AWS re:Invent

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

AIM323

Get started with Amazon SageMaker in minutes

Soji Adeshina
Applied Scientist
AWS

Arun Shankar
AI/ML Solutions Architect
AWS

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

Getting started with machine learning and challenges

How Amazon SageMaker addresses the challenges

SageMaker JumpStart

Demos

From model development to production

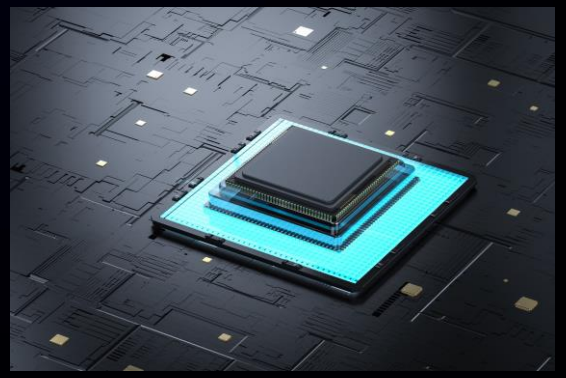
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Getting started with machine learning (ML)



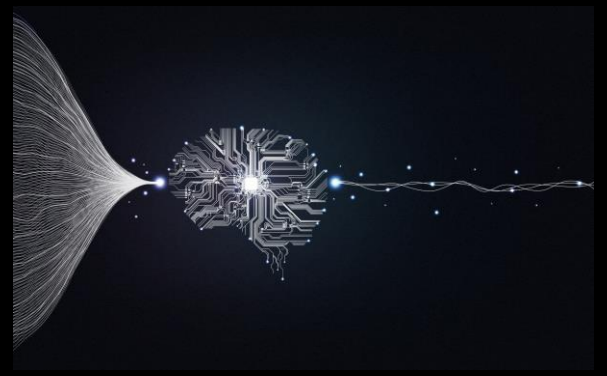
Data

+



Infrastructure

→

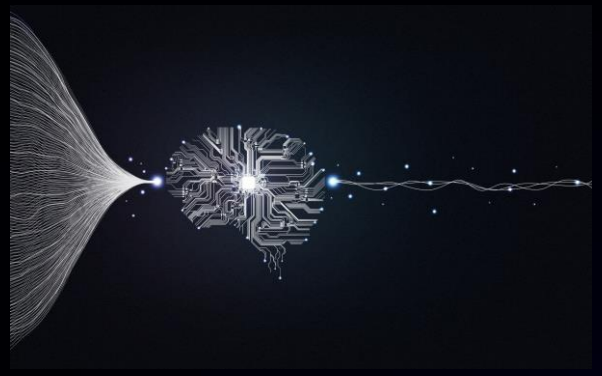


ML model



Inputs

+



ML model

→



Predictions

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Challenges getting started with ML

How do I bring my data in?

What model should I use?

How do I make off-the-shelf models work for my dataset?

How do I integrate the machine learning workflow into my use case?

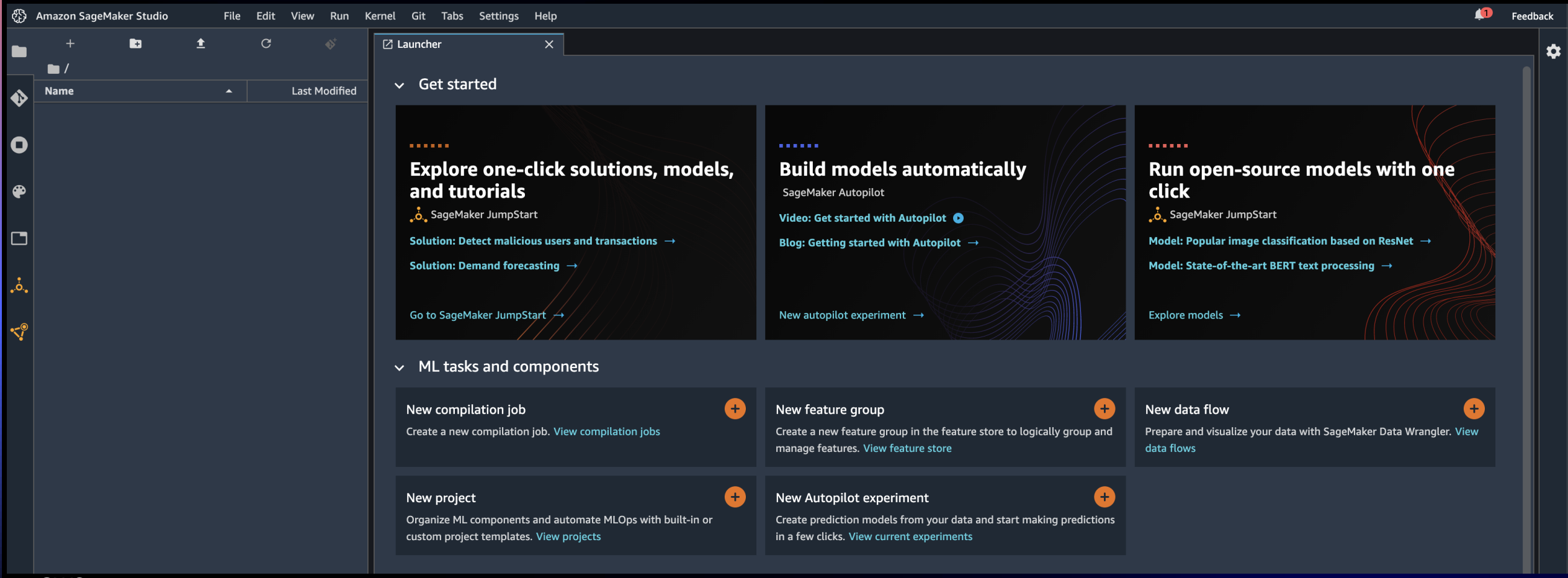
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How SageMaker addresses the challenges

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SageMaker Studio

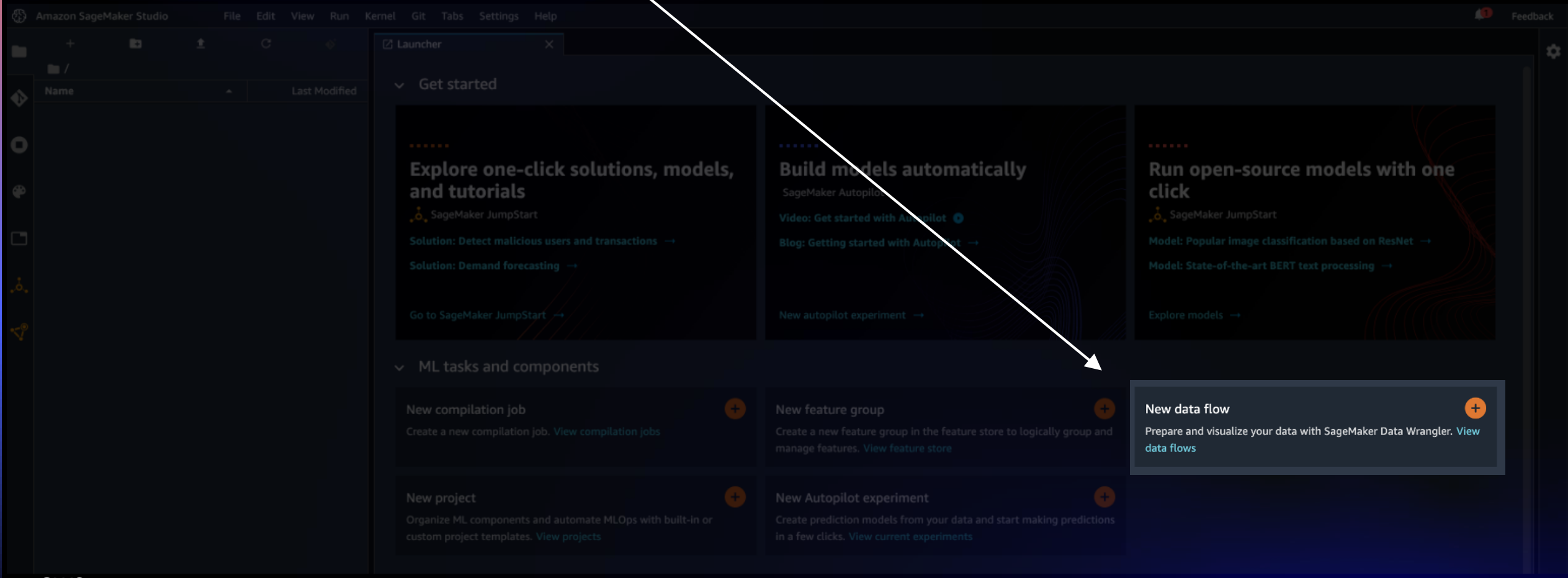
A FULLY INTEGRATED DEVELOPMENT ENVIRONMENT (IDE) FOR MACHINE LEARNING



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Studio

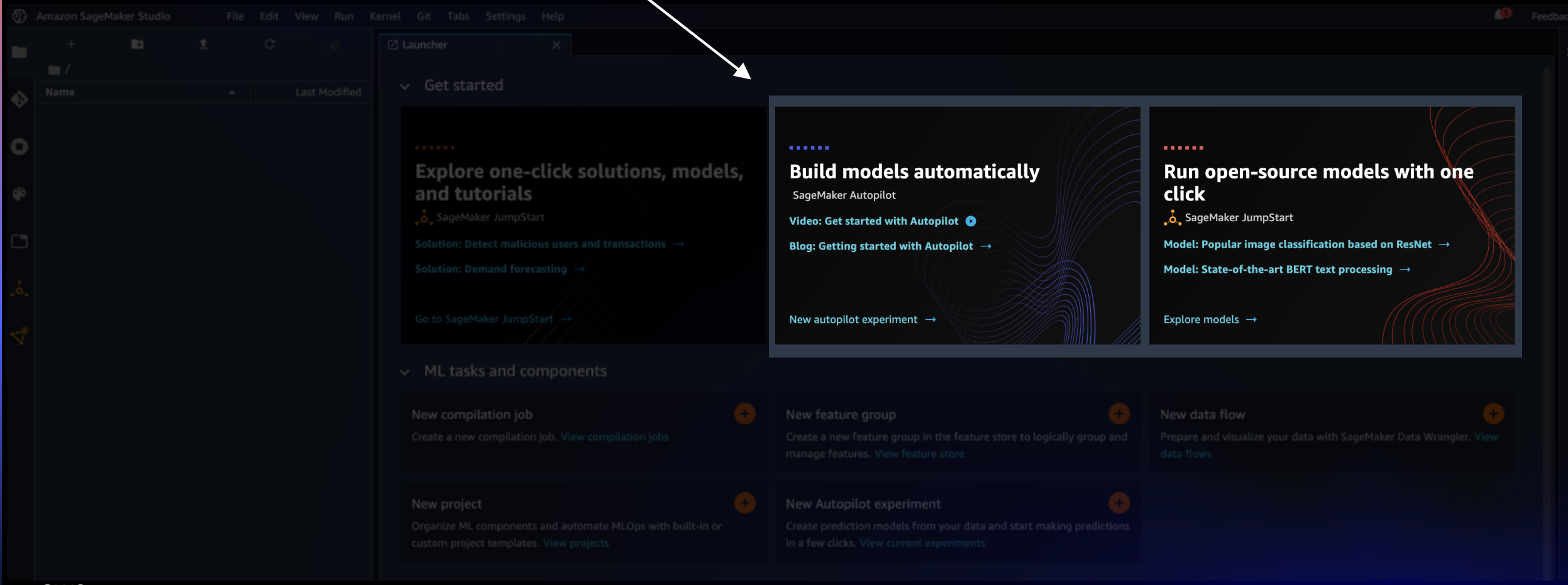
How do I bring my data in?



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Studio

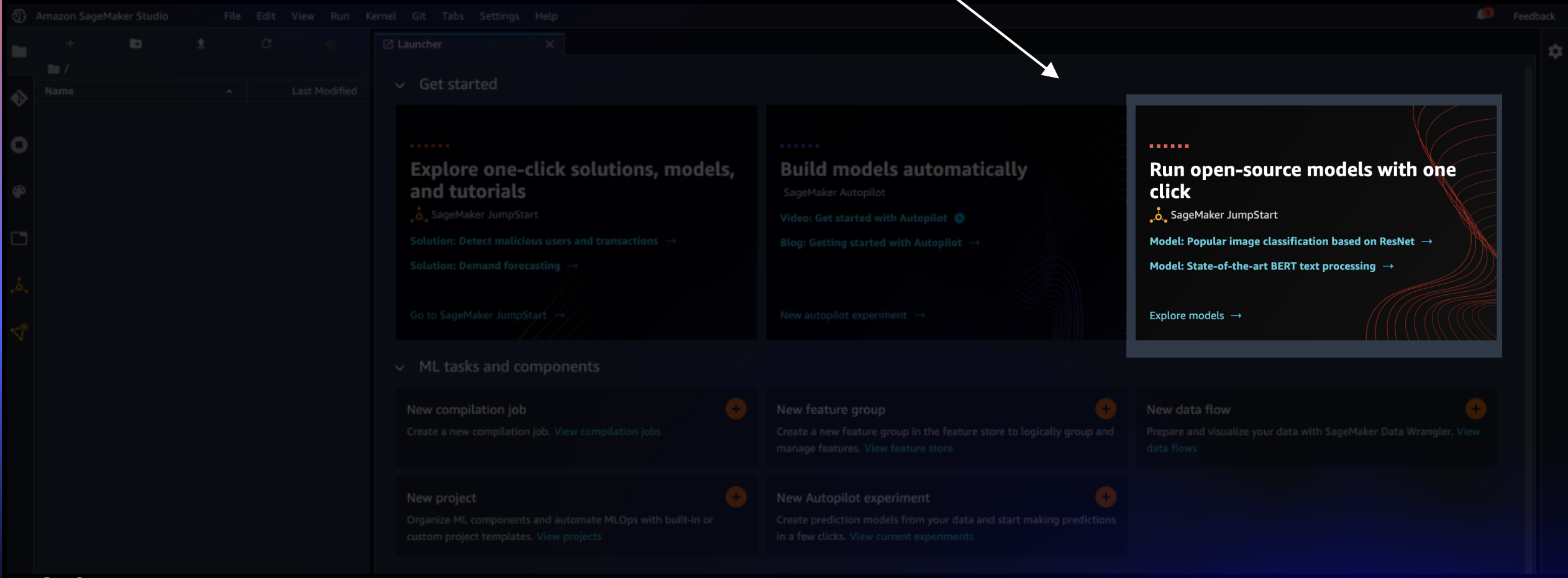
What model should I use?



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Studio

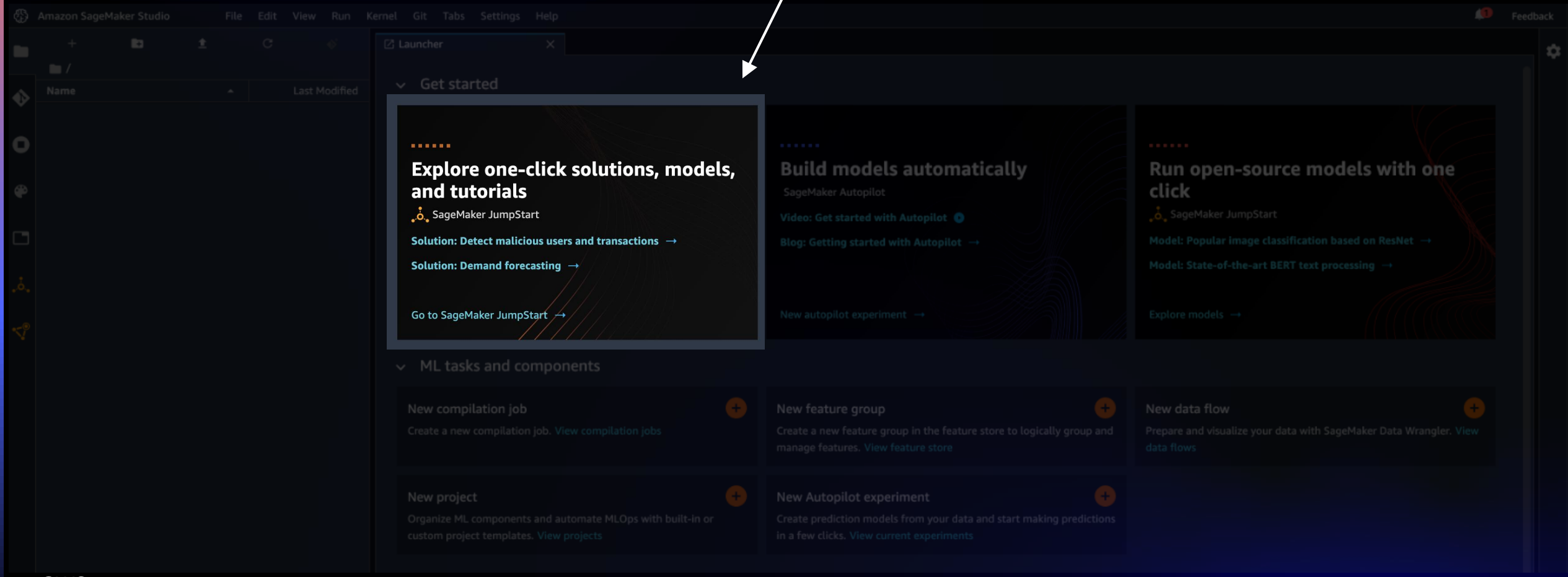
How do I make off-the-shelf models work for my dataset?



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Studio

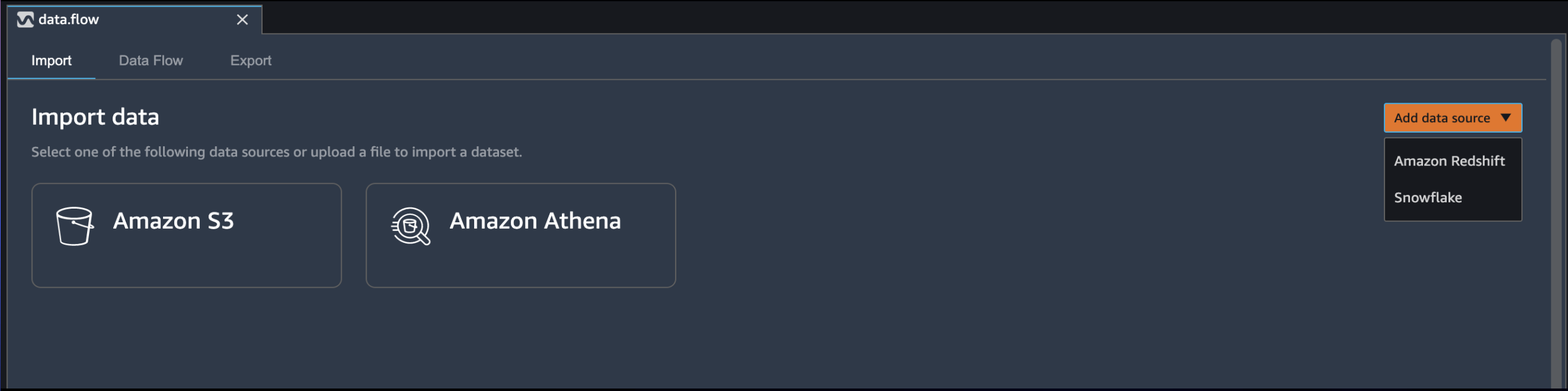
How do I integrate the machine learning workflow into my use case?



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SageMaker Data Wrangler

QUICKLY SELECT AND QUERY DATA

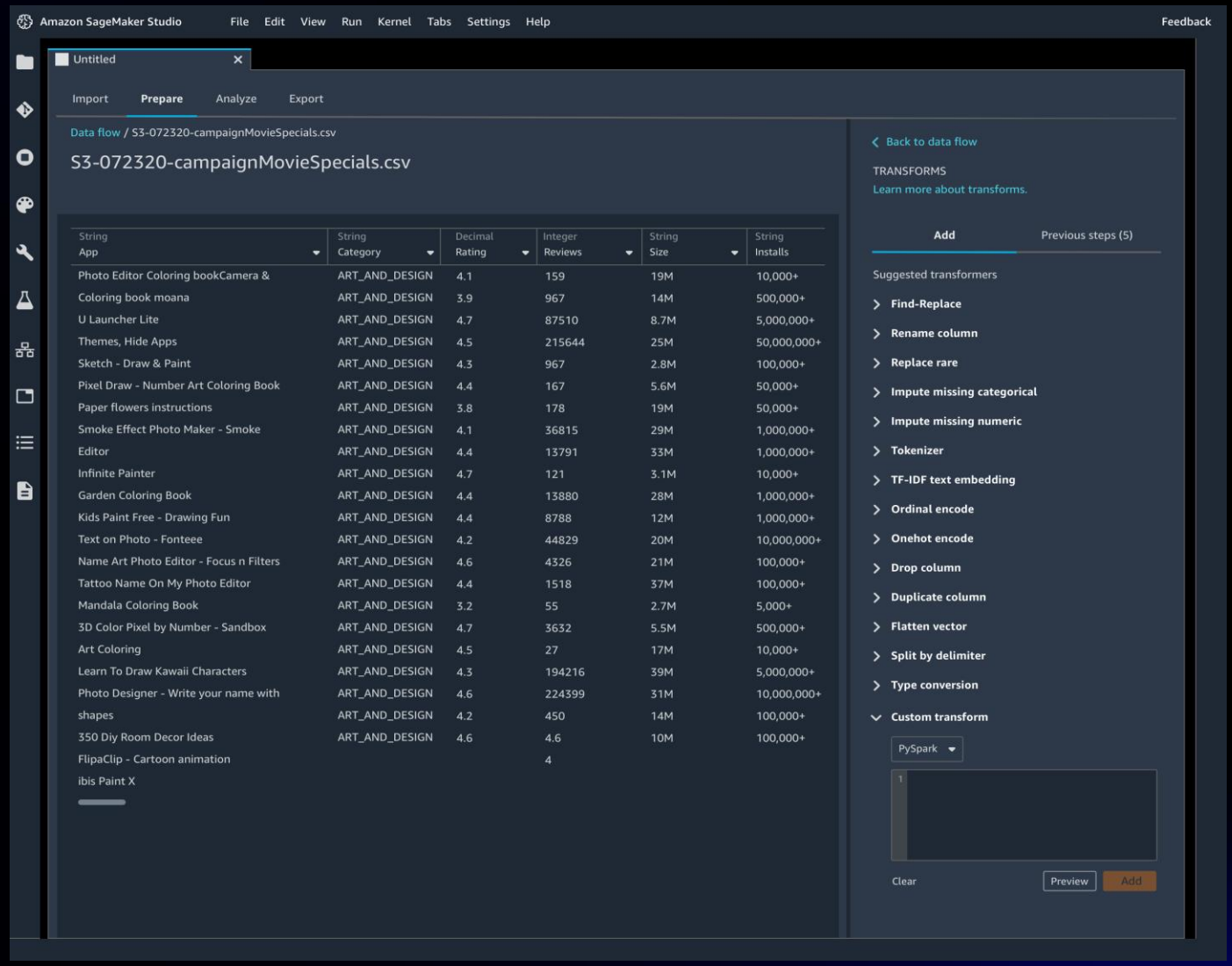


© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data Wrangler

Transform data easily

- 300+ preconfigured data transformations
- Transform your data without writing a single line of code
- Custom transformations in PySpark, SQL, and Pandas

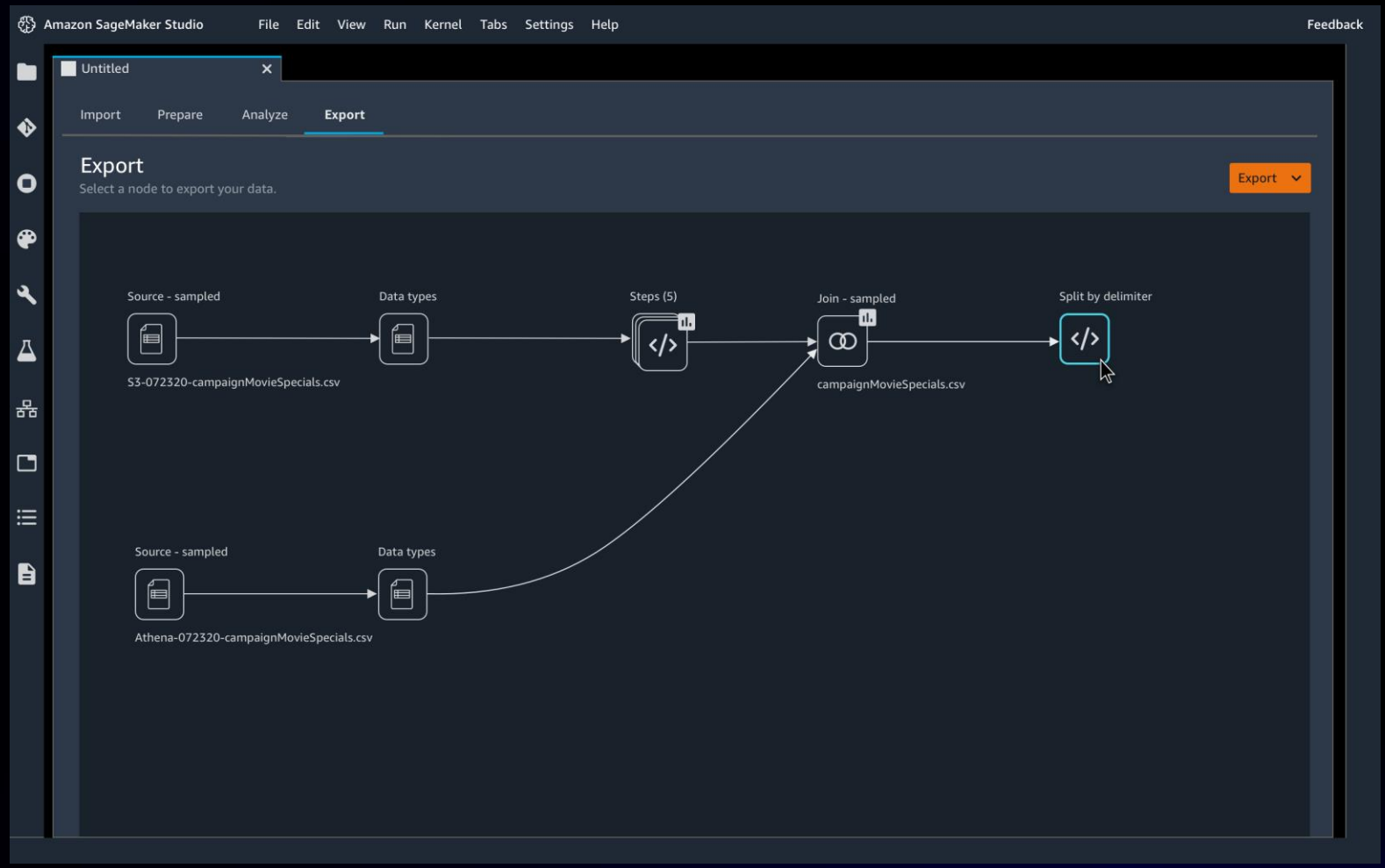


© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data Wrangler

Automate data preparation workflows

- Export your data preparation workflow to a notebook to share
- Make your data preparation workflow repeatable for production with a single click



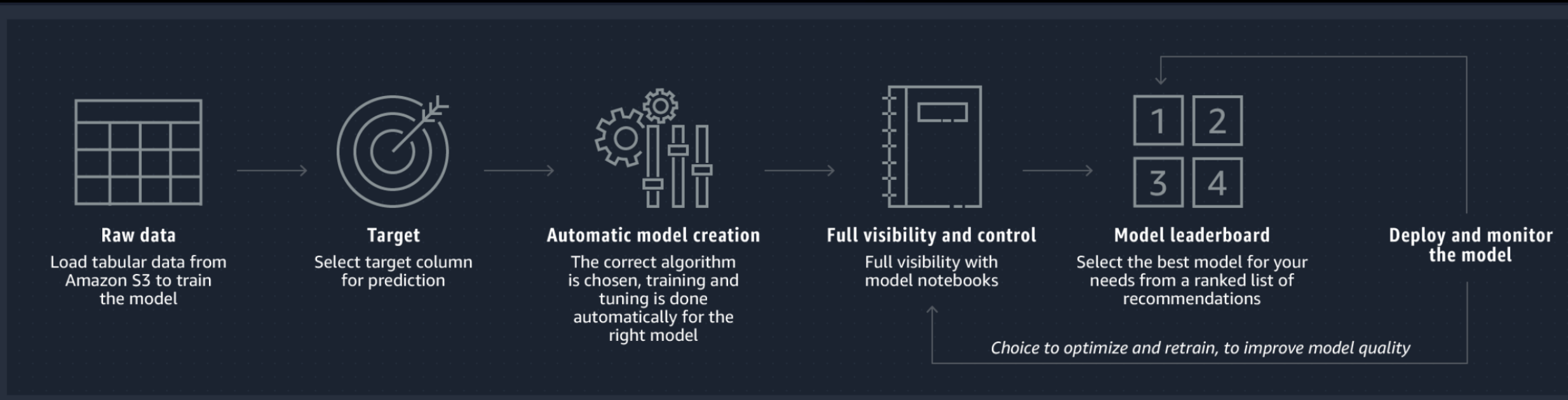
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SageMaker Autopilot

Automatic data preprocessing and feature engineering

Automatic ML model selection

Automatic notebook creation



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Autopilot

Fully transparent AutoML

Amazon SageMaker Autopilot Candidate Definition Notebook

This notebook was automatically generated by the AutoML job **airline-sampled-mc-2**. This notebook allows you to customize the candidate definitions and execute the SageMaker Autopilot workflow.

The dataset has **25** columns and the column named **DelayType** is used as the target column. This is being treated as a **MulticlassClassification** problem. The dataset also has **5** classes. This notebook will build a MulticlassClassification model that **maximizes** the **"ACCURACY"** quality metric of the trained models. The **"ACCURACY"** metric provides the percentage of times the model predicted the correct class.

As part of the AutoML job, the input dataset has been randomly split into two pieces, one for **training** and one for **validation**. This notebook helps you inspect and modify the data transformation approaches proposed by Amazon SageMaker Autopilot. You can interactively train the data transformation models and use them to transform the data. Finally, you can execute a multiple algorithm hyperparameter optimization (multi-algo HPO) job that helps you find the best model for your dataset by jointly optimizing the data transformations and machine learning algorithms.

Available Knobs Look for sections like this for recommended settings that you can change.

Contents

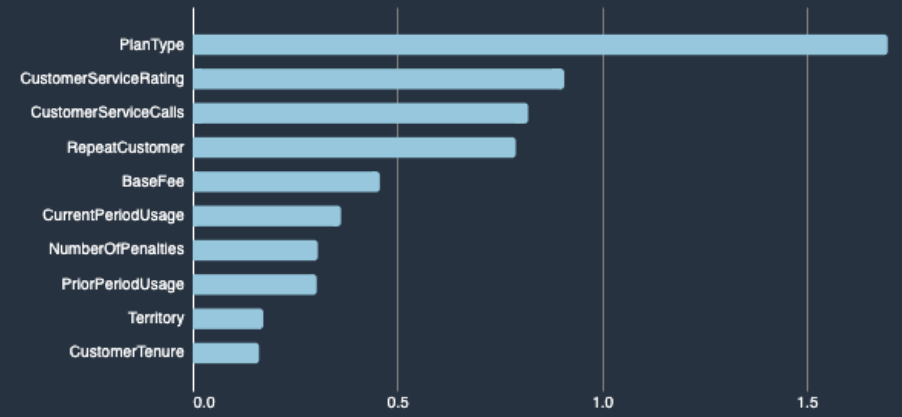
1. Sagemaker Setup
   A. Downloading Generated Candidates
   B. SageMaker Autopilot Job and Amazon Simple Storage Service (Amazon S3) Configuration
2. Candidate Pipelines
   A. Generated Candidates
   B. Selected Candidates
3. Executing the Candidate Pipelines
   A. Run Data Transformation Steps
   B. Multi Algorithm Hyperparameter Tuning
4. Model Selection and Deployment
   A. Tuning Job Result Overview
   B. Model Deployment

Feature importance

Explaining your model's predictions

Amazon SageMaker Studio helps you understand your machine learning model by portraying the importance of its features in terms of SHAP values. We plot the aggregated SHAP value for each feature across all instances of the dataset.

Predicted column: label0



Export PDF report | Download raw data

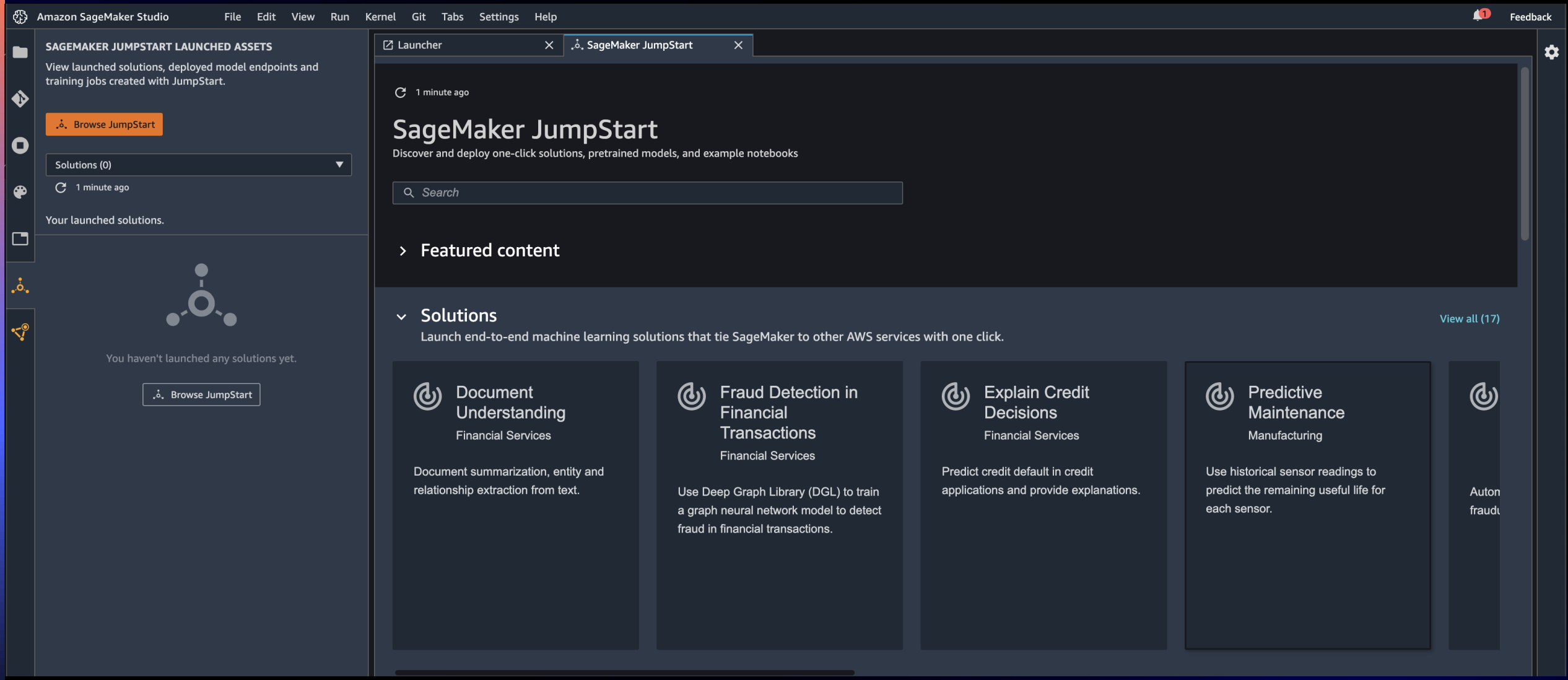
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SageMaker JumpStart

- Solutions
  - Complete ML solutions tying SageMaker to other AWS services
- Models
  - Pretrained open-source models that can be deployed as is or fine-tuned on your dataset
- SageMaker algorithms
  - Notebook tutorials showing how to use built-in SageMaker algorithms on example dataset
- Example notebooks
  - Additional notebook examples showing how to use advanced SageMaker features and functionality
- Blogs
- Video tutorials

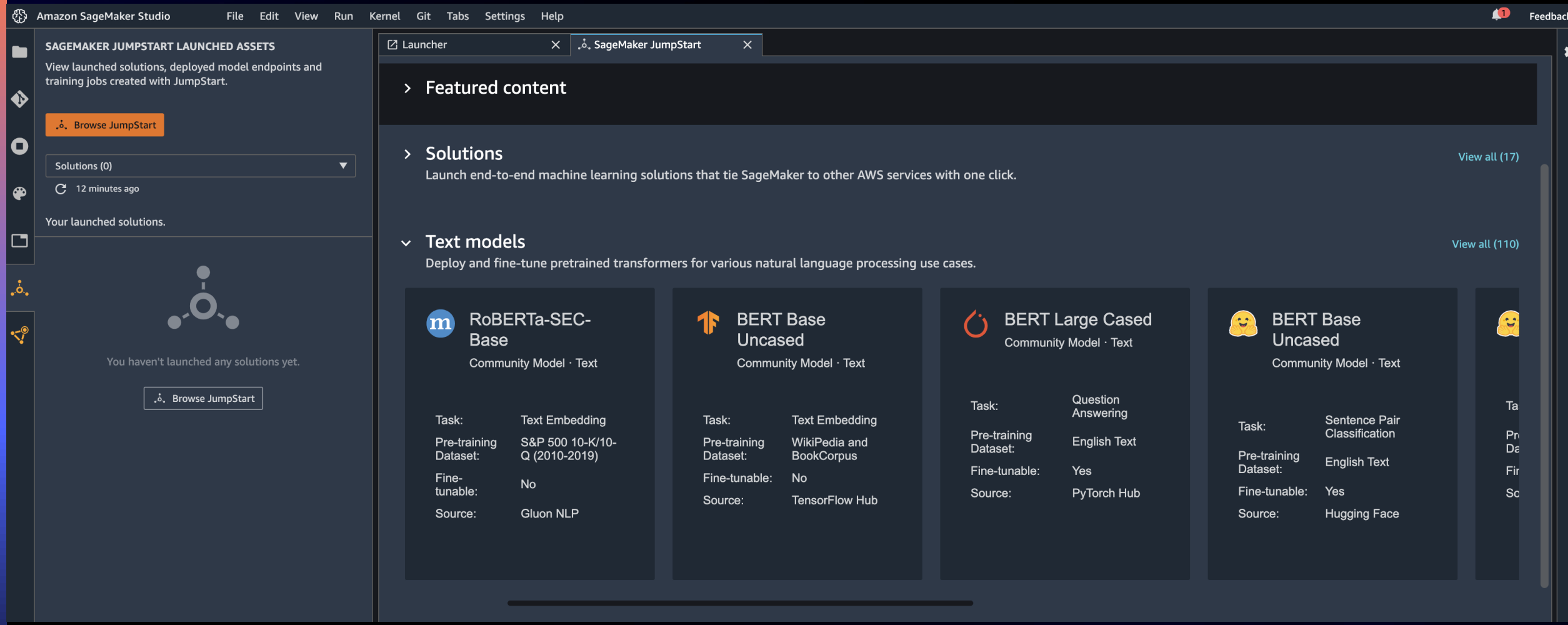
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

JumpStart solutions



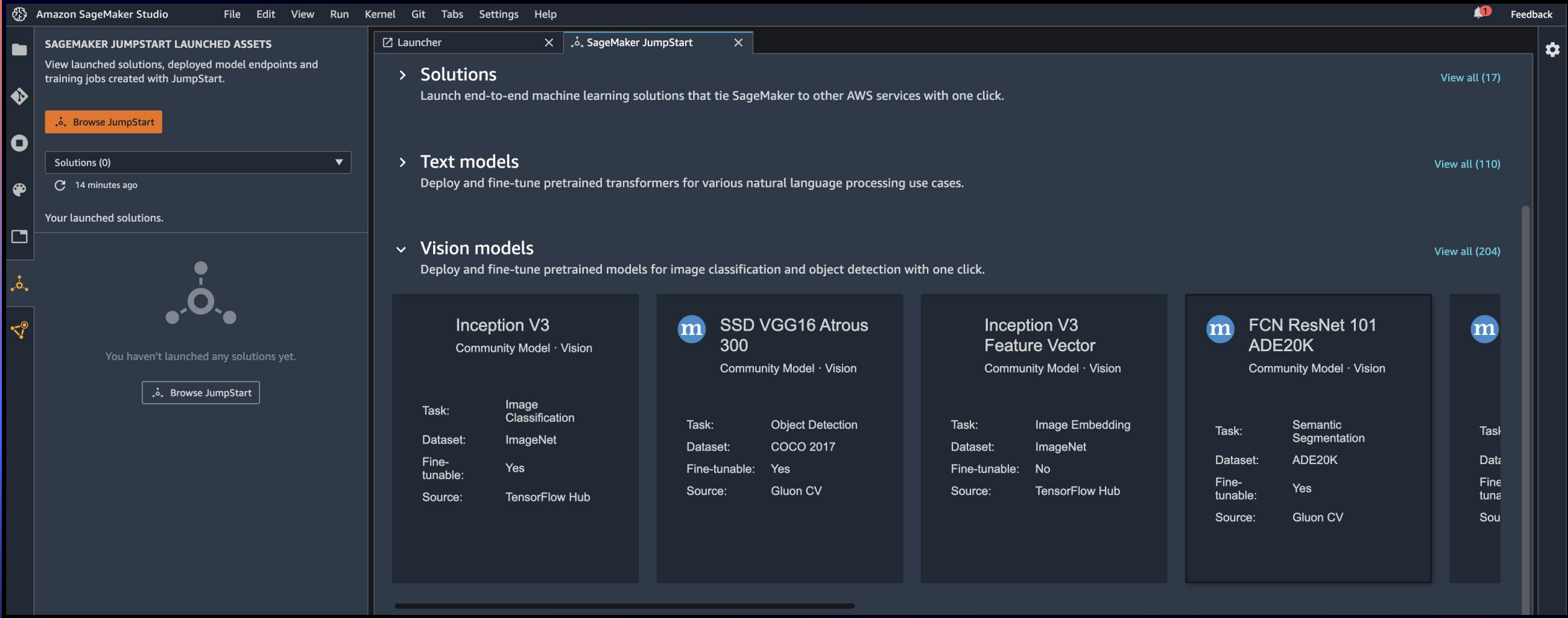
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

JumpStart text models



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

JumpStart vision models



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

From development to production

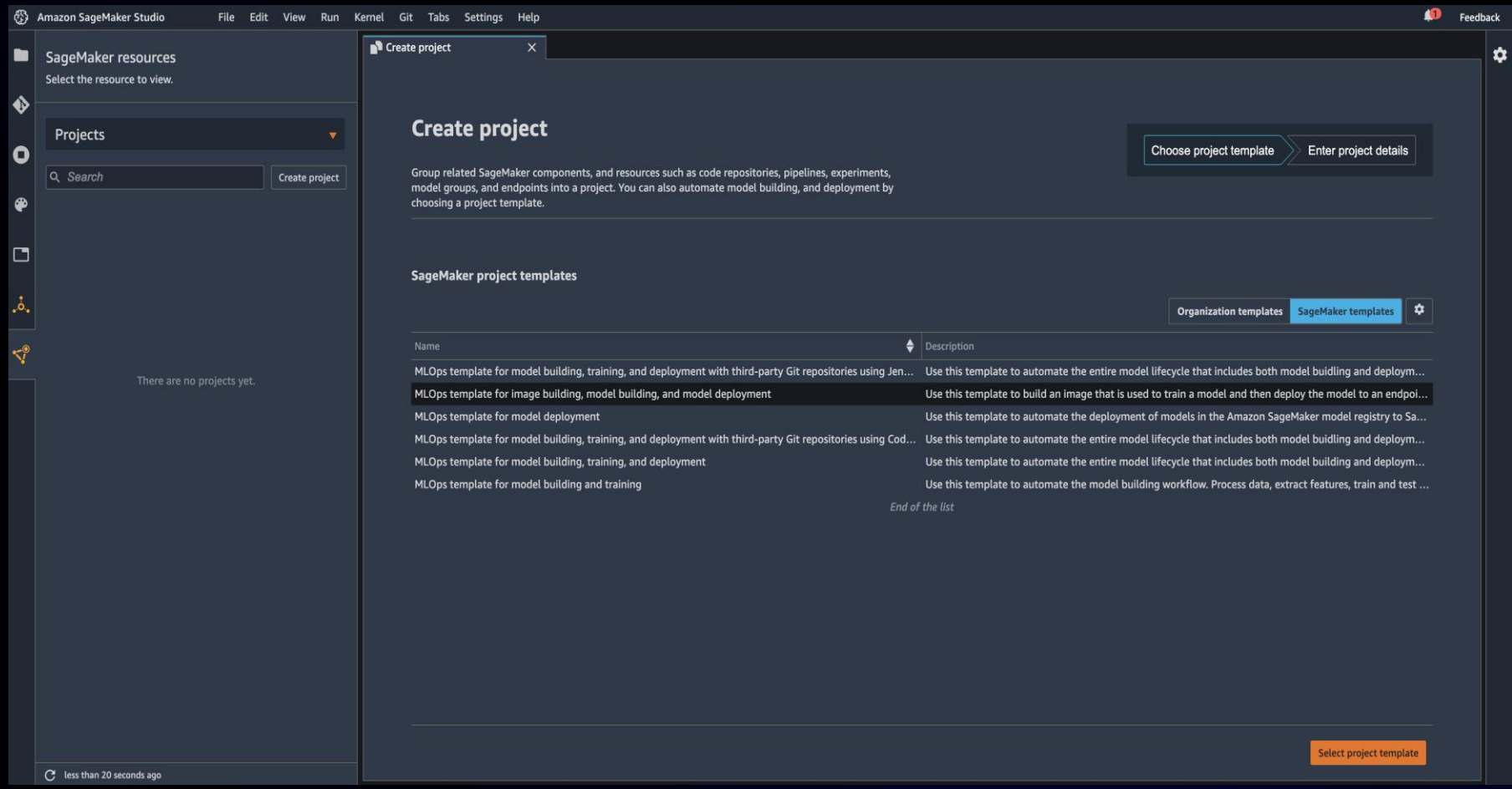
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SageMaker projects

Maintain code consistency and strict version control

Generate models based on triggers such as code check-in or data update

Adopt operational best practices for a production-quality system



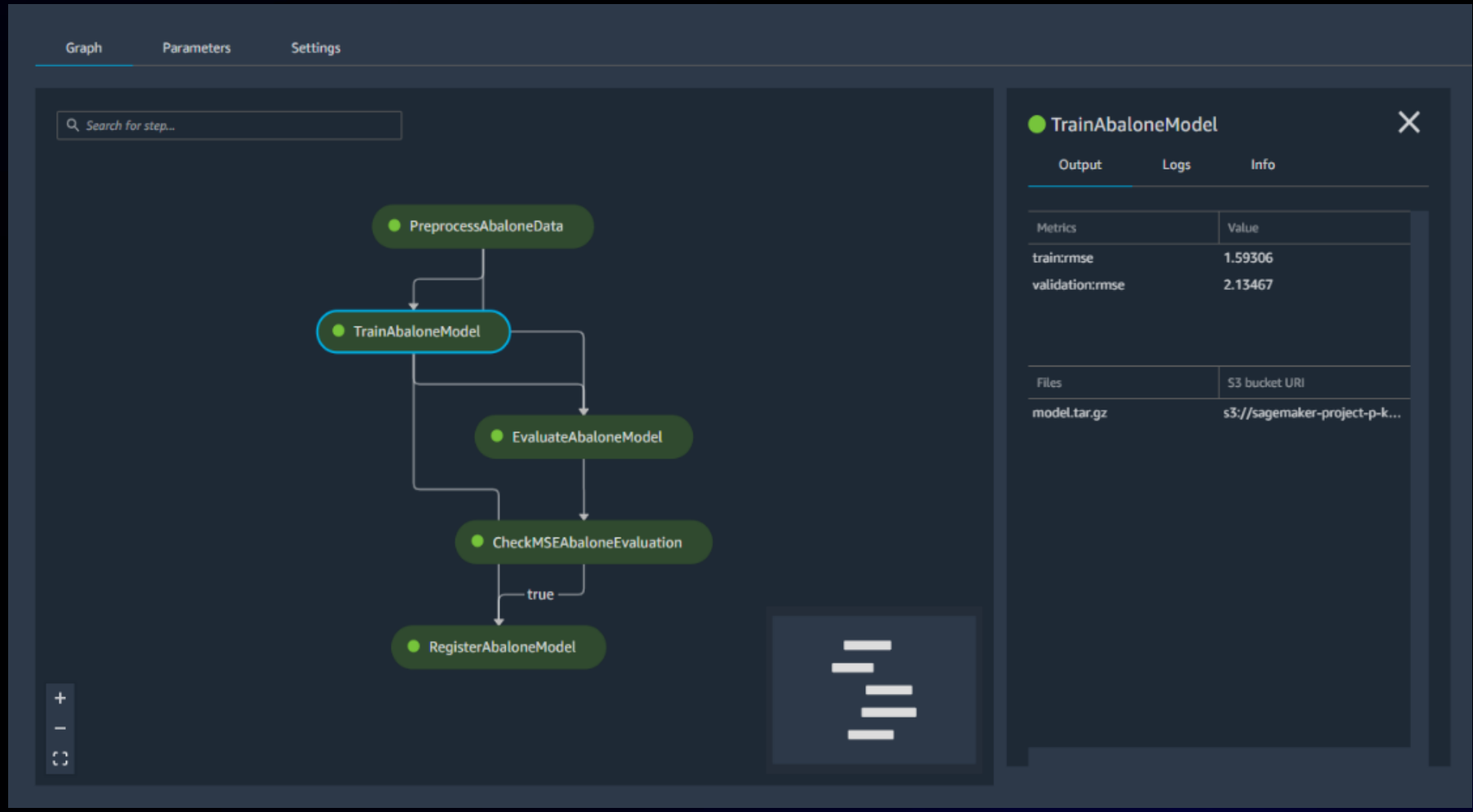
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SageMaker Pipelines

Compose, manage, and reuse ML workflows

Automatic tracking of models

CI/CD for machine learning



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Deploying SageMaker models



Inference endpoints

Real-time endpoint
for online inference



Batch transform

Run batch predictions
on large datasets



Async inference

Queue incoming requests and
process them asynchronously

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Resources

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Resources

SageMaker Studio
https://docs.aws.amazon.com/sagemaker/latest/dg/studio.html

SageMaker Data Wrangler
https://docs.aws.amazon.com/sagemaker/latest/dg/data-wrangler-getting-started.html

SageMaker Autopilot
https://docs.aws.amazon.com/sagemaker/latest/dg/autopilot-automate-model-development.html

SageMaker JumpStart
https://docs.aws.amazon.com/sagemaker/latest/dg/studio-jumpstart.html

Demos
https://github.com/arunprsh/no-code-low-code

SageMaker Pipelines
https://docs.aws.amazon.com/sagemaker/latest/dg/pipelines.html

SageMaker Inference
https://docs.aws.amazon.com/sagemaker/latest/dg/deploy-model.html

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Additional sessions

AIM321 Use AutoML to create high-quality models

AIM320 Implementing MLOps practices with Amazon SageMaker, featuring Vanguard

AIM402 Develop your ML project with Amazon SageMaker

AIM407 Train models at scale with Amazon SageMaker, featuring Aurora

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Soji Adeshina
adesojia@amazon.com

Arun Shankar
arunprsh@amazon.com

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.