Not secure | notebooksb.jarvislabs.ai:10977/#/experiments/1

**ml*flow*** 1.26.1

Experiments    Models

GitHub    Docs

## Experiments  + <

Search Experiments

Default  ✎ 🗑

Baseline_model_exp01 ✎ 🗑

Baseline_model_exp03 ✎ 🗑

Lead_scoring_mlflow... ✎ 🗑

## Baseline_model_exp01 ⧉

# MLFLOW Experiment ID-1

Share

ℹ Track machine learning training runs in experiments. Learn more    ✕

Experiment ID: 1

▸ Description  Edit

↻ Refresh | Compare | Delete | Download CSV ⬇ | ↓ Start Time ⌄ | All time ⌄

☰ ⊞ | ⚙Columns | Only show differences ⬤ ? | 🔍 metrics.rmse < 1 and params.model = "tree" | Search | ☰ Filter | Clear

Showing 11 matching runs

| | | | | | | Metrics › | | | Parameters › | |
|---|---|---|---|---|---|---|---|---|---|---|
| ↓ Start Time | Duration | Run Name | User | Source | Version | Models | AUC | Accuracy | F1 | C | CPU Jobs |
| ⊟ 13 hours ago | | Session Initi... | root | 🖥 ipykernel... | - | - | - | - | - | - | -1 |
| 13 hours ago | | Light Gradie... | root | 🖥 ipykernel... | - | 📄 sklearn | 0.821 | 0.739 | 0.762 | - | - |
| 13 hours ago | | Naive Bayes | root | 🖥 ipykernel... | - | 📄 sklearn | 0.734 | 0.663 | 0.727 | - | - |
| 13 hours ago | | Linear Discri... | root | 🖥 ipykernel... | - | 📄 sklearn | 0.773 | 0.701 | 0.728 | - | - |
| 13 hours ago | | Ridge Classif... | root | 🖥 ipykernel... | - | 📄 sklearn | 0 | 0.701 | 0.728 | - | - |
| 13 hours ago | | Logistic Reg... | root | 🖥 ipykernel... | - | 📄 sklearn | 0.784 | 0.71 | 0.74 | 1.0 | - |
| 13 hours ago | | Random For... | root | 🖥 ipykernel... | - | 📄 sklearn | 0.817 | 0.735 | 0.761 | - | - |
| 13 hours ago | | Decision Tre... | root | 🖥 ipykernel... | - | 📄 sklearn | 0.817 | 0.736 | 0.758 | - | - |
| 13 hours ago | | Extra Trees C... | root | 🖥 ipykernel... | - | 📄 sklearn | 0.818 | 0.737 | 0.758 | - | - |
| 13 hours ago | | Light Gradie... | root | 🖥 ipykernel... | - | 📄 sklearn | 0.821 | 0.739 | 0.762 | - | - |
| 13 hours ago | | Extreme Gra... | root | 🖥 ipykernel... | - | - | - | - | - | - | - |

29°C Haze    ENG US    13:58 02-11-2022

Not secure | notebooksb.jarvislabs.ai:10977/#/experiments/2

mlflow 1.26.1    Experiments    Models                                    GitHub    Docs

## Experiments

Search Experiments

Default

Baseline_model_exp01

Baseline_model_exp03

Lead_scoring_mlflow...

## Baseline_model_exp03

**MLFLOW Experiment ID - 2**

Share

ⓘ Track machine learning training runs in experiments. Learn more    ✕

Experiment ID: 2

▶ Description    Edit

⟳ Refresh | Compare | Delete | ⬇ Download CSV | ⬇ Start Time ∨ | All time ∨

☰ ⊞ | ⚙ Columns | Only show differences ⬤ | ⓘ | 🔍 metrics.rmse < 1 and params.model = "tree" | Search | ⚙ Filter | Clear

Showing 12 matching runs

| ☐ | ↓ Start Time | Duration | Run Name | User | Source | Version | Models | Metrics | | | Parameters | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | AUC | Accuracy | F1 | C | CPU Jobs |
| ☐ | ⊟ 13 hours ago | | Session Initi... | root | 💻 ipykernel... | - | - | - | - | - | - | -1 |
| ☐ | 13 hours ago | | Light Gradie... | root | 💻 ipykernel... | - | 📄 sklearn | 0.82 | 0.738 | 0.761 | - | - |
| ☐ | 13 hours ago | | Light Gradie... | root | 💻 ipykernel... | - | 📄 sklearn | 0.821 | 0.739 | 0.763 | - | - |
| ☐ | 13 hours ago | | Naive Bayes | root | 💻 ipykernel... | - | 📄 sklearn | 0.734 | 0.67 | 0.723 | - | - |
| ☐ | 13 hours ago | | Linear Discri... | root | 💻 ipykernel... | - | 📄 sklearn | 0.773 | 0.7 | 0.728 | - | - |
| ☐ | 13 hours ago | | Ridge Classif... | root | 💻 ipykernel... | - | 📄 sklearn | 0 | 0.7 | 0.728 | - | - |
| ☐ | 13 hours ago | | Logistic Reg... | root | 💻 ipykernel... | - | 📄 sklearn | 0.784 | 0.71 | 0.74 | 1.0 | - |
| ☐ | 13 hours ago | | Random For... | root | 💻 ipykernel... | - | 📄 sklearn | 0.818 | 0.735 | 0.762 | - | - |
| ☐ | 13 hours ago | | Decision Tre... | root | 💻 ipykernel... | - | 📄 sklearn | 0.817 | 0.736 | 0.758 | - | - |
| ☐ | 13 hours ago | | Extra Trees C... | root | 💻 ipykernel... | - | 📄 sklearn | 0.817 | 0.737 | 0.758 | - | - |
| ☐ | 13 hours ago | | Extreme Gra... | root | 💻 ipykernel... | - | 📄 sklearn | 0.821 | 0.738 | 0.761 | - | - |
| ☐ | 13 hours ago | | Light Gradie... | root | 💻 ipykernel... | - | 📄 sklearn | 0.821 | 0.739 | 0.763 | - | - |

mlflow 1.26.1    **Experiments**    **Models**

GitHub    Docs

Lead_scoring_mlflow_production > Lead_scoring_mlflow_production

# Lead_scoring_mlflow_production

Date: 2022-11-02 13:54:20

Source: 🖥 airflow

User: root

Duration: 9.9s

Status: FINISHED

Lifecycle Stage: active

▸ Description    Edit

▸ Parameters (21)

▸ Metrics (9)

▸ Tags

▾ Artifacts

```
▼ 📁 models
    📄 MLmodel
    📄 conda.yaml
    📄 model.pkl
    📄 python_env.yaml
    📄 requirements.txt
```

Full Path:/home/Assignment/mlruns/3/a91b626737714b87ae3dfc6d5d5ce4df/artifacts/models ⧉

⊘ LightGBM, v3 ☑
Registered on 2022/11/02

## MLflow Model

The code snippets below demonstrate how to make predictions using the logged model. This model is also registered to the model registry.

### Model schema

Input and output schema for your model. Learn more

| Name | Type |
|------|------|

No schema. See MLflow docs for how to include input and output schema with your model.

### Make Predictions

Predict on a Spark DataFrame:

```
import mlflow
logged_model = 'runs:/a91b626737714b87ae3dfc6d5d5ce4df/models'

# Load model as a Spark UDF. Override result_type if the model does not return double values.
loaded_model = mlflow.pyfunc.spark_udf(spark, model_uri=logged_model, result_type='double')
```

29°C
Haze

ENG
US

13:59
02-11-2022

# Parameters

| Name | Value |
| --- | --- |
| boosting_type | gbdt |
| class_weight | None |
| colsample_bytree | 1.0 |
| device | gpu |
| importance_type | split |
| learning_rate | 0.1 |
| max_depth | -1 |
| min_child_samples | 20 |
| min_child_weight | 0.001 |
| min_split_gain | 0.0 |
| n_estimators | 100 |
| n_jobs | -1 |
| num_leaves | 31 |
| objective | None |
| random_state | 100 |
| reg_alpha | 0.0 |

| reg_alpha | 0.0 |
| reg_lambda | 0.0 |
| silent | warn |
| subsample | 1.0 |
| subsample_for_bin | 200000 |
| subsample_freq | 0 |

▸ Metrics (9)

▸ Tags

▾ Artifacts

▾ 📁 models
  📄 MLmodel
  📄 conda.yaml
  📄 model.pkl
  📄 python_env.yaml
  📄 requirements.txt

Full Path:/home/Assignment/mlruns/3/a91b626737714b87ae3dfc6d5d5ce4df/artifacts/models 📋

⊘ LightGBM, v3 ↗
Registered on 2022/11/02

## MLflow Model

The code snippets below demonstrate how to make predictions using the logged model. This model is also registered to the model registry.

### Model schema

Input and output schema for your model. Learn more

| Name | Type |
| --- | --- |

No schema. See MLflow docs for how to include input and output schema with your model.

### Make Predictions

Predict on a Spark DataFrame:

```
import mlflow
logged_model = 'runs:/a91b626737714b87ae3dfc6d5d5ce4df/models'

# Load model as a Spark UDF. Override result_type if the model does not return double values.
loaded_model = mlflow.pyfunc.spark_udf(spark, model_uri=logged_model, result_type='double')

# Predict on a Spark DataFrame.
```

# Metrics

## Metrics (9)

| Name | Value |
|------|-------|
| AUC | 0.751 |
| False Negative | 3891 |
| False Positive | 8459 |
| Precision | 0.741 |
| Recall | 0.751 |
| True Negative | 15176 |
| True Positive | 20267 |
| f1 | 0.739 |
| test_accuracy | 0.742 |

## Tags

## Artifacts

- models
  - MLmodel
  - conda.yaml
  - model.pkl
  - python_env.yaml
  - requirements.txt

Full Path:/home/Assignment/mlruns/3/a91b626737714b87ae3dfc6d5d5ce4df/artifacts/models

⊘ LightGBM, v3
Registered on 2022/11/02

### MLflow Model

The code snippets below demonstrate how to make predictions using the logged model. This model is also registered to the model registry.

Tags

Artifacts

# Artifacts

- models
  - MLmodel
  - conda.yaml
  - model.pkl
  - python_env.yaml
  - requirements.txt

Full Path:/home/Assignment/mlruns/3/a91b626737714b87ae3dfc6d5d5ce4df/artifacts/models/MLmodel

Size: 434B

```
artifact_path: models
flavors:
  python_function:
    env: conda.yaml
    loader_module: mlflow.sklearn
    model_path: model.pkl
    python_version: 3.8.12
  sklearn:
    code: null
    pickled_model: model.pkl
    serialization_format: cloudpickle
    sklearn_version: 0.23.2
mlflow_version: 1.26.1
model_uuid: 4265ef096530470c97e55d728fd580f3
run_id: a91b626737714b87ae3dfc6d5d5ce4df
utc_time_created: '2022-11-02 08:24:26.403794'
```

Not secure | notebooksb.jarvislabs.ai:10977/#/models/LightGBM/versions/3

ml*flow* 1.26.1    Experiments    Models                                            GitHub    Docs

Registered Models > LightGBM > Version 3

# Version 3

Registered At: 2022-11-02 13:54:29                    Stage: **Production** ⌄                    Last Modified: 2022-11-02 13:56:13

Source Run: Lead_scoring_mlflow_production

▸ Description    Edit

▸ Tags

▾ Schema

| Name | Type |
| --- | --- |

No schema. See MLflow docs for how to include input and output schema with your model.
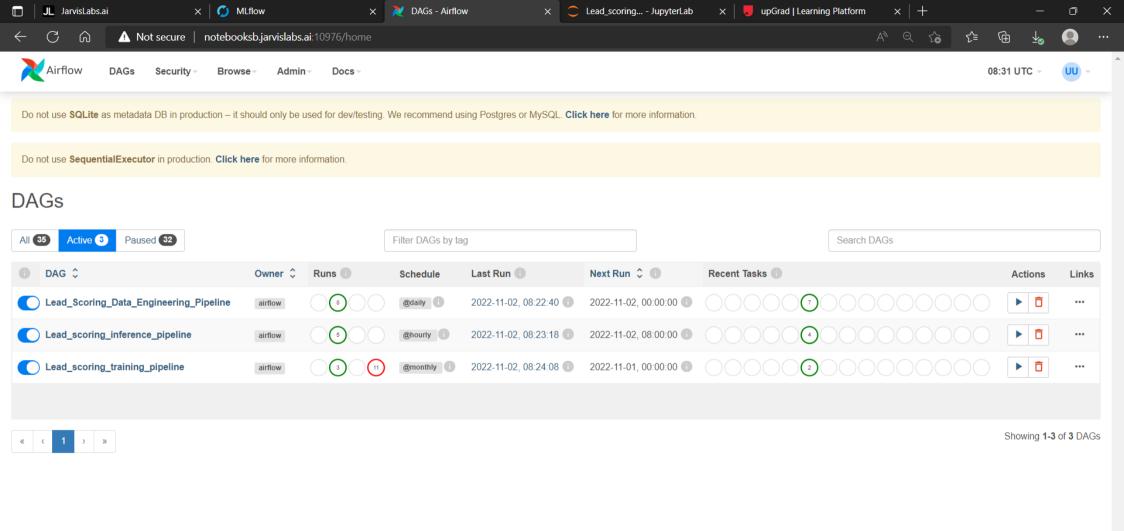
Not secure | notebooksb.jarvislabs.ai:10977/#/models/LightGBM

**ml*flow*** 1.26.1    Experiments    **Models**        GitHub    Docs

Registered Models › LightGBM

# LightGBM

Created Time: 2022-11-02 01:45:59      Last Modified: 2022-11-02 13:56:13

▸ Description   Edit

▸ Tags

▾ Versions   [ All | Active 1 ]   Compare

| | Version | Registered at ▾ | Created by | Stage | Description |
|---|---|---|---|---|---|
| ☐ ⊘ | Version 3 | 2022-11-02 13:54:29 | | **Production** | |
| ☐ ⊘ | Version 2 | 2022-11-02 13:51:32 | | None | |
| ☐ ⊘ | Version 1 | 2022-11-02 01:45:59 | | **Archived** | |

‹   **1**   ›

Not secure | notebooksb.jarvislabs.ai:10976/dags/Lead_Scoring_Data_Engineering_Pipeline/grid

## Airflow

DAGs    Security ▾    Browse ▾    Admin ▾    Docs ▾

08:31 UTC ▾    UU ▾

**DAG: Lead_Scoring_Data_Engineering_Pipeline** DAG to run data pipeline for lead scoring

Schedule: @daily ⓘ    Next Run: 2022-11-02, 00:00:00

⊞ Grid    ᵗᵗ Graph    📅 Calendar    ⌛ Task Duration    ⇄ Task Tries    ⤓ Landing Times    ☰ Gantt    ⚠ Details    <> Code    🔍 Audit Log    ▶    🗑

| 02-11-2022 08:31:55 📅 | 25 ▾ | All Run Types ▾ | All Run States ▾ | **Clear Filters** |

deferred | failed | queued | running | scheduled | skipped | success | up_for_reschedule | up_for_retry | upstream_failed | no_status

Auto-refresh ⬤    →≡

Duration    Oct 31, 00:00

00:02:13

00:01:06

00:00:00

building_db
checking_raw_data_schema
loading_data
mapping_city_tier
mapping_categorical_vars
mapping_interactions
checking_model_inputs_schema

**DAG**
**Lead_Scoring_Data_Engineering_Pipeline**

**DAG Details**

| **DAG Runs Summary** | |
|---|---|
| Total Runs Displayed | 8 |
| ■ Total success | 8 |
| First Run Start | 2022-10-25, 17:34:37 UTC |
| Last Run Start | 2022-11-02, 08:22:41 UTC |
| Max Run Duration | 00:02:13 |
| Mean Run Duration | 00:00:42 |
| Min Run Duration | 00:00:14 |
| **DAG Summary** | |
| Total Tasks | 7 |
| PythonOperators | 7 |

29°C
Haze

ENG
US

14:01
02-11-2022

Not secure | notebooksb.jarvislabs.ai:10976/dags/Lead_Scoring_Data_Engineering_Pipeline/graph

**Airflow**    DAGs    Security ▾    Browse ▾    Admin ▾    Docs ▾

08:32 UTC ▾    UU ▾

⬤ **DAG: Lead_Scoring_Data_Engineering_Pipeline** DAG to run data pipeline for lead scoring

`success`    Schedule: @daily ⓘ    Next Run: 2022-11-02, 00:00:00

▦ Grid    �ణ Graph    📅 Calendar    ⌛ Task Duration    ⇄ Task Tries    �== Landing Times    ☰ Gantt    ⚠ Details    <> Code    🔍 Audit Log    ▶    🗑

📅 2022-11-02T08:22:41Z    Runs  25 ▾    Run  manual__2022-11-02T08:22:40.171118+00:00 ▾    Layout  Left > Right ▾    Update    Find Task...

PythonOperator

deferred | failed | queued | running | scheduled | skipped | success | up_for_reschedule | up_for_retry | upstream_failed | no_status

◯ Auto-refresh    ↻

building_db → checking_raw_data_schema → loading_data → mapping_city_tier → mapping_categorical_vars → mapping_interactions → checking_model_inputs_schema

29°C Haze    🪟 🔍 🗔 🎦 🌐 🗂 ✉ 🖼    ENG US    14:02 02-11-2022

Not secure | notebooksb.jarvislabs.ai:10976/dags/Lead_scoring_training_pipeline/grid

**Airflow**  DAGs  Security ⌄  Browse ⌄  Admin ⌄  Docs ⌄    08:32 UTC ⌄  UU ⌄

**DAG: Lead_scoring_training_pipeline** Training pipeline for Lead Scoring System

Schedule: @monthly ⓘ    Next Run: 2022-11-01, 00:00:00

| ▦ Grid | 🔧 Graph | 📅 Calendar | ⏳ Task Duration | ⇄ Task Tries | ⤓ Landing Times | ☰ Gantt | ⚠ Details | <> Code | 📋 Audit Log |

02-11-2022 08:32:18 📅    25 ⌄    All Run Types ⌄    All Run States ⌄    **Clear Filters**

deferred  failed  queued  running  scheduled  skipped  success  up_for_reschedule  up_for_retry  upstream_failed  no_status

Auto-refresh ⬤

⇥☰

Duration
Sep 01, 00:00          Nov 01, 13:52
00:00:24
00:00:12
00:00:00

encoding_categorical_variables
training_model

**DAG**
**Lead_scoring_training_pipeline**

**DAG Details**

**DAG Runs Summary**

| Total Runs Displayed | 14 |
| ■ Total success | 3 |
| ■ Total failed | 11 |
| First Run Start | 2022-10-25, 18:37:08 UTC |
| Last Run Start | 2022-11-02, 08:24:09 UTC |
| Max Run Duration | 00:00:24 |
| Mean Run Duration | 00:00:16 |
| Min Run Duration | 00:00:12 |

**DAG Summary**

| Total Tasks | 2 |

29°C
Haze

ENG
US

14:02
02-11-2022

Not secure | notebooksb.jarvislabs.ai:10976/dags/Lead_scoring_inference_pipeline/grid

**Airflow**    DAGs    Security ⌄    Browse ⌄    Admin ⌄    Docs ⌄

08:33 UTC ⌄    UU ⌄

DAG: **Lead_scoring_inference_pipeline** Inference pipeline of Lead Scoring system

Schedule: @hourly ⓘ    Next Run: 2022-11-02, 08:00:00

⊞ Grid    🔧 Graph    📅 Calendar    ⧗ Task Duration    ⇄ Task Tries    ⚒ Landing Times    ≡ Gantt    ⚠ Details    <> Code    📋 Audit Log

02-11-2022 08:33:01 📅    25 ⌄    All Run Types ⌄    All Run States ⌄    **Clear Filters**

deferred    failed    queued    running    scheduled    skipped    success    up_for_reschedule    up_for_retry    upstream_failed    no_status

Auto-refresh ⬜

→≡

Duration

Nov 01, 21:00

00:01:10

00:00:35

00:00:00

encoding_categorical_variables
generating_models_prediction
checking_model_prediction_ratio
checking_input_features

**DAG**
**Lead_scoring_inference_pipeline**

**DAG Details**

| DAG Runs Summary | |
|---|---|
| Total Runs Displayed | 5 |
| ■ Total success | 5 |
| First Run Start | 2022-11-01, 21:02:40 UTC |
| Last Run Start | 2022-11-02, 08:23:19 UTC |
| Max Run Duration | 00:01:10 |
| Mean Run Duration | 00:00:45 |
| Min Run Duration | 00:00:28 |

| DAG Summary | |
|---|---|
| Total Tasks | 4 |
| PythonOperators | 4 |

29°C
Haze

ENG
US

14:03
02-11-2022

# Airflow

DAGs    Security ▾    Browse ▾    Admin ▾    Docs ▾

08:33 UTC ▾    UU ▾

## DAG: Lead_scoring_inference_pipeline    Inference pipeline of Lead Scoring system

success    Schedule: @hourly    ⓘ    Next Run: 2022-11-02, 08:00:00

▦ Grid    ⊞ Graph    📅 Calendar    ⧖ Task Duration    ⇄ Task Tries    ⤓ Landing Times    ≡ Gantt    ⚠ Details    <> Code    🔍 Audit Log    ▶    🗑

📅 2022-11-02T08:23:19Z    Runs    25 ▾    Run    manual__2022-11-02T08:23:18.228741+00:00 ▾    Layout    Left > Right ▾    Update    Find Task...

PythonOperator

deferred    failed    queued    running    scheduled    skipped    success    up_for_reschedule    up_for_retry    upstream_failed    no_status

Auto-refresh    ↻

```
encoding_categorical_variables → generating_models_prediction → checking_model_prediction_ratio → checking_input_features
```