

AiEnsured

Ensuring comprehensive validation of AI systems



An article on

MLFLOW

On Fuel Efficiency Dataset

Submitted by: Rayapureddi Venkata Sri Sai Subhash

Intern in Machine Learning May 2023-July 2023

(rayapureddi.subhash@testaing.com)

MLflow is an open-source platform designed to manage the end-to-end machine learning lifecycle. It was developed by Databricks and has become widely used in the machine learning community. MLflow provides tools and libraries to track experiments, package code, share models, and manage the deployment of machine learning models.

Key components of MLflow:

- Tracking
- Projects
- Models
- Model Registry

Benefits of using MLflow:

- Reproducibility
- Model Management
- Flexibility

Overall, MLflow provides a comprehensive framework to streamline the machine learning development process, making it easier to track experiments, reproduce results, and deploy models efficiently in real-world applications.

Implementation Of MLFLOW:

First install the mlflow in your local system.

```
Anaconda Prompt
(base) C:\Users\subha>pip install mlflow
Requirement already satisfied: mlflow in c:\users\subha\anaconda3\lib\site-packages (2.4.1)
Requirement already satisfied: pyarrow<13,>=4.0.0 in c:\users\subha\anaconda3\lib\site-packages (from mlflow) (12.0.1)
Requirement already satisfied: pandas<3 in c:\users\subha\anaconda3\lib\site-packages (from mlflow) (1.5.3)
Requirement already satisfied: matplotlib<4 in c:\users\subha\anaconda3\lib\site-packages (from mlflow) (3.7.0)
Requirement already satisfied: pytz<2024 in c:\users\subha\anaconda3\lib\site-packages (from mlflow) (2022.7)
Requirement already satisfied: numpy<2 in c:\users\subha\appdata\roaming\python\python310\site-packages (from mlflow) (1.23.5)
```

After installation connect it to the backend server to get the url.

```
Anaconda Prompt - mlflow u
(base) C:\Users\subha>mlflow ui --backend-store-uri sqlite:///mlflow.db
2023/07/30 00:05:29 INFO mlflow.store.db.utils: Creating initial MLflow database tables...
2023/07/30 00:05:29 INFO mlflow.store.db.utils: Updating database tables
INFO [alembic.runtime.migration] Context impl SQLiteImpl.
INFO [alembic.runtime.migration] Will assume non-transactional DDL.
INFO [alembic.runtime.migration] Context impl SQLiteImpl.
INFO [alembic.runtime.migration] Will assume non-transactional DDL.
2023/07/30 00:05:29 INFO mlflow.store.db.utils: Creating initial MLflow database tables...
2023/07/30 00:05:29 INFO mlflow.store.db.utils: Updating database tables
INFO [alembic.runtime.migration] Context impl SQLiteImpl.
INFO [alembic.runtime.migration] Will assume non-transactional DDL.
INFO waitress: Serving on http://127.0.0.1:5000
```

Connect the model to mlflow

```
mlflow.set_tracking_uri("sqlite:///mlflow.db")
mlflow.set_experiment("fuel efficiency")
```

Log the parameters in the model with mlflow for tracking:

```
with mlflow.start_run():
    history = model.fit(X_train, y_train, batch_size=8, epochs=50, verbose=1, validation_split=0.2)
    plt.plot(history.history['loss'])
    plt.plot(history.history['val_loss'])
    plt.title('loss')
    plt.ylabel('loss')
    plt.xlabel('epoch')
    plt.legend(['train', 'test'], loc='upper left')
    plt.show()
    y_pred = model.predict(X_test)
    #the hyperparameters changed are batch_size and epochs and the values changed are (50epochs,16);(100epochs , 32);(100epochs,16)#
    from sklearn import metrics
    mae=metrics.mean_absolute_error(y_test, y_pred)
    mse=metrics.mean_squared_error(y_test, y_pred)
    smse=np.sqrt(metrics.mean_squared_error(y_test, y_pred))
    mlflow.log_param("batch_size",32)
    mlflow.log_param("epochs",100)
    mlflow.set_tag("model","regression")
    mlflow.log_metric("mae",mae)
    print("metric : mae:{mae}")
```

Change the parameters and observe the changes in the mlflow page.

The screenshot shows the MLflow web interface for an experiment named 'fuel efficiency'. The interface includes a search bar, a list of experiments, and a table of runs. The table columns are Run Name, Created, Duration, Source, and Models. There are 11 runs listed, each with a unique name and a creation time of 26 or 27 days ago. A 'Show more columns' button is visible next to the table.

Run Name	Created	Duration	Source	Models
salty-hound-704	26 days ago	10.0s	C:\Users\...	-
angry-calf-976	26 days ago	10.5s	C:\Users\...	-
kindly-swan-400	26 days ago	10.3s	C:\Users\...	-
beautiful-mink-564	26 days ago	12.9s	C:\Users\...	-
valuable-snail-956	27 days ago	12.5s	C:\Users\...	-
powerful-shark-343	27 days ago	11.9s	C:\Users\...	-
aged-crow-884	27 days ago	12.3s	C:\Users\...	-
chill-penguin-7	27 days ago	11.8s	C:\Users\...	-
funny-hound-746	27 days ago	11.8s	C:\Users\...	-
exultant-calf-438	27 days ago	12.6s	C:\Users\...	-
flawless-ox-916	27 days ago	59ms	C:\Users\...	-

11 matching runs

Browser tabs: Email Mail, mlflow implementation.ipynb, Downloads/, MLflow

URL: 127.0.0.1:5000/#/experiments/1/runs/f3aa7047115a402db43667b86e83e09d

salty-hound-704

Run ID: f3aa7047115a402db43667b86e83e09d Date: 2023-07-03 11:40:05 Source: C:\Users\subha\anaconda3\lib\site-packages\ipykernel_launcher.py User: subha

Duration: 10.0s Status: FINISHED Lifecycle Stage: active

> Description Edit



▼ Parameters (2)

Name	Value
batch_size	32
epochs	100

▼ Metrics (1)

Name	Value
mae	8.253

▼ Tags (1)

Name	Value	Actions
model	regression	 

Name: Value: Add

▼ Artifacts

We can get the above page for every run after changing the parameters.

Browser tabs: Email Mail, mlflow implementation.ipynb, Downloads/, MLflow



URL: 127.0.0.1:5000/#/experiments/1?searchFilter=&orderByKey=attributes.start_time&orderByAsc=false&startTime=ALL&lifecycleFilter=Active&modelVersionFilter=All%20Runs&selectedColumn...



mlflow 2.4.1 Experiments Models

GitHub Docs

Experiments

Search Experiments

☐ Default  

☒ fuel efficiency  

fuel efficiency [Provide Feedback](#)

Share

Experiment ID: 1 Artifact Location: file:///C:/Users/subha/mlruns/1

> Description Edit

Table view **Chart view** Artifact view Time created State Active Refresh

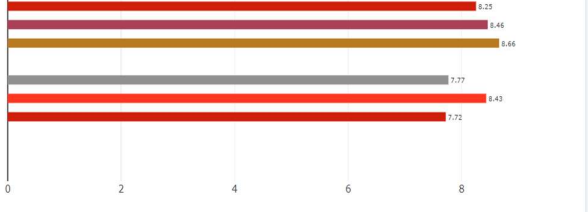
Sort: Created

Run Name
● salty-hound-704
● angry-calf-976
● kindly-swan-400
● beautiful-mink-564
● valuable-snail-956
● powerful-shark-343
● aged-crow-884
● chill-penguin-7
● funny-hound-746
● exultant-calf-438
● flawless-ox-916

11 matching runs

mae

Comparing first 10 runs



Run Name	mae
salty-hound-704	8.253
angry-calf-976	8.46
kindly-swan-400	8.66
beautiful-mink-564	7.77
valuable-snail-956	8.43
powerful-shark-343	7.72
aged-crow-884	7.72
chill-penguin-7	7.72
funny-hound-746	7.72
exultant-calf-438	7.72
flawless-ox-916	7.72

mae

Comparison first 10 runs

After the some runs of the model, we can easily track the parameters which are need to be change for the improvement of the model.

THE END