Amazon SageMaker Studio for Data Scientists

Course Duration: 24 Hours (3 Days)

Overview

The Amazon SageMaker Studio for Data Scientists course is designed to equip learners with a comprehensive understanding of AWS SageMaker Studio, a machine learning (ML) integrated development environment (IDE). Learners will explore the setup, navigation, and functionalities of Amazon SageMaker Studio, diving into data processing techniques to ensure ML-readiness and bias detection. The course covers Model development, including tuning, evaluation, and debugging, using AWS SageMaker Studio. Deployment and Inference modules teach how to effectively manage models and Automate ML workflows. Monitoring lessons focus on maintaining model quality and Detecting drifts. Finally, the course provides insights into Resource management and updates, ensuring learners can efficiently operate within SageMaker Studio. This curriculum caters to data scientists looking to harness the full potential of SageMaker Studio for end-to-end ML solutions.

Audience Profile

The Amazon SageMaker Studio for Data Scientists course provides comprehensive training in machine learning (ML) model development, deployment, and management on AWS.

- Data Scientists leveraging SageMaker for ML projects
- Machine Learning Engineers focusing on model building and deployment on AWS
- AI/ML Researchers using SageMaker for experimental purposes
- Data Analysts looking to transition into predictive analytics and machine learning
- Cloud Solutions Architects designing ML solutions on AWS
- DevOps Engineers managing ML workflows and infrastructure
- IT Professionals seeking to understand ML operations within SageMaker Studio
- Business Intelligence Professionals expanding their skill set to include ML models
- Technical Product Managers overseeing ML product development
- Software Developers integrating ML into applications using AWS services

Course Syllabus

Module 1: Amazon SageMaker Setup and Navigation

- Launch SageMaker Studio from the AWS Service Catalog
- Navigate the SageMaker Studio UI
- Demo 1: SageMaker UI Walkthrough
- Lab 1: Launch SageMaker Studio from AWS Service Catalog

Module 2: Data Processing

- Use Amazon SageMaker Studio to collect, clean, visualize, analyze, and transform data
- Set up a repeatable process for data processing
- Validate that collected data is ML-ready using SageMaker
- Detect bias in collected data and estimate baseline model accuracy
- Lab 2: Analyze and Prepare Data Using SageMaker Data Wrangler
- Lab 3: Analyze and Prepare Data at Scale Using Amazon EMR
- Lab 4: Data Processing Using SageMaker Processing and the SageMaker Python SDK
- Lab 5: Feature Engineering Using SageMaker Feature Store

Module 3: Model Development

- Use Amazon SageMaker Studio to develop, tune, and evaluate an ML model against business objectives while ensuring fairness and explainability
- Fine-tune ML models using automatic hyperparameter optimization
- Use SageMaker Debugger to detect issues during model development
- Demo 2: Autopilot
- Lab 6: Track Training and Tuning Iterations Using SageMaker Experiments
- Lab 7: Analyze, Detect, and Set Alerts Using SageMaker Debugger
- Lab 8: Identify Bias Using SageMaker Clarify

Module 4: Deployment and Inference

- Use the Model Registry to create a model group, register and manage model versions, modify model approval status, and deploy models
- Design and implement a deployment solution based on inference use case requirements
- Create, automate, and manage end-to-end ML workflows using Amazon SageMaker Pipelines
- Lab 9: Perform Inferencing with SageMaker Studio
- Lab 10: Use SageMaker Pipelines and the SageMaker Model Registry with SageMaker Studio

Module 5: Monitoring

- Configure SageMaker Model Monitor to detect issues and generate alerts for changes in data quality, model quality, bias drift, and feature attribution (explainability) drift
- Create a monitoring schedule with a predefined interval
- Demo 3: Model Monitoring

Module 6: Managing SageMaker Studio Resources and Updates

- Identify resources that accrue charges
- Understand when and how to shut down instances, notebooks, terminals, and kernels

- Learn the process of updating SageMaker Studio

Capstone Project

- Apply SageMaker Studio capabilities learned throughout the course
- Prepare, build, train, and deploy an ML model using a tabular dataset not covered in earlier labs
- Choose between basic, intermediate, and advanced versions of the instructions
- Capstone Lab: Build an End-to-End Tabular Data ML Project Using SageMaker Studio and the SageMaker Python SDK