AWS ML Associate — Section 2.3: Train and Evaluate Models

This section focuses on the iterative process of training models and evaluating their performance using the correct metrics. It covers training workflows, evaluation metrics for different problem types, and experiment tracking.

# Training Process

• Define train/validation/test splits (holdout or cross-validation)

• Run training jobs in SageMaker (built-in, framework, or custom containers)

• Use distributed training for large datasets

# Evaluation Metrics

• Classification — accuracy, precision, recall, F1-score, ROC-AUC

• Regression — MSE, RMSE, MAE, R²

• Clustering — silhouette score, inertia

• Recommendation — precision@k, recall@k, NDCG

# Experiment Tracking

• SageMaker Experiments — manage multiple runs

• Record hyperparameters, metrics, and artifacts for reproducibility

# Exam Tips

• Match metric to problem type (e.g., ROC-AUC for classification, RMSE for regression)

• Watch out for class imbalance — accuracy may be misleading

• Cross-validation is more reliable than a single holdout split

• Always log metrics and parameters for comparison across runs