

LLM Evaluation using MLflow

Duration: 03 days (24 hours)

Note: Client will require their openai key to perform labs. They can get the openai key via their credit card for pay as you go subscription:
<https://platform.openai.com/settings/organization/billing/overview>

Course Objective: This course introduces MLflow, emphasizing its core components and practical applications in managing machine learning workflows. Participants will learn to deploy and trace Large Language Models (LLMs) effectively, and gain expertise in evaluating and validating LLMs using MLflow.

Pre-requisite: Participants should have a basic understanding of machine learning concepts, proficiency in Python programming, and familiarity with model evaluation metrics.

Labs: Koenig DC (VM) will be provided for lab performance

Module 01: Introduction to MLflow

- Core components of MLflow
- Why use MLflow?
- Who uses MLFlow?
- Use cases of MLflow
- Scalability in MLflow
- Basic demo: how to get started with MLflow

Module 02: LLMs with MLflow

- MLflow tracing
- MLflow deployments servers for LLMs
- LLM Evaluation
- Prompt Engineering UI
- Native MLflow Flavors for LLMs
- LLM Tracking in MLflow

Module 03: Evaluating LLMs with MLflow

- Harnessing the power of Automation
- Evaluating with LLMs
- Evaluating with Extra Metrics
- Evaluating with a Function
- Evaluating with a Static Dataset
- Performing Model Validation
- Model Validation with Giskard's plugin
- Model Validation with Trubrics' plugin