

AWS Certified Machine Learning MLS-C01 Prep

AWS MLS-C01 Machine Learning. SageMaker, data prep, model training, and deployment.

IT Certifications Instructor: Brian Taylor • 5 sections • 43 lessons • 13.50h total

[View this course online](#) [Browse all courses](#)

About this course

This course prepares you for the AWS Certified Machine Learning Specialty (MLS-C01) exam. We cover SageMaker, data preparation, model training, evaluation, and deployment. Hands-on labs, real ML scenarios, and practice exams included.

Course curriculum

Foundations and Exam Strategy

- **Understanding the MLS-C01 Exam Blueprint (12m)**
We break down the four key exam domains and weightings to help you focus your study time effectively.
- **Setting Up Your AWS Free Tier for Labs (18m)**
Step-by-step guide to configuring your AWS environment safely for the hands-on exercises in this course.
- **Key AWS Services Overview for ML (22m)**
A high-level tour of the core AWS services you must know, from data storage to compute and ML-specific tools.
- **The AWS Well-Architected ML Framework (15m)**
Learn the core pillars of building secure, high-performing, and cost-effective ML systems on AWS.
- **Data Privacy and Compliance Basics (9m)**
Essential concepts for data governance, IAM roles, and compliance standards relevant to machine learning.
- **Calculating Total Cost of Ownership (TCO) (11m)**
How to estimate costs for ML pipelines and choose the right pricing models for different AWS services.
- **Navigating AWS Documentation and Whitepapers (8m)**
Pro tips for finding the information you need quickly during your studies and on the exam.

Data Engineering on AWS

- **Building Data Ingestion Pipelines with AWS Glue (25m)**
Real case study analysis: Setting up a serverless ETL job to prepare data for model training.
- **Optimizing Data Storage with S3 and Athena (28m)**
Learn best practices for data lakes, partitioning strategies, and querying data efficiently.
- **Batch vs. Real-time Data Processing (16m)**
Problem-solving session: Deciding when to use Kinesis, Glue, or Batch for your ML data needs.
- **Feature Engineering in the Cloud (21m)**
How to use SageMaker Processing Jobs and Scikit-learn for scalable feature transformation.
- **Managing Data Pipelines with Step Functions (19m)**
A practical guide to orchestrating complex data workflows and handling failures gracefully.
- **Data Labeling Strategies using SageMaker Ground Truth (14m)**
Walkthrough of setting up a labeling job and managing human-in-the-loop workflows.
- **Handling Large Datasets and Distributed Training Prep (23m)**
Techniques for preparing massive datasets for training, including sharding and format conversion.
- **Data Validation and Quality Checks (10m)**
Implementing automated checks to ensure your data is clean and ready for modeling.

Exploratory Data Analysis (EDA) & Preprocessing

- **Running Jupyter Notebooks on SageMaker Studio (24m)**
Comprehensive setup and overview of the SageMaker Studio environment for interactive analysis.
- **Visualizing Data with Athena and QuickSight (20m)**
How to create insightful dashboards to understand distributions and identify outliers.
- **Statistical Analysis using AWS Glue DataBrew (17m)**
A no-code approach to profiling your data and identifying anomalies before modeling.
- **Identifying and Handling Missing Data (13m)**
Practical strategies for imputation and handling null values in a cloud environment.
- **Feature Scaling and Normalization Techniques (15m)**
Applying StandardScaler and MinMaxScaler using SageMaker processing scripts.
- **Categorical Data Encoding Strategies (12m)**
When to use One-Hot, Label, or Target encoding for your specific ML problem.
- **Detecting Data Imbalance and Bias (18m)**
Tools and techniques to analyze your dataset for fairness and class imbalance issues.

Modeling: Training and Evaluation

- **Choosing the Right Algorithm: Built-in vs. Custom (22m)**
Decision framework for selecting XGBoost, Linear Learner, or bringing your own script.
- **Hyperparameter Tuning with Automatic Model Tuning (30m)**
Step-by-step guide to configuring and running hyperparameter optimization jobs.
- **Training a Model with Distributed Training (27m)**
Practical walkthrough of model training across multiple instances for speed.
- **Evaluating Model Performance: Metrics Deep Dive (19m)**
Understanding Precision, Recall, F1-score, and AUC in the context of AWS model evaluation.
- **Experiment Tracking with SageMaker Experiments (16m)**
How to log parameters, metrics, and artifacts to keep your experiments organized.
- **Debugging Training Jobs with SageMaker Debugger (21m)**
Real-time debugging of training jobs to catch issues like vanishing gradients.
- **Model Explainability with SageMaker Clarify (23m)**
Using SHAP values to understand feature importance and explain model predictions.
- **Saving and Versioning Models in Model Registry (11m)**
Best practices for organizing and managing your trained model artifacts.
- **Cross-Validation Strategies on AWS (14m)**
Implementing k-fold validation to ensure your model generalizes well.

Model Deployment and Operations (MLOps)

- **Real-time Inference Endpoints with SageMaker (26m)**
Deploying a model to a live endpoint and testing it with sample data.
- **Batch Transform Jobs for Large Datasets (20m)**
How to process huge amounts of data asynchronously without a live endpoint.

- **A/B Testing and Shadow Deployments (25m)**
Strategies for safely rolling out new model versions to production users.
- **Auto-scaling Inference Endpoints (18m)**
Configuring scaling policies to handle traffic spikes and save costs during quiet periods.
- **Monitoring Model Drift with SageMaker Model Monitor (29m)**
Setting up alerts to detect when your model's performance degrades over time.
- **Building CI/CD Pipelines for ML with CodePipeline (32m)**
Automating the training, evaluation, and deployment process from code commit to production.
- **Securing Your Inference Endpoints (15m)**
Implementing VPCs, IAM policies, and HTTPS for secure model access.
- **Troubleshooting Common Deployment Failures (13m)**
A problem-solving session on diagnosing and fixing endpoint errors.
- **Cost Optimization for Inference (17m)**
Choosing the right instance types and using Inference Recommender to lower bills.
- **Edge Deployment with SageMaker Neo (21m)**
How to optimize and deploy models to edge devices like cameras and industrial machines.
- **Building an End-to-End ML Pipeline (35m)**
Putting it all together: A complete project from data ingestion to a live, monitored endpoint.
- **Final Exam Day Tips and Common Pitfalls (8m)**
A quick review of what to expect on exam day and common mistakes to avoid.

Discover our blog

Learn more with free articles written by our professors and guest experts.

Latest article: **Flight Attendant Preparation: What to Study Before You Apply**

[Visit the blog](#) [All courses](#) [This course](#)