

Deep Learning Using TensorFlow and Apache MXNet on AWS SageMaker

MACHINE LEARNING ON THE CLOUD WITH AWS SAGEMAKER



Janani Ravi

CO-FOUNDER, LOONYCORN

www.loonycorn.com

Overview

SageMaker is a fully managed machine learning service offered by AWS

Build, train and deploy machine learning models on the AWS cloud

Integrated Jupyter notebook instance to develop models

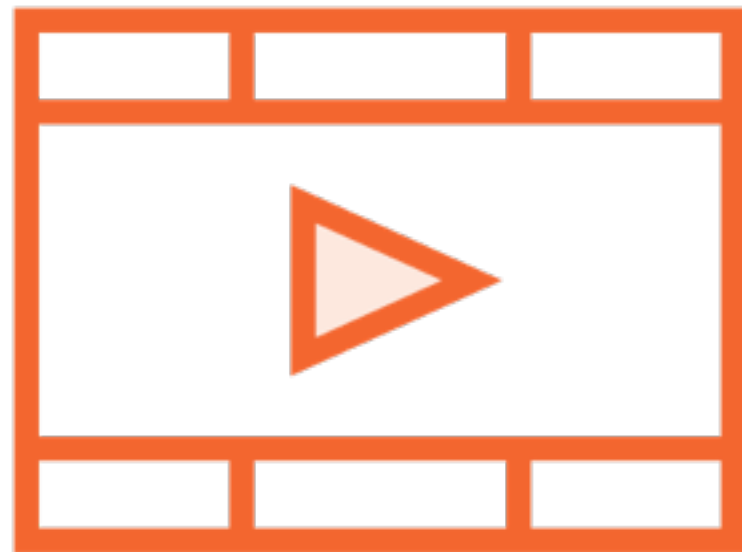
Develop new models or use built-in models on your training data

Bring your own model or container to train and host at scale

Prerequisites and Course Outline

Intermediate to advanced course
on training and deploying ML
models on AWS

Very comfortable building ML
models in Python



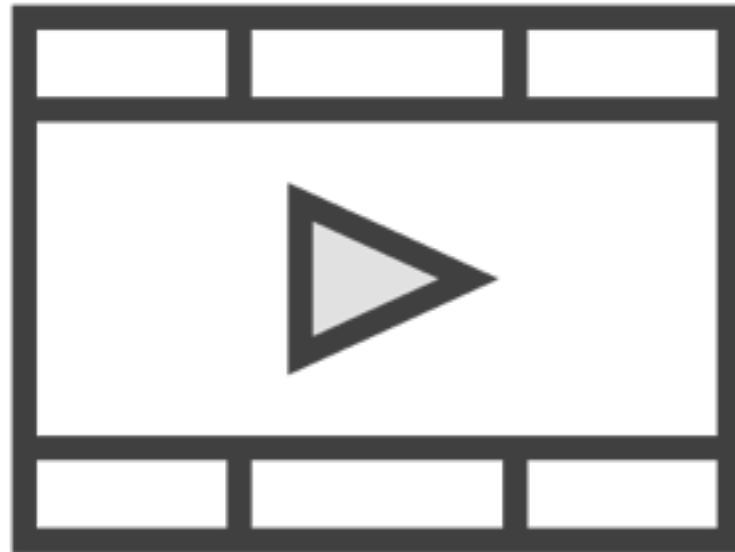
Prerequisite Courses

Understanding Machine Learning with Python

Understanding the Foundations of TensorFlow

Building Classification Models with TensorFlow

Related Courses



Getting Started with Azure Machine Learning



Software and Skills

Be very comfortable programming in Python

Be very comfortable with building ML models and neural networks

Familiar with TensorFlow or Apache MXNet

Understand the basics of cloud computing VMs, instances, distributed computing etc.



Course Outline

Introduction to AWS SageMaker

- Using SageMaker to build, train and deploy models at scale on the cloud

Using built-in models in SageMaker

- Train and deploy built-in models with your own data

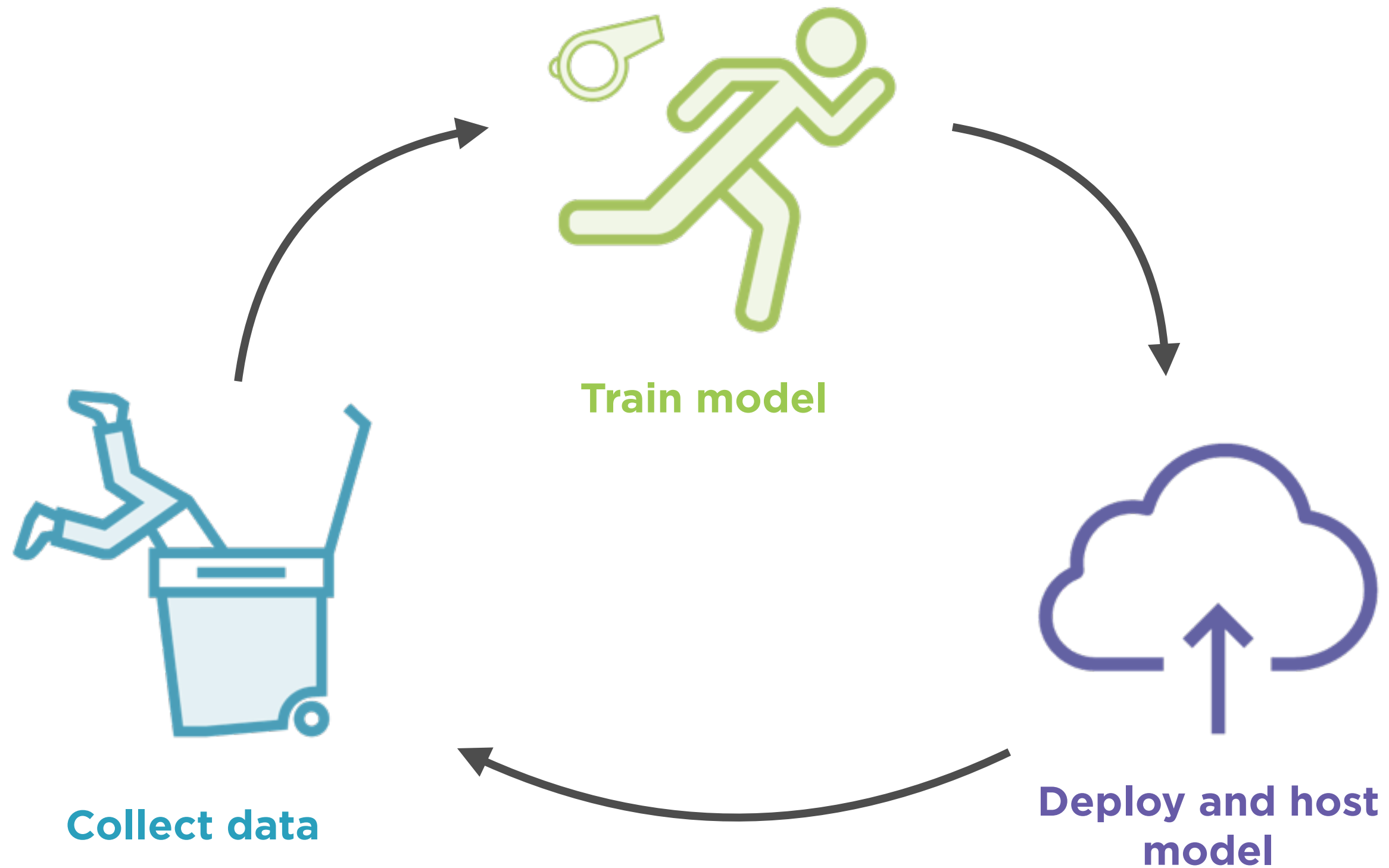
Bring your own code, model and container

- SageMaker offers support for a variety of ways to train and host models

Distributed training and autoscaling

Introducing AWS SageMaker

Machine Learning Workflow





Collect data

Data from many sources

- Database, data warehouse, sensors, streaming data

Clean the data

- remove null values, outliers

Standardize and normalize

- Convert to numeric values, standardize categories and names

Pre-process and transform

- Calculate summaries, combine attributes



Collect data

Data Preparation



SageMaker runs Jupyter notebooks on instances in the cloud to explore and prepare data



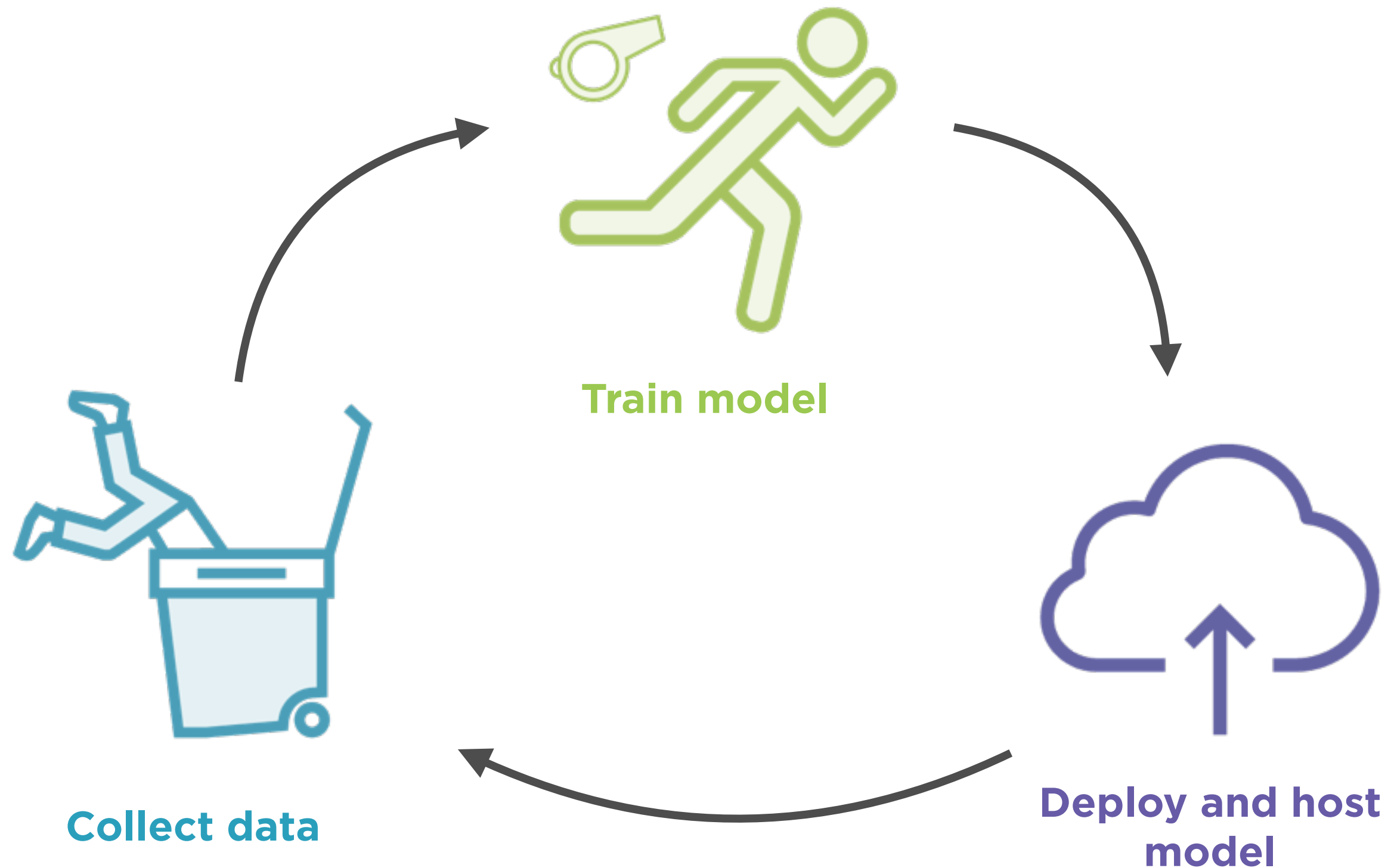
Collect data

Data Preparation

<https://aws.amazon.com/marketplace/search/results?searchTerms=data+preparation>



Machine Learning Workflow





Train model

Machine learning algorithms

- Traditional models, neural networks

Allocate compute resources

- VMs, memory, scaling parameters, GPUs/CPU

Evaluate the model

- AWS SDK for Python (boto), Jupyter notebooks



Train model

Machine learning algorithms

- Traditional models, neural networks

Allocate compute resources

- VMs, memory, scaling parameters, GPUs/CPU

Evaluate the model

- AWS SDK for Python (boto), Jupyter notebooks

ML Algorithms on SageMaker

Built-in algorithms

Out-of-the-box models hosted on containers on the AWS cloud

Bring your algorithm

Develop your own code in TensorFlow, Apache MXNet etc.

ML Algorithms on SageMaker

Built-in algorithms

Out-of-the-box models hosted on containers on the AWS cloud

Bring your algorithm

Develop your own code in TensorFlow, Apache MXNet etc.

ML Algorithms on SageMaker

Bring your algorithm

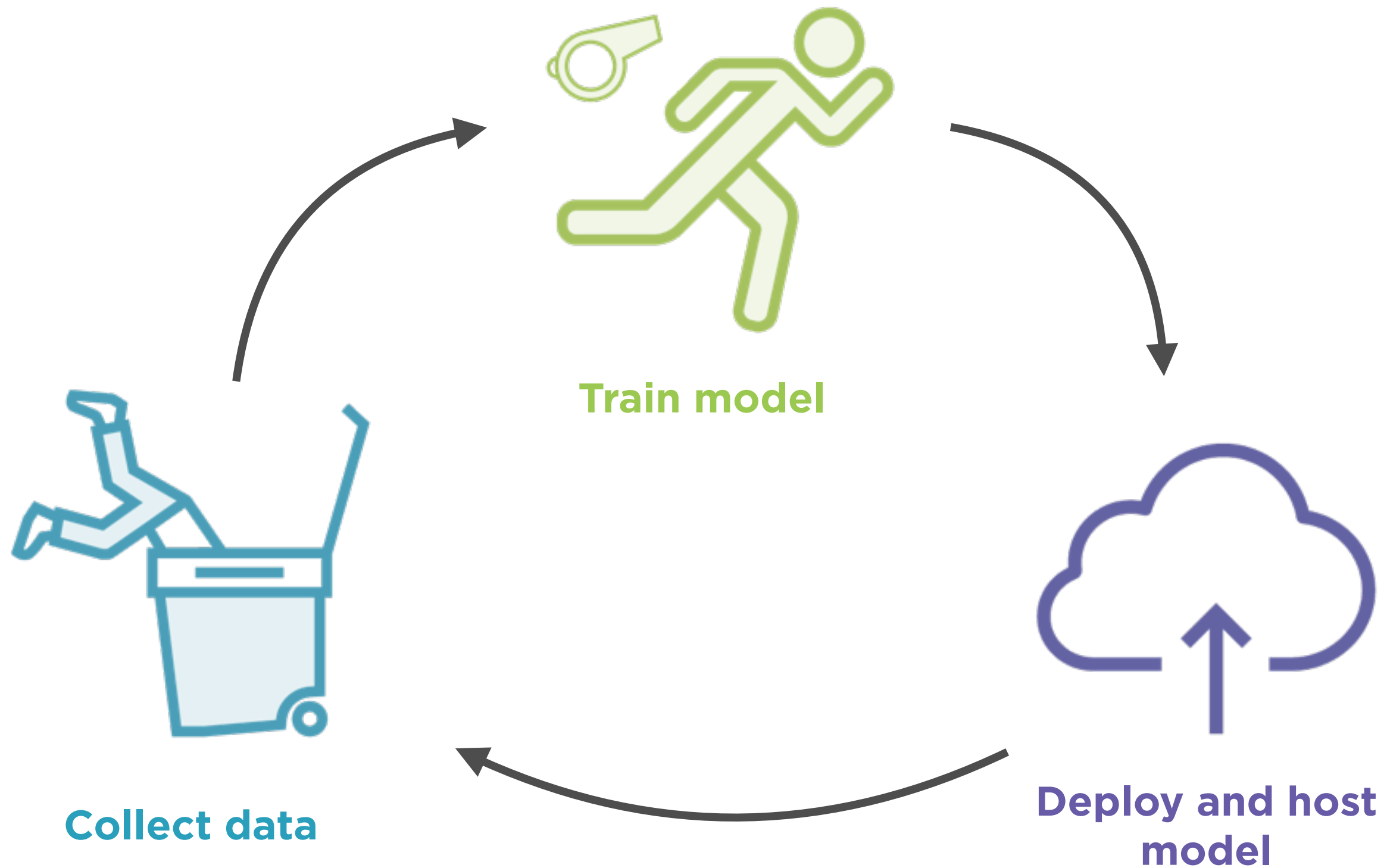
Develop your own code in TensorFlow, Apache MXNet etc.

Bring your own
code

Bring your own
model

Bring your own
container

Machine Learning Workflow





**Deploy and host
model**

Amazon hosting services

- Deploy models at scale on multiple instances

HTTPs endpoint

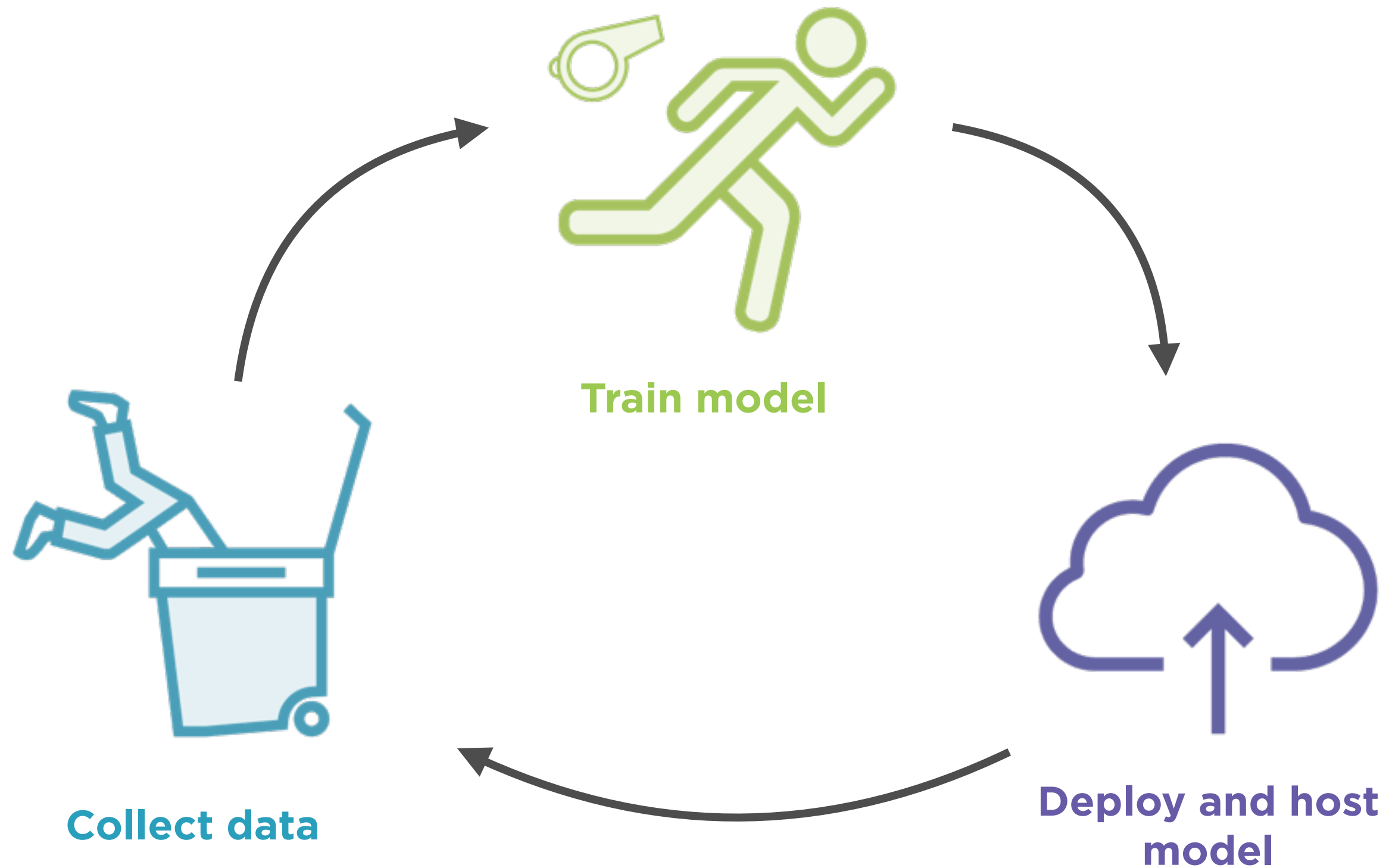
- REST API calls to get model predictions

Decoupled

- Deployment instances separate from training instances

Training a Model on SageMaker

Machine Learning Workflow

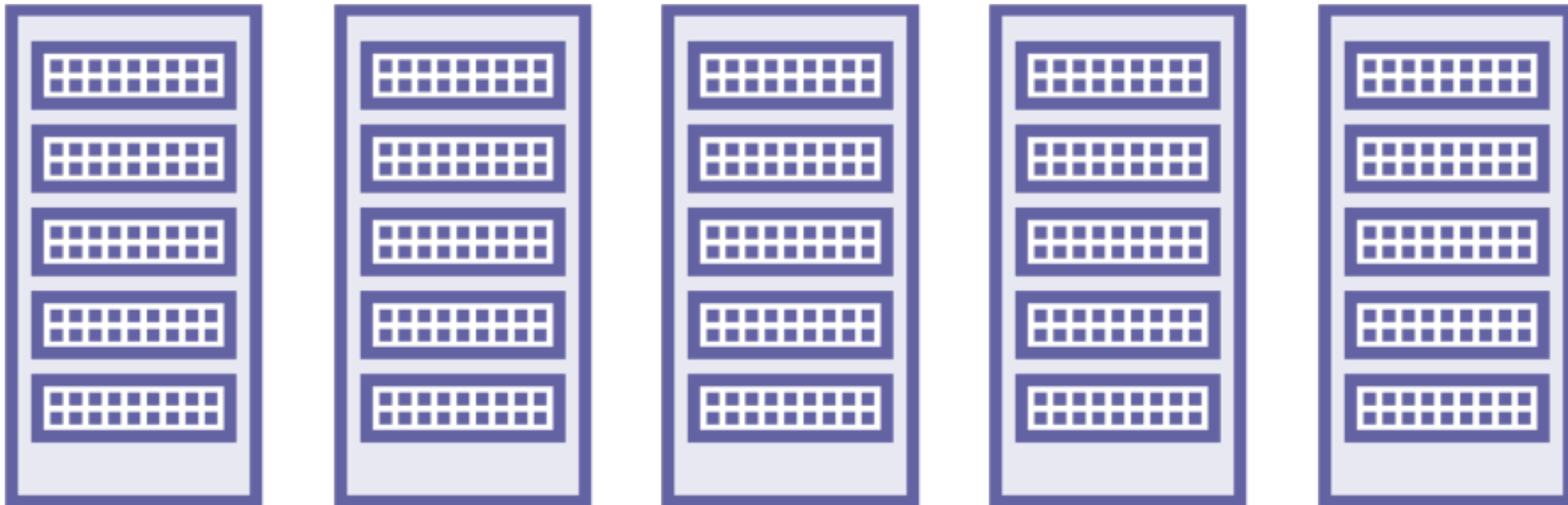


Machine Learning Workflow



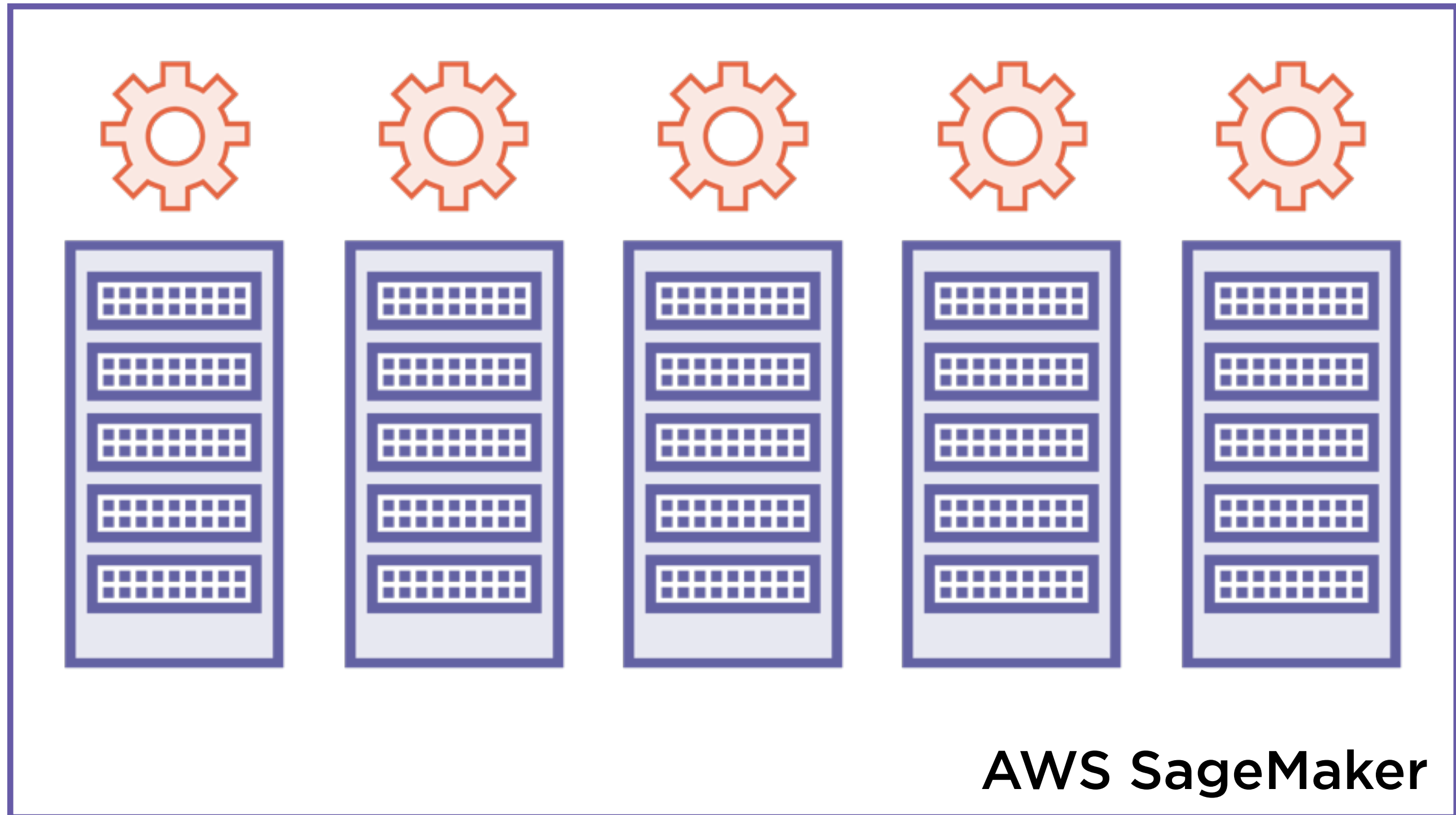
Training a Model

Compute instances for training

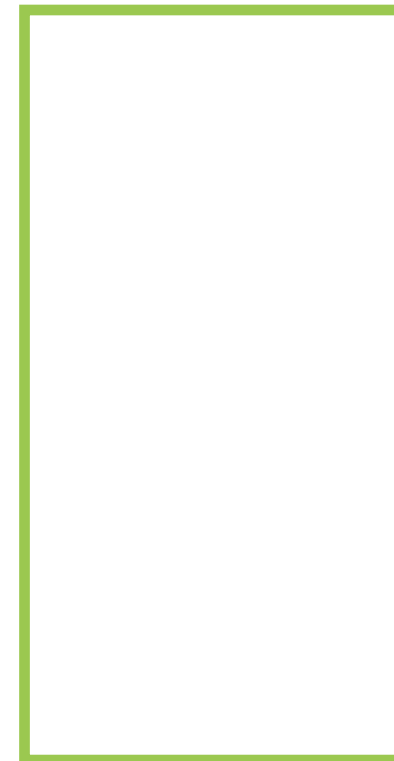
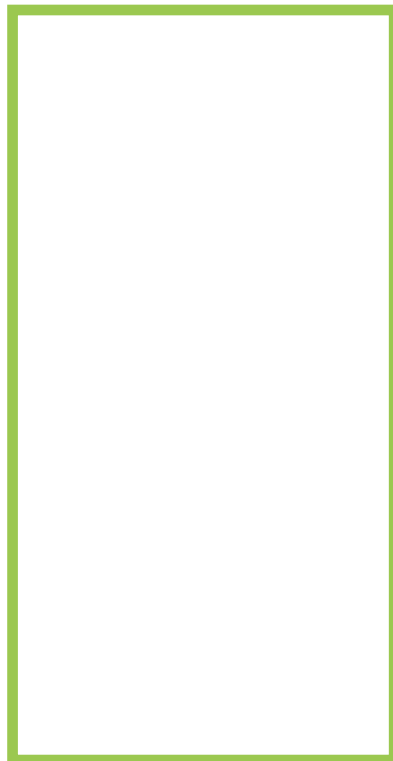


AWS SageMaker

Training a Model

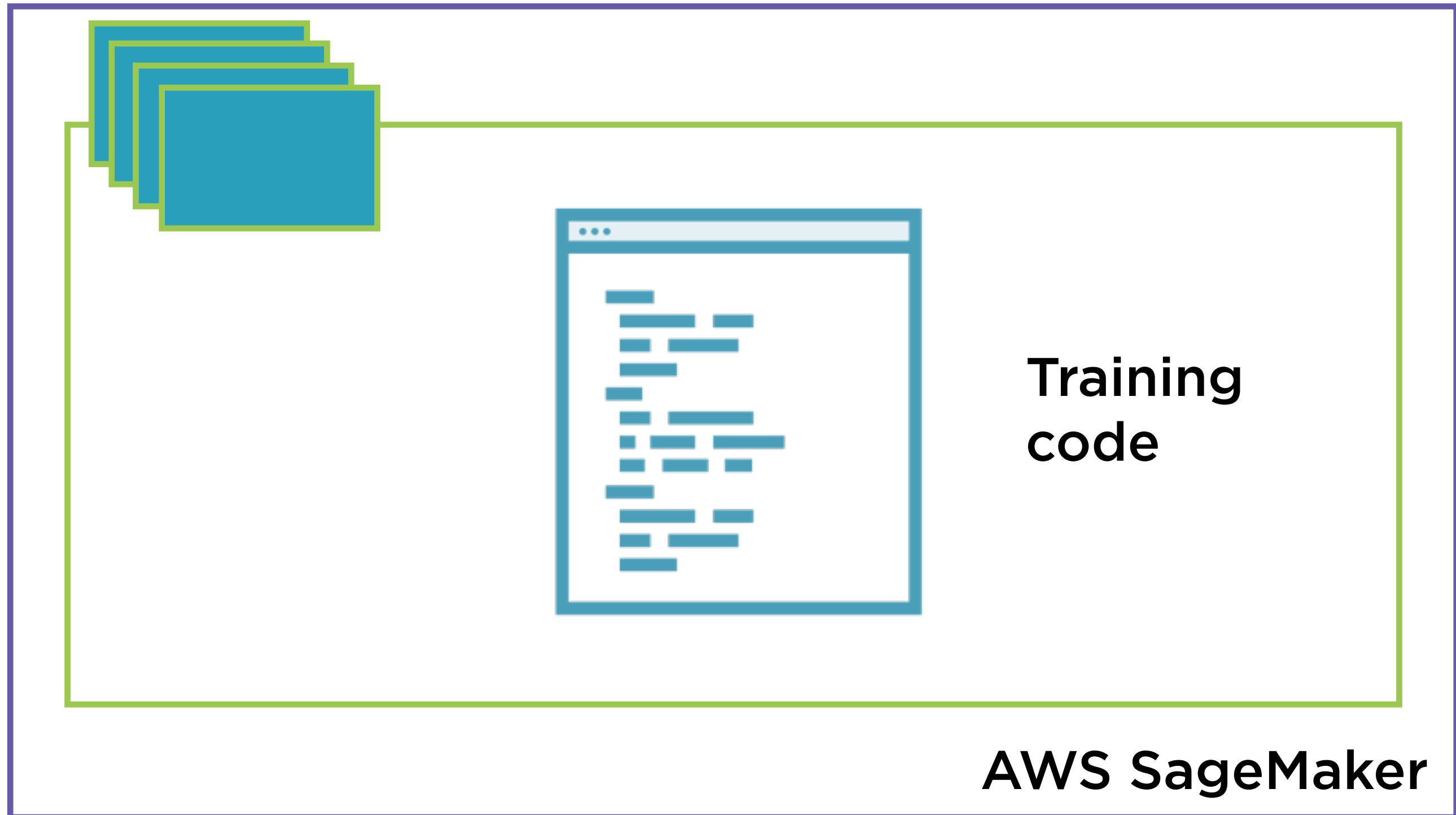


Training a Model

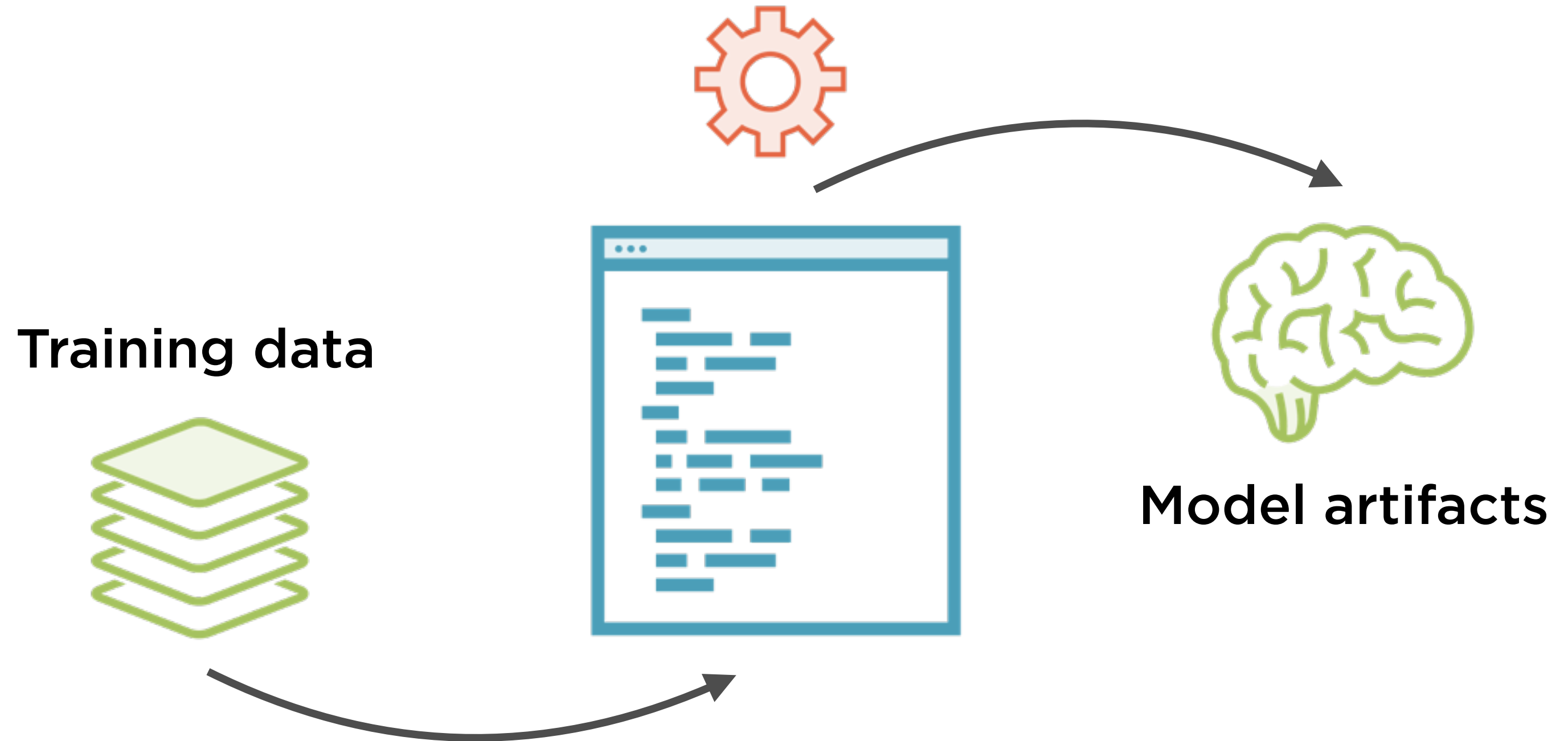


AWS SageMaker

Training a Model



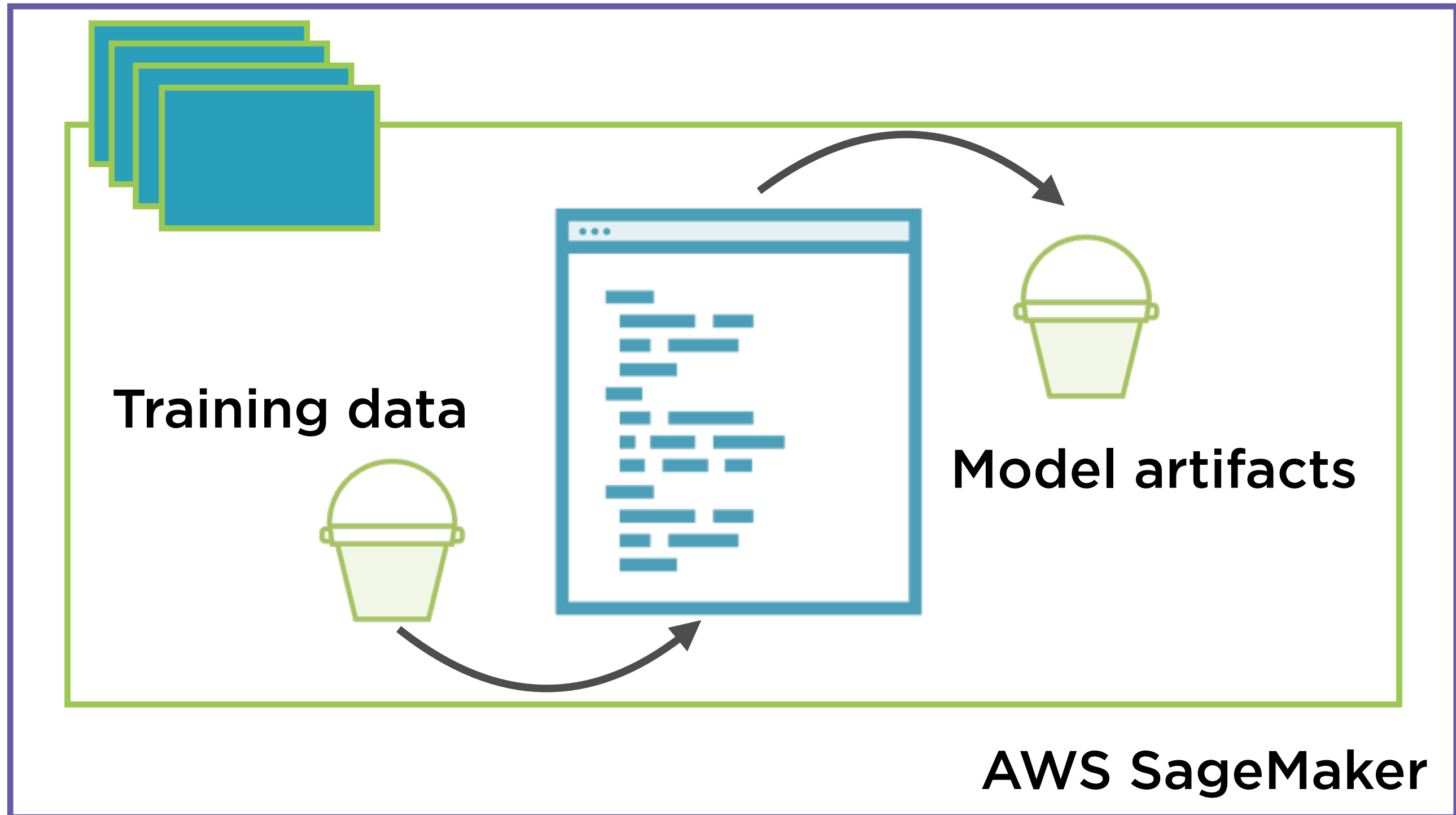
Training a Model



Training a Model



Training a Model



Machine Learning Workflow





Train model

Training Jobs

URL of S3 bucket with training data

Compute resources to run training

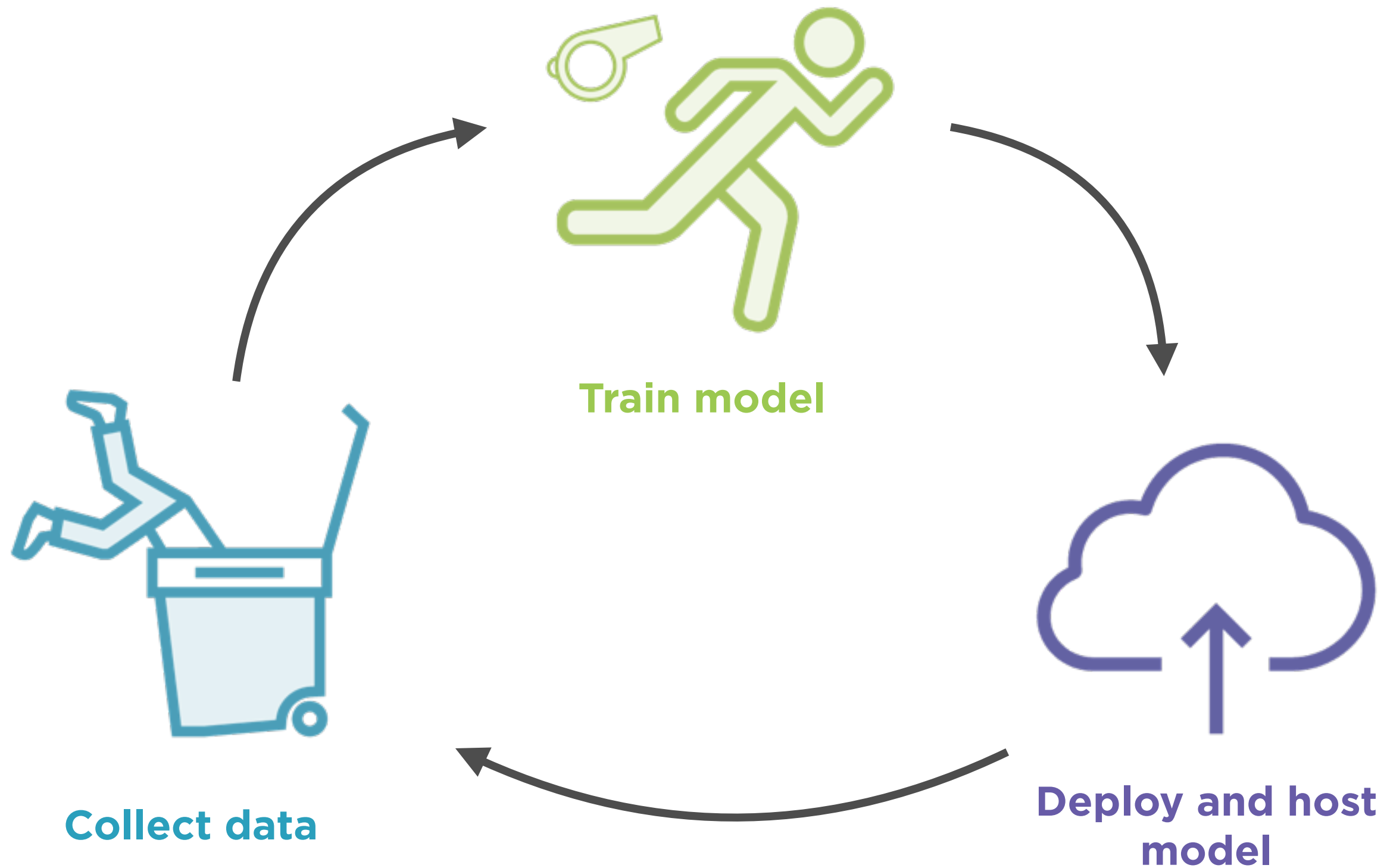
URL of S3 bucket to hold the output

Training code

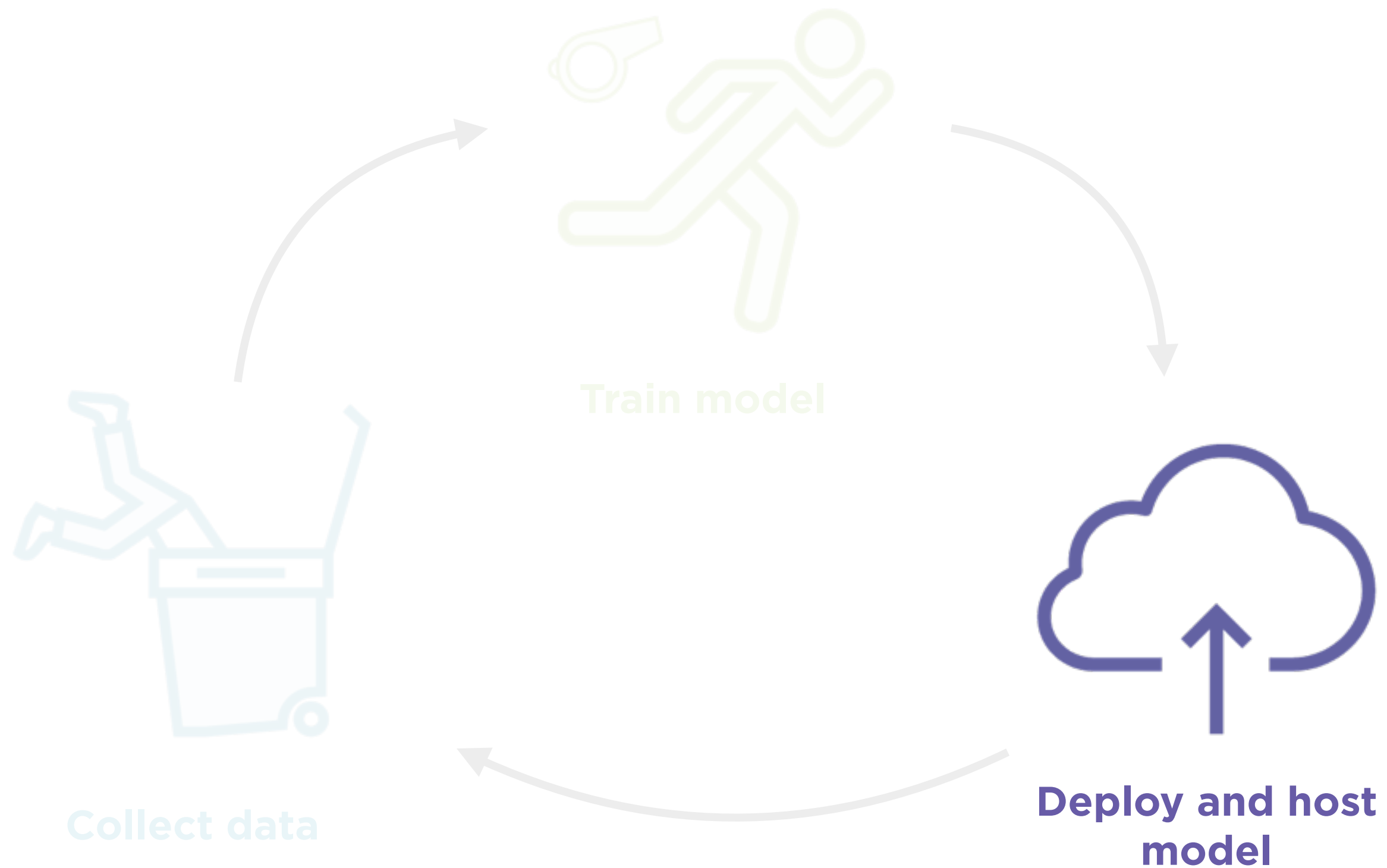
- Built-in algorithms
- ML code in Apache Spark
- Custom code in TensorFlow or Apache MXNet
- Custom code in a container

Deploying a Model on SageMaker

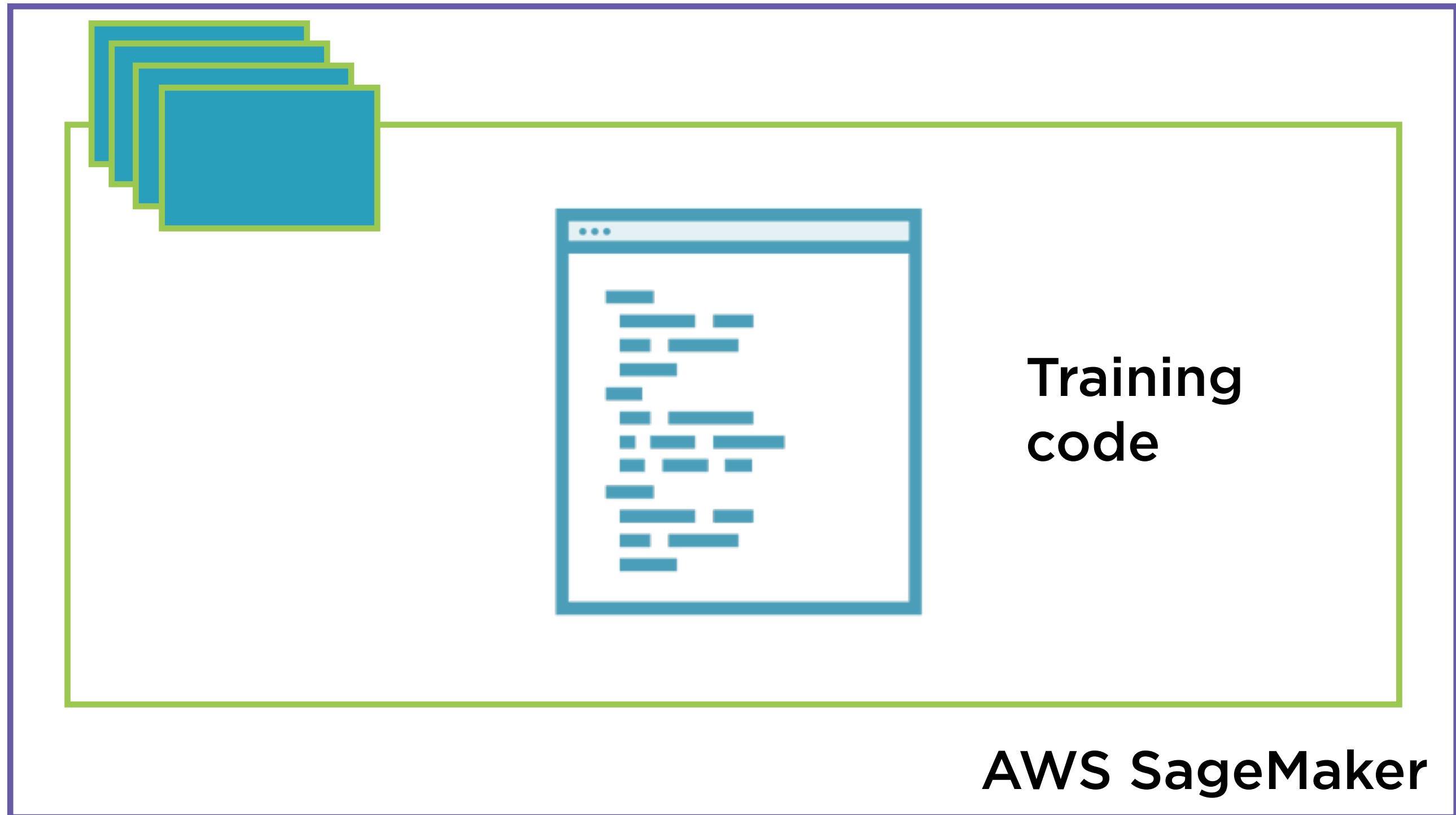
Machine Learning Workflow



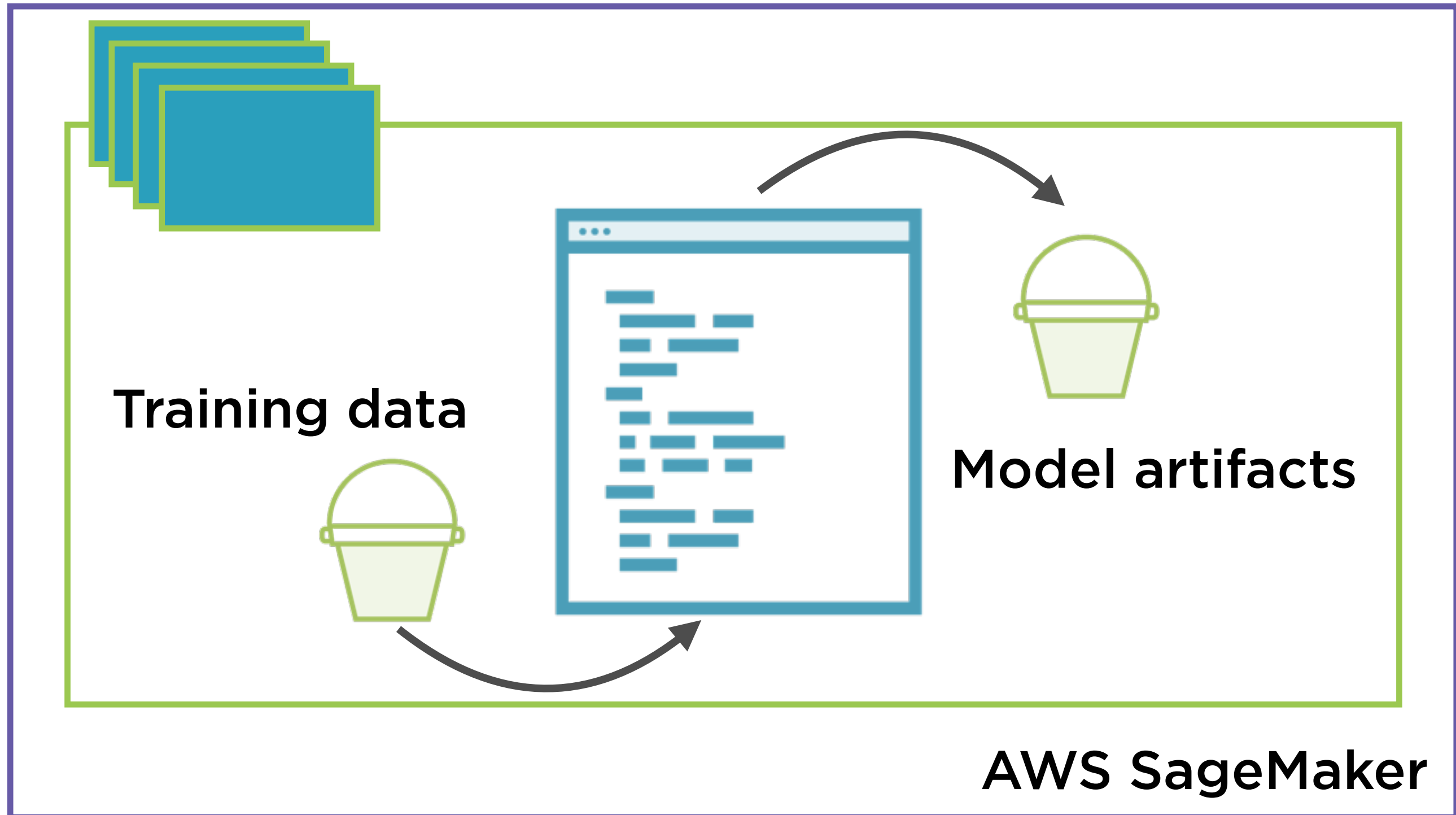
Machine Learning Workflow



Training a Model



Training a Model



Deploying a Model

Compute instances for deployment



Model
artifacts



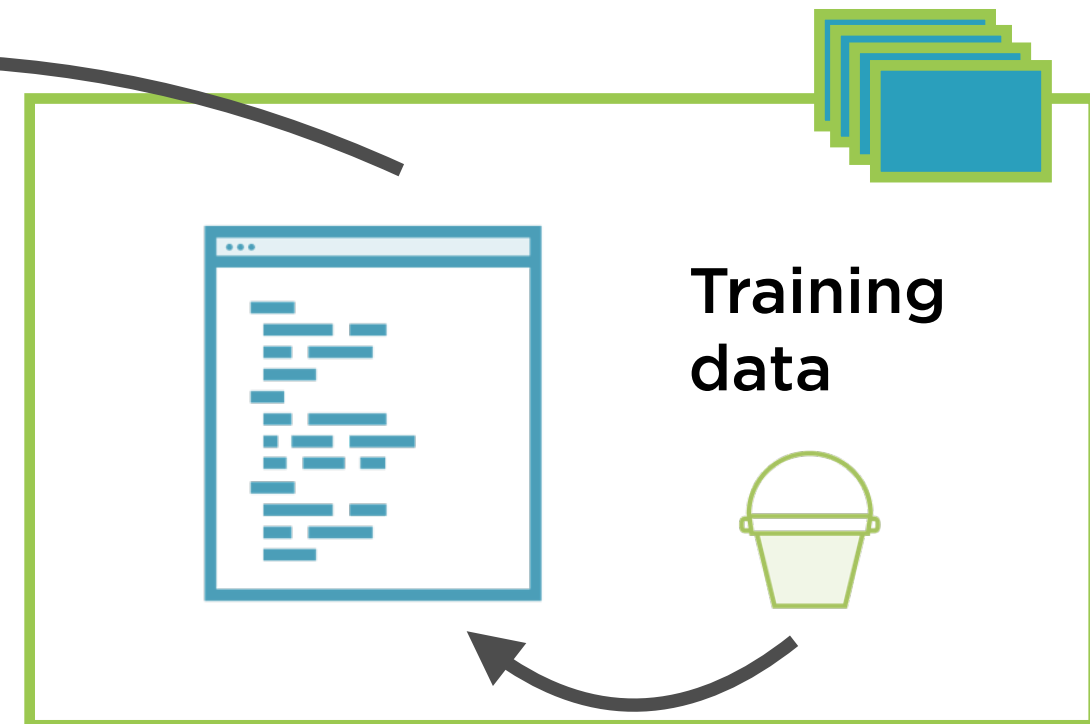
Training
data

Deploying a Model

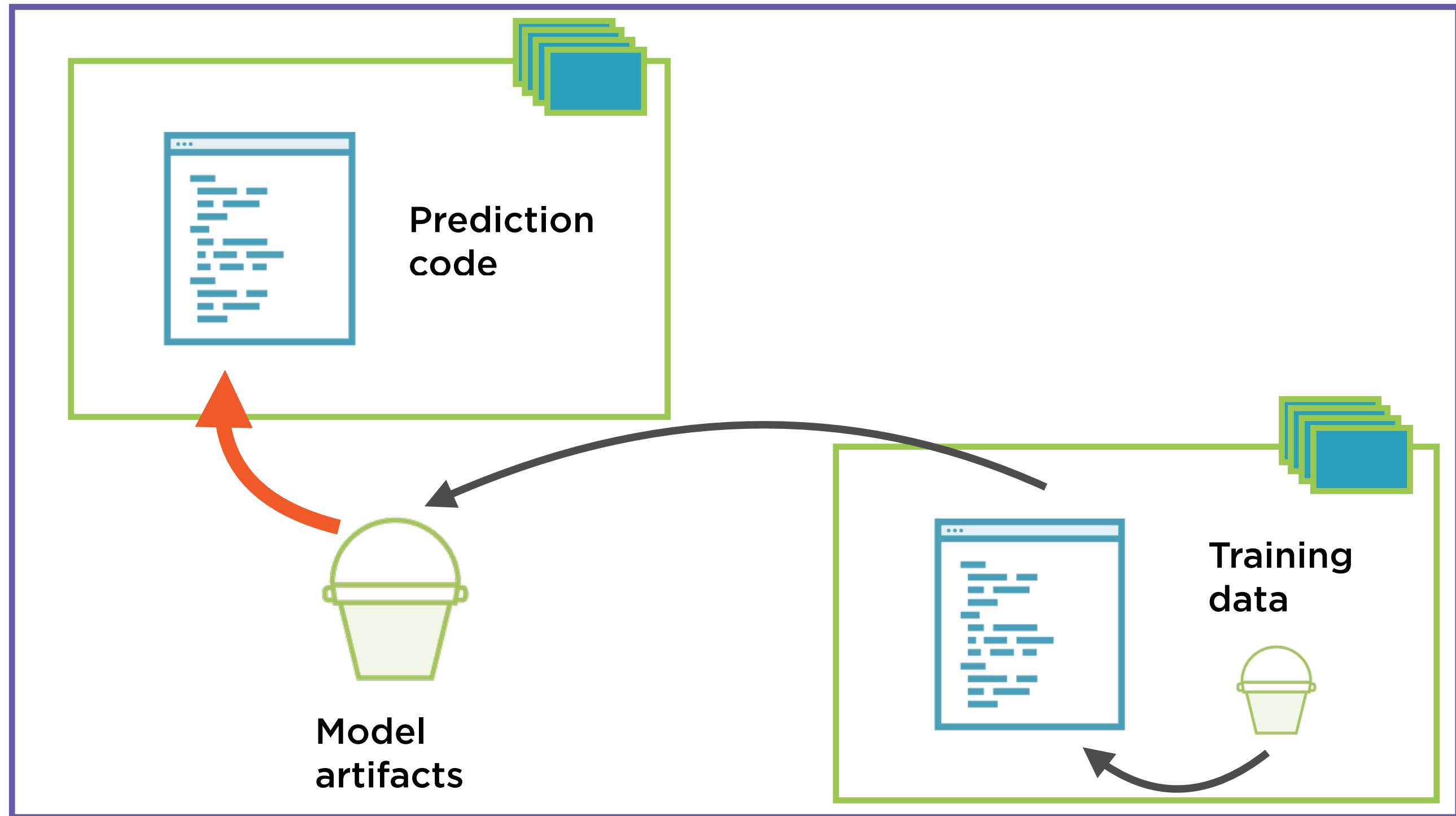
Compute instances for deployment



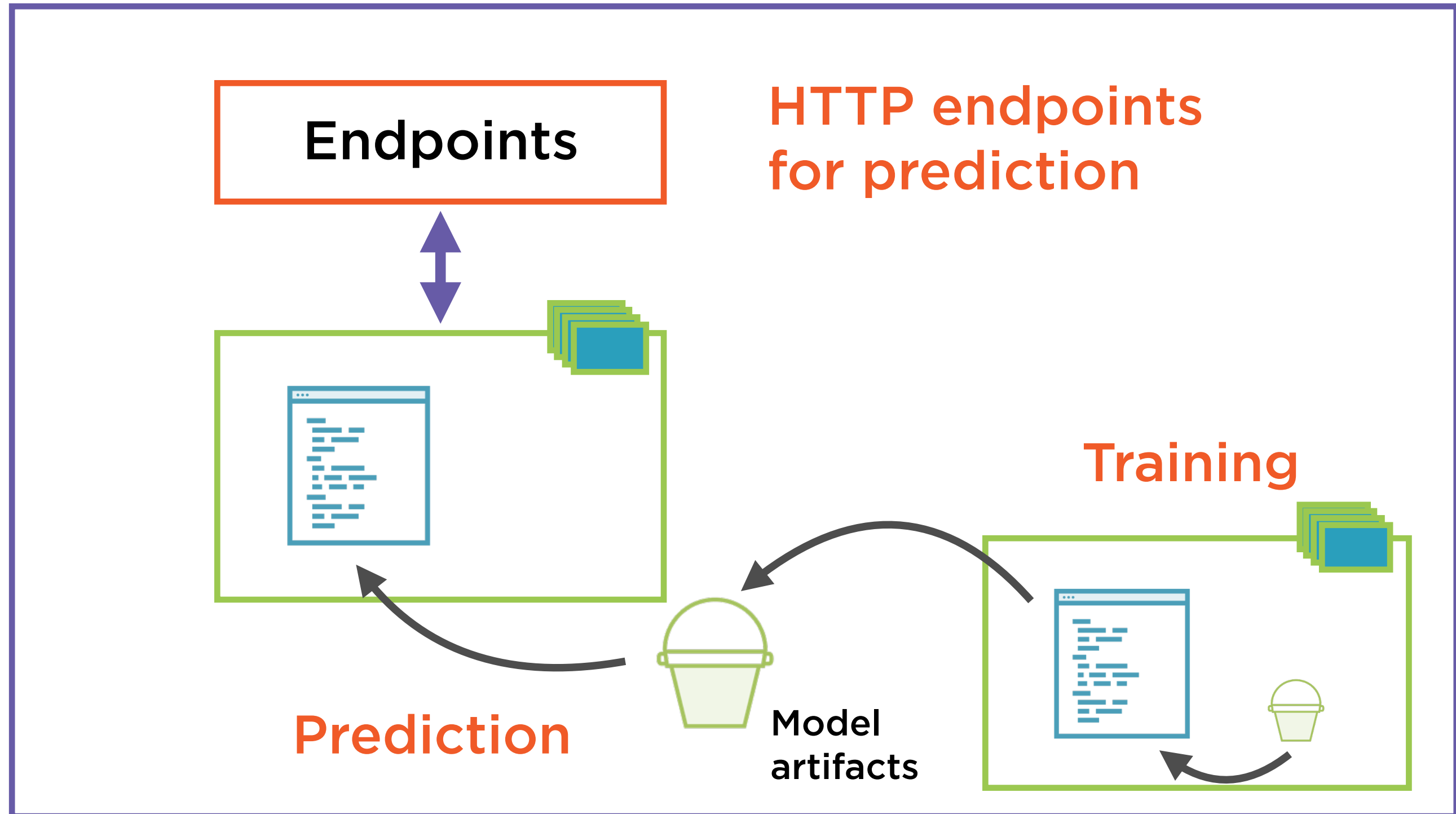
Model
artifacts



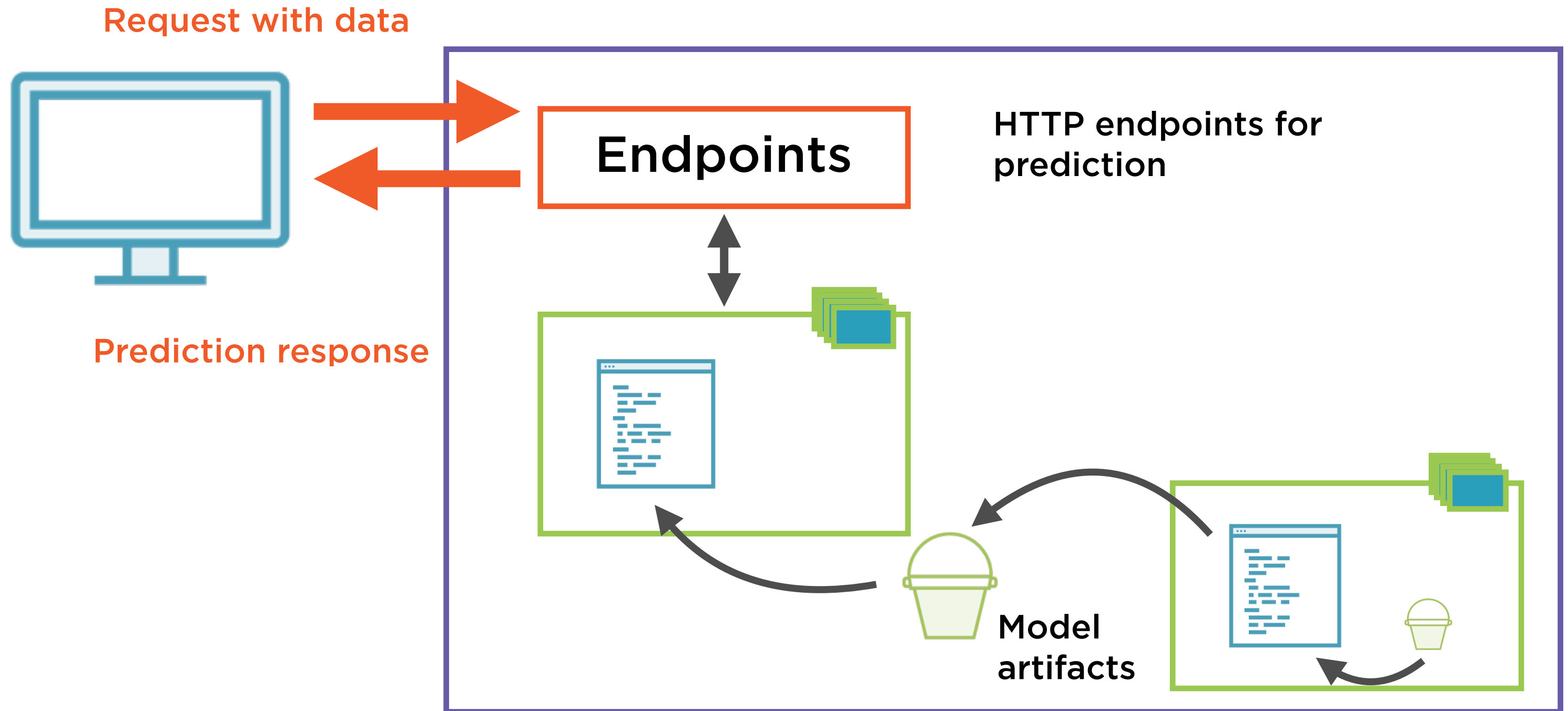
Deploying a Model



Deploying a Model



Deploying a Model



Steps in Deploying a Model

Create a model

Specify model artifacts,
give model name

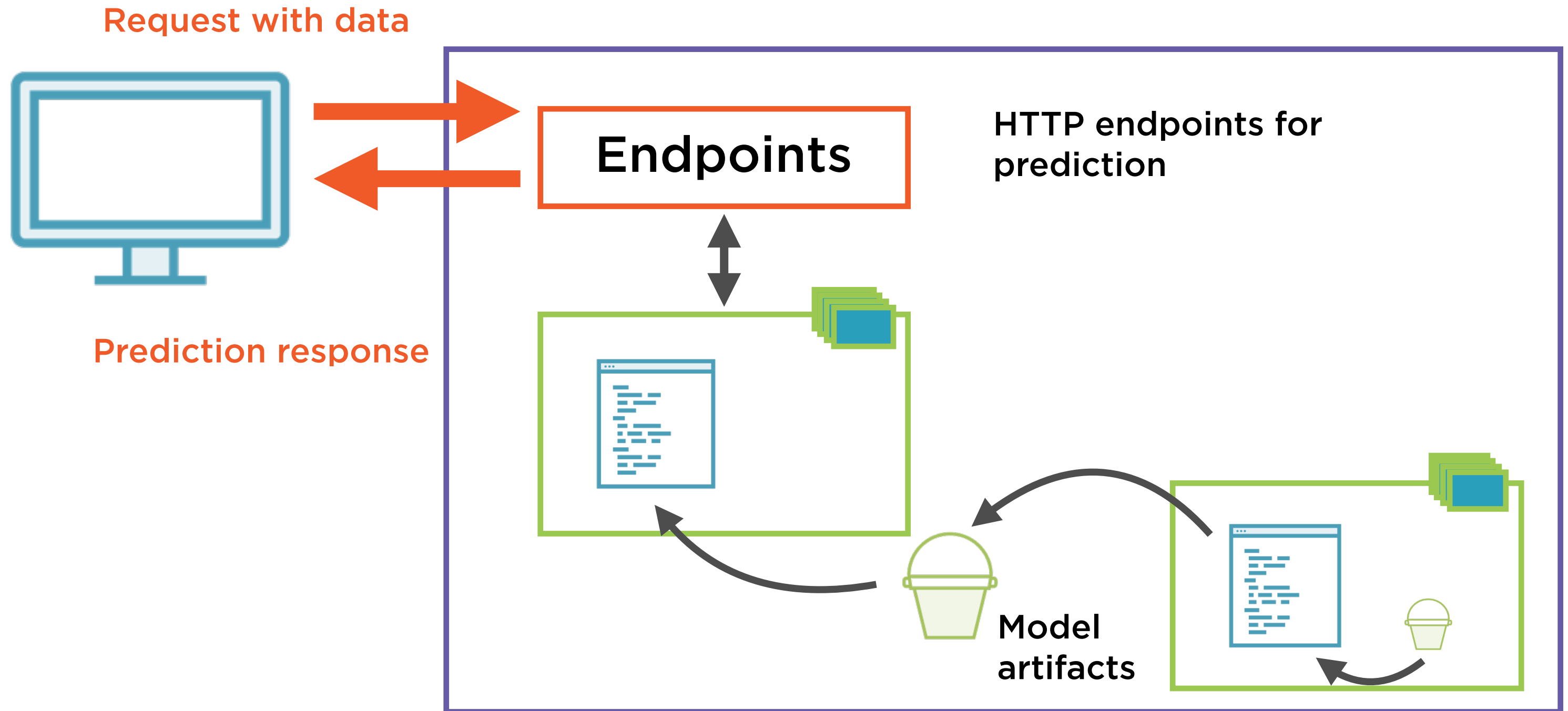
Create an endpoint configuration

Specify model name and
compute instances

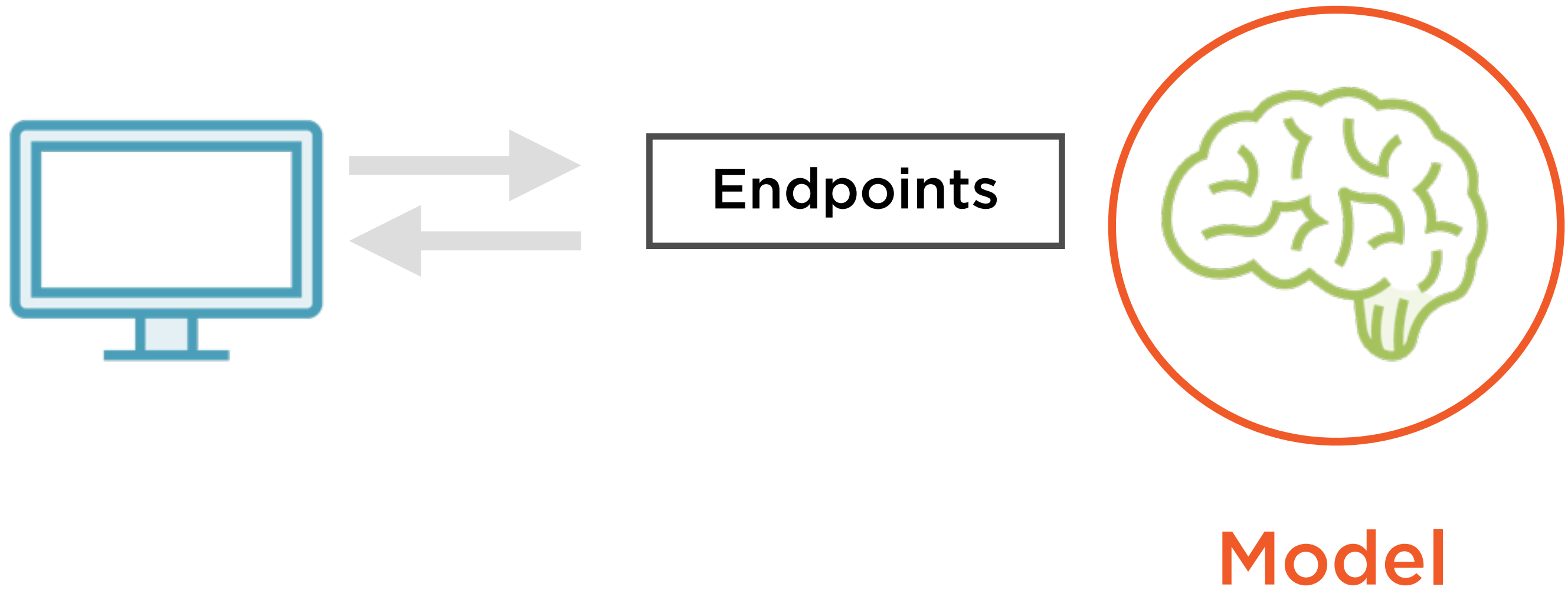
Create an HTTPs endpoint

Provide endpoint config
to SageMaker

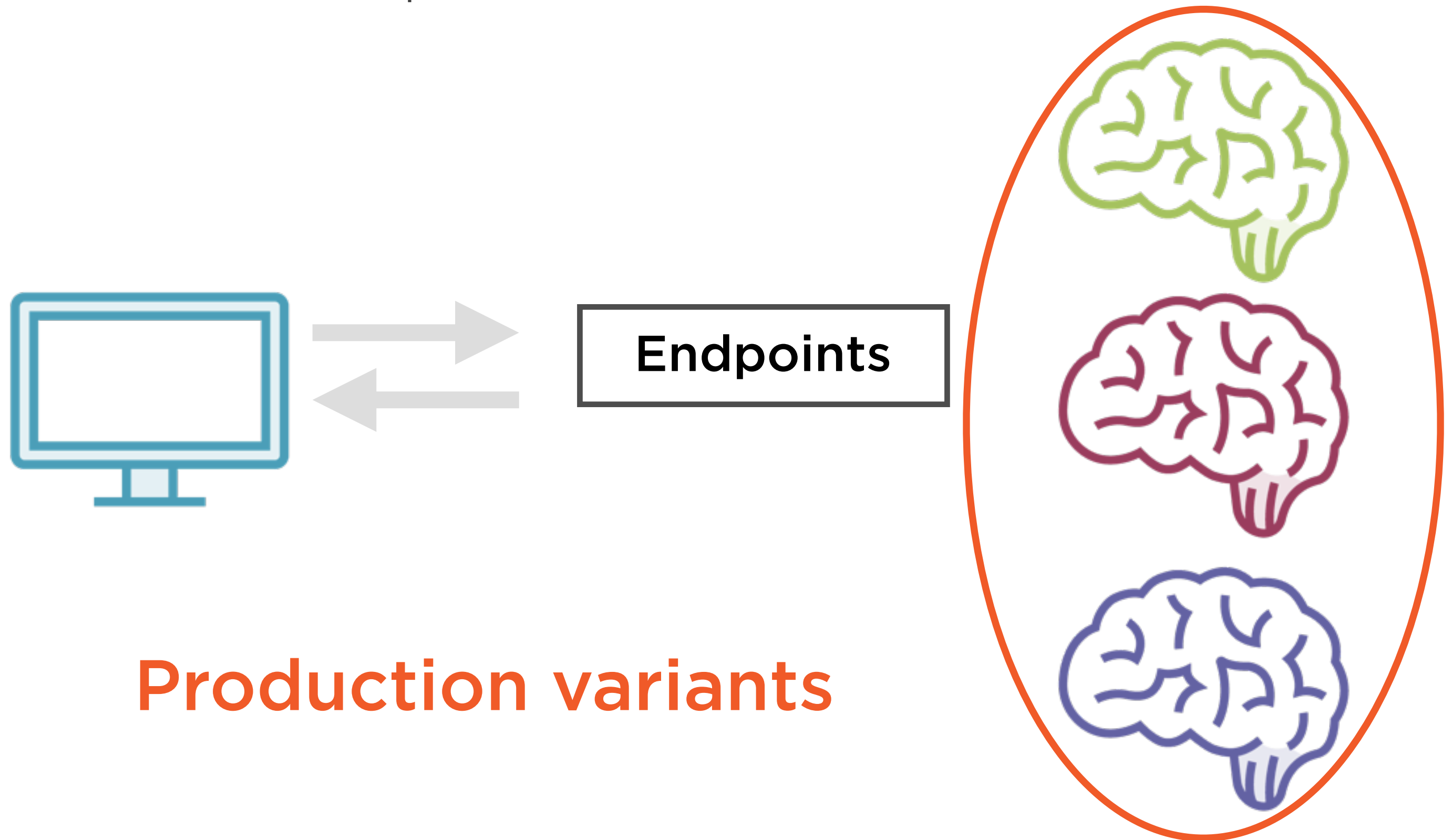
Deploying a Model



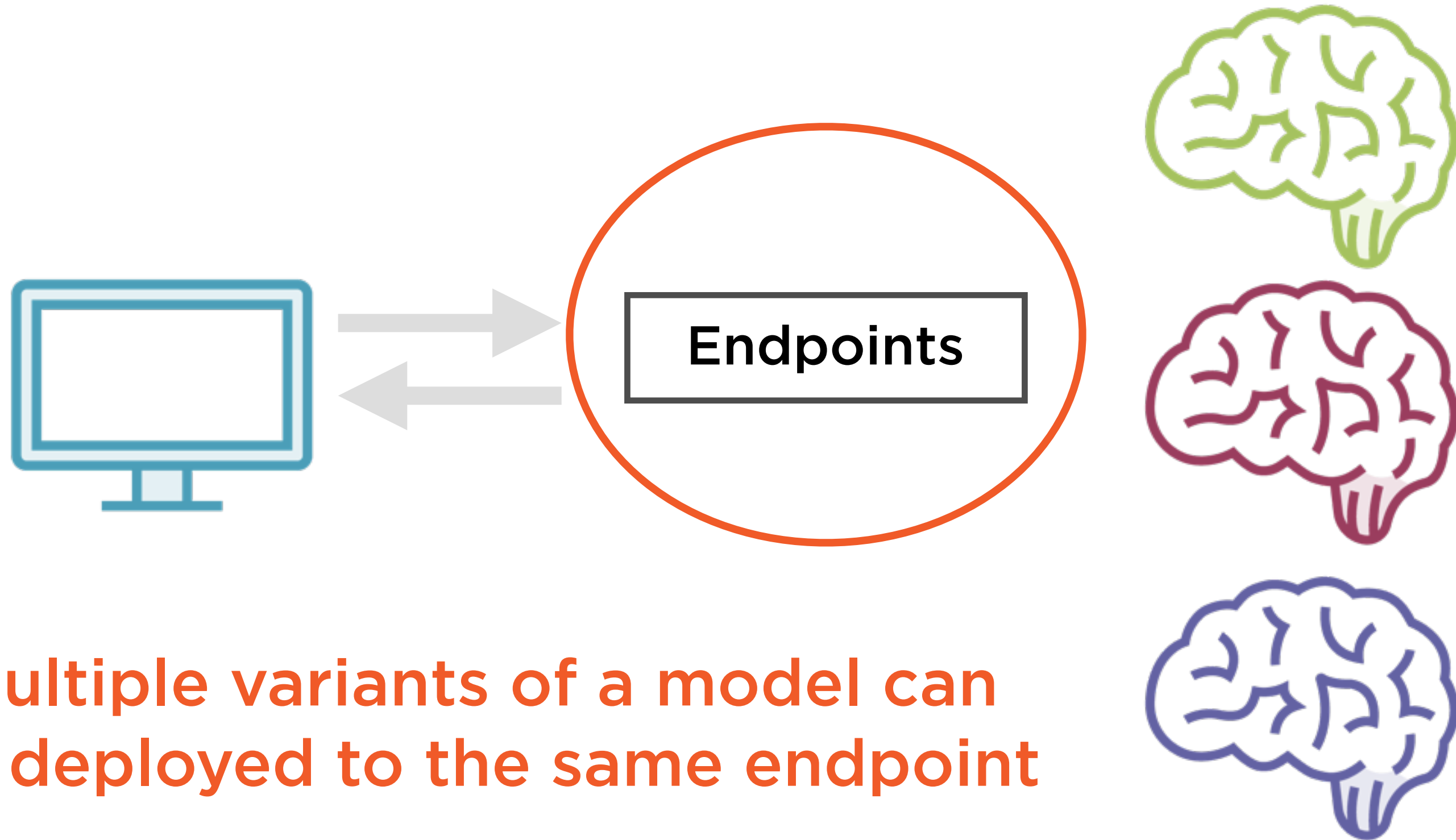
Deploying a Model



Multiple Variants of a Model

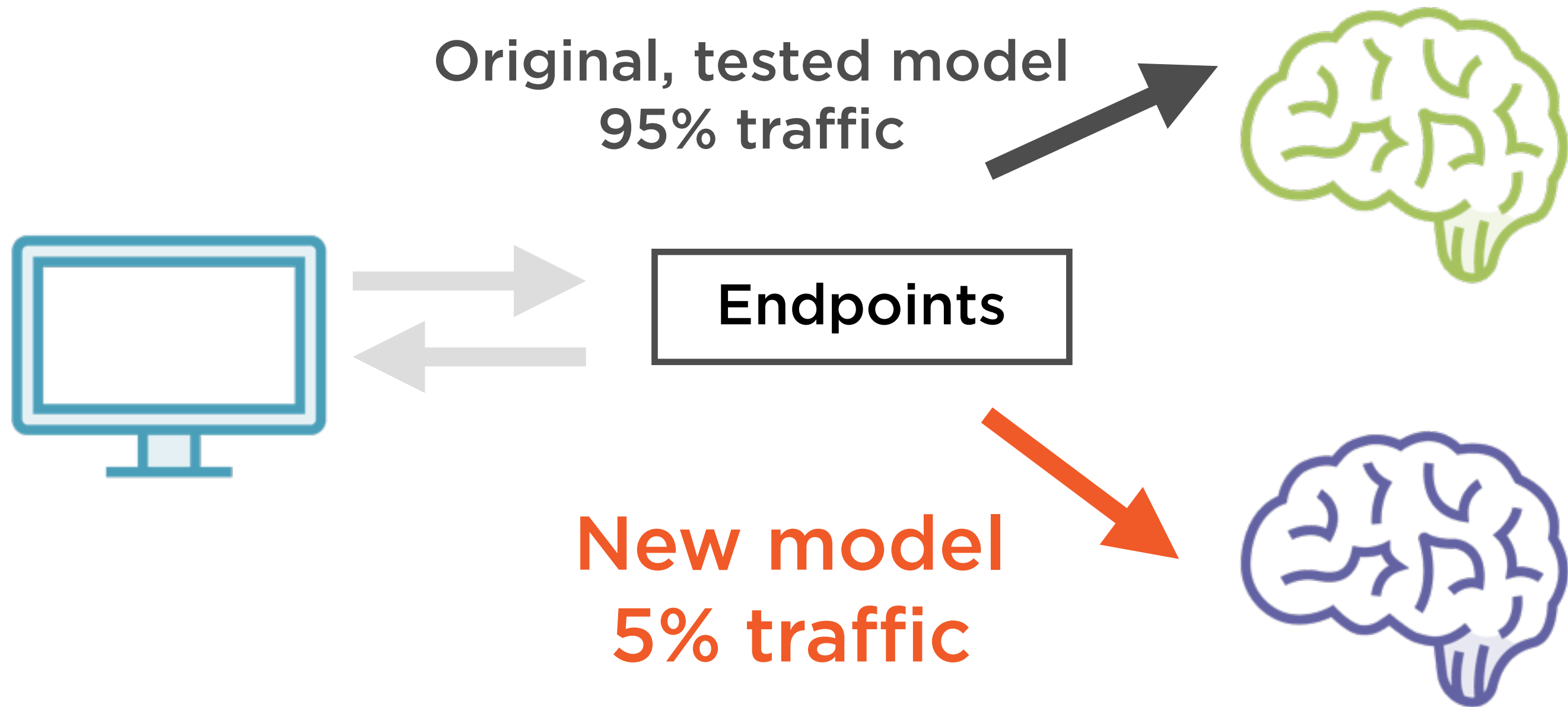


Multiple Variants of a Model

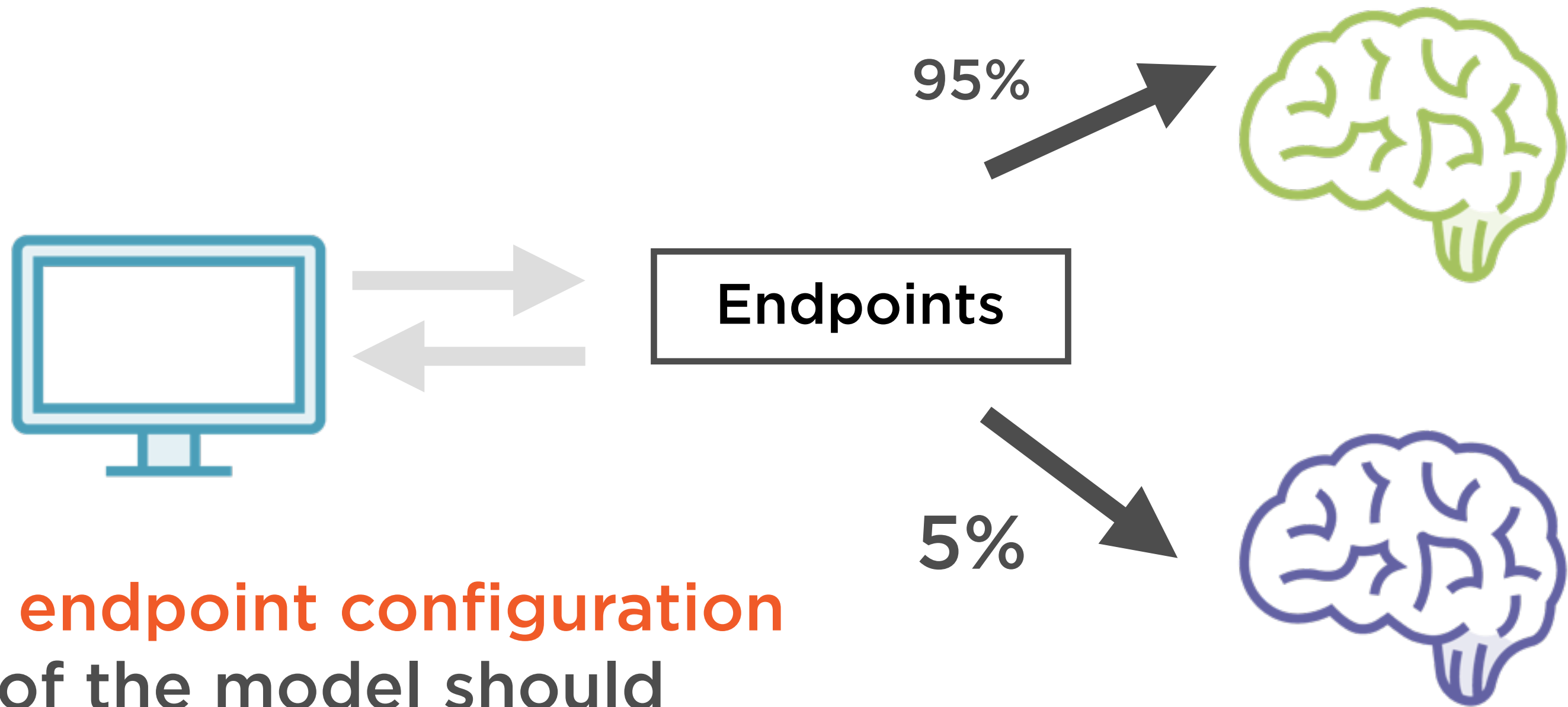


Multiple variants of a model can be deployed to the same endpoint

Test Model Variants in Production

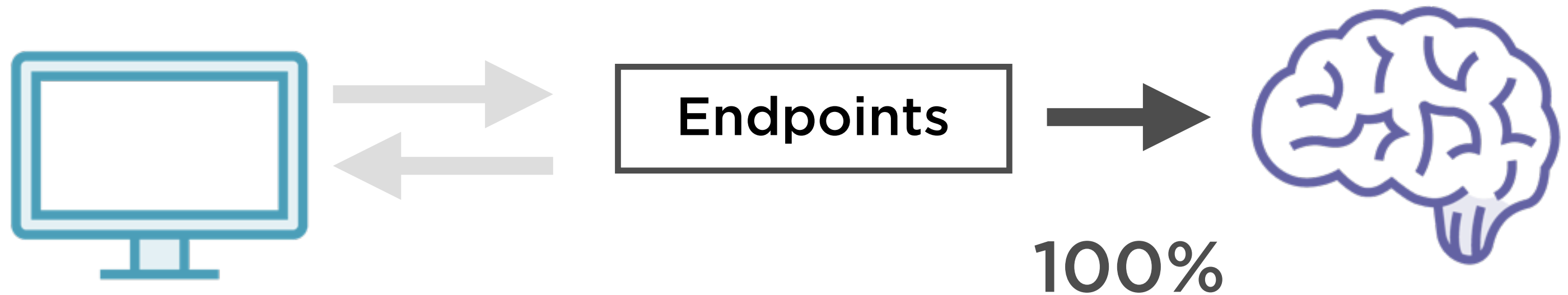


Test Model Variants in Production



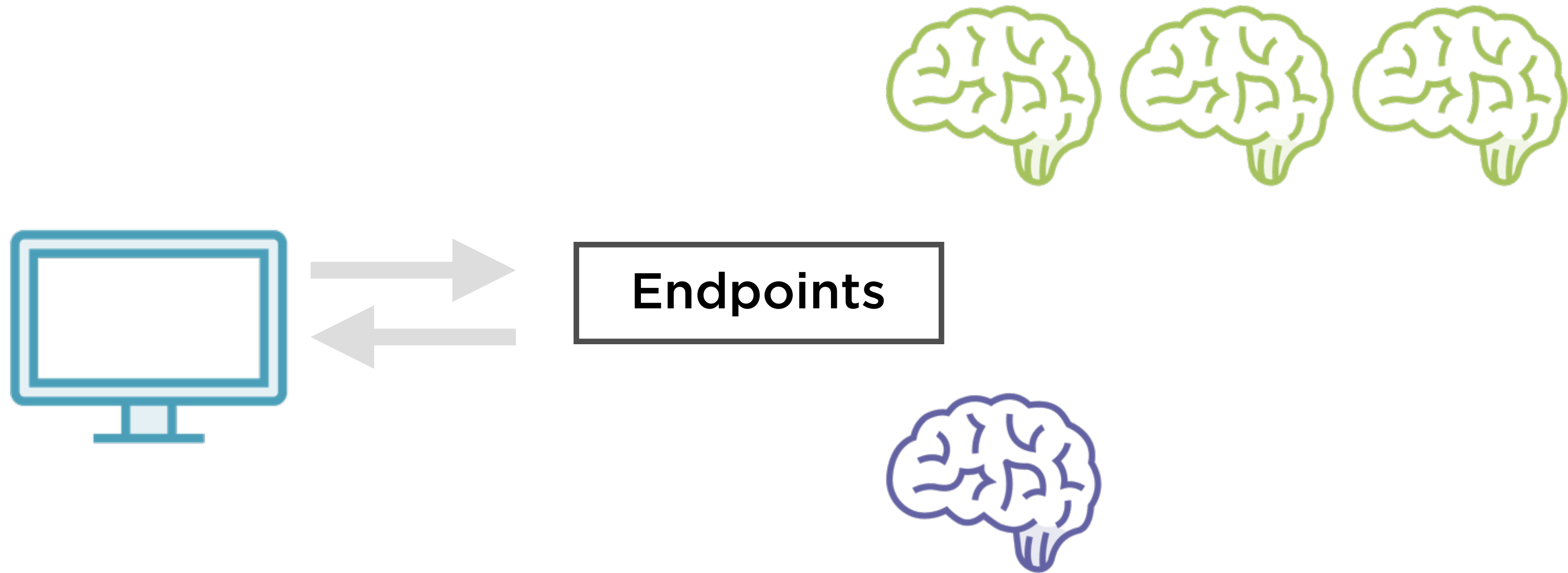
The **endpoint configuration** of the model should specify **both** variants

Deploy New Models Without Downtime



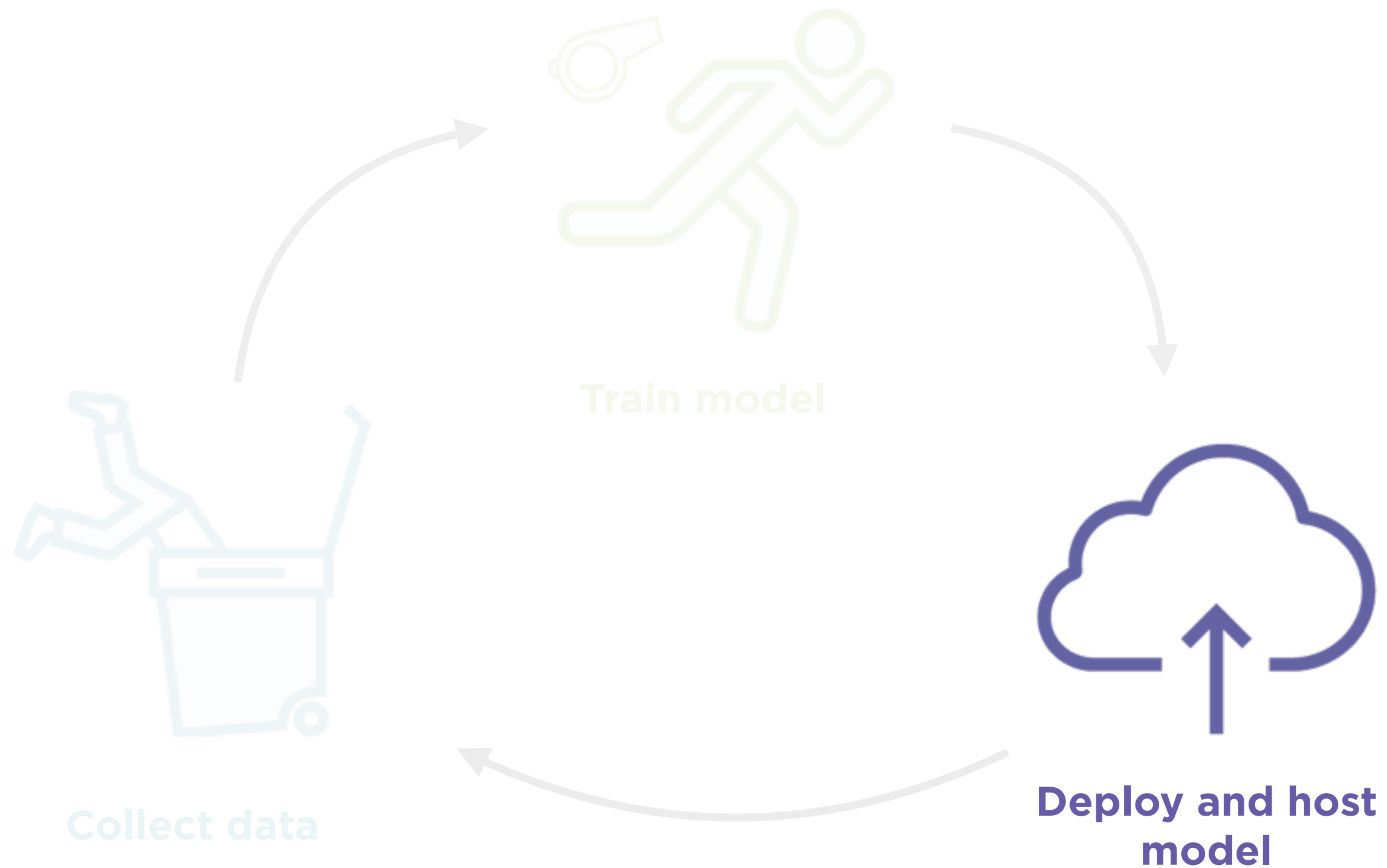
Slowly move 100% of the
traffic to the new model

Autoscale Model Variants



Autoscaling can be applied
to any production variant

Machine Learning Workflow



Deployment Jobs



**Deploy and host
model**

URL of S3 bucket with model artifacts

Compute resources to host model

**Endpoint configuration and HTTPs
endpoint**

Jupyter Notebook Instances

SageMaker offers Jupyter notebook instances to prepare, explore and preprocess data

Creating a Notebook Instance

Creates network interface

Creates this in the VPC specified by the user

Installs Anaconda packages, other libraries

TensorFlow, Apache MXNet

Provides sample code

Many example Jupyter notebooks

Launches ML compute instance

Enables traffic between this instance and your VPC

Attaches an ML storage volume

Non-persistent storage, 5GB with the compute instance and 20GB with the notebook instance

Demo

Signing in to Amazon Web Services

**Creating a notebook instance to hold
the demo Jupyter notebooks**

Demo

Setting up an Amazon S3 bucket

Used to hold training data and model artifacts

Summary

Build, train and deploy cycles using Amazon SageMaker

Develop new models or use built-in models on your training data

Bring your own model or container to train and host at scale

Integrated Jupyter notebook instance to develop models