Homework 3: MLOps System Overview

GitHub Repository

https://github.com/MattBecker92/mlops_Homework3/

System Overview

This system demonstrates an end-to-end MLOps workflow integrating **Airflow**, **MLflow**, and **FastAPI**:

1. **Airflow DAGs**
    - Orchestrate model training tasks.
    - Each DAG runs train.py to train models and log them to MLflow.
    - Models are versioned and stored in MLflow's Model Registry.
2. **MLflow**
    - Acts as the central tracking server and model registry.
    - Stores model artifacts, metrics, and parameters.
    - Provides version control for models and stages (e.g., Production).
3. **FastAPI Server**
    - Serves predictions using models loaded from MLflow.
    - Endpoints:
        - /predict: Predict Iris species from input samples.
        - /set-version: Dynamically switch the model version being served.
        - /current-version: View the currently active model version.
        - /generate-and-predict: Generate random test data and return predictions.
4. **Model Storage**
    - MLflow stores models locally or in a remote artifact store.
    - FastAPI loads models directly from MLflow using mlflow.pyfunc.