

**MLOps Engineering on AWS**

**Course Number:** AWS-144
**Duration:** 3 days

**Overview**

This live, instructor-led MLOps (Machine Learning Operations) Engineering on AWS training course teaches your team how to extend DevOps to leverage data, models, and code for successful Machine Learning (ML) deployments. This MLOps course also discusses the necessary tools and processes needed when the model prediction in production does not align with the original key performance indicators.

Accelebrate is an AWS Training Partner (ATP) and this hands-on official AWS Classroom Training course is taught by an accredited Amazon Authorized Instructor (AAI).

**Prerequisites**

All students should have taken the AWS classroom training courses [AWS Technical Essentials](file:////training/aws-technical-essentials), [DevOps Engineering in AWS](file:////training/devops-engineering-aws), and [Practical Data Science with Amazon Sagemaker](file:////training/data-science-with-amazon-sagemaker), or have the equivalent experience.

**Materials**

All AWS training students receive comprehensive courseware.

**Software Needed on Each Student PC**

A modern web browser and an Internet connection free of restrictive firewalls, so that the student can connect by SSH or Remote Desktop (RDP) into AWS virtual machines.

**Objectives**

* Describe machine learning operations
* Understand the key differences between DevOps and MLOps
* Describe the machine learning workflow
* Discuss the importance of communications in MLOps
* Explain end to end options for automation of ML workflows
* List key Amazon SageMaker features for MLOps automation
* Build an automated ML process that builds, trains, tests, and deploys models
* Build an automated ML process that retrains the model based on change(s) to the model code
* Identify elements and important steps in the deployment process
* Describe items that might be included in a model package, and their use in training or inference
* Recognize Amazon SageMaker options for selecting models for deployment, including support for ML frameworks and built-in algorithms or bring-your-own models
* Differentiate scaling in machine learning from scaling in other applications
* Determine when to use different approaches to inference
* Discuss deployment strategies, benefits, challenges, and typical use cases
* Describe the challenges when deploying machine learning to edge devices
* Recognize important Amazon SageMaker features that are relevant to deployment and inference
* Describe why monitoring is important
* Detect data drifts in the underlying input data
* Demonstrate how to monitor ML models for bias
* Explain how to monitor model resource consumption and latency
* Discuss how to integrate human-in-the-loop reviews of model results in production

**Outline**

* Introduction to MLOps
	+ Machine learning operations
	+ Goals of MLOps
	+ Communication
	+ From DevOps to MLOps
	+ ML workflow
	+ Scope
	+ MLOps view of ML workflow
	+ MLOps cases
* MLOps Development
	+ Intro to build, train, and evaluate machine learning models
	+ MLOps security
	+ Automating
	+ Apache Airflow
	+ Kubernetes integration for MLOps
	+ Amazon SageMaker for MLOps
	+ Bring your own algorithm to an MLOps pipeline
	+ Demonstration: Amazon SageMaker
	+ Intro to build, train, and evaluate machine learning models
	+ Code and serve your ML model with AWS CodeBuild
* MLOps Deployment
	+ Introduction to deployment operations
	+ Model packaging
	+ Inference
	+ Deploy your model to production
	+ SageMaker production variants
	+ Deployment strategies
	+ Deploying to the edge
	+ Conduct A/B testing
* Model Monitoring and Operations
	+ Troubleshoot your pipeline
	+ The importance of monitoring
	+ Monitoring by design
	+ Monitoring your ML model
	+ Human-in-the-loop
	+ Amazon SageMaker Model Monitor
	+ Amazon SageMaker Pipelines, Model Monitor, model registry, and Feature Store
	+ Solving the Problem(s)
* Conclusion