

Cloudera Machine Learning Engineer Certification Training (CDP-6001)

OEM: Cloudera • Duration: 3 Days (24 hrs) • Code: CDP-6001

COURSE MODULES & TOPICS

Domain 1: Cloudera Machine Learning – CML (31%)

- Workspaces, Projects, Experiments, and Accelerators for ML Projects (AMPs)
- Data Visualizations, Runtimes, and GPU configuration
- CML sessions and environment management
- End-to-end workflow: data ingestion to model serving within CML

Domain 2: Spark MLlib (22%)

- Model selection, tuning, and evaluation
- Spark ML Pipelines with estimators and transformers
- Classification, regression, and clustering algorithms
- Collaborative filtering algorithms
- Feature transformation techniques

Domain 3: Apache Spark (18%)

- DataFrames, Spark SQL, and Spark RDDs
- File types for ML workloads
- Window functions for feature engineering
- Data manipulation and ML data pipeline construction

Domain 4: Deploying Machine Learning Models (18%)

- Applications and API endpoints for model serving
- Autoscaling and performance optimization
- Model metrics, monitoring, and MLflow integration
- Model serving, versioning, and endpoint management
- Operationalizing models in production environments

Domain 5: Deep Learning & General Machine Learning (11%)

- Supervised and unsupervised learning concepts
- Algorithm selection for classification, regression, clustering
- Neural network basics and deep learning frameworks

- Overarching machine learning methodologies