

1. Model Experimentation – Screenshot of all experiments

mlflow1.26.1ExperimentsModels

Experiments+<Baseline\_model\_exp01

Search Experiments

DefaultBaseline\_model\_exp...Baseline\_model\_exp...

Track machine learning training runs in experiments. Learn more

Experiment ID: 1

DescriptionEdit

RefreshCompareDeleteDownload CSVStart TimeAll time

ColumnsOnly show differencesmetrics.rmse < 1 and params.model = "tree"SearchFilterClear

Showing 11 matching runs

								Metrics		Parameters		
	Start Time	Duration	Run Name	User	Source	Version	Models	AUC	Accuracy	F1	C	CPU Jobs
	2 minutes ago		Session Init...	root	ipykernel...	-	-	-	-	-	-	-1
	17 seconds ago		Light Gradie...	root	ipykernel...	-	sklearn	0.821	0.739	0.762	-	-
	29 seconds ago		Naive Bayes	root	ipykernel...	-	sklearn	0.734	0.663	0.727	-	-
	29 seconds ago		Linear Discri...	root	ipykernel...	-	sklearn	0.773	0.701	0.728	-	-
	29 seconds ago		Ridge Classi...	root	ipykernel...	-	sklearn	0	0.701	0.728	-	-
	30 seconds ago		Logistic Reg...	root	ipykernel...	-	sklearn	0.784	0.71	0.74	1.0	-
	30 seconds ago		Decision Tre...	root	ipykernel...	-	sklearn	0.817	0.736	0.758	-	-
	31 seconds ago		Extra Trees ...	root	ipykernel...	-	sklearn	0.818	0.737	0.758	-	-
	31 seconds ago		Random For...	root	ipykernel...	-	sklearn	0.819	0.737	0.759	-	-
	31 seconds ago		Extreme Gra...	root	ipykernel...	-	sklearn	0.821	0.738	0.762	-	-
	35 seconds ago		Light Gradie...	root	ipykernel...	-	sklearn	0.821	0.739	0.762	-	-

2. Model Experimentation – Screenshot of one experiment with all artifacts visible

mlflow1.26.1ExperimentsModels

Baseline\_model\_exp01> Light Gradient Boosting Machine

Light Gradient Boosting Machine

Date: 2023-05-16 15:14:04Source: ipykernel\_launcher.pyUser: root

Status: UNFINISHEDLifecycle Stage: activeParent Run: 0c0efc9aac6d455aafde3f761b5cb67a

DescriptionEdit

Parameters (20)

Metrics (8)

Tags (5)

Artifacts

modelMLmodelconda.yamlmodel.pklpython\_env.yamlrequirements.txtAUC.pngConfusion Matrix.pngFeature Importance.pngHoldout.html

Full Path: /mlruns/1/84a85cb161d54e11818500dc0c7e08da/artifacts/model

Register Model

MLflow Model

The code snippets below demonstrate how to make predictions using the logged model. You can also register it to the model registry to version control

Model schema

Input and output schema for your model. Learn more

NameType

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

Make Predictions

Predict on a Spark DataFrame:

import mlflowlogged\_model = 'runs:/84a85cb161d54e11818500dc0c7e08da/model'# 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.columns = list(df.columns)

### 3. Model Experimentation – Screenshot of all experiments after dropping features

The screenshot shows the mlflow Experiments page for 'Baseline\_model\_exp02'. The interface includes a sidebar with a search bar and a list of experiments. The main area displays a table of 11 runs, each with a checkbox, start time, duration, run name, user, source, version, models, and metrics (AUC, Accuracy, F1, C, CPU Jobs). The runs are sorted by start time, with the most recent run at the top. A search bar at the bottom allows filtering runs by metrics and parameters.

	Start Time	Duration	Run Name	User	Source	Version	Models	AUC	Accuracy	F1	C	CPU Jobs
<input type="checkbox"/>	2 minutes ago		Session Init...	root	ipykernel...	-	-	-	-	-	-	-1
<input type="checkbox"/>	22 seconds ago		Light Gradie...	root	ipykernel...	-	sklearn	0.821	0.739	0.762	-	-
<input type="checkbox"/>	32 seconds ago		Naive Bayes	root	ipykernel...	-	sklearn	0.734	0.67	0.723	-	-
<input type="checkbox"/>	32 seconds ago		Linear Discri...	root	ipykernel...	-	sklearn	0.773	0.7	0.728	-	-
<input type="checkbox"/>	33 seconds ago		Ridge Classi...	root	ipykernel...	-	sklearn	0	0.7	0.728	-	-
<input type="checkbox"/>	33 seconds ago		Logistic Reg...	root	ipykernel...	-	sklearn	0.784	0.71	0.74	1.0	-
<input type="checkbox"/>	33 seconds ago		Decision Tre...	root	ipykernel...	-	sklearn	0.817	0.736	0.758	-	-
<input type="checkbox"/>	34 seconds ago		Extra Trees ...	root	ipykernel...	-	sklearn	0.817	0.737	0.758	-	-
<input type="checkbox"/>	34 seconds ago		Random For...	root	ipykernel...	-	sklearn	0.818	0.737	0.759	-	-
<input type="checkbox"/>	34 seconds ago		Extreme Gra...	root	ipykernel...	-	sklearn	0.821	0.738	0.762	-	-
<input type="checkbox"/>	37 seconds ago		Light Gradie...	root	ipykernel...	-	sklearn	0.821	0.739	0.762	-	-

### 4. Model Experimentation – Screenshot of one experiment with all artifacts visible after dropping features

The screenshot shows the mlflow Model page for 'Light Gradient Boosting Machine'. The page displays the model's metadata, including its source, user, and lifecycle stage. A list of artifacts is shown on the left, including 'MLmodel', 'conda.yaml', 'model.pkl', 'python\_env.yaml', 'requirements.txt', 'AUC.png', 'Confusion Matrix.png', 'Feature Importance.png', and 'Holdout.html'. The main area contains code snippets for making predictions using the logged model, both on a Spark DataFrame and a Pandas DataFrame.

Full Path: `runs:/462882847c1b48c5980d40fcb84516e8/artifacts/model` [Register Model](#)

#### MLflow Model

The code snippets below demonstrate how to make predictions using the logged model. You can also register it to the model registry to version control.

##### Model schema

Input and output schema for your model. [Learn more](#)

Name	Type
No schema. See <a href="#">MLflow docs</a> for how to include input and output schema with your model.	

##### Make Predictions

Predict on a Spark DataFrame:

```
import mlflow
logged_model = 'runs:/462882847c1b48c5980d40fcb84516e8/model'

# 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.
columns = list(df.columns)
df.withColumn('predictions', loaded_model(*columns)).collect()
```

Predict on a Pandas DataFrame:

```
import mlflow
logged_model = 'runs:/462882847c1b48c5980d40fcb84516e8/model'
```

5. Model Experimentation - Metrics of Model

mlflow1.26.1ExperimentsModels

Baseline\_model\_exp02 > Light Gradient Boosting Machine

Light Gradient Boosting Machine

Date: 2023-05-16 15:16:29

Source: ipykernel\_launcher.py

User: root

Status: UNFINISHED

Lifecycle Stage: active

Parent Run: ee6d015368bc4699b6c5a9d11f4244e8

Description

Edit

Parameters (20)

Metrics (8)

Name	Value
AUC	0.821
Accuracy	0.739
F1	0.762
Kappa	0.477
MCC	0.485
Prec.	0.702
Recall	0.834
TT	1.678

Tags (5)

Artifacts

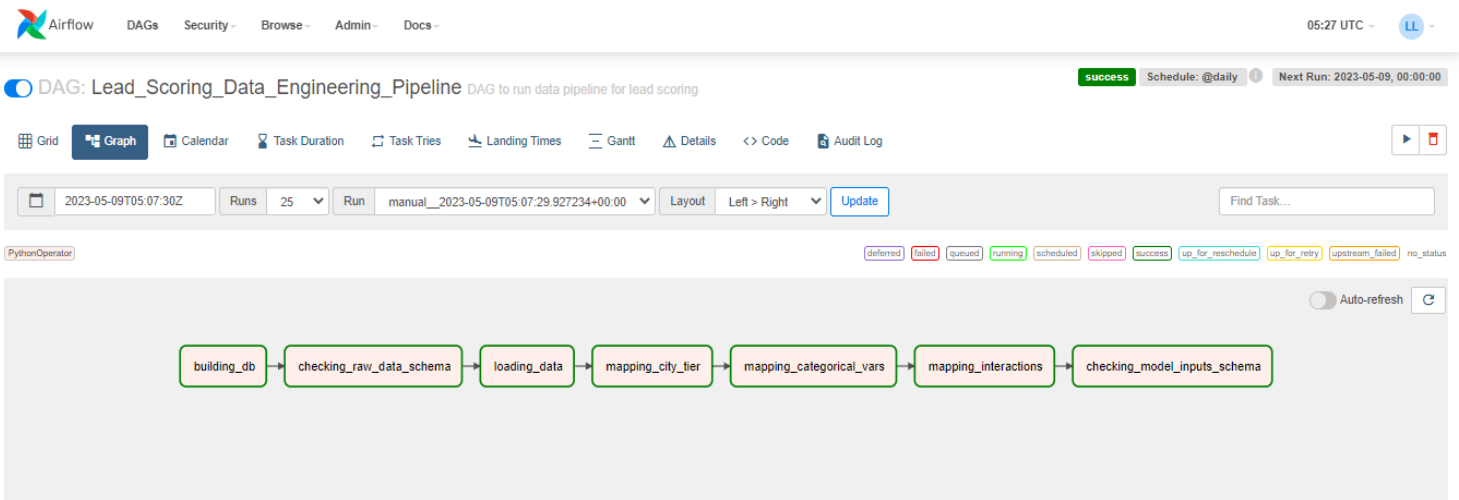
model

MLmodel

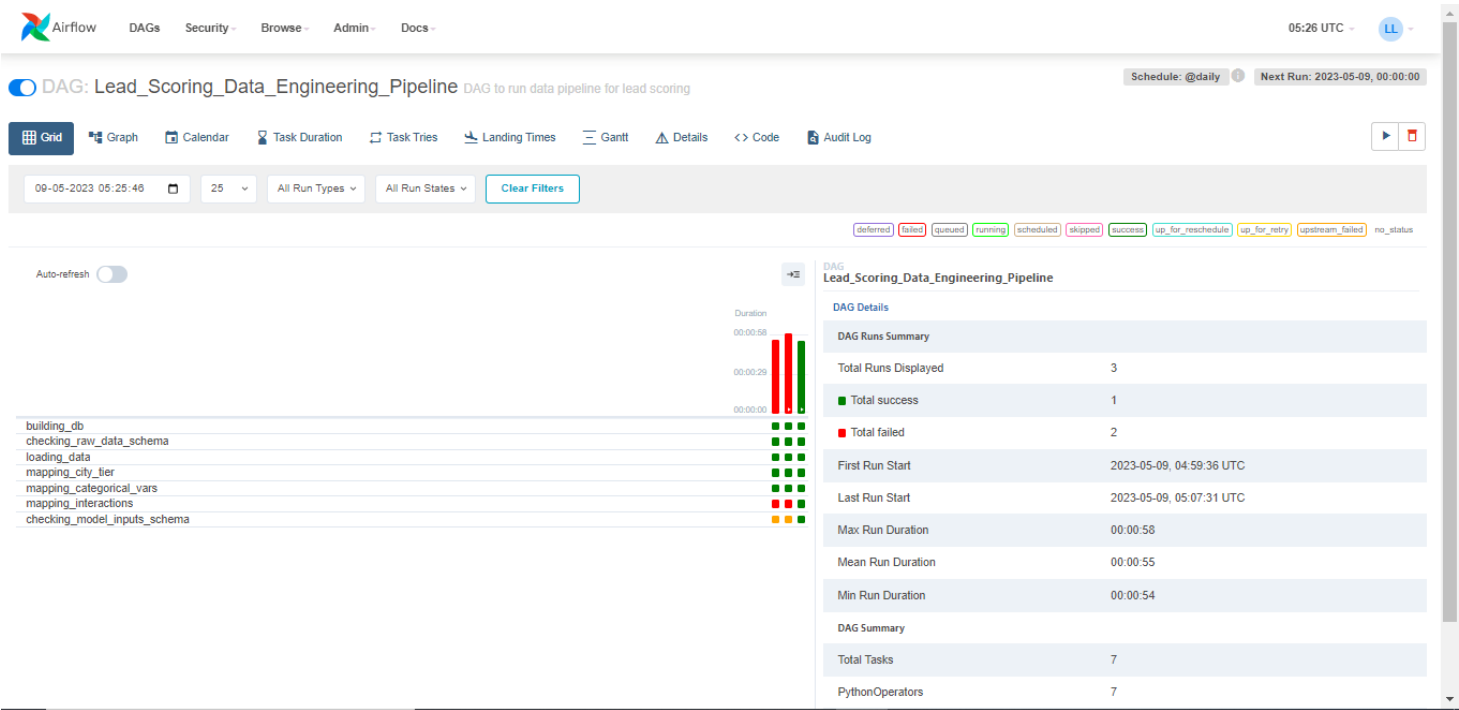
Full Path: ./miruns/2/4652882847c1b48c5980d40fcb84516e8/artifacts/model

Register Model

6. Data Pipeline – Screenshot of successful execution Airflow DAG in graph



7. Data Pipeline – Screenshot of airflow UI Grid



8. Training Pipeline – Screenshot of all experiments with all artifacts visible

mlflow1.26.1

ExperimentsModels

Default > run\_LightGB\_model

run\_LightGB\_model

Date: 2023-05-13 10:22:46

Source: airflow

User: root

Duration: 5.4s

Status: FINISHED

Lifecycle Stage: active

Description

Parameters (23)

Metrics (13)

Tags

Artifacts

models

MLmodel

conda.yaml

model.pkl

python\_env.yaml

requirements.txt

Full Path: mlruns/0/3eb9bef0df6c40ffa8743d7302779c36/artifacts/models

LightGBM, v2

Registered on 2023/05/13

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 <a href="#">MLflow docs</a> for how to include input and output schema with your model.	

Make Predictions

Predict on a Spark DataFrame:

```
import mlflow
logged_model = 'runs:/3eb9bef0df6c40ffa8743d7302779c36/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.
columns = list(df.columns)
```

9. Training Pipeline – Screenshot of model registry with model name and stage as ‘production’

mlflow1.26.1

ExperimentsModels

Registered Models

Share and manage machine learning models. [Learn more](#)

Create Model

Search by model name

Search

Filter

Clear

Name	Latest Version	Staging	Production	Last Modified	Tags
LightGBM	Version 2	-	Version 2	2023-05-13 10:33:29	-

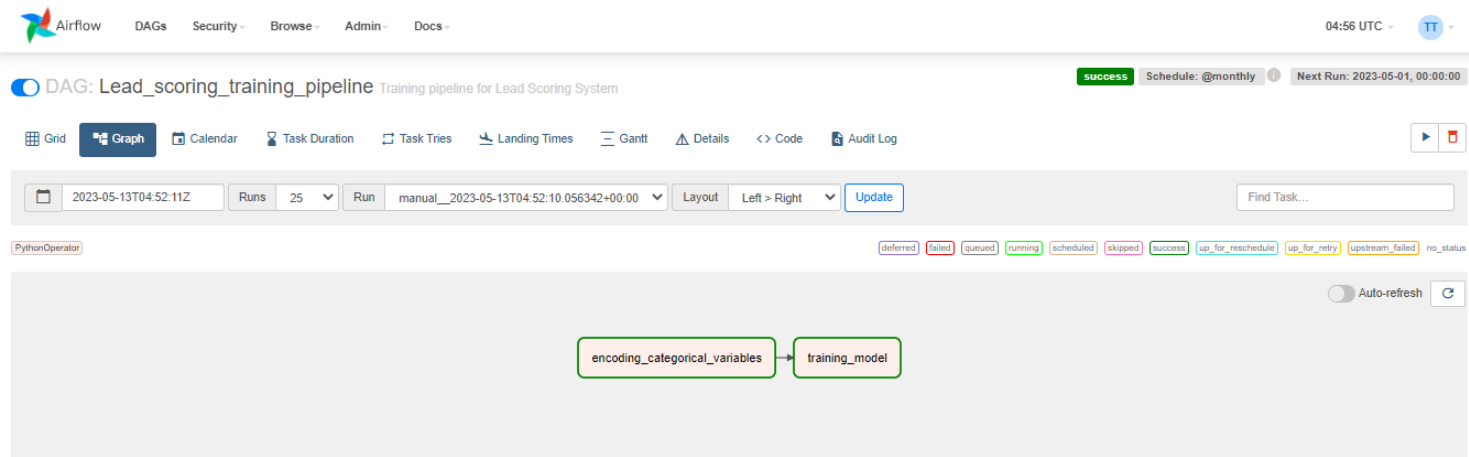
<

1

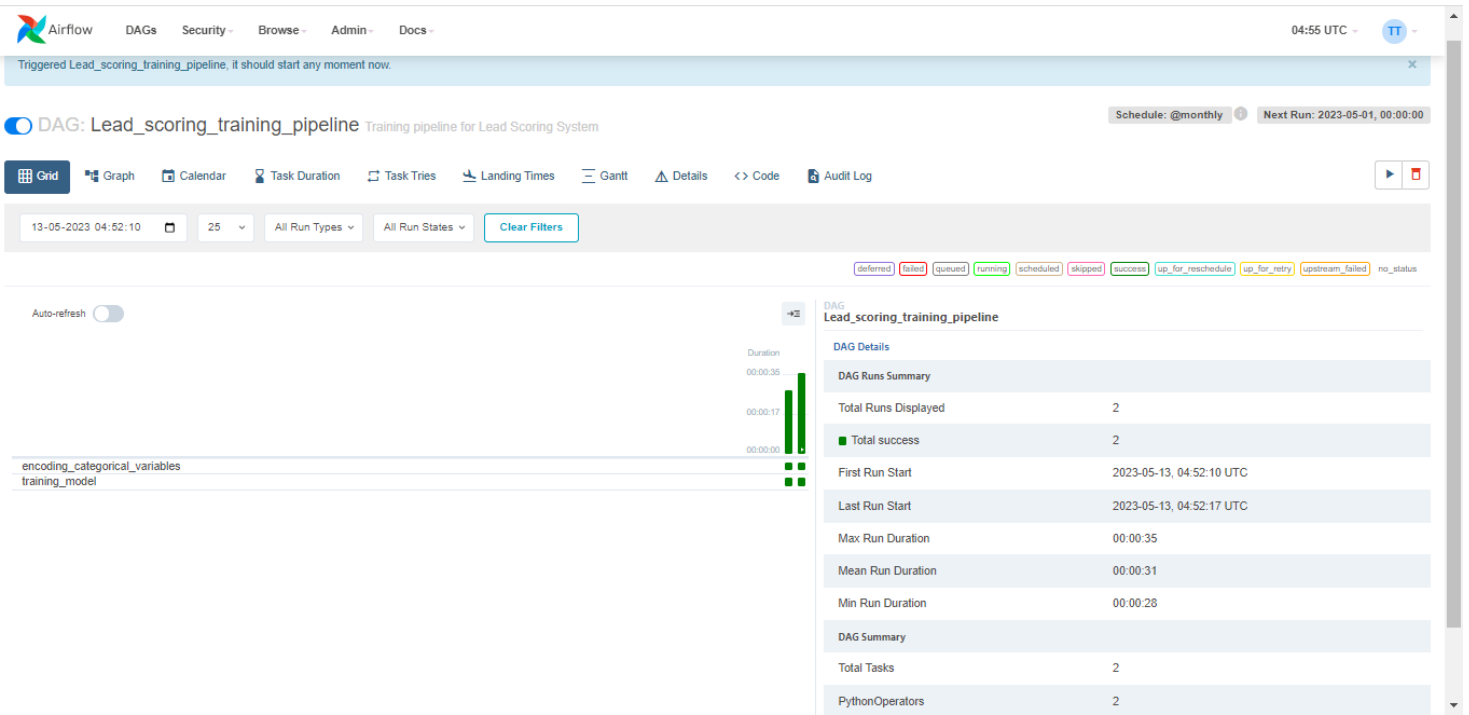
>

10 / page

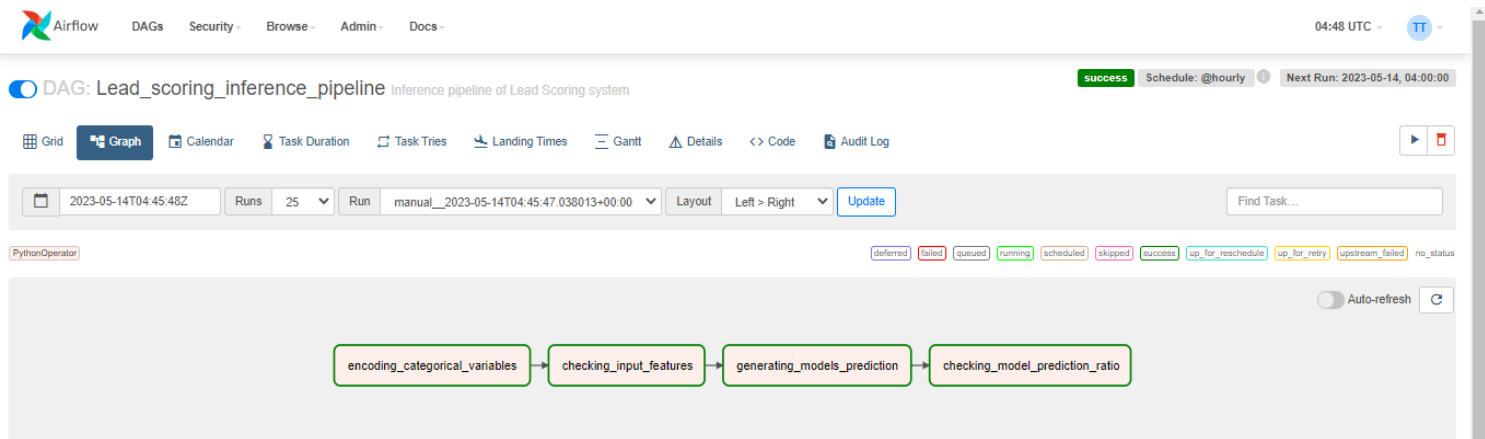
10. Training Pipeline - Screenshot of successful execution Airflow DAG in graph



11. Training Pipeline – Screenshot of airflow UI Grid



## 12. Inference Pipeline - Screenshot of successful execution Airflow DAG in graph



## 13. Inference Pipeline – Screenshot of airflow UI Grid

