Day 1: Introduction to Databricks on AWS & Fundamentals (4 hours)

1. **Introduction to Databricks**
    - What is Databricks?
    - Differences between Databricks on AWS vs. Azure vs. GCP
    - Databricks workspace, clusters, notebooks, and jobs
    - Pricing and cost optimization
2. **Databricks & AWS Integration**
    - Setting up Databricks on AWS
    - Connecting Databricks with S3
    - IAM roles and permissions for secure access
    - Networking and security best practices
3. **Working with Databricks Notebooks**
    - Basics of Python & Scala in Databricks
    - Using SQL in Databricks
    - Introduction to MLflow for tracking

Day 2: Delta Lake & Data Engineering in Databricks (4 hours)

1. **Introduction to Delta Lake**
    - What is Delta Lake?
    - Delta Lake vs. traditional data lakes
    - ACID transactions in Delta Lake
2. **Implementing Delta Lake on AWS**
    - Creating Delta tables
    - Schema evolution & time travel
    - Optimizing Delta Lake performance
    - Using Databricks Auto Loader
3. **Data Engineering in Databricks**
    - ETL vs. ELT in Databricks
    - Using Spark SQL & PySpark for data processing
    - Handling semi-structured & unstructured data

Day 3: Orchestration with Airflow & Transformation with DBT (4 hours)

1. **Apache Airflow & Databricks Integration**
    - Introduction to Airflow
    - DAGs, tasks, and scheduling basics
    - Connecting Airflow to Databricks via API
2. **Building Workflows with Airflow**
    - Creating & triggering Databricks jobs from Airflow
    - Using Airflow for dependency management
    - Error handling and retry mechanisms
3. **Introduction to DBT (Data Build Tool)**
    - What is DBT?
    - DBT workflow for SQL-based transformations
    - Connecting DBT with Databricks
    - Running DBT models in Databricks

Day 4: Advanced Topics & Real-World Project (4 hours)

1. **Performance Optimization & Best Practices**
    - Databricks Delta Optimizations (Z-Ordering, Caching, Indexing)
    - Query tuning & cluster configurations
    - Auto-scaling & cost control in AWS
2. **Machine Learning & Advanced Analytics in Databricks**
    - Introduction to MLflow for ML model tracking
    - Training & deploying models in Databricks
    - Integrating Databricks with AWS services (SageMaker, Redshift)
3. **End-to-End Project: Real-World Data Pipeline**
    - Extract data from AWS S3
    - Process & transform data using Delta Lake
    - Orchestrate the pipeline with Airflow
    - Use DBT for transformations
    - Visualize data using Databricks SQL

Day 5: Performance Tuning & Optimization in Databricks (4 hours)

4. **Optimizing Databricks Clusters**
    - Choosing the right cluster size & instance types
    - Auto-scaling & cluster termination policies
    - Spot instances vs. on-demand instances for cost optimization
5. **Optimizing Delta Lake Performance**
    - Z-Ordering and Data Skipping
    - File Compaction (Optimize & Vacuum commands)
    - Bloom filters for fast queries
    - Partitioning strategies
6. **Query & Job Optimization Techniques**
    - Broadcast joins vs. Shuffle joins
    - Catalyst optimizer & query execution plans
    - Using caching & materialized views
    - Parallel processing & resource utilization
7. **Monitoring & Debugging Performance Issues**
    - Using Databricks Spark UI for performance tuning
    - Understanding job execution metrics & logs
    - Troubleshooting slow queries and job failures
    - Best practices for production environments