python/R corporate data + Environment to practice

Data scientists
I want experiment ML and move them into production

Data science communities

What aspiring data scientists need

→ Jupyter notebook environment	Based on JupyterLab
→ Easy to get started	Free, no cloud infrastructure setup
→ Satisfactory compute	CPU (T3.Xl) and GPU (G4D.XL)
→ Time to code	Save ML project, pick up where left
→ Version control management	Integrated with Git
→ Supportive community	Integrated with GitHub
→ Full support of shell commands	Terminal access

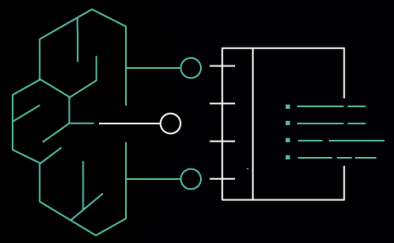
aws

What is Amazon SageMaker Studio Lab

A JUPYTER NOTEBOOK SERVICE TO HELP CUSTOMERS MASTER THEIR SKILLS

Amazon SageMaker Studio Lab

A no-charge, no-configuration service that enable data scientists to learn and experiment with machine learning



Create an account with an email address – free

No setup or configuration required

15 GBs to save your work projects.

As many compute sessions as you need – CPU (12 hrs)/GPU (4 hrs)

Access any notebook on GitHub

Migrate to SageMaker Studio when ready

aws

Provides a beginners view to the stack

Amazon SageMaker

PREPARE

SageMaker Ground Truth
Label training data for machine learning

SageMaker Data Wrangler
Aggregate and prepare data for machine learning

SageMaker Processing
Built-in Python, BYO R/Spark

SageMaker Feature Store
Store, update, retrieve, and share features

SageMaker Clarify
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
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 libraries
Training for large datasets and models

SageMaker Debugger
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 and Kubeflow Integration
Simplify Kubernetes-based machine learning

Multi-model endpoints
Reduce cost by hosting multiple models per instance

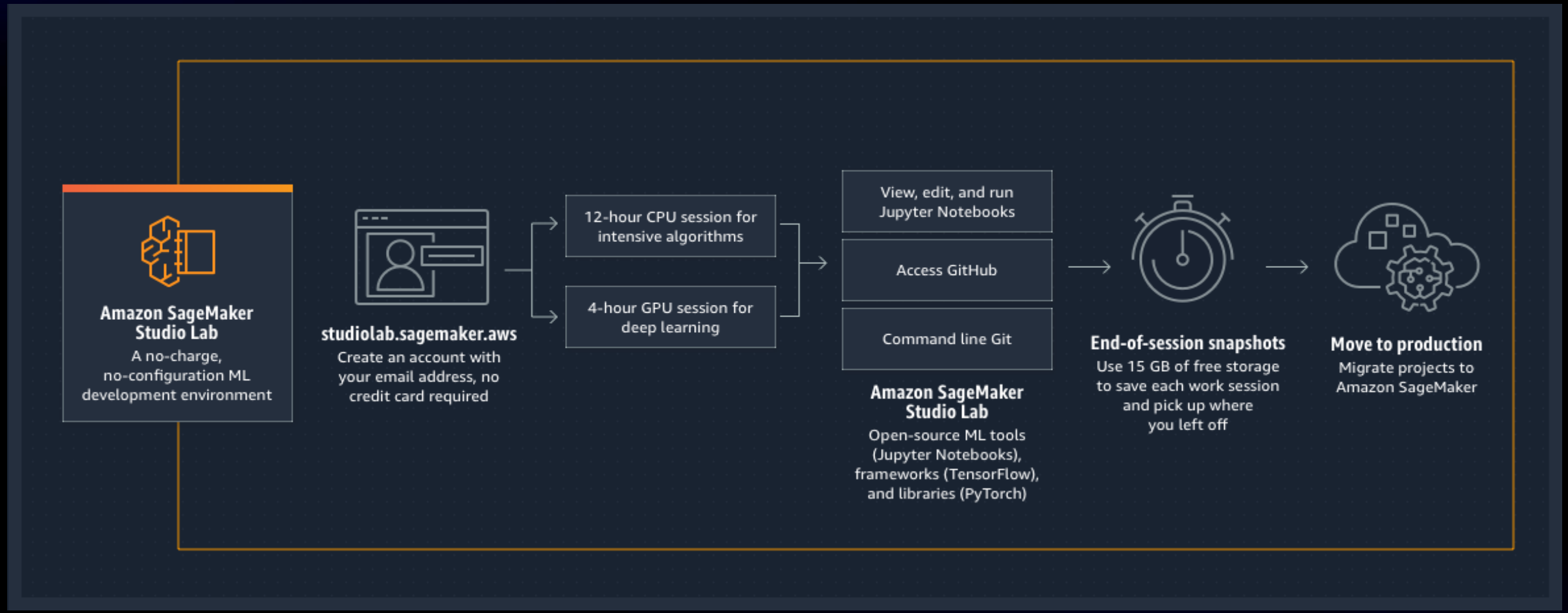
SageMaker Model Monitor
Maintain accuracy of deployed models

SageMaker Edge Manager
Manage and monitor models on edge devices

SageMaker Pipelines
Workflow orchestration and automation

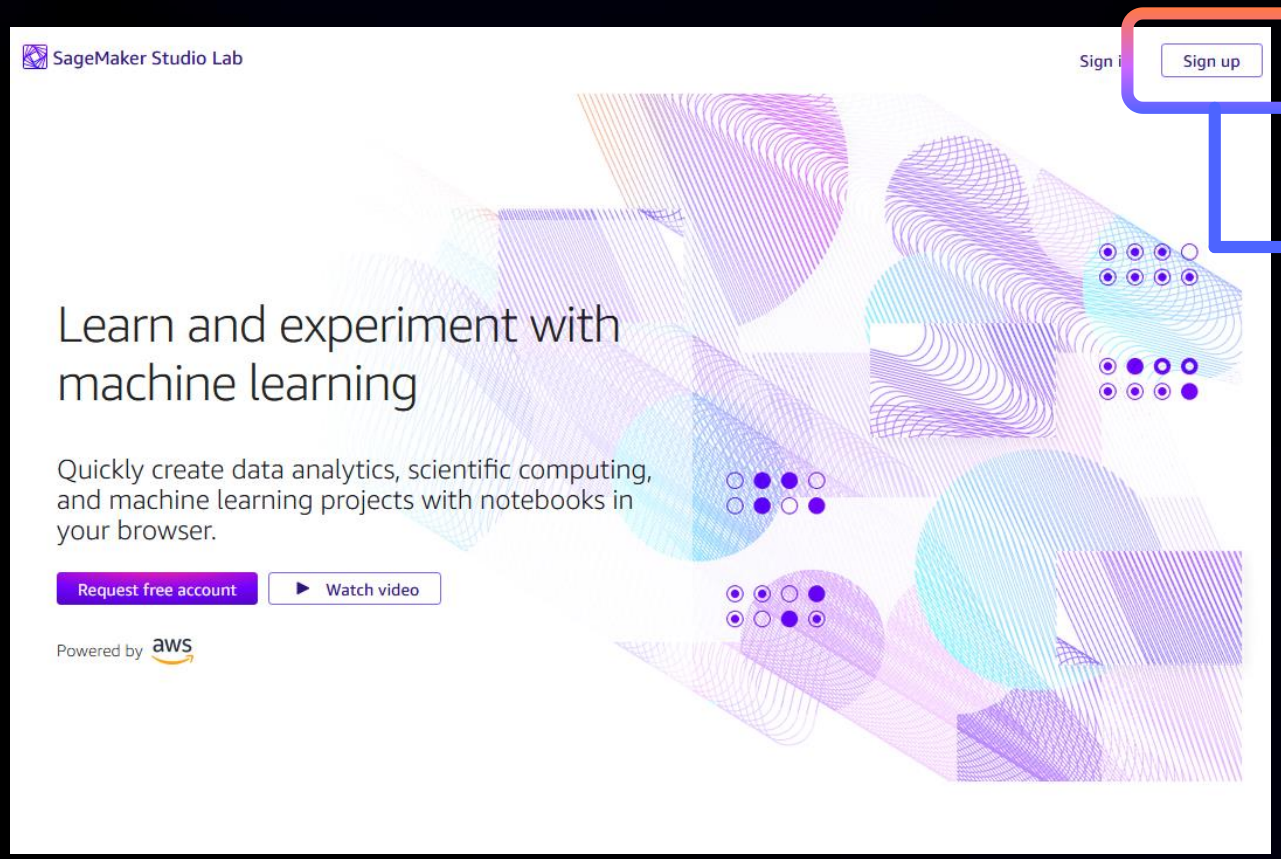
SageMaker Studio
Integrated development environment (IDE) for ML

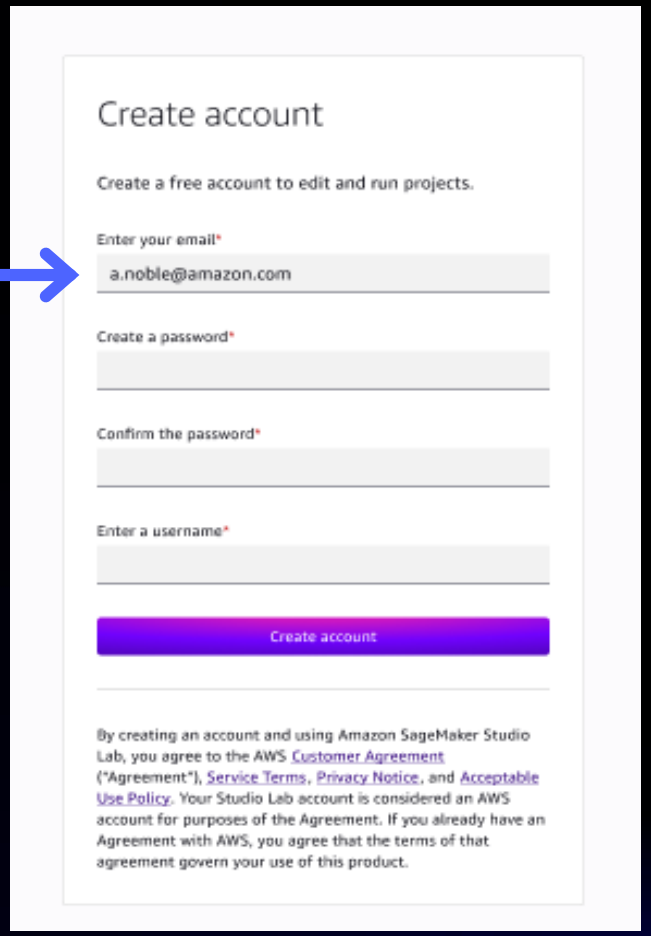
How does it work?



https://studiolab.sagemaker.aws/

Request and register for an account; no credit card or AWS account required





Project page

Status of your project

- Time remaining
- Select runtime type
- Start or stop instance

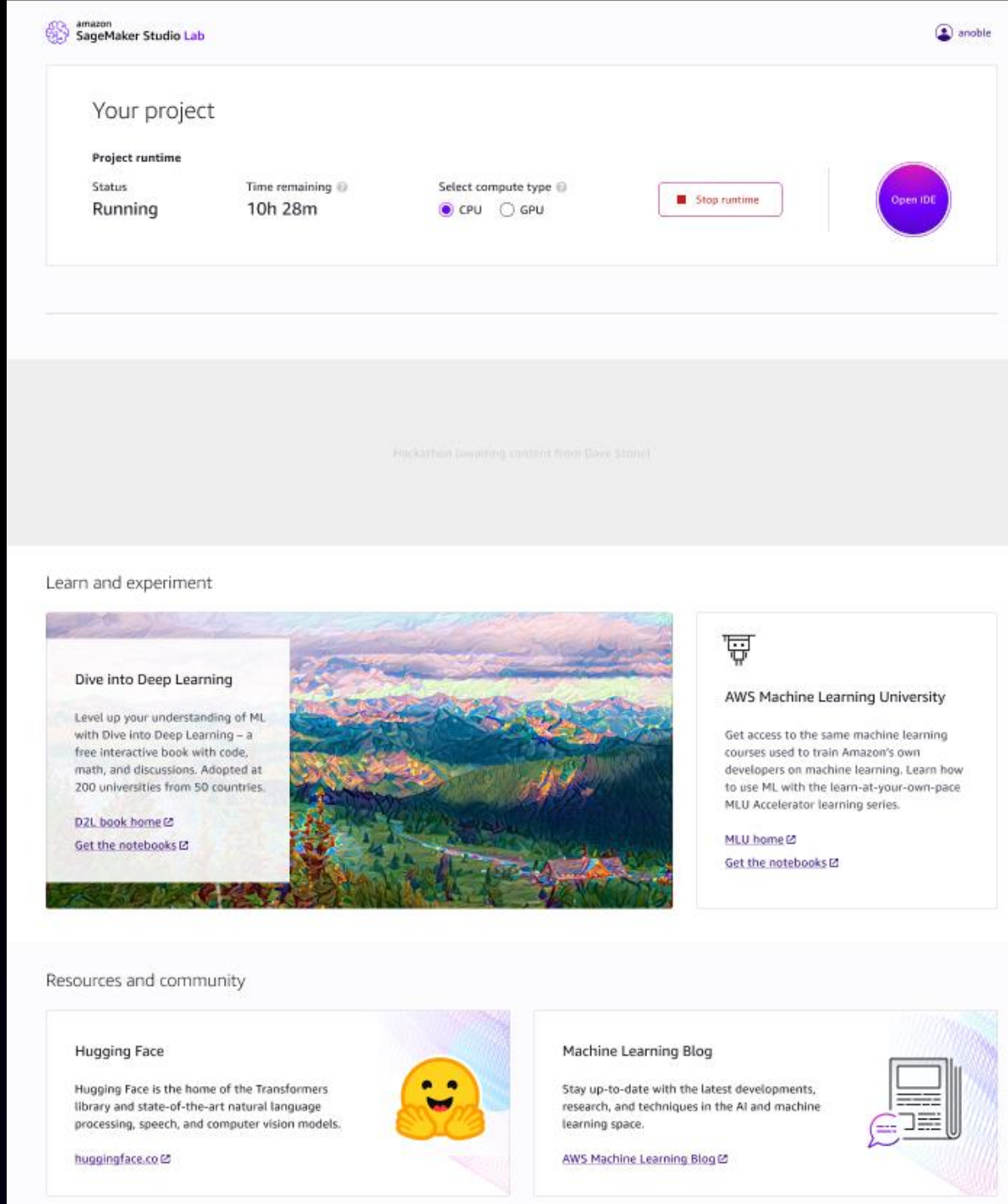
Start a user session

- CPU—12 hours
- GPU—4 hours

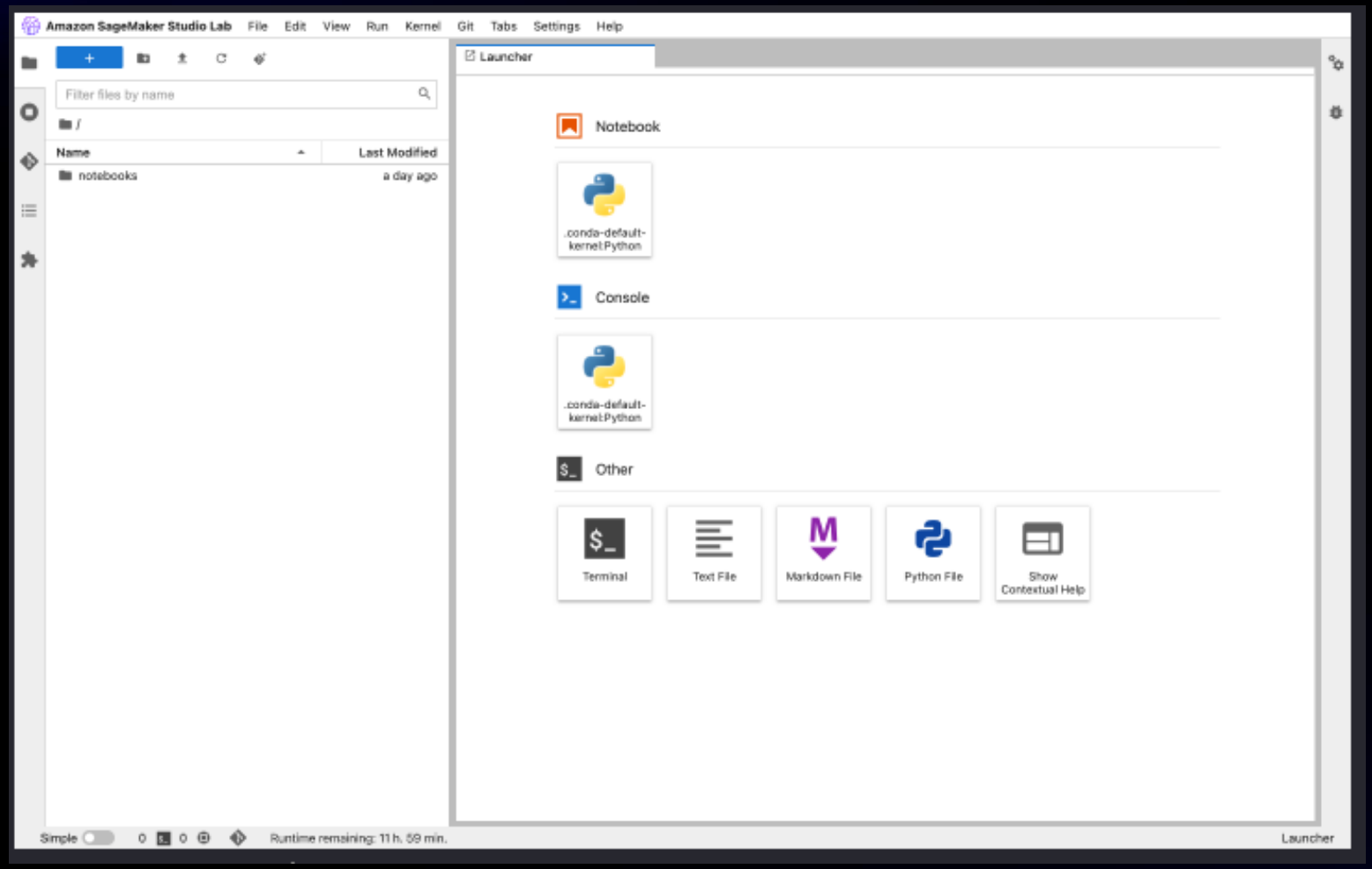
Open project (in a new browser tab)

Assets to get you started

- Dive into Deep Learning
- AWS Machine Learning University
- Hugging Face
- Popular blogs
- Community links



Notebook development environment



Familiar JupyterLab experience

- Terminal access
- Git/GitHub

Your ML environment on AWS

- Compute dedicated to you
- 12 hours CPU/4 hours GPU
- Install the libraries you want
- Dedicated 15 GB for your project
- Unlimited user sessions
- Pick up where you left off

Beta customer quotes

“At Hugging Face, our mission is to democratize state-of-the-art machine learning (ML). With Amazon SageMaker Studio Lab, AWS is doing just that by enabling anyone to learn and experiment with ML through a web browser, without the need for a high-powered PC or a credit card to get started. This makes ML more accessible and easier to share with the community. We are excited to be part of this launch and contribute Hugging Face transformers examples and resources to make ML even more accessible!”

What universities are saying

"Amazon SageMaker Studio Lab will help my students learn the building blocks of machine learning by removing the cloud configuration steps required to get started. Now, in my natural language processing classes, students have more time to enhance their skills."

Sanjiv Das, Professor of Finance and Data Science at Santa Clara University

"One of the hardest parts about programming with machine learning is configuring the environment to build. Students usually have to choose the compute instances, security polices, and provide a credit card. My students needed Amazon SageMaker Studio Lab to abstract away all of the complexity of setup and provide a free powerful sandbox to experiment. This lets them write code immediately without needing to spend time configuring the ML environment."

Dan Roth, Distinguished Professor of Computer and Information Science at the University of Pennsylvania

AWS Heroes

Kesha Williams is an AWS Machine Learning Hero and a Principal AWS Training Architect with A Cloud Guru

Cyrus Wong, Data Scientist at Hong Kong IVE, AWS Machine Learning Hero

Mike Chambers, AI/ML Teacher at https://mikegchambers.teachable.com/

aws

Recap and highlights

- Accessing SageMaker Studio Lab: https://studiolab.sagemaker.aws/
    - You do have to request access/approvals are sent via email
    - Re:Invent attendees get priority access
- One account = one user session at a time, as many times as you want
- One user session is 12 max hours for CPU and max 4 hours for GPU
    - You can leave training jobs running
- JupyterLab environment is open
    - Pip install your own libraries with % not !
    - At least 15 GB to save your project

aws

Announcing the global AWS Disaster Response Hackathon

Improve disaster response with machine learning

December 1, 2021 through February 7, 2022

Total of $54,000 USD in prizes – join today!

AWSDisasterResponse.devpost.com

aws

Thank you!

Michele Monclova
monclom@amazon.com

Emily Webber
egwebber@amazon.com

https://studiolab.sagemaker.aws/