AWS Certified Machine Learning Engineer - Associate Exam Guide

# Summary of Sections 1.1 to 4.3

# Domain 1: Data Preparation (28%)

## 1.1 Ingest and store data

Knowledge: Data formats (CSV, JSON, Parquet, ORC, Avro, RecordIO), core storage (S3, EFS, FSx), streaming (Kinesis, Flink, Kafka), tradeoffs.

Skills: Extract with correct AWS options, choose formats by access pattern, ingest to Data Wrangler & Feature Store, merge with Glue/Spark, troubleshoot ingestion, decide storage on cost/performance.

## 1.2 Transform data & perform feature engineering

Knowledge: Cleaning (outliers, missing values, deduplication), feature engineering (scaling, splitting, binning, normalization, log transforms), encodings (one-hot, label, tokenization), tools (Data Wrangler, Glue, DataBrew), streaming transforms (Lambda, Spark).

Skills: Transform with Glue/EMR/Data Wrangler, manage features in Feature Store, label and validate with Ground Truth/Mechanical Turk.

## 1.3 Ensure data integrity & prepare for modeling

Knowledge: Bias metrics (class imbalance, difference in proportions), imbalance strategies (resampling, synthetic data), encryption, anonymization/masking, compliance (PII/PHI, residency).

Skills: Validate with DataBrew/Data Quality, mitigate bias with Clarify, prep with dataset splits/shuffling/augmentation, configure load into EFS/FSx.

# Domain 2: ML Model Development (26%)

## 2.1 Choose a modeling approach

Knowledge: Algorithm capabilities, AI services (Translate, Transcribe, Rekognition, Bedrock), interpretability, SageMaker built-ins.

Skills: Assess data & feasibility, select algorithms/templates (JumpStart, Bedrock), weigh cost, pick AI services for business needs.

## 2.2 Train and refine models

Knowledge: Training (epochs, steps, batch size), time reduction (early stopping, distributed), model size factors, performance improvement, regularization (dropout, L1/L2), HPO (random/Bayesian), hyperparameters, integrating external models.

Skills: Train with built-ins/frameworks (TF, PyTorch), fine-tune pre-trained (JumpStart, Bedrock), run AMT, prevent over/underfitting, ensemble methods, reduce model size (pruning, compression), manage versions with Model Registry.

## 2.3 Analyze model performance

Knowledge: Metrics (confusion matrix, accuracy, F1, precision, recall, RMSE, ROC, AUC), baselines, over/underfitting detection, Clarify metrics, convergence issues.

Skills: Interpret metrics, tradeoff performance vs cost/time, reproducible experiments, compare shadow vs production, interpret with Clarify, debug with Model Debugger.

## 2.4 Tune models

Knowledge: Hyperparameters vs parameters, optimization (grid, random, Bayesian), HPO tools, Autopilot.

Skills: Define objective metric, avoid overfitting, early stopping.

# Domain 3: Deployment & Orchestration (22%)

## 3.1 Select deployment infrastructure

Knowledge: Deployment best practices (versioning, rollback), endpoint types (serverless, async, batch, real-time), compute (CPU/GPU), container options, SageMaker Neo for edge.

Skills: Evaluate performance, cost, latency, choose compute env, orchestrators (Airflow, Pipelines), multi-model/multi-container, deployment target (SageMaker, ECS, EKS, Lambda), strategy (real-time vs batch).

## 3.2 Create and script infrastructure

Knowledge: On-demand vs provisioned, scaling policies, IaC (CloudFormation, CDK), containers/EKS/ECS, SageMaker endpoint auto scaling.

Skills: Apply best practices (Spot, auto scaling, Lambda behind endpoints), automate provisioning (CloudFormation, CDK), BYOC with SageMaker, configure VPC endpoints, deploy/host via SDK, choose scaling metrics (latency, CPU).

## 3.3 Automate ML workflows (CI/CD)

Knowledge: CodePipeline, CodeBuild, CodeDeploy, version control (Git), CI/CD principles, deployment strategies (blue/green, canary, linear).

Skills: Configure Code\* services, apply GitFlow, automate with EventBridge/Pipelines, build retraining mechanisms, add automated tests (unit/integration/E2E).

# Domain 4: Monitoring, Maintenance & Security (24%)

## 4.1 Monitor model inference

Knowledge: Model drift, monitoring data quality and performance, monitoring design.

Skills: Monitor with Model Monitor, detect anomalies/errors, detect distribution shifts (Clarify), A/B test models.

## 4.2 Monitor & optimize infrastructure and costs

Knowledge: Infra KPIs (utilization, throughput, scalability, fault tolerance), monitoring tools (X-Ray, CloudWatch Logs/Lambda Insights), CloudTrail, instance types, cost tools (Cost Explorer, Budgets, Trusted Advisor), tagging.

Skills: Configure CloudWatch/alarms, CloudTrail trails, dashboards (QuickSight), monitor with EventBridge, right-size with Inference Recommender/Compute Optimizer, troubleshoot latency/scaling, optimize cost with Spot/On-Demand/Reserved/Savings Plans.

## 4.3 Secure AWS resources

Knowledge: IAM roles, policies, SageMaker security, network access controls, CI/CD security best practices.

Skills: Enforce least privilege, configure IAM, monitor/audit/log, troubleshoot security issues, secure VPCs/subnets/security groups.