

Diamond Dataset ML Pipeline Orchestration and Experiment Tracking.

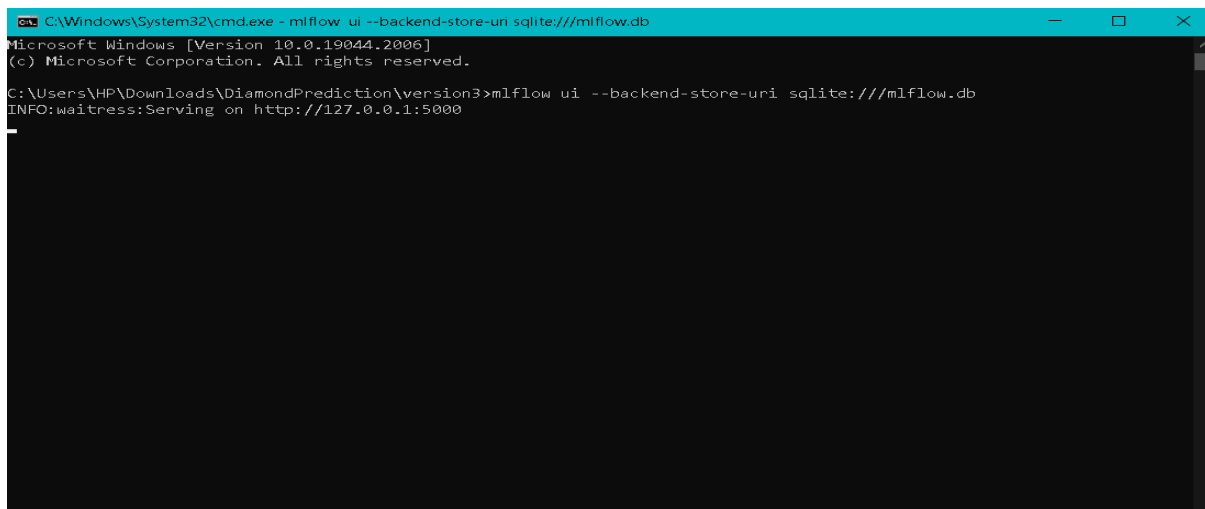
Experiment Tracking: -

Experiment tracking is the process of saving all experiment related information that you care about for every experiment you run. This “metadata you care about” will strongly depend on your project, but it may include:

- Scripts used for running the experiment
- Environment configuration files
- Versions of the data used for training and evaluation
- Parameter configurations
- Evaluation metrics
- Model weights
- Performance visualizations

Experiment tracking focuses on the iterative model development phase when you try many things to get your model performance to the level you need.

- In this project I used 'MLFlow' tool for experiment tracking.

A screenshot of a Windows command prompt window. The title bar is blue and shows the path 'C:\Windows\System32\cmd.exe - mlflow ui --backend-store-uri sqlite:///mlflow.db'. The command prompt shows the following text: 'Microsoft Windows [Version 10.0.19044.2006] (c) Microsoft Corporation. All rights reserved. C:\Users\HP\Downloads\DiamondPrediction\version3>mlflow ui --backend-store-uri sqlite:///mlflow.db INFO:waitress:Serving on http://127.0.0.1:5000'. The rest of the window is black.

Command for running MLFlow.

The screenshot shows the MLFlow Experiments dashboard for an experiment named 'Diamonds Price Prediction'. The interface includes a sidebar with a search bar and a list of experiments. The main panel displays the experiment details, including a description, a table of runs, and a list of metrics.

Experiment ID: 1

Description Edit

Refresh Compare Delete Download CSV Created All time

Columns Only show differences metrics.rmse < 1 and params.model = "tree" Search Filter Clear

Showing 62 matching runs

	Created	Duration	Run Name	User	Source	Version	Models	best_cv_score	mean_fit_time	mean_scc
<input type="checkbox"/>	2 days ago	25.1min	honorable-a...	HP	workflow...	-	sklearn, 1 more	-351.1	-	-
<input type="checkbox"/>	2 days ago	24.7min	learned-har...	HP	workflow...	-	-	0.192	1.226	-
<input type="checkbox"/>	2 days ago	24.7min	colorful-cat...	HP	workflow...	-	-	0.232	0.412	-
<input type="checkbox"/>	2 days ago	24.7min	chill-sponge...	HP	workflow...	-	-	0.217	1.26	-
<input type="checkbox"/>	2 days ago	24.7min	unleashed-r...	HP	workflow...	-	-	0.188	0.64	-
<input type="checkbox"/>	2 days ago	24.7min	youthful-de...	HP	workflow...	-	-	0.203	0.901	-
<input type="checkbox"/>	2 days ago	24.7min	thoughtful-a...	HP	workflow...	-	-	0.192	0.519	-
<input type="checkbox"/>	2 days ago	24.7min	smiling-duc...	HP	workflow...	-	-	0.218	0.873	-
<input type="checkbox"/>	2 days ago	24.7min	sassy-cod-553	HP	workflow...	-	-	0.184	1.272	-
<input type="checkbox"/>	2 days ago	24.7min	youthful-do...	HP	workflow...	-	-	0.241	0.584	-

MLFlow Dashboard

ML Pipeline Orchestration: -

ML Pipeline Orchestration basically means the workflow of our Machine Learning Project in which we convert python jupyter notebook to Python Script for Production ready code.

Pipelines in machine learning are an infrastructural medium for the medium for the entire ML workflow. Pipelines help automate the overall MLOps workflow, from data gathering, EDA, data arguments, to model building and deployment. After the deployment, it also supports reproduction, tracking and monitoring.

ML pipelines help improve the performance and management of the entire model, resulting in quick and easy deployment.

Machine Learning Orchestration tool :- ML orchestration tools are used to automate and manage workflows and pipeline infrastructure, with a simple, collaborative interface, Along with management and creation of custom workflows and their pipelines, these tools also help us track and monitor models for further analysis.

- In this project I used 'Prefect' tool for pipeline orchestration.

```
C:\Windows\System32\cmd.exe - prefect orion start

(env) C:\Users\HP\Downloads\DiamondPrediction\version3>prefect orion start

[ASSEMBLY] [ORION]

Configure Prefect to communicate with the server with:

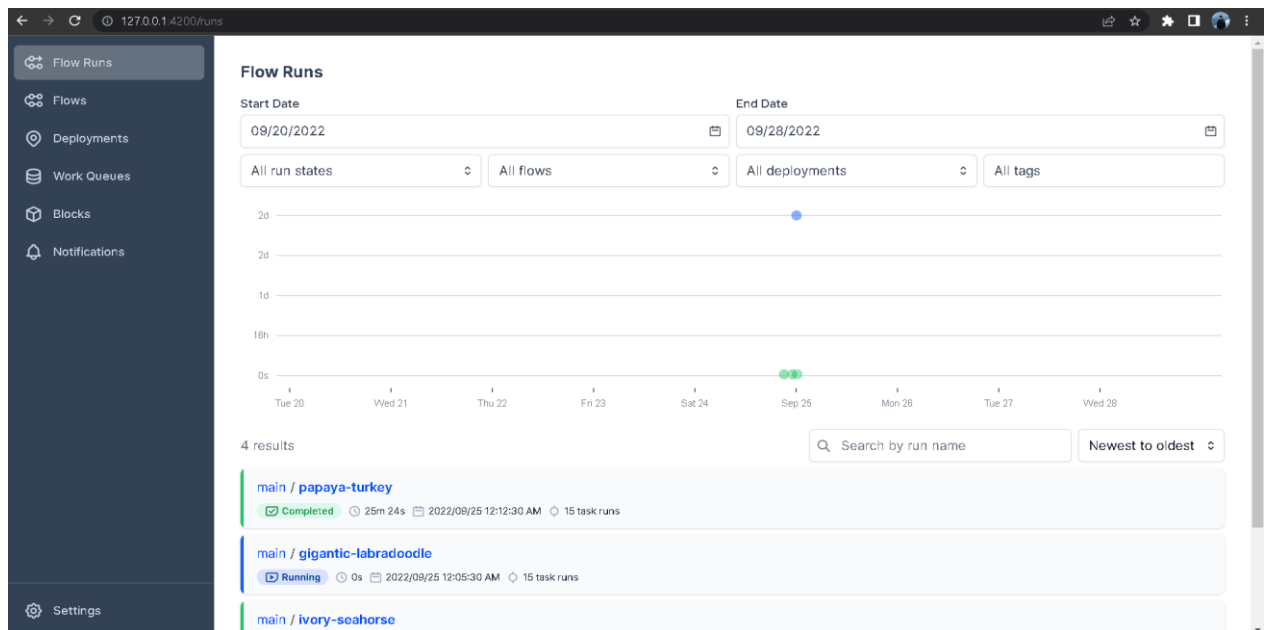
    prefect config set PREFECT_API_URL=http://127.0.0.1:4200/api

View the API reference documentation at http://127.0.0.1:4200/docs

Check out the dashboard at http://127.0.0.1:4200

INFO: Started server process [24020]
INFO: Waiting for application startup.
INFO: Application startup complete.
INFO: Uvicorn running on http://127.0.0.1:4200 (Press CTRL+C to quit)
```

cmd for running prefect



Prefect dashboard