

Assignment 2

▼ Model Evaluation Suite

Model Evaluation Suite

The suite is composed of the following checks: Train Test Performance, Model Inference Time.

Each check may contain conditions (which will result in pass ✓ / fail ✖ / warning ! / error ?!) as well as other outputs such as plots or tables.

Suites, checks and conditions can all be modified. Read more about [custom suites](#).

▼ Didn't Pass

Status	Check	Condition	More Info
✖	Train Test Performance	Train-Test scores relative degradation is less than 0.2	2 scores failed. Found max degradation of 47.8% for metric Neg MAE

Train Test Performance

Summarize given model performance on the train and test datasets based on selected scorers. [Read More...](#)

Conditions Summary

Status	Condition	More Info
✖	Train-Test scores relative degradation is less than 0.2	2 scores failed. Found max degradation of 47.8% for metric Neg MAE

Additional Outputs

Additional Outputs

Neg RMSE

Dataset	Neg RMSE
Train	-52
Test	-72

Neg MAE

Dataset	Neg MAE
Train	-38
Test	-52

R2

Dataset	R2
Train	0.9
Test	0.8

Dataset

Train

Test

[Go to top](#)

▼ Passed

Status	Check	Condition	More Info
✓	Model Inference Time - Train Dataset	Average model inference time for one sample is less than 0.1	Found average inference time (seconds): 5.762e-05
✓	Model Inference Time - Test Dataset	Average model inference time for one sample is less than 0.1	Found average inference time (seconds): 2.01e-05

Model Inference Time - Train Dataset

Measure model average inference time (in seconds) per sample. [Read More...](#)

Conditions Summary

Status	Condition	More Info
✓	Average model inference time for one sample is less than 0.1	Found average inference time (seconds): 5.762e-05

Additional Outputs

Average model inference time for one sample (in seconds): 5.762e-05

Note - data sampling: Data is sampled from the original dataset, running on 1000 samples out of 10718. Sample size can be controlled with the "n_samples" parameter.

[Go to top](#)

Model Inference Time - Test Dataset

Model Inference Time - Test Dataset

Measure model average inference time (in seconds) per sample. [Read More...](#)

Conditions Summary

Status	Condition	More Info
✓	Average model inference time for one sample is less than 0.1	Found average inference time (seconds): 2.01e-05

Additional Outputs

Average model inference time for one sample (in seconds): 2.01e-05

[Go to top](#)

Assignment 3

mlflow2.9.2ExperimentsModels

Experiments

Search Experiments

☐ Default

☐ Kubeflow Pipeline test r...

☒ week1-lgbm-bike-dem...

week1-lgbm-bike-demand

Provide Feedback

Share

Experiment ID: 4Artifact Location: s3://mlflow/4

Description Edit

metrics.rmse < 1 and params.model = "tree"

Time created

State: Active

Sort: Created

Columns

Table

Chart

Evaluation

Experimental

Run Name	Created	Duration	Source	Models
puzzled-flea-645	2 minutes ago	1.5min	ipykerne...	lightgbm

mlflow2.9.2ExperimentsModels

week1-lgbm-bike-demand

puzzled-flea-645

Run ID: 2bca3afc91714024afa76fcb9c88995bDate: 2024-11-02 10:21:17Source: ipykernel_launcher.pyUser: RuperlDuration: 1.5minStatus: FINISHEDLifecycle Stage: active

Description Edit

Datasets

Parameters (3)

Metrics

Tags

Artifacts

Full Path:s3://mlflow/4/2bca3afc91714024afa76fcb9c88995b/artifacts/evaluation_result.htmlSize: 7.36MB

Model Evaluation Suite

The suite is composed of the following checks: Train Test Performance, Model Inference Time. Each check may contain conditions (which will result in pass / fail / warning / error ?) as well as other outputs such as plots or tables. Suites, checks and conditions can all be modified. Read more about custom suites.

Didn't Pass

Status	Check	Condition	More Info
X	Train Test Performance	Train-Test scores relative degradation is less than 0.2 2 scores failed. Found max degradation of 47.8% for metric Neg MAE	

Full Path:s3://mlflow/4/2bca3afc91714024afa76fcb9c88995b/artifacts/evaluation_result.htmlSize: 7.36MB

Additional Outputs

Neg RMSE

Neg MAE

R2

Train

Test

0

-10

-20

-30

-40

-50

-60

-70

0

-10

-20

-30

-40

-50

0

0.2

0.4

0.6

0.8

Dataset

Train

Test

[Go to top](#)

- > Parameters (3)
- > Metrics
- > Tags
- > Artifacts

Week1LgbmBikeDemand

- ML_model
- conda.yaml
- model.pkl
- python_env.yaml
- requirements.txt
- evaluation_result.html

Full Path:s3://mlflow/4/2bca3af-c91714024afa76fcb9c88995b/artifacts/evaluation_result.html
Size: 7.36MB

Passed

Status	Check	Condition	More Info
✓	Model Inference Time - Train Dataset	Average model inference time for one sample is less than 0.1	Found average inference time (seconds): 4.447e-05
✓	Model Inference Time - Test Dataset	Average model inference time for one sample is less than 0.1	Found average inference time (seconds): 3.018e-05

Model Inference Time - Train Dataset

Measure model average inference time (in seconds) per sample. [Read More...](#)

Conditions Summary

Status	Condition	More Info
✓	Average model inference time for one sample is less than 0.1	Found average inference time (seconds): 4.447e-05

Additional Outputs

Average model inference time for one sample (in seconds): 4.447e-05

Note - data sampling: Data is sampled from the original dataset, running on 1000 samples out of 10718. Sample size can be controlled with the "n_samples" parameter.

[Go to top](#)

mlflow 2.9.2

ExperimentsModels

GitHubDocs

Registered Models >

Week1LgbmBikeDemand

Created Time: 2024-11-02 11:18:50Last Modified: 2024-11-02 11:18:50

> Description [Edit](#)

> Tags

> Versions

AllActive 0Compare

New model registry UI ☐

Version	Registered at	Created by	Stage	Description
Version 1	2024-11-02 11:18:50		None	

1

Assignment 4

week1-igbm-bike-demand

sedate-sponge-781

Run ID: 8a266fdee8424d54b6da4881273162bf

Date: 2024-11-06 14:06:54

Source: `ipykernel_launcher.py`

User: ilupeil

Duration: 24.9min

Status: FINISHED

Lifecycle Stage: [active](#)

- Datasets

- Parameters (3)

- Metrics

› [Tags](#)

- ▼ Artifacts

Full Path:s3://mlflow/45a266f5ee8424dd4b9da4681273162M/artifacts/model_comparison.html
Size: 7.32MB

Multi Model Performance Report

Summarize performance scores for multiple models on test datasets. [Read More...](#)

Additional Outputs

