



ROVER

ROVER CONSULTING & Training Services

Mastering Databricks

Comprehensive Training Programs

Unlock the full potential of Databricks with our expert-led training programs. From advanced data engineering techniques to cutting-edge machine learning applications, our comprehensive courses are designed to elevate your skills and drive data excellence.

16 Hours Training



Duration: 16 Hours

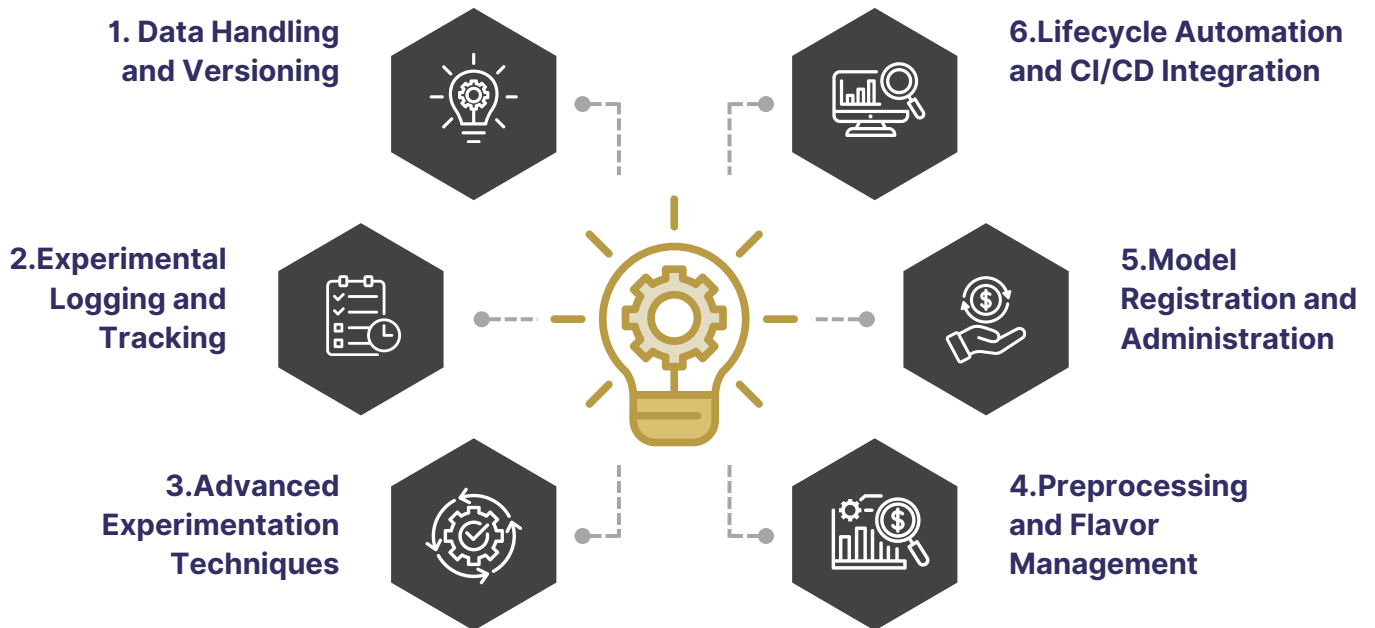
Machine Learning Essentials

A 2-Day Professional Training on Databricks

Databricks Machine Learning Professional covers Experimentation and Data Management by teaching Delta table operations, versioning, and Feature Store management, along with using MLflow for logging, tracking, and advanced experimentation techniques. In Model Lifecycle Management, it addresses MLflow flavors, model registration, and automation with Webhooks and Databricks Jobs. Model Deployment Strategies include batch, streaming, and real-time methods, focusing on optimization and pipeline conversion. Finally, Solution and Data Monitoring involves drift analysis, monitoring with statistical tests, and strategies for managing model drift to ensure performance.



High Level Overview



This overview addresses essential practices for managing and optimizing the machine learning lifecycle. It begins with data handling and versioning, ensuring consistency and traceability across experiments. The discussion then explores experimental logging and tracking, crucial for maintaining detailed records of model performance.

Advanced experimentation techniques are highlighted to enhance model development, along with preprocessing and flavor management for streamlined data preparation. Additionally, the guide covers model registration and administration, facilitating organized model management, and concludes with strategies for lifecycle automation and CI/CD integration, ensuring continuous delivery and efficient deployment in machine learning pipelines.

High Level Overview



This overview focuses on the deployment strategies and ongoing monitoring crucial for maintaining effective machine learning models in production. It begins with an examination of batch, streaming, and real-time deployment methods, each tailored to different operational needs.

The discussion then shifts to drift analysis, which identifies shifts in model performance over time, followed by techniques for drift monitoring and testing to detect and respond to these changes. Finally, comprehensive drift management strategies are presented, ensuring models remain accurate and reliable in dynamic environments.

Course Curriculum



MODULE 1: EXPERIMENTATION AND DATA MANAGEMENT

1. Data Handling and Versioning

- 1.1. Execute read and write operations on Delta tables
- 1.2. Access historical versions of Delta tables and perform version restoration
- 1.3. Create, update, and merge Feature Store tables for machine learning workflows

2. Experimental Logging and Tracking

- 2.1. Log parameters, models, and evaluation metrics manually using MLflow
- 2.2. Programmatically retrieve and utilize data, metadata, and models from MLflow experiments

3. Advanced Experimentation Techniques

- 3.1. Conduct MLflow experiment tracking with model signatures and input examples
- 3.2. Configure and utilize nested runs and enable autologging with Hyperopt
- 3.3. Record and analyze artifacts including SHAP plots, custom visualizations, feature data, images, and metadata

MODULE 2: MODEL LIFECYCLE MANAGEMENT

1. Preprocessing and Flavor Management

- 1.1. Understand MLflow flavors and their benefits, including the pyfunc flavor
- 1.2. Integrate preprocessing logic and contextual information within custom model classes and objects

2. Model Registration and Administration

- 2.1. Navigate Model Registry functionalities, including model registration, metadata management, and stage transitions
- 2.2. Register new models and versions programmatically and manage model lifecycle stages (e.g., transition, archive, delete)

3. Lifecycle Automation and CI/CD Integration

- 3.1. Automate model lifecycle processes using Model Registry Webhooks and Databricks Jobs
- 3.2. Create Jobs triggered by model stage changes and integrate Webhooks for automated workflows
- 3.3. Manage HTTP webhooks, including creation, listing, and deletion

Course Curriculum



MODULE 4: SOLUTION AND DATA MONITORING

1. Drift Analysis

- 1.1. Differentiate between label drift, feature drift, and concept drift
- 1.2. Assess scenarios likely to exhibit feature drift and label drift and evaluate their impact on model performance

2. Drift Monitoring and Testing

- 2.1. Utilize summary statistics and robust statistical tests for monitoring numeric and categorical feature drift
- 2.2. Apply Jensen-Shannon divergence, Kolmogorov-Smirnov, and chi-square tests for drift detection and analysis

3. Comprehensive Drift Management

- 3.1. Develop workflows for measuring and addressing concept drift and feature drift
- 3.2. Implement model retraining and assess model performance on updated data to address detected drift

MODULE 3: MODEL DEPLOYMENT STRATEGIES

1. Batch Deployment

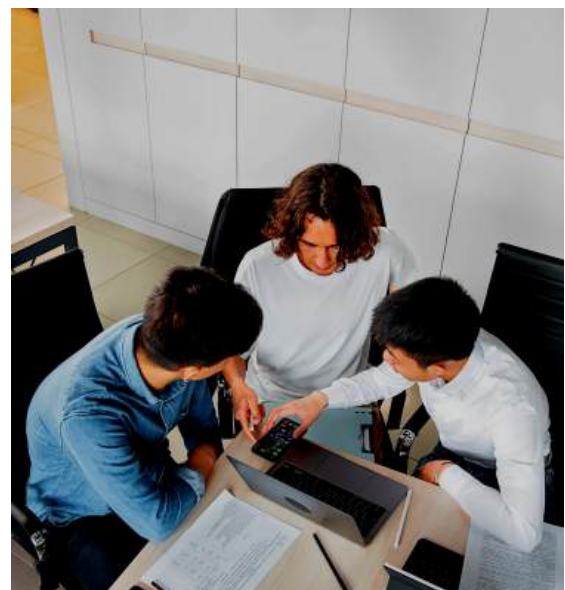
- 1.1. Implement batch deployment for precomputed predictions and assess its use cases
- 1.2. Optimize prediction performance with z-ordering and partitioning strategies
- 1.3. Load models using `load_model` and deploy single-node models in parallel with `spark_udf`

2. Streaming Deployment

- 2.1. Utilize Structured Streaming for continuous data processing and real-time inference
- 2.2. Transform batch deployment pipelines into streaming pipelines for real-time data and predictions

3. Real-Time Deployment

- 3.1. Deploy real-time inference for scenarios requiring rapid predictions
- 3.2. Leverage cloud-based RESTful services and containers for production-grade real-time deployments



For India Region

Pricing

Course fees

₹ 3,00,000

Travel and Accommodation cost
(* Applicable for In-person training)

₹1,00,000





Contact Us !



CONSULTING & Training Services

Contact Us Today

 9052776606

 info@rovertek-ai.com



ROVER

ROVER CONSULTING INDIA PRIVATE LIMITED

**Thank
You**

