# Mlflow

MLflow is an open-source platform designed to manage the entire machine learning lifecycle. It provides solutions for experiment tracking, reproducibility, and deployment.

**Core Components of MLflow**

1. MLflow Tracking:
- Tracks experiments, metrics, parameters, and artifacts.
- Allows comparison of different model runs.
- Stores results locally or in a central tracking server.
1. MLflow Projects:
- Standardizes the way to organize machine learning code.
- Uses a directory structure and MLproject file for reproducibility.
1. MLflow Models:
- A standardized format for packaging machine learning models.
- Enables deployment across multiple platforms such as Docker, REST API, or cloud environments.
1. MLflow Registry:
- A model repository to manage model lifecycle stages: staging, production, and archived.
- Provides model versioning and annotations.

**Advantages of MLflow**

- Centralized tracking of experiments.
- Model reproducibility and deployment.
- Framework-agnostic: Works with TensorFlow, PyTorch, Scikit-learn, and others.
- Simplifies collaboration by sharing experiment data.

# DagsHub

DagsHub is a platform for versioning, collaboration, and tracking in machine learning. It integrates with Git for version control, DVC (Data Version Control) for data tracking, and MLflow for experiment tracking.

**Key Features of DagsHub**

1. Version Control for Code and Data:
- Built on Git and DVC for tracking code, datasets, and machine learning models.
- Enables branching, merging, and collaboration similar to GitHub.
1. Experiment Tracking:
- Direct integration with MLflow for tracking and comparing experiments.
1. Visualization and Insights:
- Visualize metrics, parameters, and data pipelines directly from the UI.
1. Collaboration:
- Share experiments, datasets, and pipelines with teams.

- Comment and discuss results and changes.

**Advantages of DagsHub**

- Centralized platform for machine learning projects.
- Seamless integration with Git, DVC, and MLflow.
- Facilitates reproducibility and collaboration.
- Reduces the overhead of setting up individual tools.

## How Mlflow and DagsHub related?

- DagsHub uses MLflow's tracking capabilities, making it easy to visualize and compare experiment runs.
- It serves as a central repository for MLflow logs, enabling cloud-based access and collaboration.
- While MLflow focuses on the technical tracking of models, DagsHub enhances collaboration by combining MLflow with Git, DVC, and team tools.