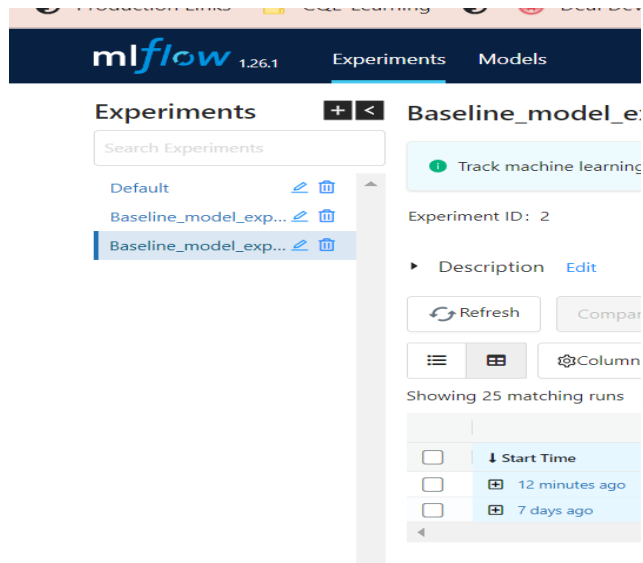


# ML ops Assignment

## Screenshot of ML Flow and Airflow

### 1. List of all Experiments



### 2. All Experiments Before dropping Features.

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"

Showing 23 matching runs

	Start Time	Duration	Run Name	User	Source	Version	Models
<input type="checkbox"/>	22 minutes ago		Session Initi...	root	ipykernel...	-	-
<input type="checkbox"/>	17 minutes ago		Light Gradie...	root	ipykernel...	-	sklearn
<input type="checkbox"/>	19 minutes ago		Naive Bayes	root	ipykernel...	-	sklearn
<input type="checkbox"/>	19 minutes ago		Linear Discr...	root	ipykernel...	-	sklearn
<input type="checkbox"/>	19 minutes ago		Ridge Classif...	root	ipykernel...	-	sklearn
<input type="checkbox"/>	19 minutes ago		Logistic Reg...	root	ipykernel...	-	sklearn
<input type="checkbox"/>	19 minutes ago		Decision Tre...	root	ipykernel...	-	sklearn
<input type="checkbox"/>	19 minutes ago		Extra Trees C...	root	ipykernel...	-	sklearn
<input type="checkbox"/>	19 minutes ago		Random For...	root	ipykernel...	-	sklearn
<input type="checkbox"/>	19 minutes ago		Extreme Gra...	root	ipykernel...	-	sklearn
<input type="checkbox"/>	19 minutes ago		Light Gradie...	root	ipykernel...	-	sklearn
<input type="checkbox"/>	7 days ago		Session Initi...	root	ipykernel...	-	-
<input type="checkbox"/>	8 days ago		Session Initi...	root	ipykernel...	-	-

3. One of the model's parameters.

Not secure | notebooksc.jarvislabs.ai:10964/#/experiments/

Production Links CQE-Learning Deal Devils - ISS > O... My Times

Description Edit

Parameters (20)

Metrics (8)

Name	Value
AUC	0.816
Accuracy	0.736
F1	0.761
Kappa	0.472
MCC	0.482
Prec.	0.701
Recall	0.833
TT	2.41

4. Experiments after Dropping features .

Experiment ID: 2

Description Edit

Refresh Compare Delete Download CSV Start Time All time

Columns Only show differences Search Filter Clear

Showing 25 matching runs

	Start Time	Duration	Run Name	User	Source	Version	Models	Metrics
								AUC Accuracy F1
	22 minutes ago		Session Init...	root	ipykernel...	-	-	- - -
	16 minutes ago		Light Gradie...	root	ipykernel...	-	sklearn	0.819 0.738 0.76
	20 minutes ago		Light Gradie...	root	ipykernel...	-	sklearn	0.821 0.738 0.763
	20 minutes ago		Naive Bayes	root	ipykernel...	-	sklearn	0.734 0.675 0.726
	20 minutes ago		Linear Discrim...	root	ipykernel...	-	sklearn	0.773 0.701 0.728
	20 minutes ago		Ridge Classif...	root	ipykernel...	-	sklearn	0 0.701 0.728
	20 minutes ago		Logistic Reg...	root	ipykernel...	-	sklearn	0.784 0.71 0.74
	20 minutes ago		Decision Tre...	root	ipykernel...	-	sklearn	0.816 0.736 0.757
	20 minutes ago		Extra Tree C...	root	ipykernel...	-	sklearn	0.817 0.736 0.757
	20 minutes ago		Random For...	root	ipykernel...	-	sklearn	0.818 0.738 0.759
	20 minutes ago		Light Gradie...	root	ipykernel...	-	sklearn	0.821 0.738 0.763
	20 minutes ago		Extreme Gra...	root	ipykernel...	-	-	- - -
	7 days ago		Session Init...	root	ipykernel...	-	-	- - -

5. One of Model's Parameters

Baseline\_model\_exp02 > Light Gradient Boosting Machine

Light Gradient Boosting Machine

Date: 2023-02-19 09:05:03 Source: ipykernel\_la

Status: UNFINISHED Lifecycle Stage: active

Description Edit

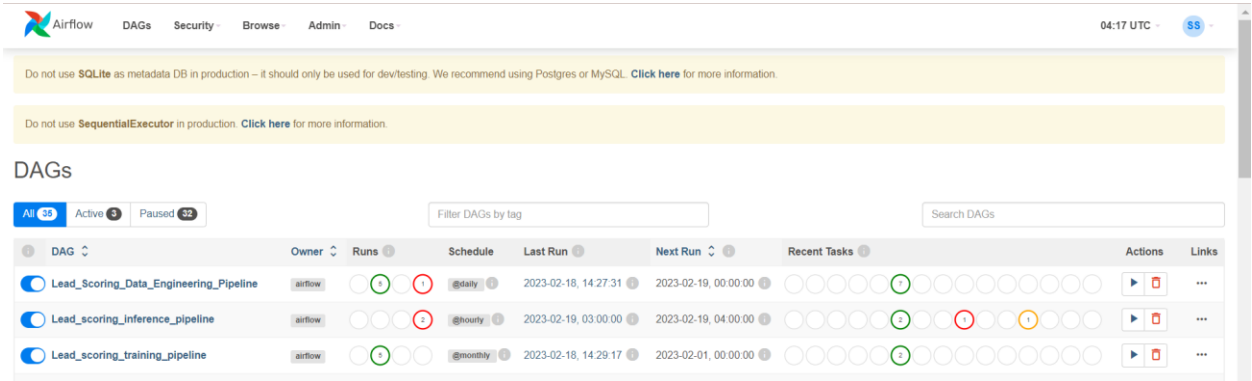
Parameters (20)

Metrics (8)

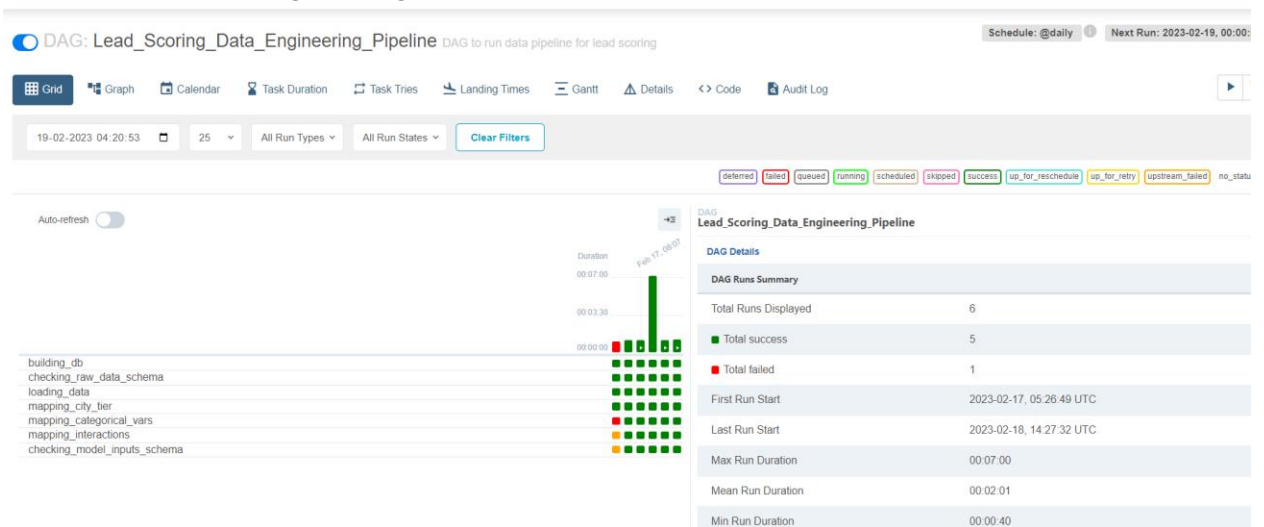
Name	Value
AUC	0.819
Accuracy	0.738
F1	0.76
Kappa	0.475
MCC	0.482
Prec.	0.704
Recall	0.826
TT	1.7

# Airflow Pipelines

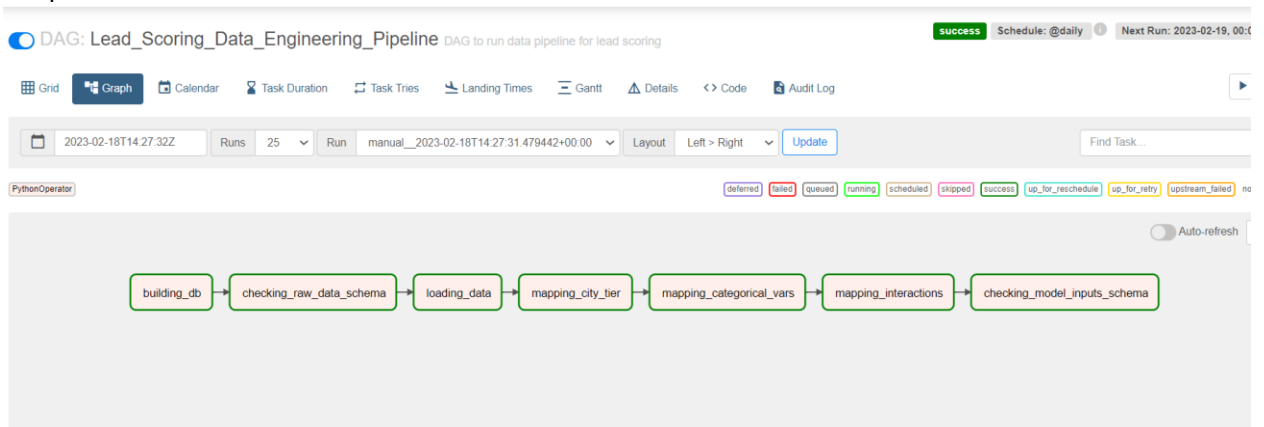
## 1. Screenshot of Airflow UI



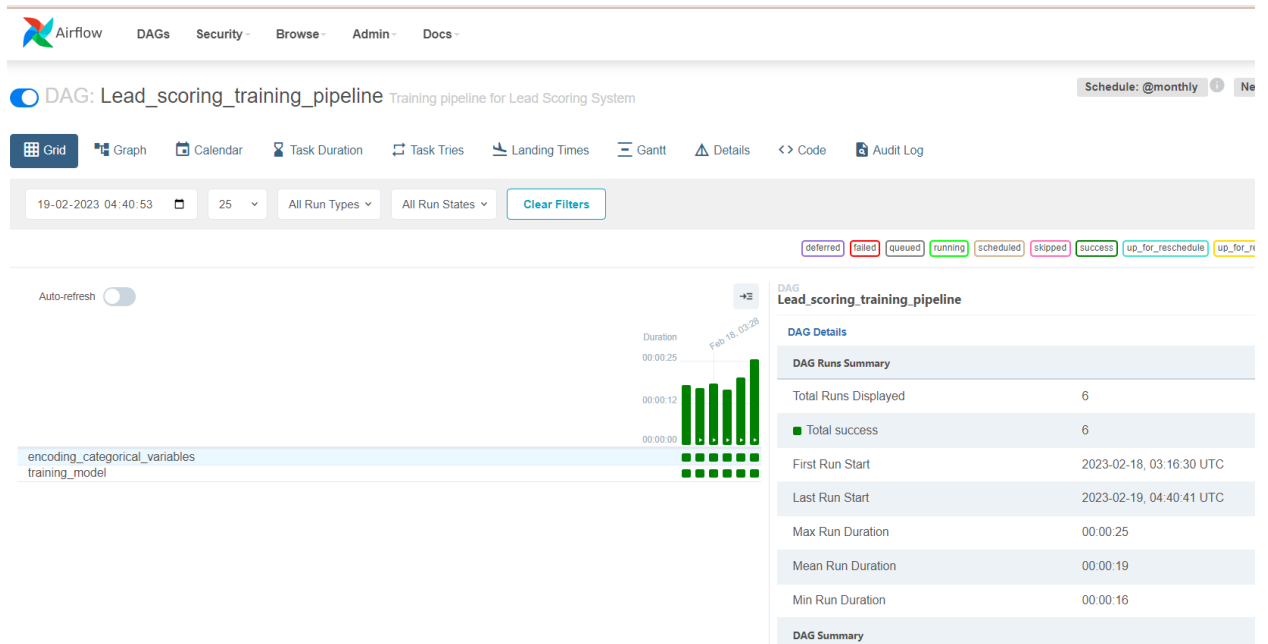
## 2. Screenshot of Data Engineering Pipeline



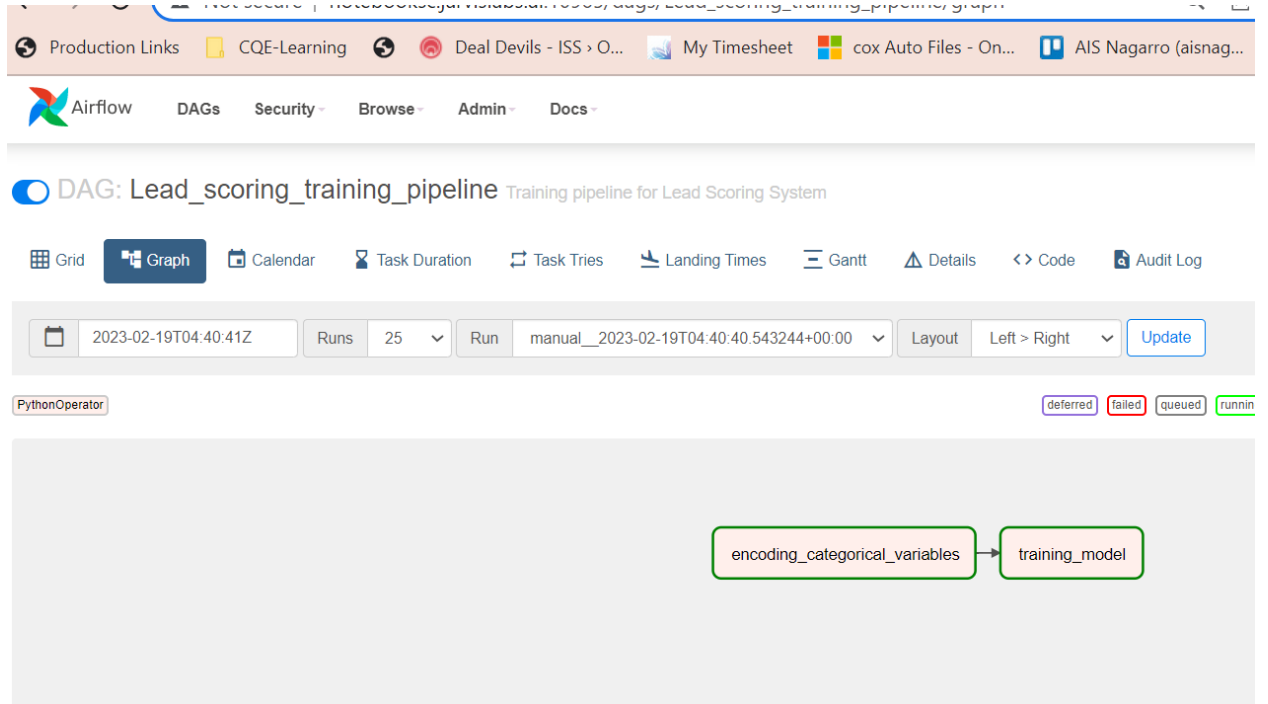
## Graph View of DAGs



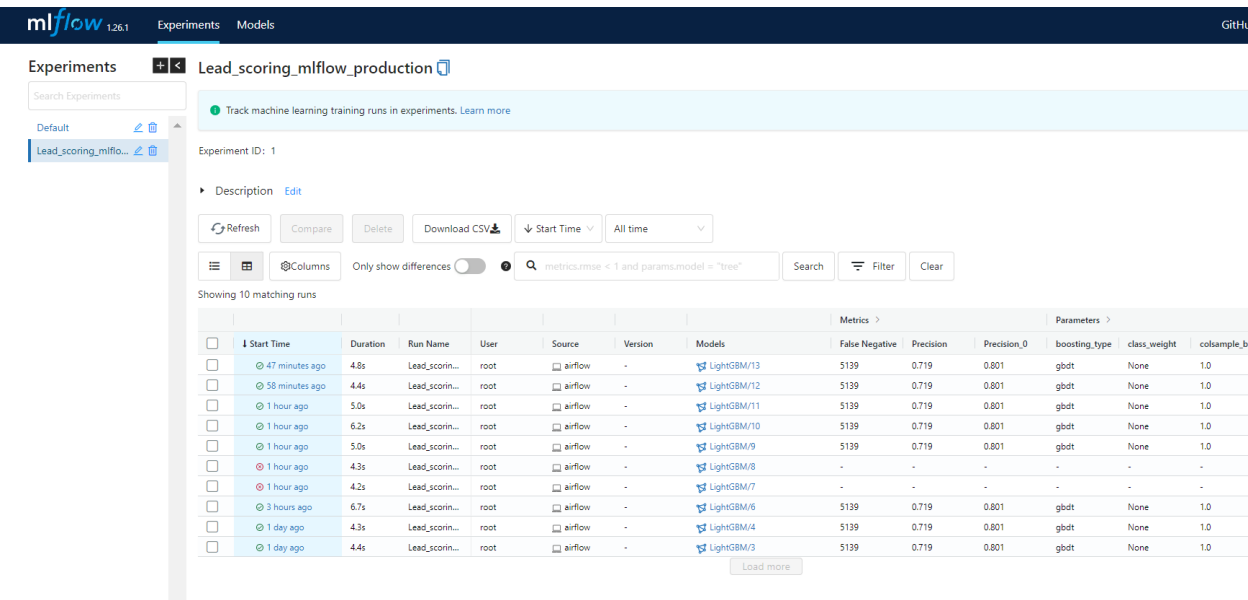
## Screenshot of Data Training Pipeline



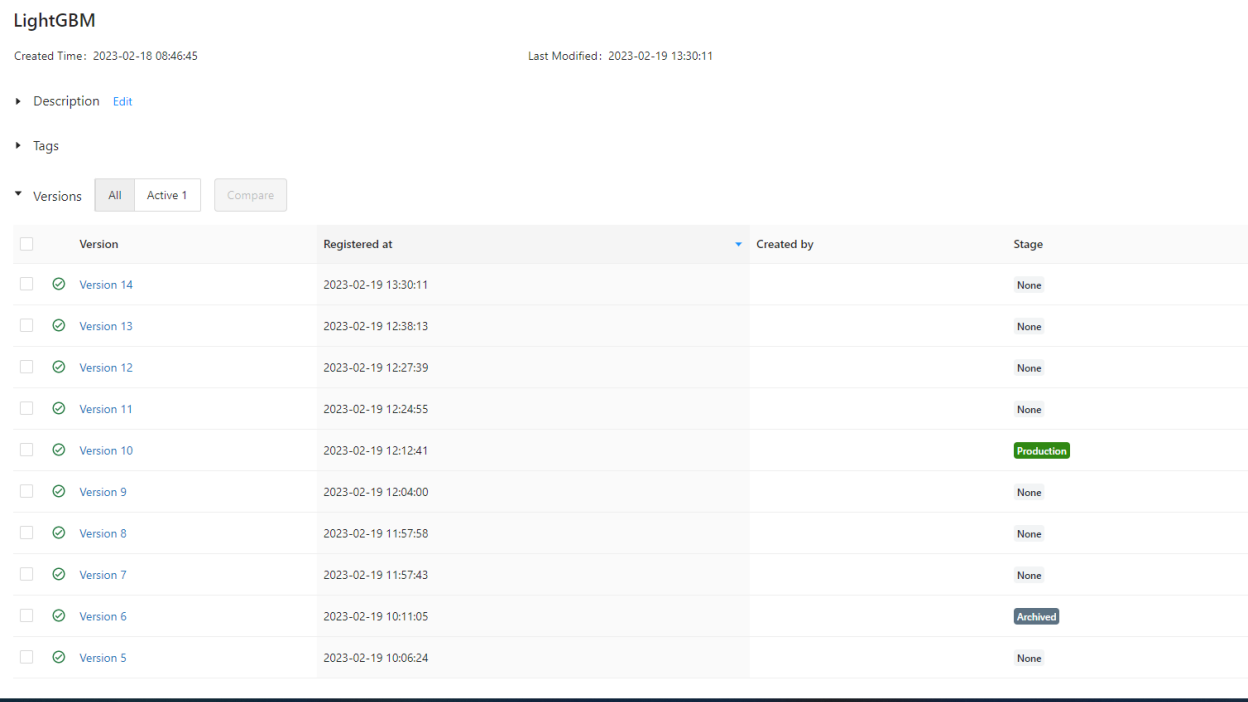
## DAGs Screenshot



# Screenshot ML Flow -Training Pipeline



## Model in Production:



mlflow1.26.1

ExperimentsModels

Registered Models

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

Create Model

Search by model name

Name	Latest Version	Staging	Production	Last Modified
LightGBM	Version 13	-	Version 10	2023-02-19 13:21:14

Artifacts:

Artifacts

Full Path: /miruns/1/2aafc448491340f8a53dbfba23833b2f/artifacts/models

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:/2aafc448491340f8a53dbfba23833b2f/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:/2aafc448491340f8a53dbfba23833b2f/models'

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

# Predict on a Pandas DataFrame.
import pandas as pd
loaded_model.predict(pd.DataFrame(data))
```

Lead\_scoring\_mlflow\_production > Lead\_scoring\_mlflow\_production

Lead\_scoring\_mlflow\_production

Date: 2023-02-19 12:12:36

Source: airflow

User: root

Duration: 6:2s

Status: FINISHED

Lifecycle Stage: active

Description [Edit](#)

Parameters (20)

Metrics (12)

Tags

Artifacts

Full Path: /miruns/1/2aafc448491340f8a53dbfba23833b2f/artifacts/models

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:

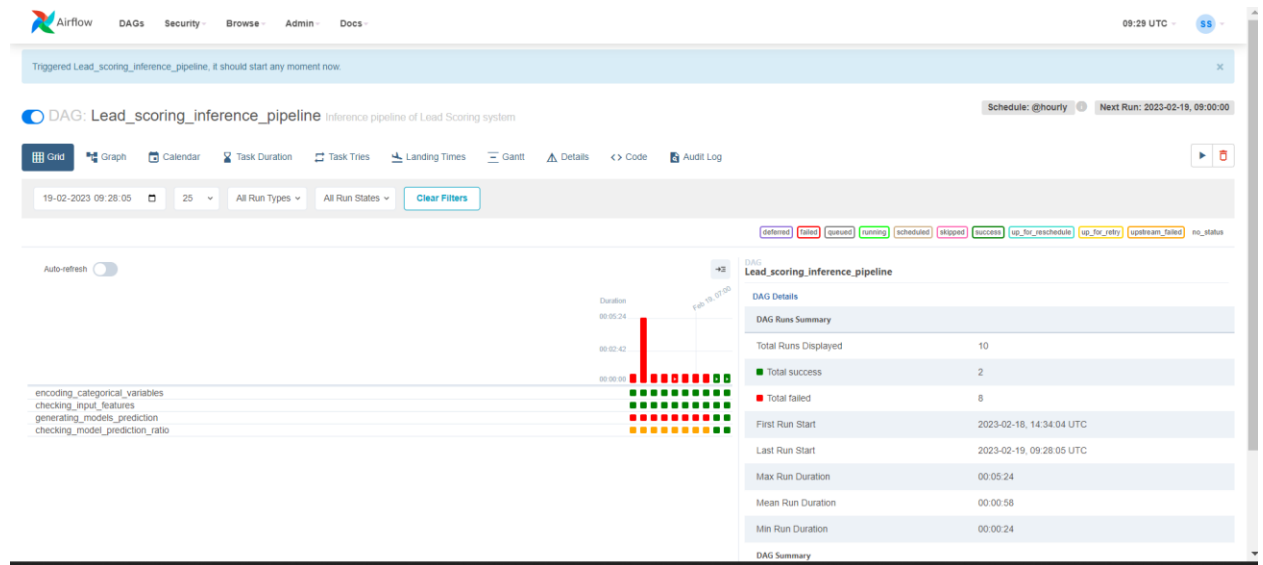
```
import mlflow
logged_model = 'runs:/2aafc448491340f8a53dbfba23833b2f/models'

# 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:

## Model Airflow-Inference Pipeline



## DAG Graph

