

MLOps Task Report

1. What is MLOps?

MLOps stands for Machine Learning Operations. MLOps is a core function of Machine Learning engineering, focused on streamlining the process of taking machine learning models to production, and then maintaining and monitoring them.

2. MLFlow

MLflow is a platform to streamline machine learning development, including tracking experiments, packaging code into reproducible runs, and sharing and deploying models. Main work of MLFlow are:

- Experiment Tracking
- Model Management

3. Prefect

Prefect is an open-sourced framework to build workflows in Python. Prefect makes it easy to build, run, and monitor data pipelines at scale.

- Workflow Management.
- Scheduling and Monitoring Tasks.
- Manage ML Pipelines

4. Why Experiment tracking is important?

Experimental tracking allows data scientists to identify factors that would affect a model's performance, compare results and select the optimal version. It is very important to stay organized throughout the iterative learning process, even if your models do not make it into production

“Save everything in one place and never lose your progress again.”

5. ML Flow Implementation Steps:

- Step 1 - Import MLFlow
 - ◆ `import mlflow`
- Step 2 - Set the tracker and experiment
 - ◆ `mlflow.set_tracking_uri(DATABASE_URI)`
 - ◆ `mlflow.set_experiment("EXPERIMENT_NAME")`
- Step 3 - Start a experiment run
 - ◆ with `mlflow.start_run()`:
- Step 4 - Logging the metadata
 - ◆ `mlflow.set_tag(KEY, VALUE)`
 - ◆ `mlflow.log_param(KEY, VALUE)`
 - ◆ `mlflow.log_metric(KEY, VALUE)`
- Step 5 - Logging the model and other files (2 ways)
 - ◆ `mlflow.<FRAMEWORK>.log_model(MODEL_OBJECT, artifact_path="PATH")`
 - ◆ `mlflow.log_artifact(LOCAL_PATH, artifact_path="PATH")`

6. Prefect Implementation Steps:

Creating and activating a Virtual Environment

- In order to install prefect, create a virtual environment:
 - ◆ `$ python -m venv virenv`
- Enter the Virtual Environment using below mentioned command:
 - ◆ `$.\virenv\Scripts\activate`

Installing Prefect 2.0

- ◆ `$ pip install prefect`
- OR if you have Prefect 1, upgrade to Prefect 2 using this command:
 - ◆ `$ pip install -U prefect`
- OR to install a specific version:
 - ◆ `$ pip install prefect==2.4`

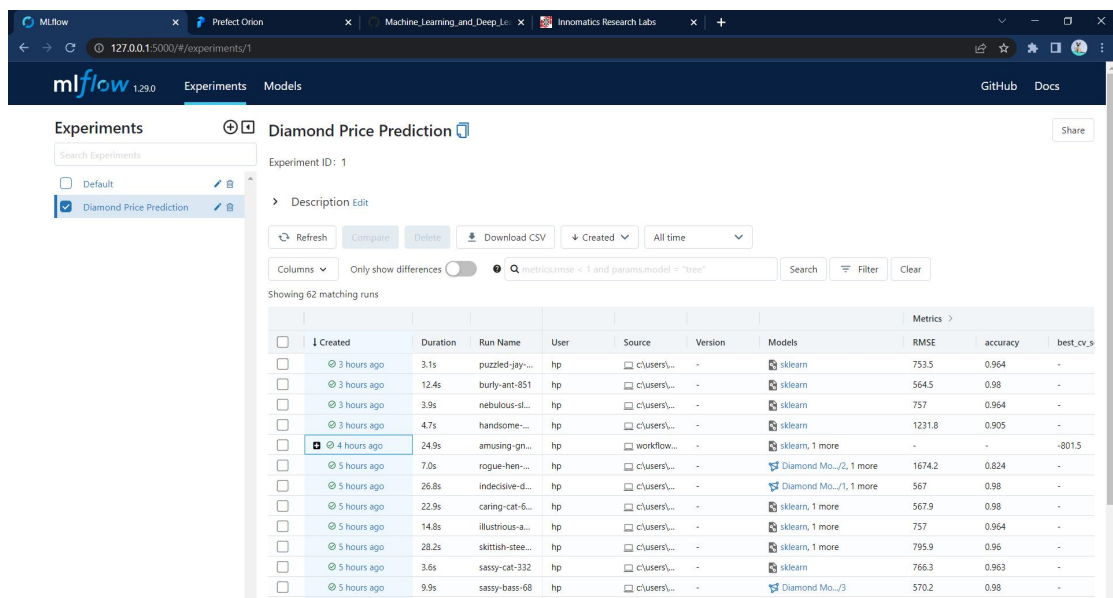
Check Prefect Version

- ◆ `$ prefect version`

Running Prefect Dashboard

- ◆ `$ prefect orion start`

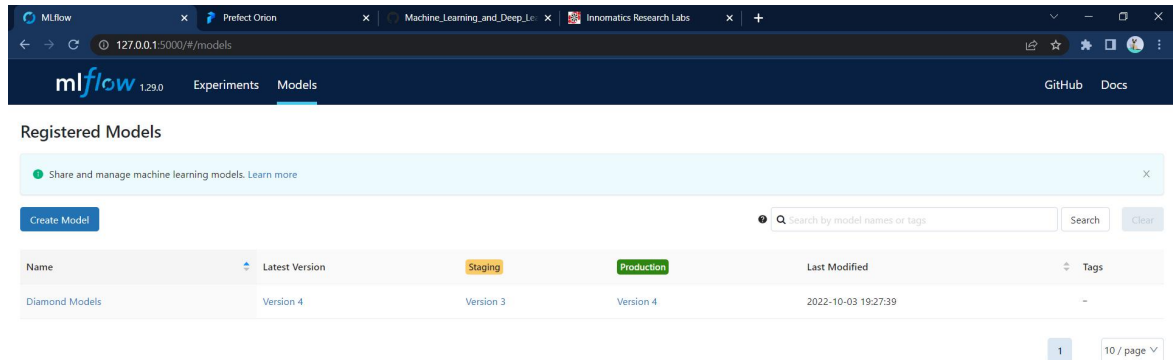
7. Implemented MLFlow Experiment Tracking:



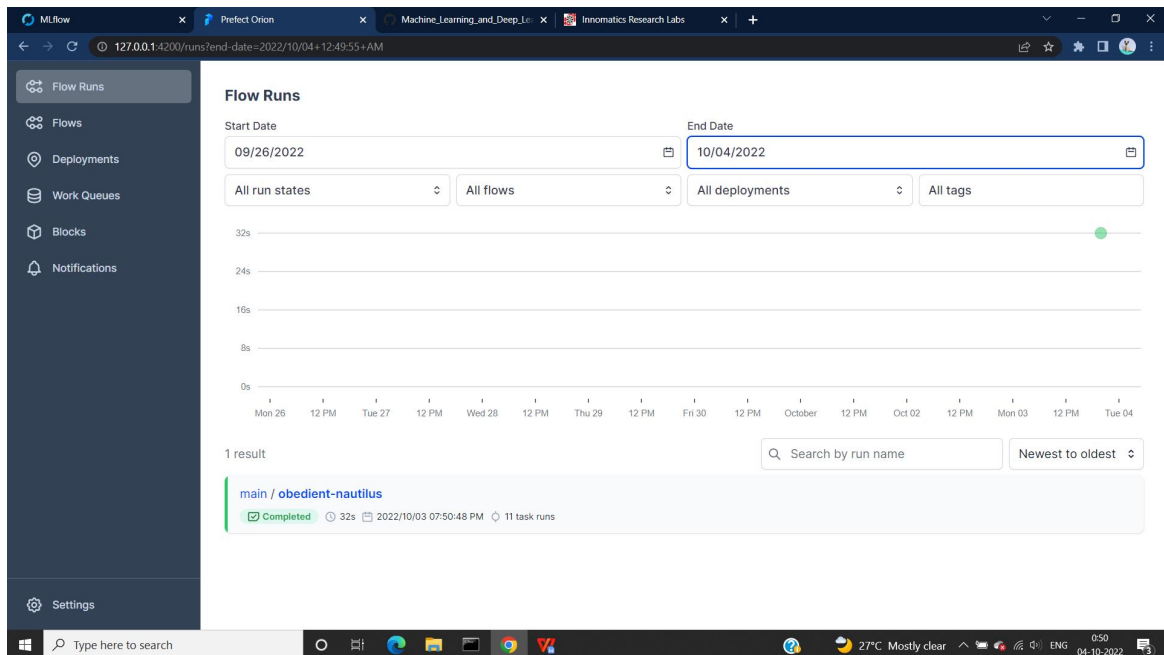
The screenshot displays the MLFlow web interface for experiment tracking. The main view is for the 'Diamond Price Prediction' experiment, showing a list of 62 runs. The interface includes a sidebar with the experiment name, a search bar, and a table of runs with columns for Created, Duration, Run Name, User, Source, Version, Models, and Metrics (RMSE, accuracy, best_cv_s). The table is sorted by Created time, showing runs from 3 hours ago to 5 hours ago. The metrics column shows RMSE values ranging from 753.5 to 766.3, accuracy values from 0.964 to 0.983, and best_cv_s values from -801.5 to -.

Created	Duration	Run Name	User	Source	Version	Models	RMSE	accuracy	best_cv_s
3 hours ago	3.1s	puzzled-jay...	hp	c:\users\...	-	sklearn	753.5	0.964	-
3 hours ago	12.4s	burly-ant-851	hp	c:\users\...	-	sklearn	564.5	0.98	-
3 hours ago	3.9s	nebulous-sl...	hp	c:\users\...	-	sklearn	757	0.964	-
3 hours ago	4.7s	handsome-...	hp	c:\users\...	-	sklearn	1231.8	0.905	-
4 hours ago	24.9s	amusing-gn...	hp	workflow...	-	sklearn, 1 more	-	-	-801.5
5 hours ago	7.0s	rogue-hen-...	hp	c:\users\...	-	Diamond Mo./2, 1 more	1674.2	0.824	-
5 hours ago	26.8s	indecisive-d...	hp	c:\users\...	-	Diamond Mo./1, 1 more	567	0.98	-
5 hours ago	22.9s	caring-cat-6...	hp	c:\users\...	-	sklearn, 1 more	567.9	0.98	-
5 hours ago	14.8s	illustrious-a...	hp	c:\users\...	-	sklearn, 1 more	757	0.964	-
5 hours ago	28.2s	skittish-ste...	hp	c:\users\...	-	sklearn, 1 more	795.9	0.96	-
5 hours ago	3.6s	sassy-cat-332	hp	c:\users\...	-	sklearn	766.3	0.963	-
5 hours ago	9.9s	sassy-bass-68	hp	c:\users\...	-	Diamond Mo./3	570.2	0.98	-

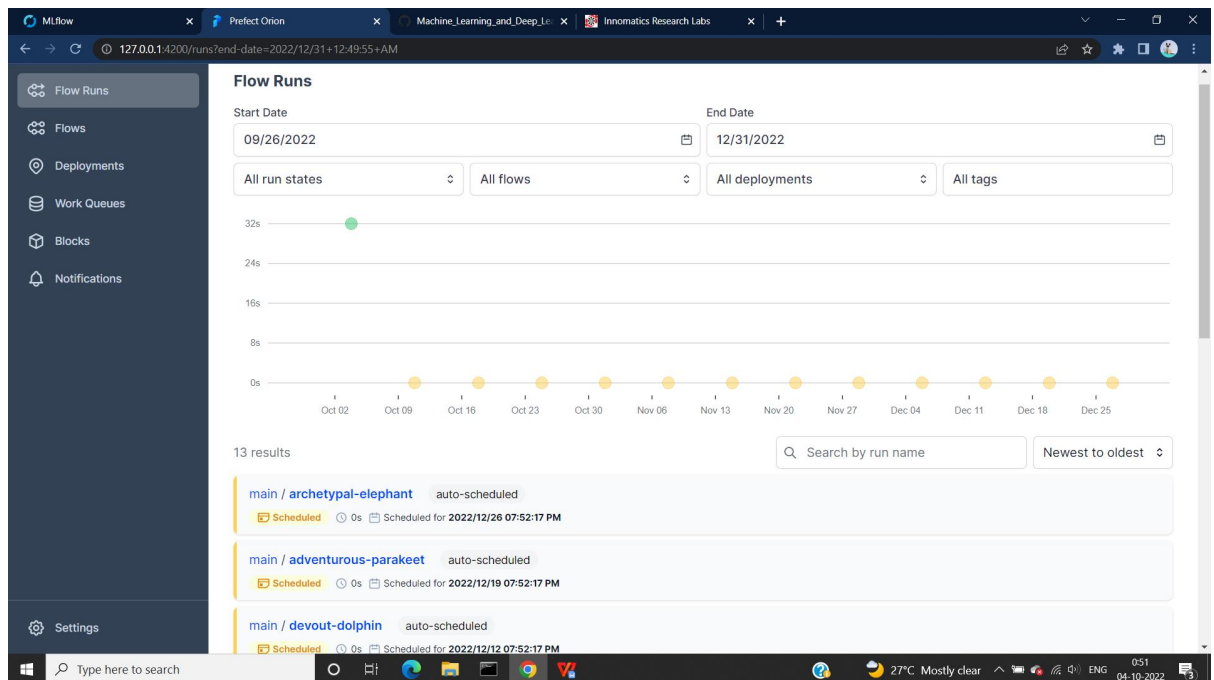
8. Implemented MLFlow Model Management:



9. Implemented Prefect (Flow Runs):



10. Implemented Prefect (Auto-Scheduling):



Github link: https://github.com/dhreeti414/Innomatics_Internship_July_2022

LinkedIn link: <https://www.linkedin.com/in/dhreetikesharwani/>