

Peer Assignment

Sheikh Muhammad Sabih (2303.KHI.DEG.010)

M Humza Moeen (2303.KHI.DEG.019)

Assignment (3.4)

Write a component that will log metadata of your Classification model that you trained on the day dedicated to Supervised Learning. Remember to include all metadata that are important to track for this problem.

Solution

1. Setting up MLFlow tracking server

We also specify artifact root and backend store URI. This makes it possible to store models.

After running this command tracking server will be accessible at `localhost:5006`

```
%%bash --bg
```

```
mlflow server --host 0.0.0.0 \  
--port 5006 \  
--backend-store-uri sqlite:///mlflow.db \  
--default-artifact-root ./mlruns
```

[7] ✓ 0.0s

Sheikh Muhammad Sabih (2303.KHI.DEG.010)
M Humza Moeen (2303.KHI.DEG.019)

2. Now make MLProject file.

```
%cat MLproject
✓ 0.1s

name: basic_mlflow

# this file is used to configure Python package dependencies.
# it uses Anaconda, but it can be also alternatively configured to use pip.
conda_env: conda.yml

# entry points can be ran using `mlflow run <project_name> -e <entry_point_name>
entry_points:
  # download_data:
    # you can run any command using MLFlow
    # command: "bash download_data.sh"
  # MLproject file has to have main entry_point. It can be toggled without using -e option.
main:
  # parameters is a key-value collection.
  parameters:
    file_name:
      type: str
      default: "day.csv"
    max_k:
      type: int
      default: 10
  command: "python train.py"
```

3. Now we are **Train** the model.

```
%%bash
source mlflow_env_vars.sh
mlflow run .
```

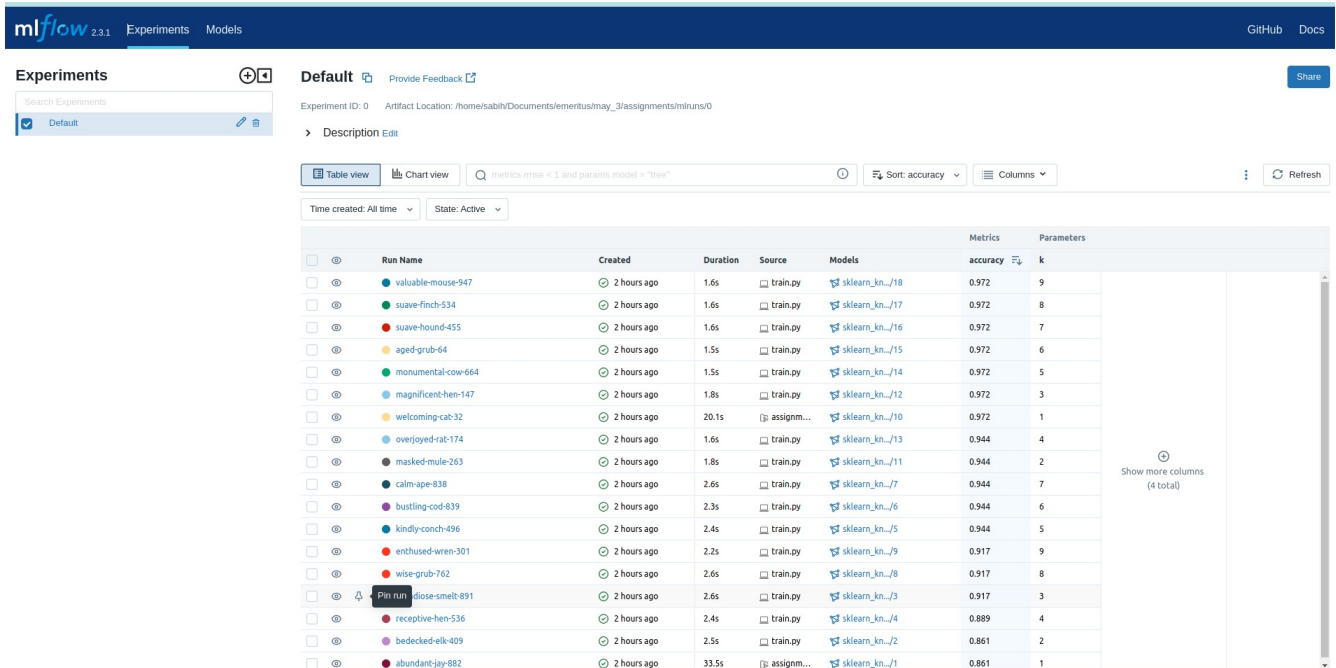
[12] ✓ 21.4s

```
... 2023/05/08 08:43:34 INFO mlflow.utilsconda: Conda environment mlflow-dd0fbdd40ba98798131458f29496394bd1a3fb
2023/05/08 08:43:34 INFO mlflow.projects.utils: == Created directory /tmp/tmpkg6uz9zc for downloading remot
2023/05/08 08:43:34 INFO mlflow.projects.backend.local: == Running command 'source /home/sabih/anaconda3/bi
{'data': array([[1.423e+01, 1.710e+00, 2.430e+00, ..., 1.040e+00, 3.920e+00,
    1.065e+03],
 [1.320e+01, 1.780e+00, 2.140e+00, ..., 1.050e+00, 3.400e+00,
    1.050e+03],
 [1.316e+01, 2.360e+00, 2.670e+00, ..., 1.030e+00, 3.170e+00,
    1.185e+03],
 ...,
 [1.327e+01, 4.280e+00, 2.260e+00, ..., 5.900e-01, 1.560e+00,
    8.350e+02],
 [1.317e+01, 2.590e+00, 2.370e+00, ..., 6.000e-01, 1.620e+00,
    8.400e+02],
 [1.413e+01, 4.100e+00, 2.740e+00, ..., 6.100e-01, 1.600e+00,
```

Peer Assignment

Sheikh Muhammad Sabih (2303.KHI.DEG.010)

M Humza Moeen (2303.KHI.DEG.019)



The screenshot shows the mlflow Experiments page. At the top, there's a navigation bar with 'mlflow 2.3.1', 'Experiments', and 'Models'. Below this, the 'Experiments' section is active, showing a search bar and a 'Default' tab. The main content area displays a table of runs. The table has columns for 'Run Name', 'Created', 'Duration', 'Source', 'Models', 'Metrics', and 'Parameters'. The 'Metrics' column shows 'accuracy' and 'k'. The 'Parameters' column shows 'k'. The table lists 20 runs, each with a unique name and a corresponding accuracy and k value. The runs are sorted by accuracy in descending order. The last run in the list is 'abundant-jay-882' with an accuracy of 0.861 and k=1.

Run Name	Created	Duration	Source	Models	Metrics	Parameters
valuable-mouse-947	2 hours ago	1.6s	train.py	sklearn_kn../18	0.972	9
suave-finch-534	2 hours ago	1.6s	train.py	sklearn_kn../17	0.972	8
suave-hound-455	2 hours ago	1.6s	train.py	sklearn_kn../16	0.972	7
aged-grub-64	2 hours ago	1.5s	train.py	sklearn_kn../15	0.972	6
monumental-cow-664	2 hours ago	1.5s	train.py	sklearn_kn../14	0.972	5
magnificent-hen-147	2 hours ago	1.8s	train.py	sklearn_kn../12	0.972	3
welcoming-cat-32	2 hours ago	20.1s	assignment...	sklearn_kn../10	0.972	1
overjoyed-rat-174	2 hours ago	1.6s	train.py	sklearn_kn../13	0.944	4
masked-mule-263	2 hours ago	1.8s	train.py	sklearn_kn../11	0.944	2
calm-ape-838	2 hours ago	2.6s	train.py	sklearn_kn../7	0.944	7
bustling-cod-839	2 hours ago	2.3s	train.py	sklearn_kn../6	0.944	6
kindly-conch-496	2 hours ago	2.4s	train.py	sklearn_kn../5	0.944	5
enthusied-wren-301	2 hours ago	2.2s	train.py	sklearn_kn../9	0.917	9
wise-grub-762	2 hours ago	2.6s	train.py	sklearn_kn../8	0.917	8
disse-smelt-891	2 hours ago	2.6s	train.py	sklearn_kn../3	0.917	3
receptive-hen-536	2 hours ago	2.4s	train.py	sklearn_kn../4	0.889	4
bedecked-elk-409	2 hours ago	2.5s	train.py	sklearn_kn../2	0.861	2
abundant-jay-882	2 hours ago	33.5s	assignment...	sklearn_kn../1	0.861	1

4. Inspecting stored models The trained models are stored in `mlruns/0`.

```
%%bash
last_model_path=$(ls -tr mlruns/0/ | tail -1)
cat mlruns/0/$last_model_path/artifacts/knn/MLmodel

[13] ✓ 0.0s

... artifact_path: knn
flavors:
python_function:
env:
conda: conda.yaml
virtualenv: python_env.yaml
loader_module: mlflow.sklearn
model_path: model.pkl
predict_fn: predict
python_version: 3.11.3
sklearn:
code: null
pickled_model: model.pkl
serialization_format: cloudpickle
sklearn_version: 1.2.2
mlflow_version: 2.3.1
model_uuid: 236205d23ce845649b8d7d3079bad324
run_id: 8e509c2c290c426f9b293de1005ee636
utc_time_created: '2023-05-08 03:43:50.294035'
```

Peer Assignment

Sheikh Muhammad Sabih (2303.KHI.DEG.010)

M Humza Moeen (2303.KHI.DEG.019)

5. Now we have to **serve** our model.

```
%%bash --bg
source mlflow_env_vars.sh
mlflow --version
mlflow models serve -m models:/sklearn_knn/Production -p 5007 --env-manager=conda
```

[18] ✓ 0.0s

6. Now to check our model we have to **predict** it.

```
%%bash
data='[[14.23,1.71,2.43,15.6,127.0,2.80,3.06,0.28,2.29,5.64,1.04,3.92,1065.0]]'
echo $data

curl -d "{\"inputs\": $data}" -H 'Content-Type: application/json' 127.0.0.1:5007/invocations
```

✓ 0.0s

```
[[14.23,1.71,2.43,15.6,127.0,2.80,3.06,0.28,2.29,5.64,1.04,3.92,1065.0]]
% Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
           Dload  Upload  Total      Spent    Left     Speed
100   104  100    20  100    84   4960   20833  --:--:--  --:--:--  --:--:-- 34666
{"predictions": [0]}
```

```
%%bash
data='[[13.05,1.73,2.04,12.4,92.0,2.72,3.27,0.17,2.91,7.2,1.12,2.91,1150.0]]'
echo $data
```

```
curl -d "{\"instances\": $data}" -H 'Content-Type: application/json' 127.0.0.1:5007/invocations
```

✓ 0.0s

```
[[14.23,1.71,2.43,15.6,127.0,2.80,3.06,0.28,2.29,5.64,1.04,3.92,1065.0]]
% Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
           Dload  Upload  Total      Spent    Left     Speed
100   107  100    20  100    87   4956   21561  --:--:--  --:--:--  --:--:-- 35666
{"predictions": [0]}
```

Peer Assignment

Sheikh Muhammad Sabih (2303.KHI.DEG.010)

M Humza Moeen (2303.KHI.DEG.019)

```
%%bash
data='[[ 12.93, 2.81, 2.7 , 21.0 , 96.0 , 1.54, 0.5 , 0.53, 0.75, 4.6 , 0.77, 2.31,600.0]]'
echo $data

curl -d '{"instances\: $data}" -H 'Content-Type: application/json' 127.0.0.1:5007/invocations
✓ 0.0s

[[ 12.93, 2.81, 2.7 , 21.0 , 96.0 , 1.54, 0.5 , 0.53, 0.75, 4.6 , 0.77, 2.31,600.0]]
% Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
                                 Dload  Upload   Total   Spent    Left   Speed
100  139  100    20  100   119    8054  47925  --:--:-- --:--:-- --:--:-- 69500
{"predictions": [2]}
```