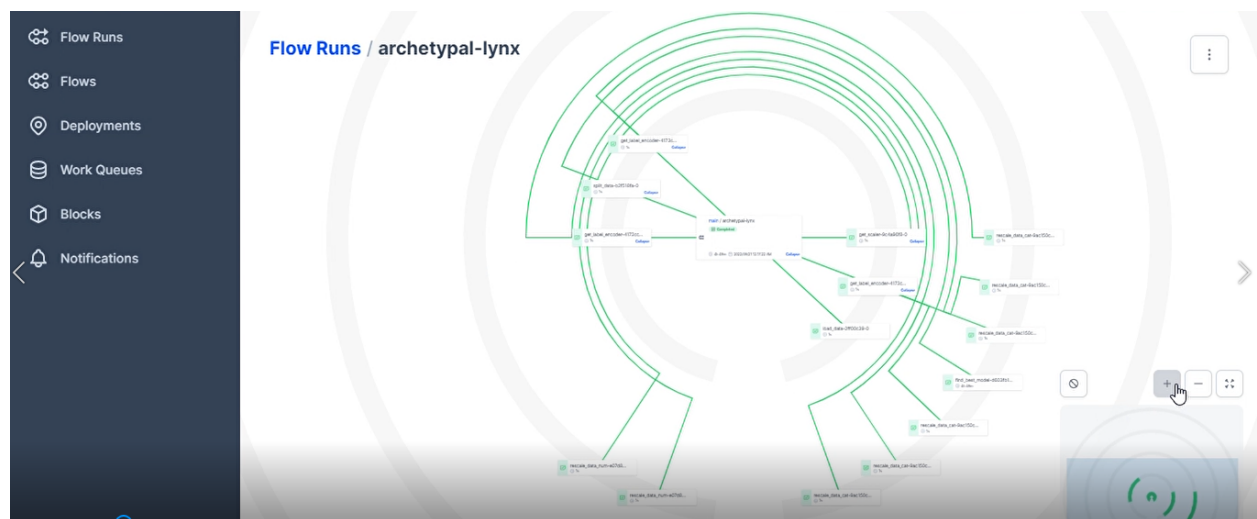


## Task:

## MLOPs Task - Diamond Dataset ML Pipeline Orchestration and Experiment Tracking

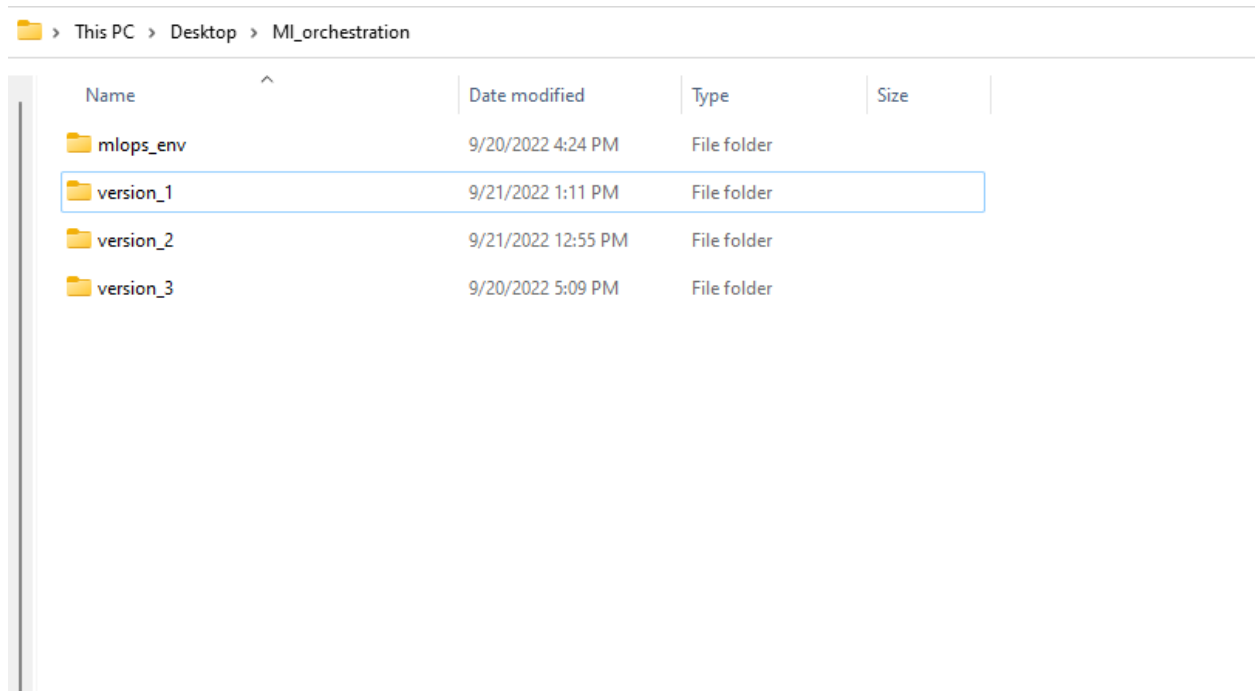
**Description:**

production ready pipeline using Workflow Orchestration tool i.e. Prefect. And to schedule the task runs every one week.

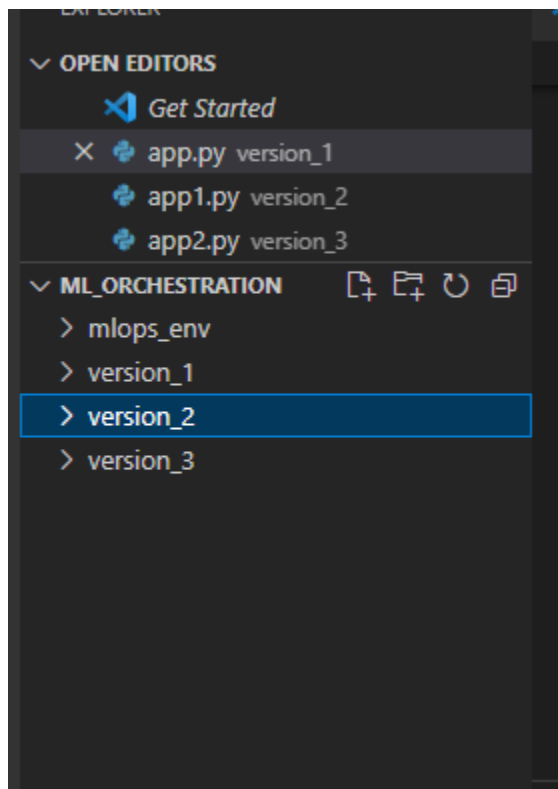


### Step1:

I have created three folders named version\_1 , version\_2 and version\_3 in a Seperate folder named ML\_orchestration.



In VS-Code I have created .py files for each folder.



## **Step-2:**

Created a Python VirtualEnv by running below codes in CMD in current working directory(ML\_Orchestration)

```
python -m venv mlops_env /  
.\mlops_env\Scripts\activate => activate virtual environment in Windows
```

Installed these packages in virtualenv:

```
pip install matplotlib seaborn plotly sklearn mlflow prefect pandas
```

## **Step3:**

After creating a virtual environment with necessary packages I ran the .py files using the below command for each file.

```
(mlops_env) C:\Users\mrjdh\Desktop\ML_orchestration>cd version_1
```

```
(mlops_env)
```

```
C:\Users\mrjdh\Desktop\ML_orchestration\version_1>Python app.py
```

Python app.py(name of first file containing the code Breaking the Jupyter Notebook to Python Script (Basic Code without workflow management)

```
(mlops_env) C:\Users\mrjdh\Desktop\ML_orchestration\version_1>cd..
```

```
(mlops_env) C:\Users\mrjdh\Desktop\ML_orchestration>cd version_2
```

```
(mlops_env)
```

```
C:\Users\mrjdh\Desktop\ML_orchestration\version_2>Python app1.py
```

Python app1.py(name of the file which contains the code with Prefect Workflow - Defining the workflow and running them)

```
(mlops_env) C:\Users\mrjdh\Desktop\ML_orchestration\version_2>cd..
```

```
(mlops_env) C:\Users\mrjdh\Desktop\ML_orchestration>cd version_3
```

```
(mlops_env)
```

```
C:\Users\mrjdh\Desktop\ML_orchestration\version_3>Python app2.py
```

Python app2.py(name of the third file which contains the code for Deployment and Scheduling tasks

## Experiment Tracking and Model Management using MLFlow

Run below mentioned commands to install mlflow on your system:

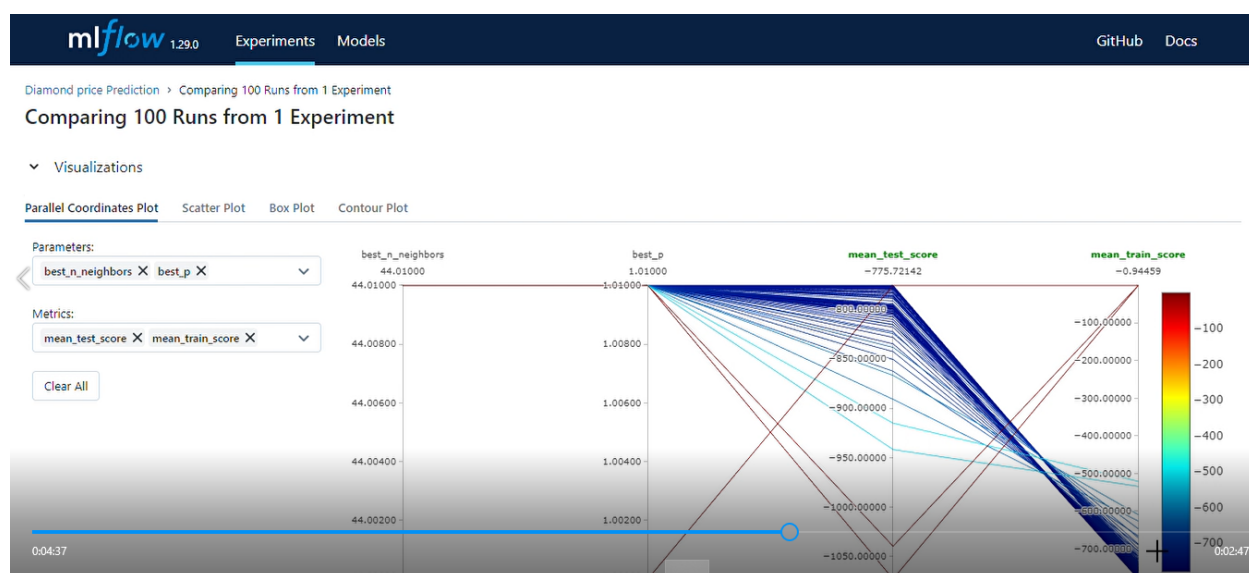
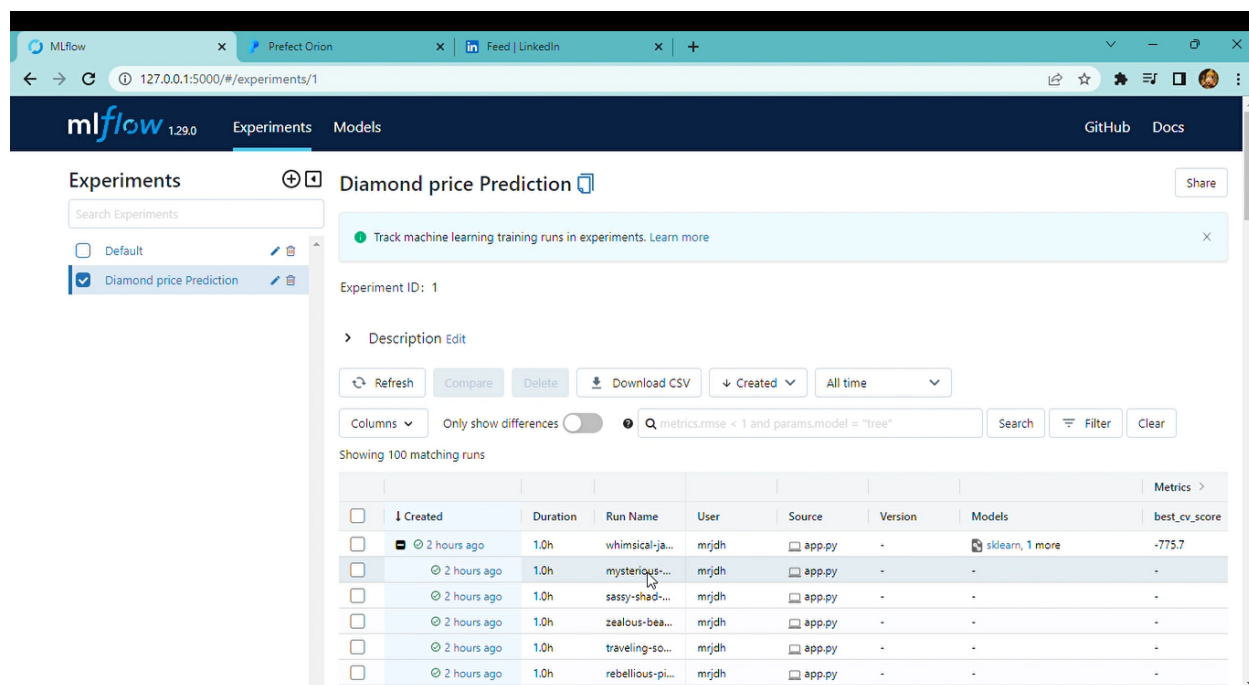
```
pip install mlflow(skip this step if already installed)
mlflow ui --backend-store-uri sqlite:///mlflow.db
```

Add the below commands in each .py files to track the experiments.

```
from sklearn.model_selection import mlflow
```

```
# WORKFLOW
def main(path: str):
    mlflow.set_tracking_uri("sqlite:///mlflow.db")
    mlflow.set_experiment("Diamond price Prediction")
    # Define Parameters
```

## MLFlow Interface for Tracking Experiments



# Managing Machine Learning Workflows using Prefect 2.0

## Installing Prefect 2.0

```
$ pip install prefect==2.4
```

## Running Prefect Dashboard

```
$ prefect orion start
```

\_ \ \_ \ \_	\_	\_ / \_	\_				
\_ / / \_		\_		( \_			
\_		\_ \ \_	\_		\_ \ \_		

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/>

## Deployment of Prefect Flow

I have added the below code in the app2.py file.

```
from prefect.deployments import Deployment
from prefect.orion.schemas.schedules import IntervalSchedule
from datetime import timedelta
```

```
deployment = Deployment.build_from_flow(
    flow=main,
    name="model_training",
    schedule=IntervalSchedule(interval=timedelta(minutes=5)),
    work_queue_name="ml"
)
```

```
deployment.apply()
```

## Running an Agent

```
$ prefect agent start --work-queue "ml"
```

MLflow

Prefect Orion

Feed | LinkedIn

127.0.0.1:4200/work-queue/716c1415-3851-416e-aa64-550e827c9090

Flow Runs

Flows

Deployments

Work Queues

Blocks

Notifications

Settings

Work Queues / ml

Work queue is ready to go!

Work queues define the work to be done and agents poll a specific work queue for new work.

Copy

prefect agent start --work-queue "ml"

Upcoming Runs

main / tidy-anteater

auto-scheduled

Scheduled

0s

Scheduled for 2022/09/28 01:04:27 PM

main / silent-myna

auto-scheduled

Scheduled

0s

Scheduled for 2022/10/05 01:04:27 PM

main / judicious-leopard

auto-scheduled

Description

None

Work Queue ID

716c1415-3851-416e-aa64-550e827c9090

Flow Run Concurrency

None

Created

2022/09/20 10:50:14 PM