AWS machine learning with data analytics

1. AWS Sagemaker MLOPS using Python (20 hrs)

1.1 Introduction to MLOps

1.1.1 Processes

1.1.2 People

1.1.3 Technology

1.1.4 Security and governance

1.1.5 MLOps maturity mode

1.2 Repeatable MLOps: Repositories

1.2.1 Managing data for MLOps

1.2.2 Version control of ML models

1.2.3 Code repositories in ML

1.3 Introduction to Machine Learning and the ML Pipeline

1.3.1 End-to-end orchestration with AWS Step Functions

1.3.2 End-to-end orchestration with SageMaker Projects

1.3.3 Governance and security

1.4 Introduction to Amazon SageMaker

1.5 Problem Formulation

1.6 Preprocessing

1.7 Model Training and Evaluation

1.8 Deployment Pipeline

1.9 MLOps: Monitoring

1.9.1 The importance of monitoring in ML

1.9.2 Operations considerations for model monitoring

1.9.3 Remediating problems identified by monitoring ML solutions

2. AWS Datalake formation (8 hr)

2.1 Introduction to data lakes

2.2 Data ingestion, cataloging, and preparation

2.3 Data processing and analytics

2.4 Building a data lake with AWS Lake Formation

2.5 Terminologies

2.6 Lake Formation Integrations

2.7 Configure Data Lake

2.8 Using crawlers in AWS Glue

2.9 Query Data

3. AWS Glue and Databrew (8 hrs)

3.1 Introduction to AWS Glue

3.2 Understanding Data Transformation with AWS Glue

3.3 Working with Data Sources and Targets

3.4 Managing and Monitoring AWS Glue Jobs

3.5 Working with AWS Glue DataBrew

3.6 Data Profiling

3.7 Data Quality

3.8 Transformation

3.9 Feature Engineering

4. Amazon Redshift (8 hrs)

4.1 Warehouse creation and connecting with Query Editor

4.2 Table creation, one time data load

4.3 Continuous load with stored procedures

4.4 Data sharing from data producer (serverless) to consumer (provisioned cluster)

4.5 Create Machine Learning model with Redshift ML features

4.6 User Redshift Spectrum features to query data in S3 data lake

4.7 Redshift monitoring and audit logging