


Lead_scoring_model_experimentation > Light Gradient Boosting Machine

Light Gradient Boosting Machine

Date: 2023-04-17 01:52:28

Source:  ipykernel_launcher.py

User: root

Status: UNFINISHED

Lifecycle Stage: active

Parent Run: ad765291dca2491592264083da42a31c









► Description [Edit](#)

▼ Parameters (20)









Name	Value
boosting_type	gbdt
class_weight	None
colsample_bytree	1.0
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

► Parameters (20)

▼ Metrics (8)

Name	Value
AUC 	0.821
Accuracy 	0.738
F1 	0.762
Kappa 	0.476
MCC 	0.485
Prec. 	0.702
Recall 	0.833
TT 	1.28

▼ Tags (5)

Name	Value	Actions
Run ID	59f7cc098efb4acd965a811f38d267c0	 
Run Time	6.62	 
Source	compare_models	 
URI	6c69436b	 

lead_scoring... - JupyterLab

MLflow

← → ↻ ⚠ Not secure | notebooks.jarvislabs.ai:10960/#/experiments/2

🔖 ☆ 🏠 🗖

⌵

Experiments

Search Experiments

Default

Lead_scoring_model...

Lead_scoring

Lead_scoring

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.mean < 1 and params.model = "tree"

Search

⏴ Filter

🗑 Clear

Showing 12 matching runs

								Metrics			Parameters			Tags	
	Start Time	Duration	Run Name	User	Source	Version	Models	AUC	Accuracy	F1	C	CPU Jobs	Categorical Feat	Source	
<input type="checkbox"/>	17 minutes ago		Session Init...	root	ipykernel...	-	-	-	-	-	-	-1	5	setup	
<input type="checkbox"/>	32 minutes ago		Light Gradie...	root	ipykernel...	-	sklearn	0.821	0.739	0.76	-	-	-	tune_model	
<input type="checkbox"/>	32 minutes ago		Light Gradie...	root	ipykernel...	-	sklearn	0.821	0.739	0.762	-	-	-	create_model	
<input type="checkbox"/>	32 minutes ago		Naive Bayes	root	ipykernel...	-	sklearn	0.734	0.673	0.725	-	-	-	compare_m...	
<input type="checkbox"/>	32 minutes ago		Linear Discr...	root	ipykernel...	-	sklearn	0.773	0.7	0.727	-	-	-	compare_m...	
<input type="checkbox"/>	32 minutes ago		Ridge Classi...	root	ipykernel...	-	sklearn	0	0.7	0.727	-	-	-	compare_m...	
<input type="checkbox"/>	32 minutes ago		Logistic Reg...	root	ipykernel...	-	sklearn	0.784	0.71	0.74	1.0	-	-	compare_m...	
<input type="checkbox"/>	32 minutes ago		Decision Tre...	root	ipykernel...	-	sklearn	0.817	0.736	0.758	-	-	-	compare_m...	
<input type="checkbox"/>	32 minutes ago		Extra Trees C...	root	ipykernel...	-	sklearn	0.818	0.736	0.758	-	-	-	compare_m...	
<input type="checkbox"/>	32 minutes ago		Random For...	root	ipykernel...	-	sklearn	0.819	0.737	0.759	-	-	-	compare_m...	
<input type="checkbox"/>	32 minutes ago		Extreme Gra...	root	ipykernel...	-	sklearn	0.821	0.738	0.761	-	-	-	compare_m...	
<input type="checkbox"/>	32 minutes ago		Light Gradie...	root	ipykernel...	-	sklearn	0.821	0.739	0.762	-	-	-	compare_m...	

Load more

lead_scoring... - JupyterLab

MLflow

← → ↻ ⚠ Not secure | notebooks.jarvislabs.ai:10960/#/experiments/2/runs/a3605a6102194a9197d58ad022adae2c

🔖 ☆ 🏠 🗖

⌵

mlflow 1.26.1

Experiments

Models

GitHub

Docs

Lead_scoring > Light Gradient Boosting Machine

Light Gradient Boosting Machine

Date: 2023-04-17 02:11:22

Status: UNFINISHED

Source: ☐ ipykernel_launcher.py

Lifecycle Stage: active

User: root

Parent Run: f5dff63496384b48a144ec4643fe000

Description [Edit](#)

Parameters (23)

Name	Value
bagging_fraction	0.887437593014418
bagging_freq	1
boosting_type	gbdt
class_weight	None
colsample_bytree	1.0
feature_fraction	0.7588282465072769
importance_type	split
learning_rate	0.037577891518562063
max_depth	-1
min_child_samples	53
min_child_weight	0.001
min_split_gain	0.09545503921499346

read_scoring... - jupyterlab

MLflow

Not secure | notebooks.jarvislabs.ai:10960/#/experiments/2/runs/a3605a6102194a9197d58ad022adae2c

min_split_gain	0.09545503921499346
n_estimators	229
n_jobs	-1
num_leaves	224
objective	None
random_state	42
reg_alpha	1.074527356155175
reg_lambda	1.7978270329637185e-06
silent	warn
subsample	1.0
subsample_for_bin	200000
subsample_freq	0

Metrics (8)

Name	Value
AUC 🔗	0.821
Accuracy 🔗	0.739
F1 🔗	0.76
Kappa 🔗	0.477
MCC 🔗	0.484
Prec. 🔗	0.705
Recall 🔗	0.825
TT 🔗	3.62







































































































































Tags (5)

Name	Value	Actions
Run ID	a3605a6102194a9197d58ad022adae2c	🔗 🗑️
Run Time	98.05	🔗 🗑️
Source	tune_model	🔗 🗑️
URI	2938dab1	🔗 🗑️

Artifacts

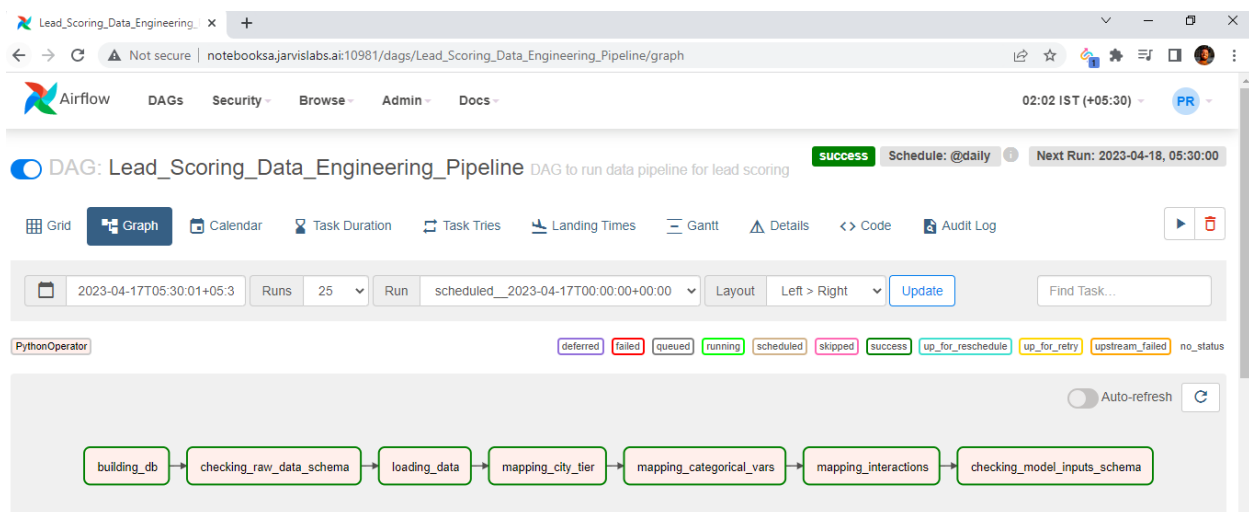
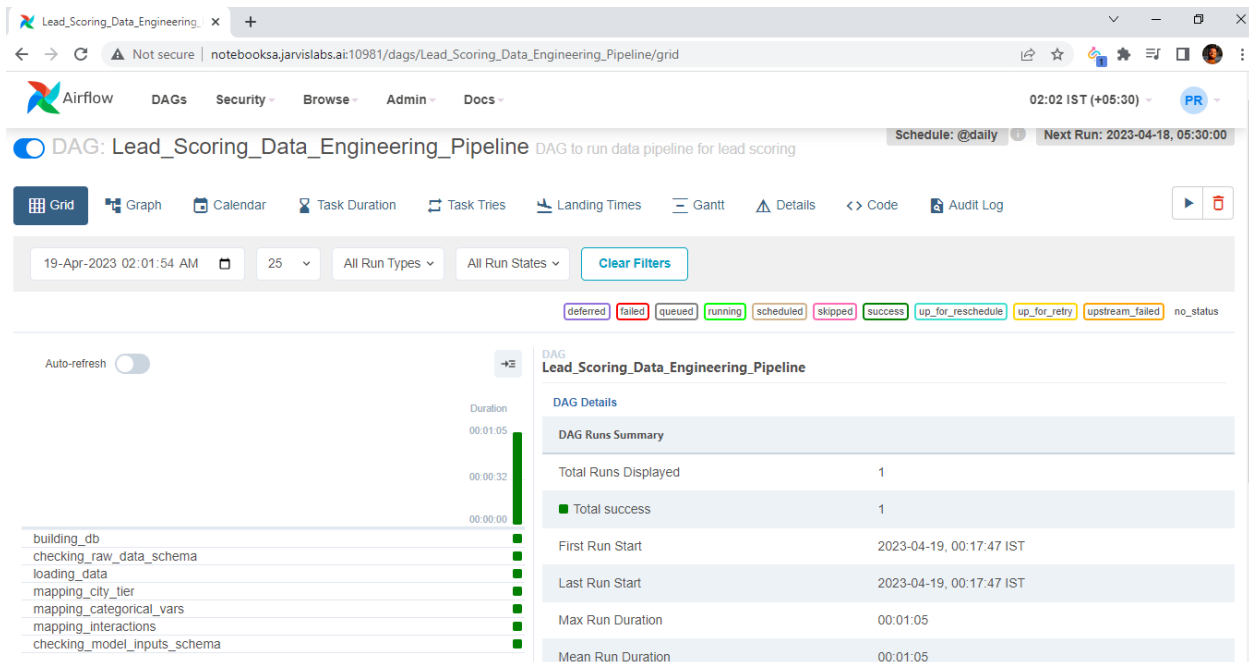
The screenshot shows the Apache Airflow web interface. The browser's address bar indicates the URL is `notebooks.jarvislabs.ai:1081/home`. The navigation bar at the top contains the Airflow logo and several menu items: `DAGs`, `Security`, `Browse`, `Admin`, and `Docs`. The time displayed is 18:18 UTC. Two yellow warning banners are present at the top of the main content area. The first banner states: "Do not use **SQLite** as metadata DB in production – it should only be used for dev/testing. We recommend using Postgres or MySQL. [Click here](#) for more information." The second banner states: "Do not use **SequentialExecutor** in production. [Click here](#) for more information."

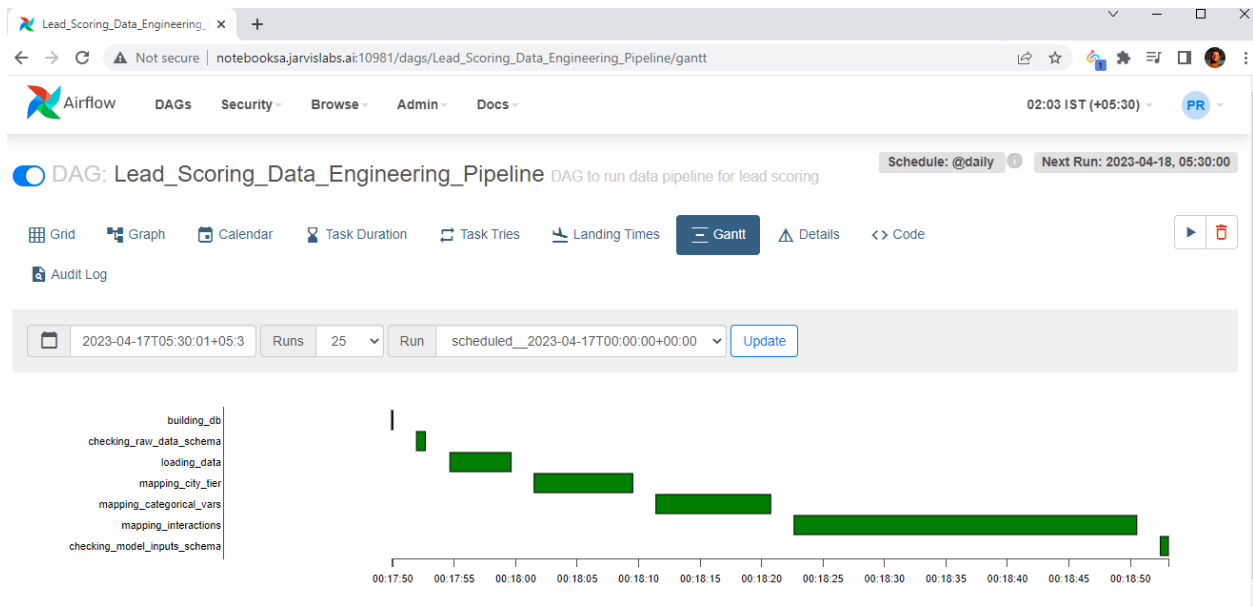
DAGs

All 21		Active 9	Paused 12	Filter DAGs by tag		Search DAGs									
DAG ↕	Owner ↕	Runs	Schedule	Last Run	Next Run ↕	Recent Tasks	Actions		Links						
 Lead_Scoring_Data_Engineering_Pipeline	airflow	   	@daily	2023-04-17, 00:00:00		               	 	...							
 example_bash_operator example example2	airflow	   	00:00:00	2023-04-17, 00:00:00		               	 	...							
 example_branch_datetime_operator example	airflow	   	@daily	2023-04-17, 00:00:00		               	 	...							
 example_branch_datetime_operator_2 example	airflow	   	@daily	2023-04-17, 00:00:00		               	 	...							
 example_branch_dop_operator_v3 example	airflow	   	00:00:00	2023-04-18, 18:13:00		               	 	...							
 example_branch_labels	airflow	   	@daily	2023-04-17, 00:00:00		        <									

DAGs

All 33		Active 1	Paused 32	Filter DAGs by tag		Search DAGs				
DAG ↕	Owner ↕	Runs 1	Schedule	Last Run 1	Next Run ↕ 1	Recent Tasks 1				
<input checked="" type="checkbox"/> Lead_Scoring_Data_Engineering_Pipeline	airflow	<input checked="" type="radio"/> 1	@daily 1	2023-04-17, 05:30:00 1	2023-04-18, 05:30:00 1	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/> 7
<input type="checkbox"/> example_bash_operator	airflow	<input type="radio"/>	0 0 * * * 1		2023-04-17, 05:30:00 1	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> example_branch_datetime_operator	airflow	<input type="radio"/>	@daily 1		2023-04-17, 05:30:00 1	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> example_branch_datetime_operator_2	airflow	<input type="radio"/>	@daily 1		2023-04-17, 05:30:00 1	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> example_branch_dop_operator_v3	airflow	<input type="radio"/>	* * * * * 1		2023-04-19, 02:00:00 1	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>





Lead_Scoring_Data_Engineering_ x +

Not secure | notebooksajarvislabs.ai:10981/dags/Lead_Scoring_Data_Engineering_Pipeline/details

Airflow DAGs Security Browse Admin Docs 02:03 IST (+05:30) PR

DAG: Lead_Scoring_Data_Engineering_Pipeline DAG to run data pipeline for lead scoring Schedule: @daily

Grid Graph Calendar Task Duration Task Tries Landing Times Gantt Details <> Code Audit Log

DAG Details

success 7

Schedule Interval	@daily
Catchup	False
Started	True
End Date	None
Max Active Runs	0 / 16
Concurrency	16
Default Args	{'owner': 'airflow', 'retries': 1, 'retry_delay': datetime.timedelta(seconds=5), 'start_date': DateTime(2022, 7, 30, 0, 0, 0, tzinfo=Timezone('UTC'))}
Tasks Count	7

Lead_Scoring_Data_Engineering_

+

Not secure | notebooksajarvislabs.ai:10981/audit_log?dag_id=Lead_Scoring_Data_Engineering_Pipeline&root=

🔍 ⭐ ⚙️ 📄 👤

Airflow

DAGs

Security

Browse

Admin

Docs

02:03 IST (+05:30) PR

DAG: Lead_Scoring_Data_Engineering_Pipeline DAG to run data pipeline for lead scoring

Grid

Graph

Calendar

Task Duration

Task Tries

Landing Times

Gantt

Details

<> Code

Audit Log

▶

🗑️

Show

100

 entries

Search:

MLflow

+

Not secure | notebooksajarvislabs.ai:10980/#/experiments/1

🔍 ⭐ ⚙️ 📄 👤

1.26.1

Experiments

Models

GitHub Docs

Experiments

+

<

Lead_scoring_mlflow_production

Share

Search Experiments

Default

Lead_scoring_mlflow...

🟢 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 1 matching run

Load more

Lead_scoring_mlflow_production > run_LightGB

run_LightGB

Date: 2023-04-19 01:58:51

Source:  ipykernel_launcher.py

User: root

Duration: 10.9s

Status: FINISHED

Lifecycle Stage: active

▸ Description [Edit](#)

▼ Parameters (20)

Name	Value
boosting_type	gbdt
class_weight	None
colsample_bytree	1.0
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	42
reg_alpha	0.0
reg_lambda	0.0
silent	warn
subsample	1.0
subsample_for_bin	200000

MLflow

Not secure | notebooksajarvislabs.ai:10980/#/experiments/1/runs/53033d51705b4260ac94e895a625a7eb

silent	warn
subsample	1.0
subsample_for_bin	200000
subsample_freq	0

▼ Metrics (1)

Name	Value
auc	0.747

► Tags

▼ Artifacts

models

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

Full Path: mlruns/1/53033d51705b4260ac94e895a625a7eb/artifacts/models/requirements.txt

Size: 287B

```
mlflow
cffi==1.15.0
cloudpickle==2.1.0
colorama==0.4.4
configparser==5.2.0
```

MLflow

Not secure | notebooksajarvislabs.ai:10980/#/models

mlflow 1.26.1 Experiments Models GitHub Docs

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 1	-	-	2023-04-19 01:59:02	-

< 1 > 10 / page

MLflow

Not secure | notebooksajarvislabs.ai:10980/#/models/LightGBM

mlflow 1.26.1 Experiments Models GitHub Docs

Registered Models > LightGBM

LightGBM

Created Time: 2023-04-19 01:59:02 Last Modified: 2023-04-19 01:59:02

► Description [Edit](#)

► Tags

▼ Versions All Active 0 Compare

Version	Registered at	Created by	Stage	Description
<input type="checkbox"/> Version 1	2023-04-19 01:59:02		None	

< 1 >

MLflow 1.26.1 Experiments Models GitHub Docs

Registered Models > LightGBM > Version 1

Version 1

Registered At: 2023-04-19 01:59:02 Stage: None Last Modified: 2023-04-19 01:59:02

Source Run: [run_LightGB](#)

- Description [Edit](#)
- Tags
- Schema

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

Airflow DAGs Security Browse Admin Docs 02:31 IST (+05:30) PR

Do not use SequentialExecutor in production. [Click here to learn more.](#)

DAGs

All 34 Active 1 Paused 33

Variables Configurations Connections Plugins Providers Pools XComs

DAG	Provider	Schedule	Last Run	Next Run	Recent Tasks
<input checked="" type="checkbox"/> Lead_Scoring_Data_Engineering_Pipeline	airflow	@daily	2023-04-17, 05:30:00	2023-04-18, 05:30:00	7
<input type="checkbox"/> Lead_scoring_training_pipeline	airflow	@monthly		2023-03-01, 05:30:00	
<input type="checkbox"/> example_bash_operator	airflow	0 0 * * *	2023-04-17, 05:30:00		
<input type="checkbox"/> example_branch_datetime_operator	airflow	@daily	2023-04-17, 05:30:00		
<input type="checkbox"/> example_branch_datetime_operator_2	airflow	@daily	2023-04-17, 05:30:00		
<input type="checkbox"/> example_branch_dop_operator_v3	airflow	* * * * *	2023-04-19, 02:28:00		
<input type="checkbox"/> example_branch_labels	airflow	@daily	2023-04-17, 05:30:00		

Lead_scoring_training_pipeline - x +

Not secure | notebooksajarvislabs.ai:10981/dags/Lead_scoring_training_pipeline/grid?root=

Airflow DAGs Security Browse Admin Docs 02:58 IST (+05:30) PR

DAG: Lead_scoring_training_pipeline Training pipeline for Lead Scoring System Schedule: @monthly Next Run: 2023-04-01, 05:30:00

Grid Graph Calendar Task Duration Task Tries Landing Times Gantt Details Code Audit Log

19-Apr-2023 02:56:28 AM 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 00:00:32 00:00:16 00:00:00

encoding_categorical_variables training_model

DAG: Lead_scoring_training_pipeline

DAG Details

DAG Runs Summary

Total Runs Displayed	1
Total success	1
First Run Start	2023-04-19, 02:55:22 IST
Last Run Start	2023-04-19, 02:55:22 IST
Max Run Duration	00:00:32

Lead_scoring_training_pipeline - x +

Not secure | notebooksajarvislabs.ai:10981/dags/Lead_scoring_training_pipeline/graph?root=

Airflow DAGs Security Browse Admin Docs 02:58 IST (+05:30) PR

DAG: Lead_scoring_training_pipeline Training pipeline for Lead Scoring System success Schedule: @monthly Next Run: 2023-04-01, 05:30:00

Grid Graph Calendar Task Duration Task Tries Landing Times Gantt Details Code Audit Log

2023-03-01T05:30:01+05:3 Runs 25 Run scheduled__2023-03-01T00:00:00:00 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 training_model

Lead_scoring_training_pipeline - x +

Not secure | notebooksajarvislabs.ai:10981/dags/Lead_scoring_training_pipeline/gantt?root=

Airflow DAGs Security Browse Admin Docs 02:59 IST (+05:30) PR

DAG: Lead_scoring_training_pipeline Training pipeline for Lead Scoring System Schedule: @monthly Next Run: 2023-04-01, 05:30:00

Grid Graph Calendar Task Duration Task Tries Landing Times Gantt Details Code Audit Log

2023-03-01T05:30:01+05:3 Runs 25 Run scheduled__2023-03-01T00:00:00:00 Update

encoding_categorical_variables training_model

02:55:30 02:55:35 02:55:40 02:55:45 02:55:50

Lead_scoring_training_pipeline - x

+

Not secure | notebooksajarvislabs.ai:10981/dags/Lead_scoring_training_pipeline/details

🔗 ⭐ ⚙️ 📄 👤 ⋮

Airflow

DAGs

Security

Browse

Admin

Docs

02:59 IST (+05:30) PR

DAG: Lead_scoring_training_pipeline

Training pipeline for Lead Scoring System

Schedule: @monthly

Grid

Graph

Calendar

Task Duration

Task Tries

Landing Times

Gantt

Details

<> Code

Audit Log

DAG Details

SUCCESS 2

Lead_scoring_training_pipeline - x

+

Not secure | notebooksajarvislabs.ai:10981/audit_log?dag_id=Lead_scoring_training_pipeline&root=

🔗 ⭐ ⚙️ 📄 👤 ⋮

Airflow

DAGs

Security

Browse

Admin

Docs

02:59 IST (+05:30) PR

DAG: Lead_scoring_training_pipeline

Training pipeline for Lead Scoring System

Schedule: @monthly

Grid

Graph

Calendar

Task Duration

Task Tries

Landing Times

Gantt

Details

<> Code

Audit Log

Dag Audit Log

This view displays selected events and operations that have been taken on this dag. The included and excluded events are set in the Airflow configuration, which by default remove view only actions. For a full list of events regarding this DAG, click [here](#).

Show

100

 entries

Search:

Time	Task ID	Event	Logical Date	Owner	Details
2023-04-19, 02:55:51	training_model	success	2023-03-01 00:00:00+00:00	airflow	None
2023-04-19, 02:55:37	training_model	running	2023-03-01 00:00:00+00:00	airflow	None
2023-04-19, 02:55:37	training_model	cli_task_run	None	root	{'host_name': '67cc4321afa8', 'full_command': '/opt/conda/bin/airflow, 'tasks', 'run', 'Lead_scoring_training_pipeline', 'training_model', 'scheduled__2023-03-01T00:00:00+00:00 local', '-subdir', 'DAGS_FOLDER/Lead_scoring_training_pipeline/.ipynb_checkpoints/lead_scoring_training_checkpoint.py']}

MLflow 1.26.1 Experiments Models GitHub Docs

Experiments

Search Experiments

Default

Lead_scoring_mlflow...

Lead_scoring_mlflow_production

Share

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 ☐ [Search](#) [Filter](#) [Clear](#)

Showing 2 matching runs

		Duration	Run Name	User	Source	Version	Models	Metrics	Param
<input type="checkbox"/>	↓ Start Time							auc	boostin
<input type="checkbox"/>	4 minutes ago	9.7s	run_LightGB	root	airflow	-	LightGBM/2	0.747	gbdt
<input type="checkbox"/>	1 hour ago	10.9s	run_LightGB	root	ipykernel...	-	LightGBM/1	0.747	gbdt

MLflow 1.26.1 Experiments Models GitHub Docs

Lead_scoring_mlflow_production > run_LightGB

run_LightGB

Date: 2023-04-19 02:55:42 Source: airflow User: root

Duration: 9.7s Status: FINISHED Lifecycle Stage: active

Description [Edit](#)

Parameters (20)

Name	Value
boosting_type	gbdt
class_weight	None
colsample_bytree	1.0
importance_type	split
learning_rate	0.1
max_depth	-1

Registered Models > LightGBM

LightGBM



Created Time: 2023-04-19 01:59:02

Last Modified: 2023-04-19 02:55:51

Description [Edit](#)

Tags

Versions

AllActive 0

Compare

<input type="checkbox"/>	Version	Registered at	Created by	Stage	Description
<input type="checkbox"/>	Version 2	2023-04-19 02:55:51		None	
<input type="checkbox"/>	Version 1	2023-04-19 01:59:02		None	

Registered Models > LightGBM > Version 2

Version 2



Registered At: 2023-04-19 02:55:51

Stage: Production

Last Modified: 2023-04-19 03:12:09

Source Run: run_LightGB

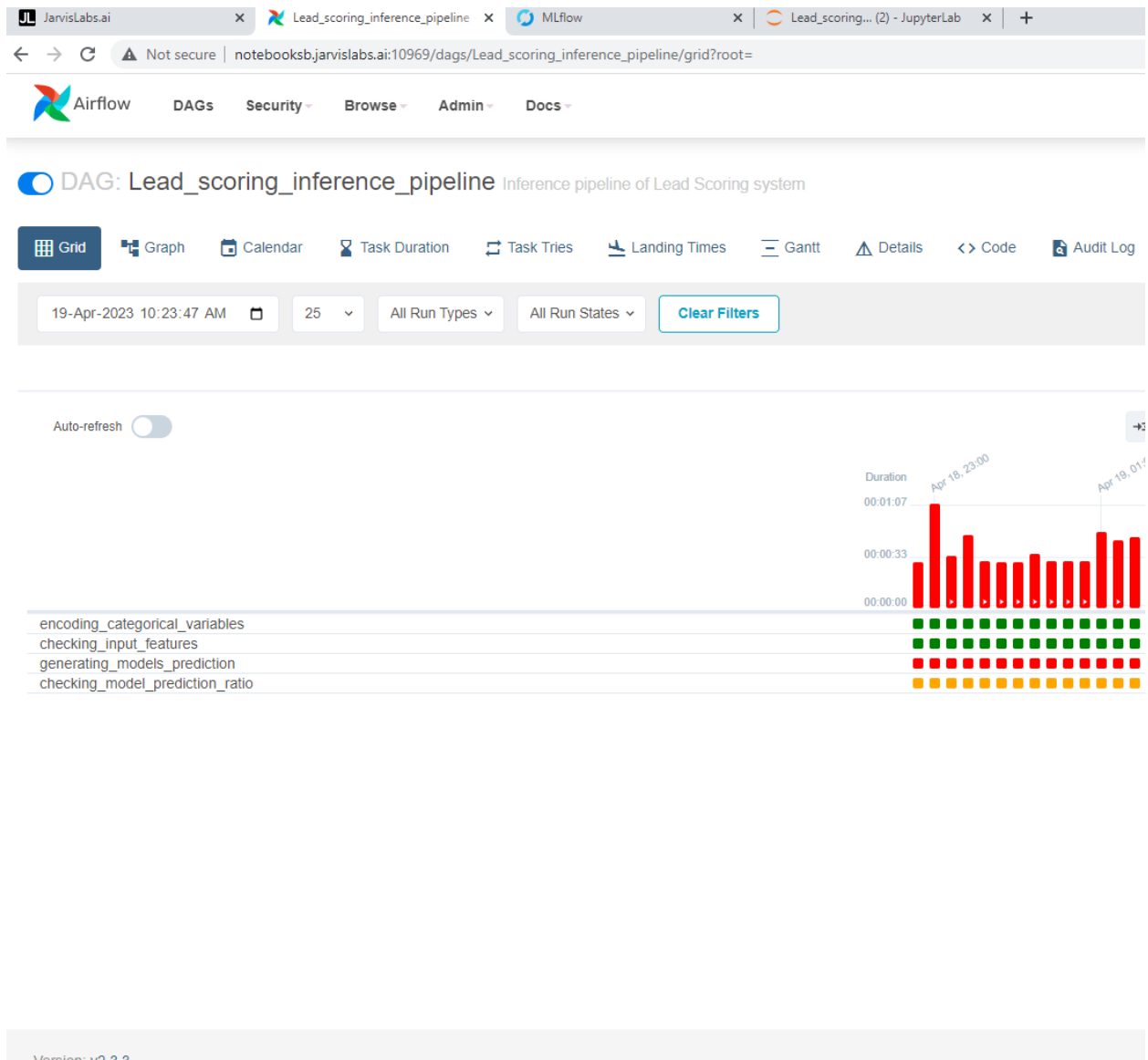
Description [Edit](#)

Tags

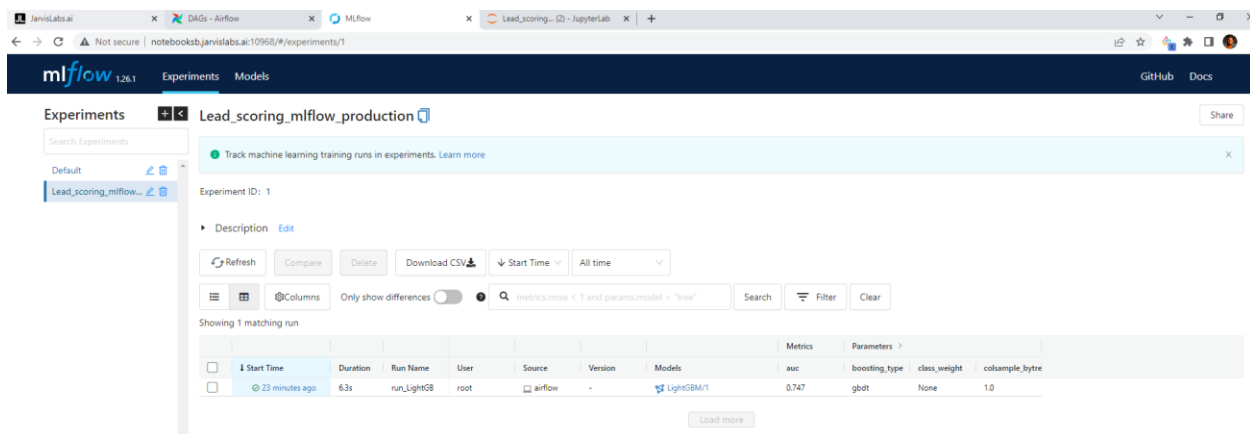
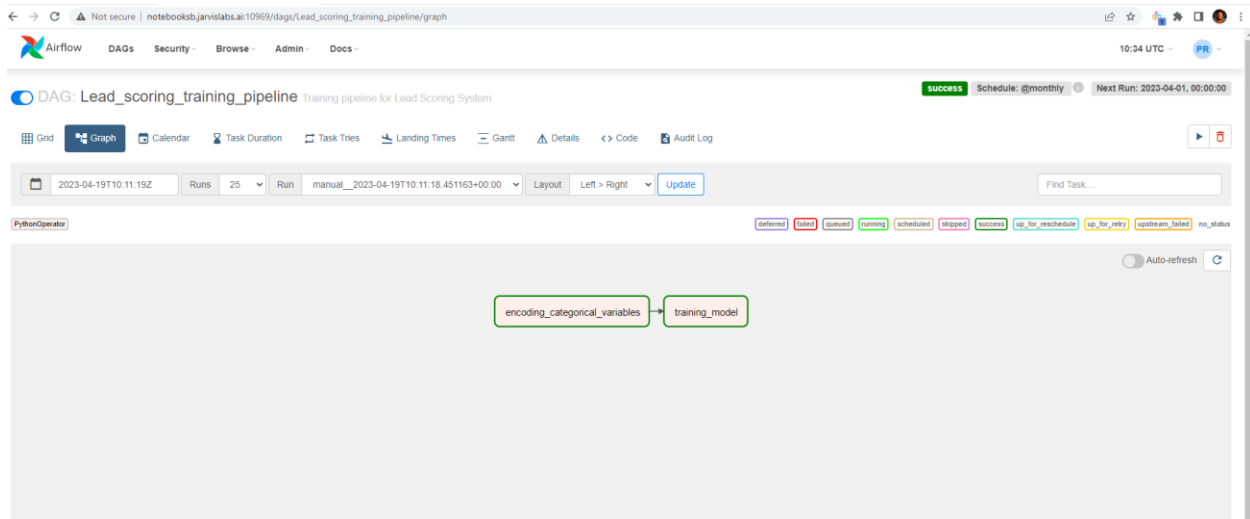
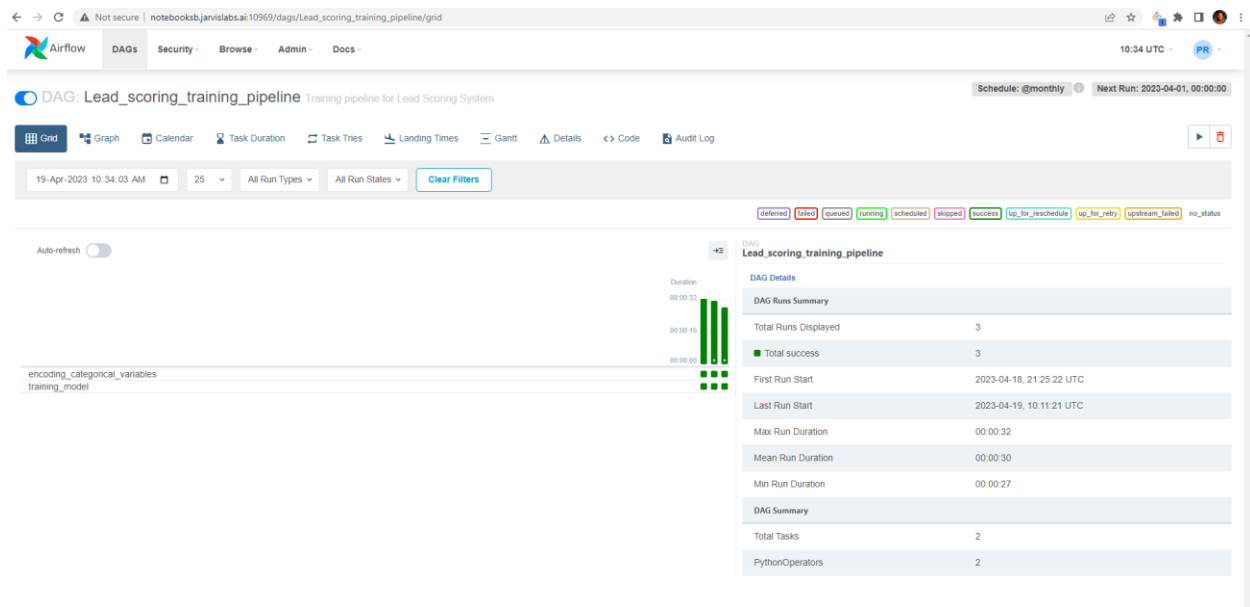
Schema

Name	Type
------	------

No schema. See [MLflow docs](#) for how to include input and output schema with your model.



Ran Fresh Training pipeline:



mlflow1.26.1

ExperimentsModels

GitHubDocs

Lead_scoring_mlflow_production>run_LightGB

run_LightGB

Date: 2023-04-19 15:41:41

Source: airflow

User: root

Duration: 6.3s

Status: FINISHED

Lifecycle Stage: active

DescriptionEdit

Parameters (20)

Metrics (1)

Tags

Artifacts

models

MLmodel

conda.yaml

model.pkl

python_env.yaml

requirements.txt

Full Path: mlruns/1/4e2630dc2829480ea2fd5f1ae85695a/artifacts/models

LightGBM, v1.27.0Registered on 2023/04/19

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:/4e2630dc2829480ea2fd5f1ae85695a/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)
df.withColumn('predictions', loaded_model(*columns)).collect()
```

Predict on a Pandas DataFrame:

Parameters (20)

Name	Value
boosting_type	gbdt
class_weight	None
colsample_bytree	1.0
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	42
reg_alpha	0.0
reg_lambda	0.0
silent	warn
subsample	1.0
verbose	0

subsample_freq 0

Metrics (1)

Name	Value
auc	0.747

Tags

Artifacts

models

MLmodel

conda.yaml

model.pkl

python_env.yaml

requirements.txt

Full Path/mlruns/1/4e2630dc2829480ea2fdfs5fae85695a/artifacts/models

LighGBM v1

Registered on 2023/04/19

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:/4e2630dc2829480ea2fdfs5fae85695a/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)
df.withColumn('predictions', loaded_model(*columns)).collect()
```

Predict on a Pandas DataFrame:

```
import mlflow
logged_model = 'runs:/4e2630dc2829480ea2fdfs5fae85695a/models'

# Load model as a PyFuncModel.
loaded_model = mlflow.pyfunc.load_model(logged_model)

# Predict on a Pandas DataFrame.
import pandas as pd
```

mlflow 1.26.1 Experiments Models GitHub Docs

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 1	-	Version 1	2023-04-19 15:42:46	-

< 1 > 10 / page

mlflow 1.26.1 Experiments Models GitHub Docs

Registered Models > LightGBM

LightGBM

Created Time: 2023-04-19 15:41:47 Last Modified: 2023-04-19 15:42:46

Description Edit

Tags

Versions All Active 1 Compare

Version	Registered at	Created by	Stage	Description
<input type="checkbox"/> Version 1	2023-04-19 15:41:47		Production	

< 1 >

mlflow 1.26.1 Experiments Models

Registered Models > LightGBM > Version 1

Version 1

Registered At: 2023-04-19 15:41:47 Stage: Production Last Modified: 2023-04-19 15:42:46

Source Run: run_LightGBM

Description Edit

Tags

Schema

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

Airflow DAG: Lead_scoring_inference_pipeline Inference pipeline of Lead Scoring system Schedule: @hourly Next Run: 2023-04-19, 10:00:00

Grid Graph Calendar Task Duration Task Times Landing Times Gantt Details Code Audit Log

19-Apr-2023 10:37:41 AM 25 All Run Types All Run States Clear Filters

Auto-refresh

Duration 00:01:07 00:00:33 00:00:00

Apr 18, 23:00 Apr 18, 01:00

encoding_categorical_variables
checking_input_features
generating_models_prediction
checking_model_prediction_ratio

Lead_scoring_inference_pipeline

DAG Details

DAG Runs Summary

Total Runs Displayed	15
Total success	1
Total failed	14
First Run Start	2023-04-18, 23:21:31 UTC
Last Run Start	2023-04-19, 10:22:40 UTC
Max Run Duration	00:01:07
Mean Run Duration	00:00:38
Min Run Duration	00:00:29

DAG Summary

Total Tasks	4
PythonOperators	4

Airflow DAG: Lead_scoring_inference_pipeline Inference pipeline of Lead Scoring system success Schedule: @hourly Next Run: 2023-04-19, 10:00:00

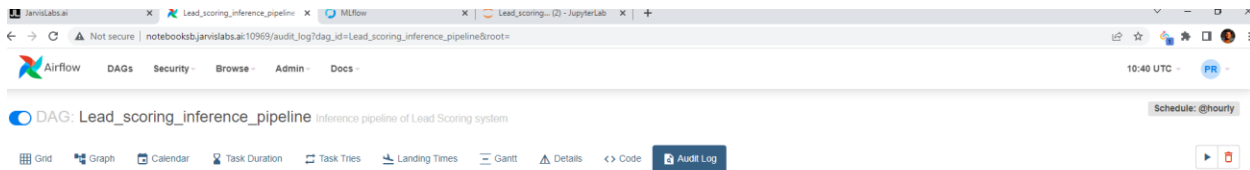
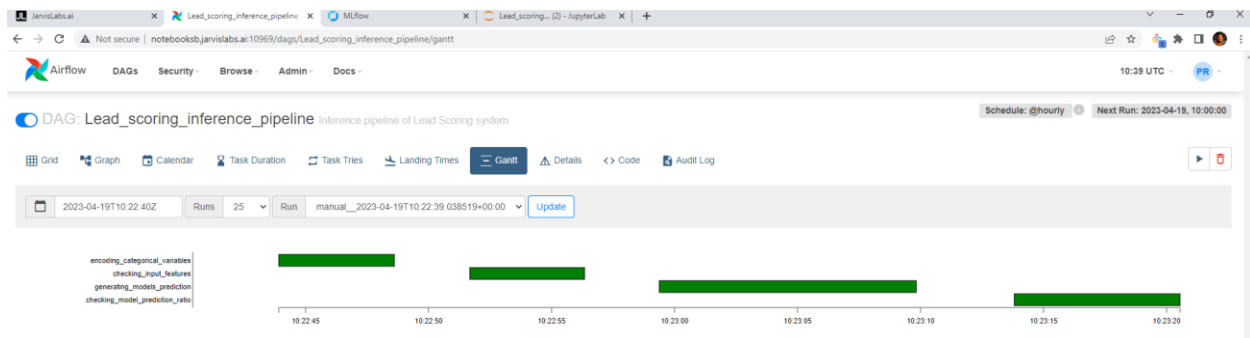
Grid Graph Calendar Task Duration Task Times Landing Times Gantt Details Code Audit Log

2023-04-19T10:22:40Z Runs 25 Run manual__2023-04-19T10:22:39.038519+00:00 Layout Left > Right Update Find Task...

PythonOperator

Auto-refresh

encoding_categorical_variables → checking_input_features → generating_models_prediction → checking_model_prediction_ratio



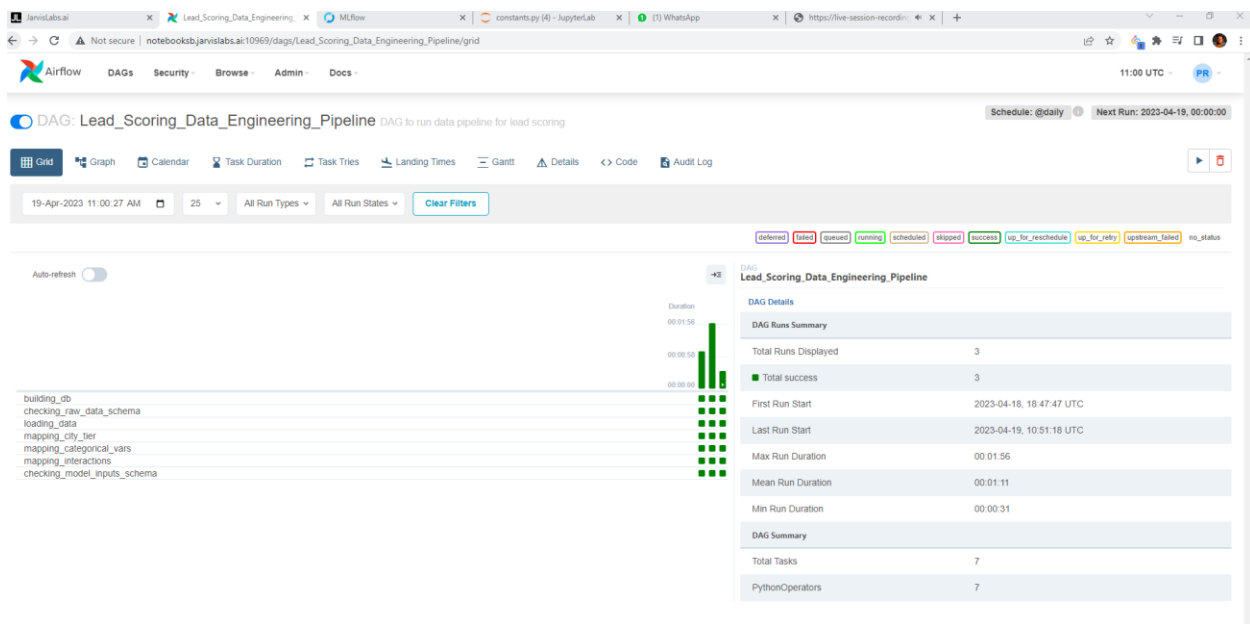
Dag Audit Log

This view displays selected events and operations that have been taken on this dag. The included and excluded events are set in the Airflow configuration, which by default remove view only actions. For a full list of events regarding this DAG, click [here](#).

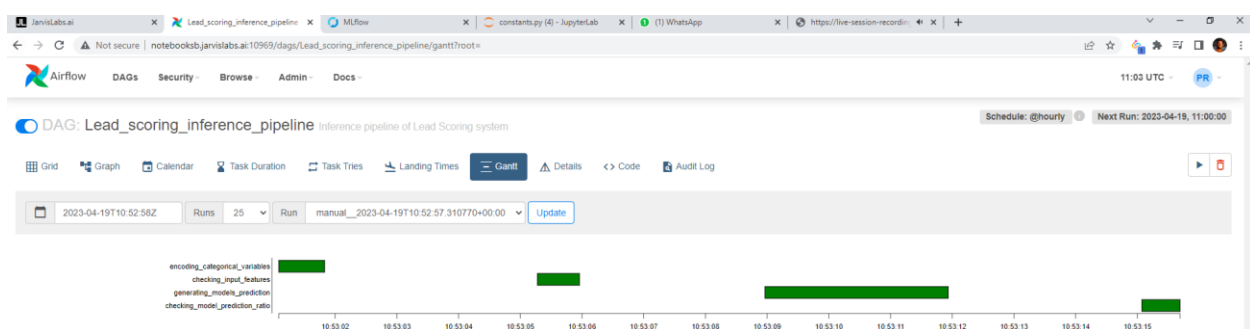
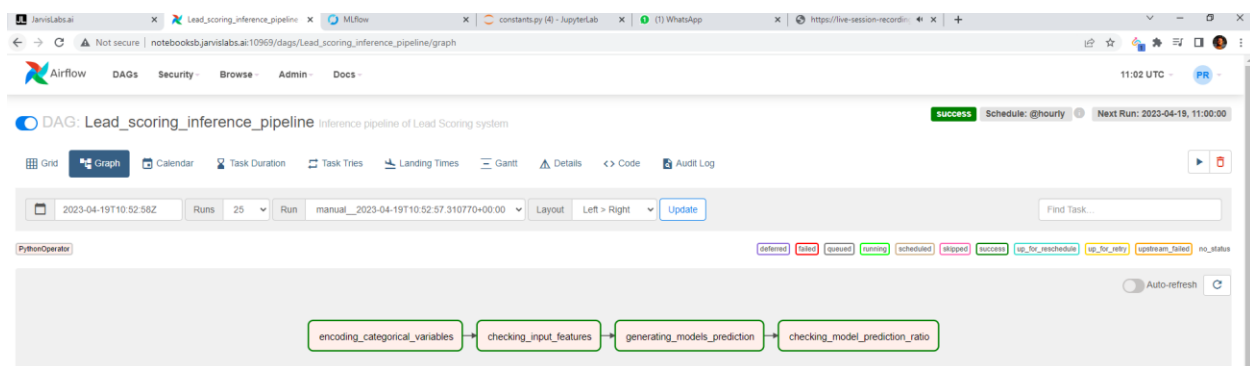
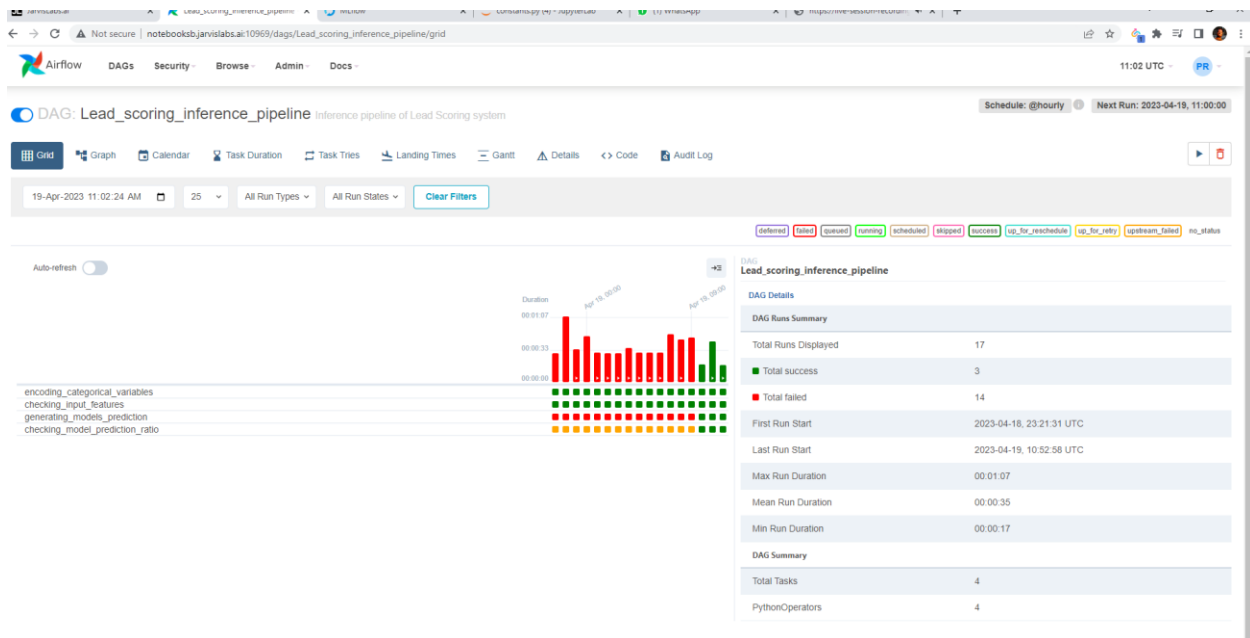
Show 100 entries

Search:

Time	Task ID	Event	Logical Date	Owner	Details
2023-04-19, 10:23:20	checking_model_prediction_ratio	success	2023-04-19 10:22:39.038519+00:00	airflow	None
2023-04-19, 10:23:13	checking_model_prediction_ratio	running	2023-04-19 10:22:39.038519+00:00	airflow	None
2023-04-19, 10:23:13	checking_model_prediction_ratio	cli_task_run	None	root	["host_name": "884056527f6d", "full_command": "T:/opt/conda/bin/airflow, 'tasks', 'run', 'Lead_scoring_inference_pipeline', 'checking_model_prediction_ratio', 'manual_2023-04-19T10:22:39.038519+00:00', '-local', '-subdir', 'DAGS_FOLDER/Lead_scoring_inference_pipeline/lead_scoring_inference_pipeline.py']"]
2023-04-19, 10:23:12	checking_model_prediction_ratio	cli_task_run	None	root	["host_name": "884056527f6d", "full_command": "T:/opt/conda/bin/airflow, 'tasks', 'run', 'Lead_scoring_inference_pipeline', 'checking_model_prediction_ratio', 'manual_2023-04-19T10:22:39.038519+00:00', '-local', '-subdir', 'DAGS_FOLDER/Lead_scoring_inference_pipeline/lead_scoring_inference_pipeline.py']"]
2023-04-19, 10:23:09	generating_models_prediction	success	2023-04-19 10:22:39.038519+00:00	airflow	None
2023-04-19, 10:22:59	generating_models_prediction	running	2023-04-19 10:22:39.038519+00:00	airflow	None
2023-04-19, 10:22:59	generating_models_prediction	cli_task_run	None	root	["host_name": "884056527f6d", "full_command": "T:/opt/conda/bin/airflow, 'tasks', 'run', 'Lead_scoring_inference_pipeline', 'generating_models_prediction', 'manual_2023-04-19T10:22:39.038519+00:00', '-local', '-subdir', 'DAGS_FOLDER/Lead_scoring_inference_pipeline/lead_scoring_inference_pipeline.py']"]
2023-04-19, 10:22:57	generating_models_prediction	cli_task_run	None	root	["host_name": "884056527f6d", "full_command": "T:/opt/conda/bin/airflow, 'tasks', 'run', 'Lead_scoring_inference_pipeline', 'generating_models_prediction', 'manual_2023-04-19T10:22:39.038519+00:00', '-local', '-subdir', 'DAGS_FOLDER/Lead_scoring_inference_pipeline/lead_scoring_inference_pipeline.py']"]
2023-04-19, 10:22:56	checking_input_features	success	2023-04-19 10:22:39.038519+00:00	airflow	None



Show	100	entries	Search	<input type="text"/>	
Time	Task ID	Event	Logical Date	Owner	Details
2023-04-19, 11:00:59	checking_raw_data_schema	task	2023-04-19 10:51:15.153684+00:00	pkir	["dag_id", "Lead_Scoring_Data_Engineering_Pipeline", "task_id", "checking_raw_data_schema", "execution_date", "2023-04-19T10:51:15.153684+00:00"]
2023-04-19, 11:00:52	checking_raw_data_schema	rendered_templates	2023-04-19 10:51:15.153684+00:00	pkir	["dag_id", "Lead_Scoring_Data_Engineering_Pipeline", "task_id", "checking_raw_data_schema", "execution_date", "2023-04-19T10:51:15.153684+00:00"]
2023-04-19, 10:51:49	checking_model_inputs_schema	running	2023-04-19 10:51:15.153684+00:00	airflow	None
2023-04-19, 10:51:49	checking_model_inputs_schema	cli_task_run	None	root	["host_name", "88405d52765", "full_command", "['opticonda/bin/airflow', 'tasks', 'run', 'Lead_Scoring_Data_Engineering_Pipeline', 'checking_model_inputs_schema', 'manual_2023-04-19T10:51:15.153684+00:00', '-local', '-subdir', 'DAGS_FOLDER/Lead_scoring_data_pipeline/pynb_checkpoints/lead_scoring_data_pipeline-checkpoint.py']"]
2023-04-19, 10:51:49	checking_model_inputs_schema	success	2023-04-19 10:51:15.153684+00:00	airflow	None
2023-04-19, 10:51:48	checking_model_inputs_schema	cli_task_run	None	root	["host_name", "88405d52765", "full_command", "['opticonda/bin/airflow', 'tasks', 'run', 'Lead_Scoring_Data_Engineering_Pipeline', 'checking_model_inputs_schema', 'manual_2023-04-19T10:51:15.153684+00:00', '-local', '-subdir', 'DAGS_FOLDER/Lead_scoring_data_pipeline/pynb_checkpoints/lead_scoring_data_pipeline-checkpoint.py']"]
2023-04-19, 10:51:47	mapping_interactions	success	2023-04-19 10:51:15.153684+00:00	airflow	None
2023-04-19, 10:51:42	mapping_interactions	running	2023-04-19 10:51:15.153684+00:00	airflow	None
2023-04-19, 10:51:42	mapping_interactions	cli_task_run	None	root	["host_name", "88405d52765", "full_command", "['opticonda/bin/airflow', 'tasks', 'run', 'Lead_Scoring_Data_Engineering_Pipeline', 'mapping_interactions', 'manual_2023-04-19T10:51:15.153684+00:00', '-local', '-subdir', 'DAGS_FOLDER/Lead_scoring_data_pipeline/pynb_checkpoints/lead_scoring_data_pipeline-checkpoint.py']"]



← → ↻ ⚠ Not secure | notebooks.jarvislabs.ai:10969/audit_log?dag_id=Lead_scoring_inference_pipeline&root=

Airflow

DAGs

Security

Browse

Admin

Docs

11:53 UTC

PR

DAG: Lead_scoring_inference_pipeline Inference pipeline of Lead Scoring system Schedule: @hourly

Grid

Graph

Calendar

Task Duration

Task Tries

Landing Times

Garnt

Details

<> Code

Audit Log

▶

0

Dag Audit Log

This view displays selected events and operations that have been taken on this dag. The included and excluded events are set in the Airflow configuration, which by default remove view only actions. For a full list of events regarding this DAG, click here.

Show 100 entries Search:

Time	Task ID	Event	Logical Date	Owner	Details
2023-04-19, 11:00:18	checking_model_prediction_ratio	success	2023-04-19 10:00:00+00:00	airflow	None
2023-04-19, 11:00:17	checking_model_prediction_ratio	running	2023-04-19 10:00:00+00:00	airflow	None
2023-04-19, 11:00:17	checking_model_prediction_ratio	cli_task_run	None	root	["host_name": "884056527f6d", "full_command": "T:/opt/conda/bin/airflow, 'tasks', 'run', 'Lead_scoring_inference_pipeline', 'checking_model_prediction_ratio', 'scheduled__2023-04-19T10:00:00+00:00', '-local', '-subdir', 'DAGS_FOLDER/Lead_scoring_inference_pipeline/lead_scoring_inference_pipeline.py']"]
2023-04-19, 11:00:16	checking_model_prediction_ratio	cli_task_run	None	root	["host_name": "884056527f6d", "full_command": "T:/opt/conda/bin/airflow, 'tasks', 'run', 'Lead_scoring_inference_pipeline', 'checking_model_prediction_ratio', 'scheduled__2023-04-19T10:00:00+00:00', '-local', '-subdir', 'DAGS_FOLDER/Lead_scoring_inference_pipeline/lead_scoring_inference_pipeline.py']"]
2023-04-19, 11:00:14	generating_models_prediction	success	2023-04-19 10:00:00+00:00	airflow	None
2023-04-19, 11:00:12	generating_models_prediction	cli_task_run	None	root	["host_name": "884056527f6d", "full_command": "T:/opt/conda/bin/airflow, 'tasks', 'run', 'Lead_scoring_inference_pipeline', 'generating_models_prediction', 'scheduled__2023-04-19T10:00:00+00:00', '-local', '-subdir', 'DAGS_FOLDER/Lead_scoring_inference_pipeline/lead_scoring_inference_pipeline.py']"]
2023-04-19, 11:00:11	generating_models_prediction	running	2023-04-19 10:00:00+00:00	airflow	None
2023-04-19, 11:00:10	generating_models_prediction	cli_task_run	None	root	["host_name": "884056527f6d", "full_command": "T:/opt/conda/bin/airflow, 'tasks', 'run', 'Lead_scoring_inference_pipeline', 'generating_models_prediction', 'scheduled__2023-04-19T10:00:00+00:00', '-local', '-subdir', 'DAGS_FOLDER/Lead_scoring_inference_pipeline/lead_scoring_inference_pipeline.py']"]
2023-04-19, 11:00:09	checking_input_features	success	2023-04-19 10:00:00+00:00	airflow	None
2023-04-19, 11:00:08	checking_input_features	running	2023-04-19 10:00:00+00:00	airflow	None
2023-04-19, 11:00:08	checking_input_features	cli_task_run	None	root	["host_name": "884056527f6d", "full_command": "T:/opt/conda/bin/airflow, 'tasks', 'run', 'Lead_scoring_inference_pipeline', 'checking_input_features', 'scheduled__2023-04-19T10:00:00+00:00', '-local', '-subdir', 'DAGS_FOLDER/Lead_scoring_inference_pipeline/lead_scoring_inference_pipeline.py']"]
2023-04-19, 11:00:06	checking_input_features	cli_task_run	None	root	["host_name": "884056527f6d", "full_command": "T:/opt/conda/bin/airflow, 'tasks', 'run', 'Lead_scoring_inference_pipeline', 'checking_input_features', 'scheduled__2023-04-19T10:00:00+00:00', '-local', '-subdir', 'DAGS_FOLDER/Lead_scoring_inference_pipeline/lead_scoring_inference_pipeline.py']"]