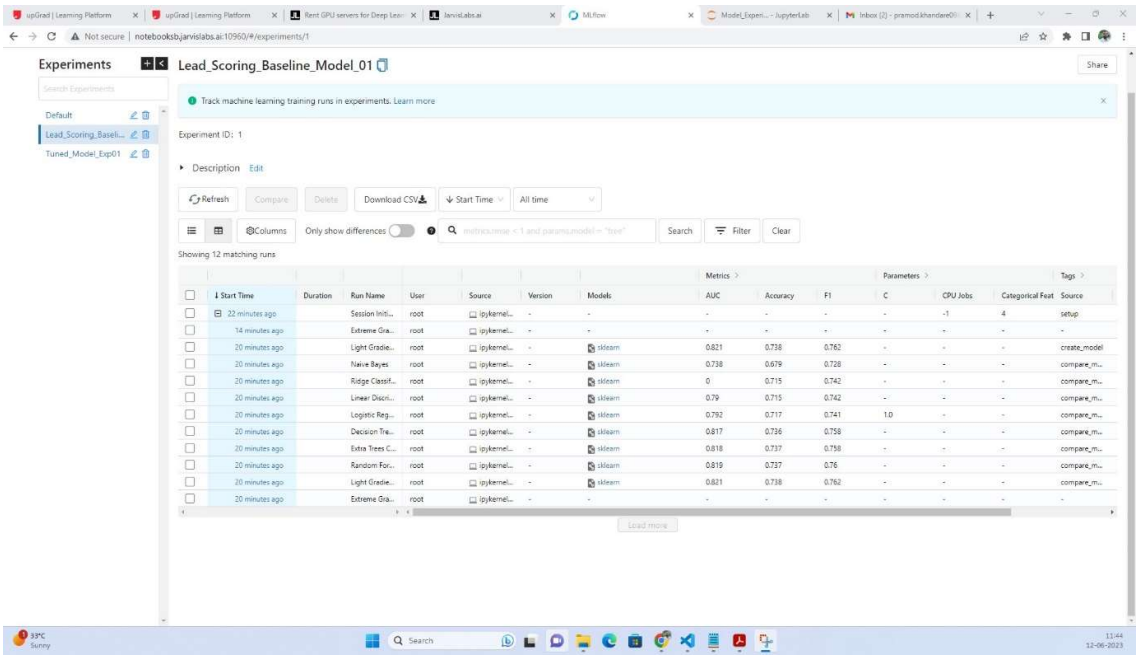
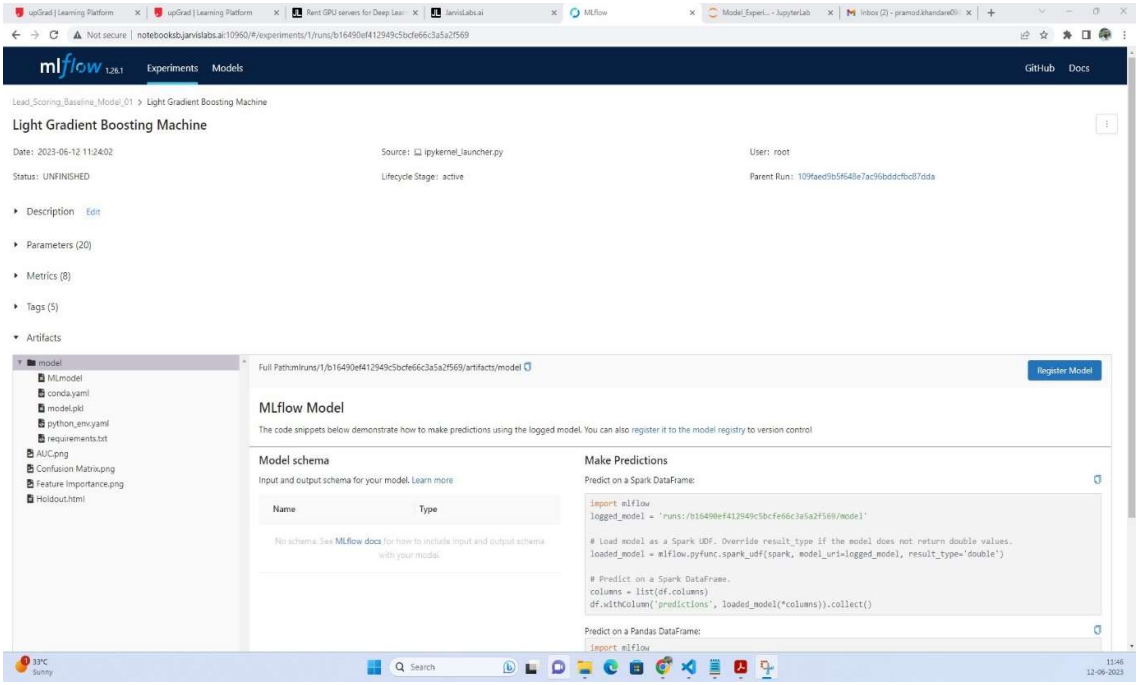
MODEL EXPERIMENTATION:

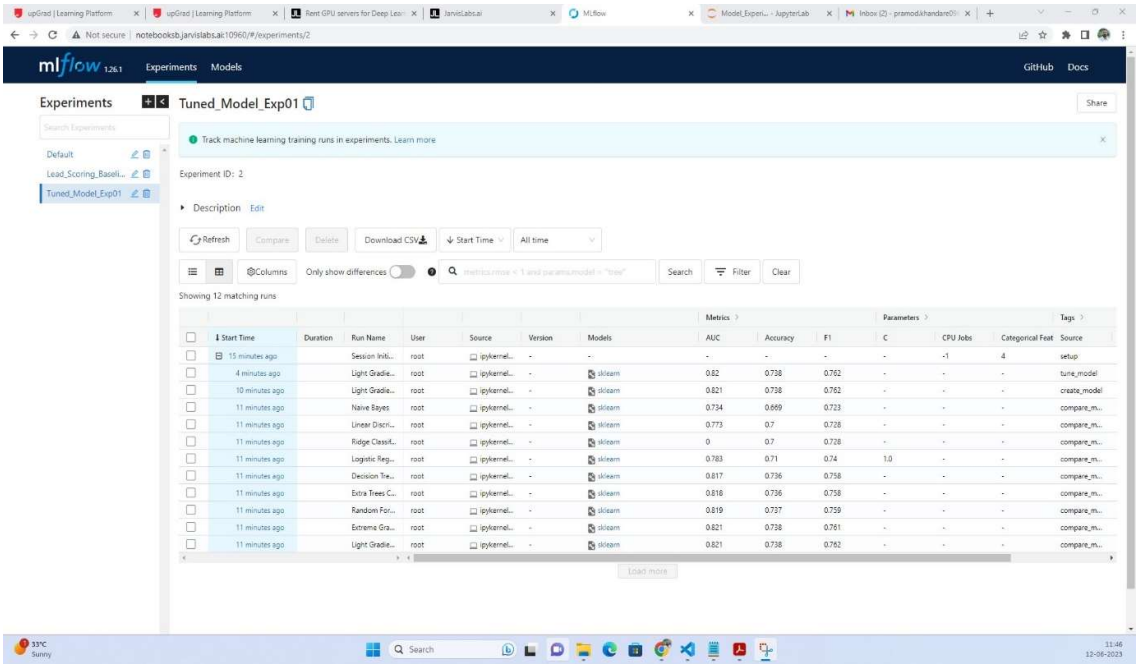
Screenshot of mlflow ui- all experiments



Screenshot of mlflow ui-- screenshot of one experiment with all the artifacts visible

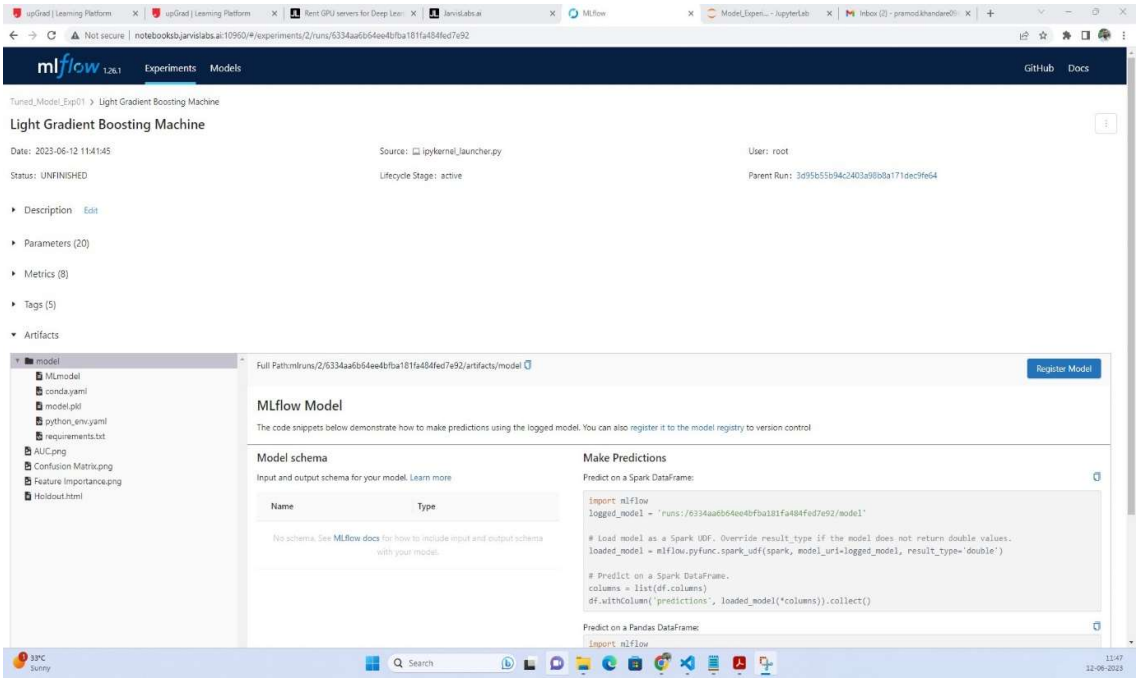


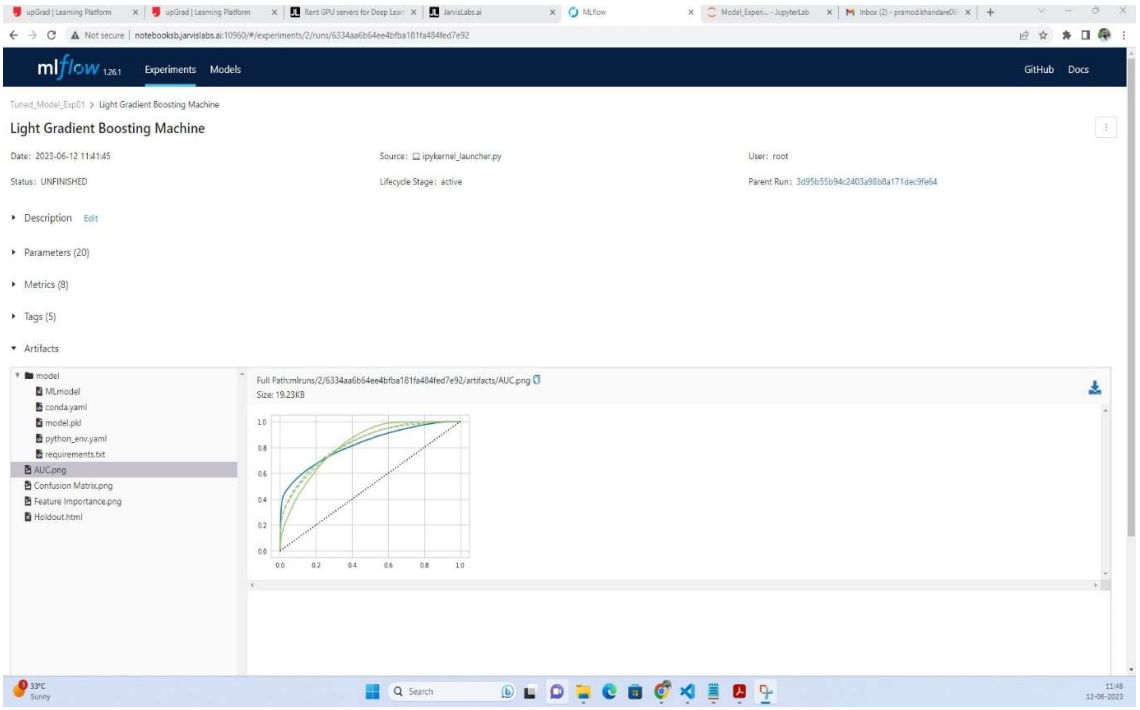
Screenshot of mlflow ui after dropping features—all experiments

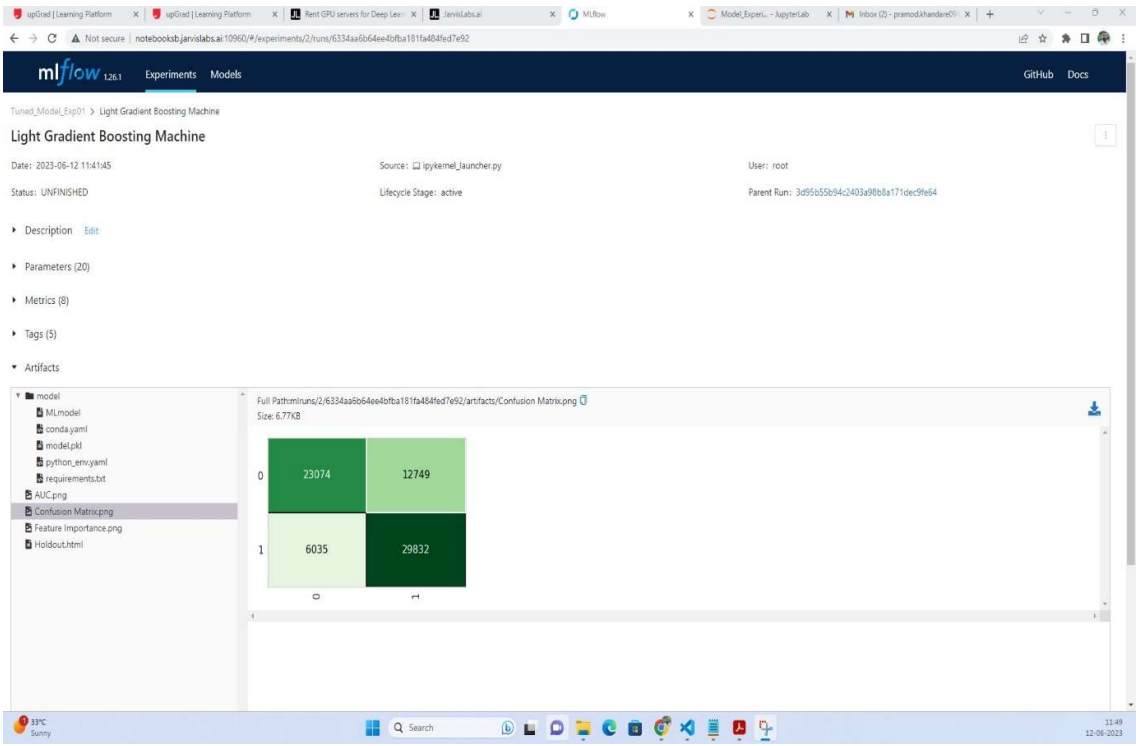


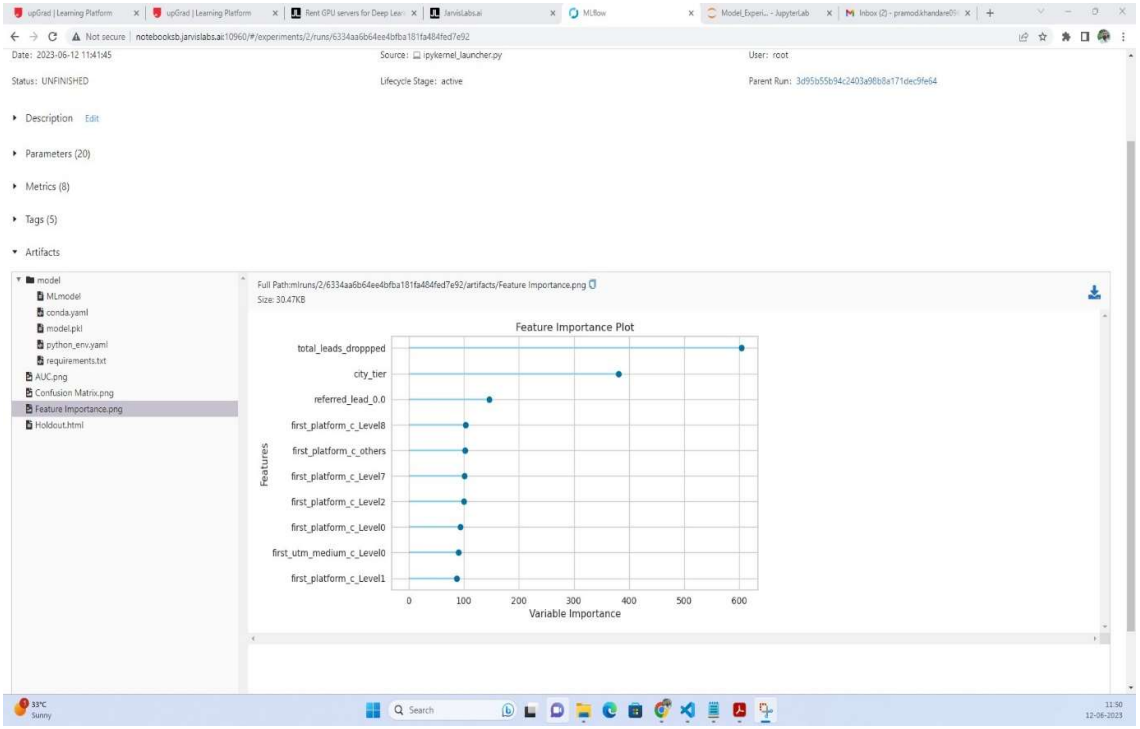
Screenshot of mlflow ui after dropping features- one experiment with all artifacts

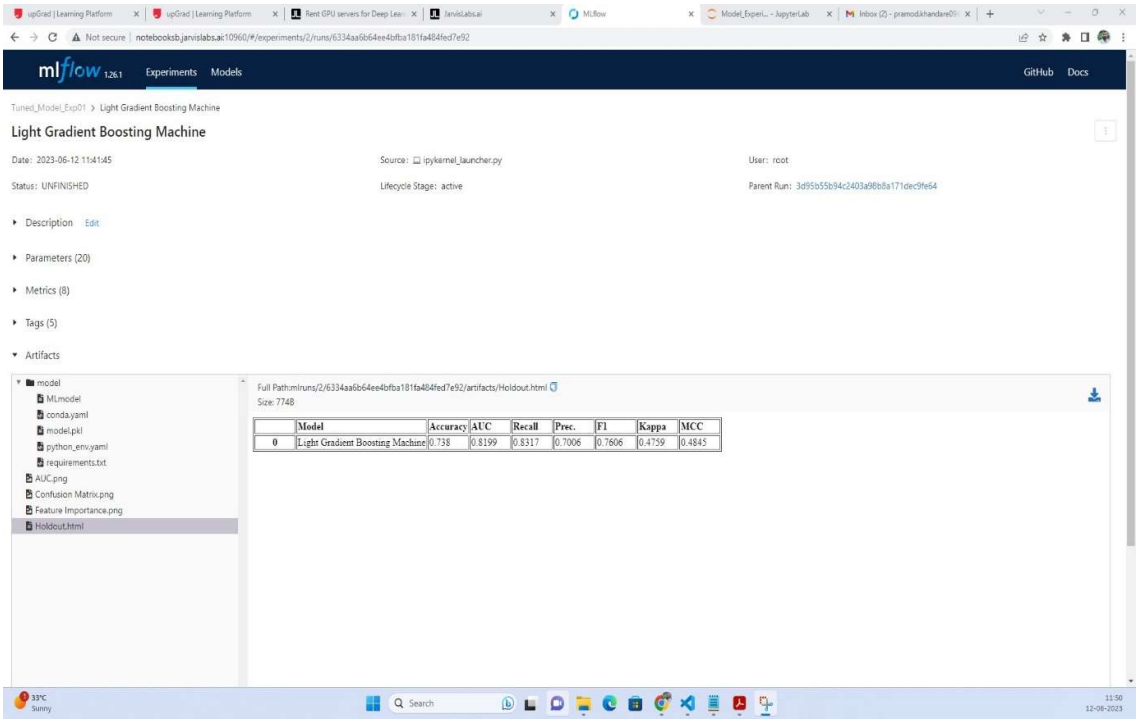
(Model, AUC, Confusion Matrix, Feature Importance, Holdout.html)




mlflow 1.26.1
Experiments
Models
GitHub
Docs
Tuned_Model_Exp01 > Light Gradient Boosting Machine
Light Gradient Boosting Machine
Date: 2023-06-12 11:41:45
Source: ipykernel_launcher.py
User: root
Status: UNFINISHED
Lifecycle Stage: active
Parent Run: 3d95b55b94c2403a98b8a171dec9fe64
Description
Edit
Parameters (20)
Metrics (8)
Tags (5)
Artifacts
model
MLmodel
conda.yaml
model.pkl
python_env.yaml
requirements.txt
AUC.png
Confusion Matrix.png
Feature Importance.png
Holdout.html
Full Path:mlruns/2/6334aa6b64ee4bfba181fa484fed7e92/artifacts/AUC.png
Size: 19.23KB


mlflow 1.26.1
Experiments
Models
GitHub
Docs
Tuned_Model_Exp01 > Light Gradient Boosting Machine
Light Gradient Boosting Machine
Date: 2023-06-12 11:41:45
Source: ipykernel_launcher.py
User: root
Status: UNFINISHED
Lifecycle Stage: active
Parent Run: 3d95b55b94c2403a98b8a171dec9fe64
Description
Edit
Parameters (20)
Metrics (8)
Tags (5)
Artifacts
model
MLmodel
conda.yaml
model.pkl
python_env.yaml
requirements.txt
AUC.png
Confusion Matrix.png
Feature Importance.png
Holdout.html
Full Path:mlruns/2/6334aa6b64ee4bfba181fa484fed7e92/artifacts/Confusion Matrix.png
Size: 6.77KB
23074
12749
6035
29832

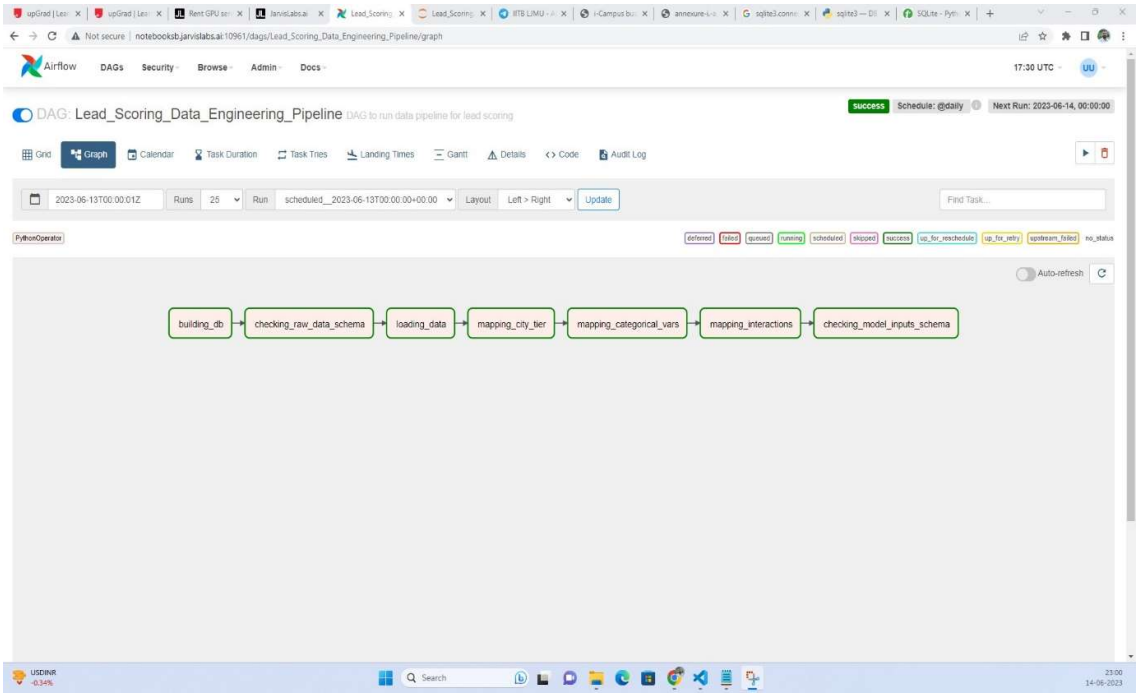




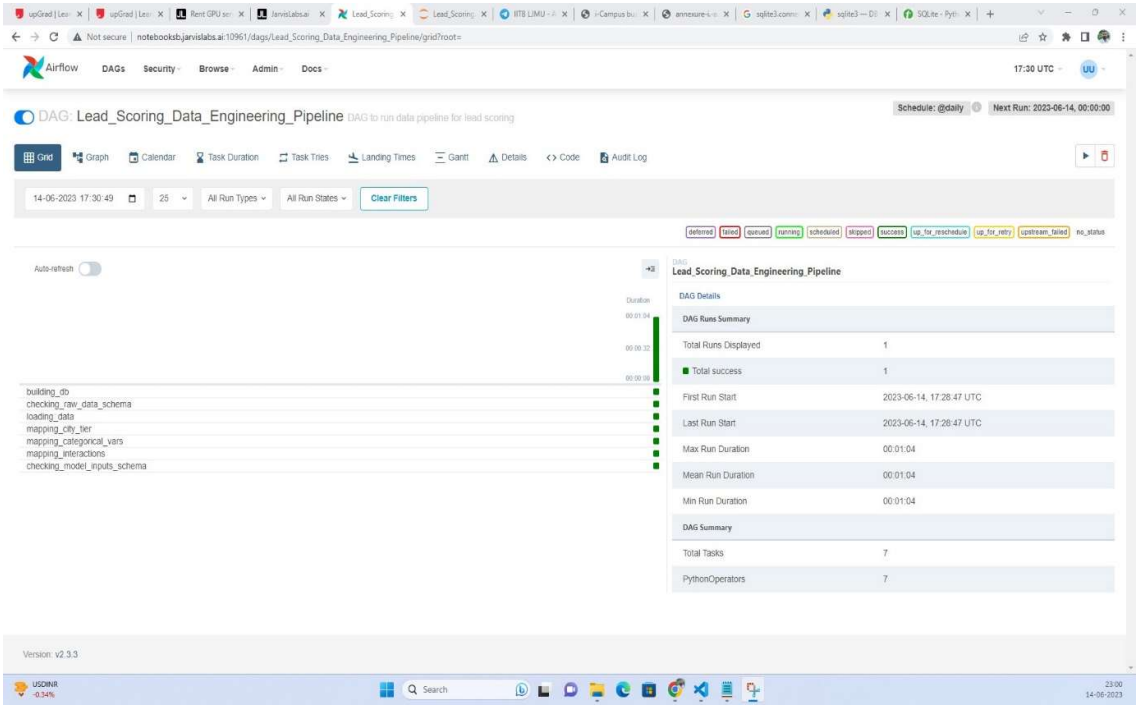
	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	Light Gradient Boosting Machine	0.738	0.8199	0.8317	0.7006	0.7606	0.4759	0.4845

DATA PIPELINE

Screenshot of successful execution Airflow DAG in graph

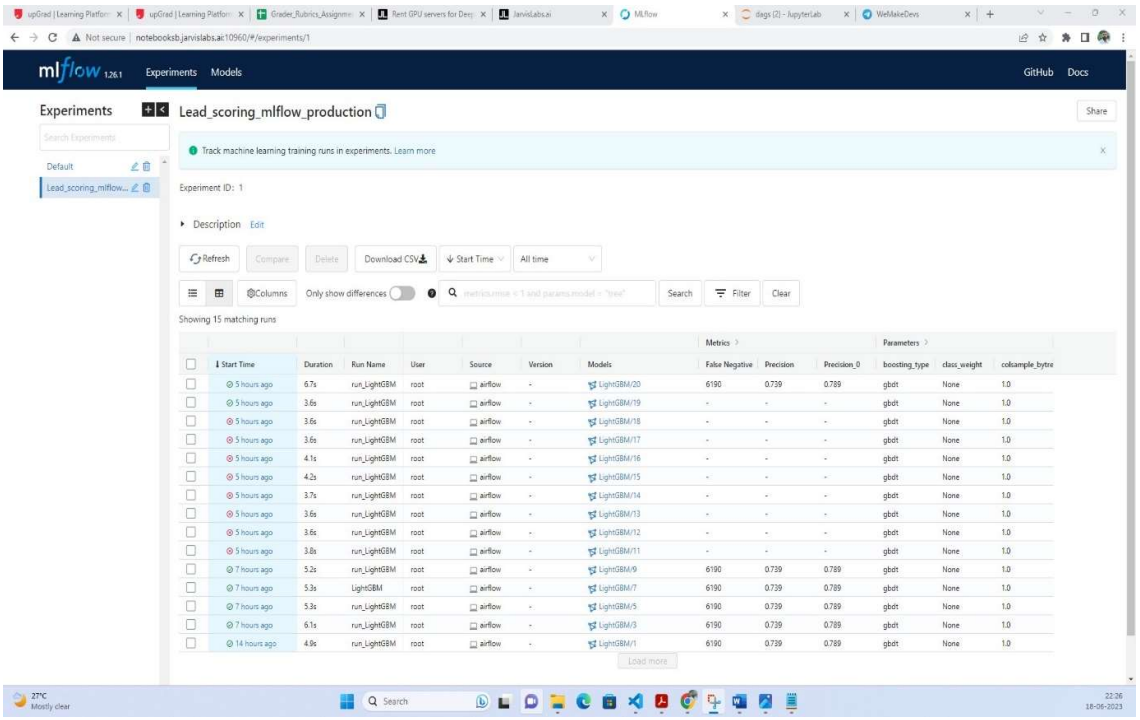


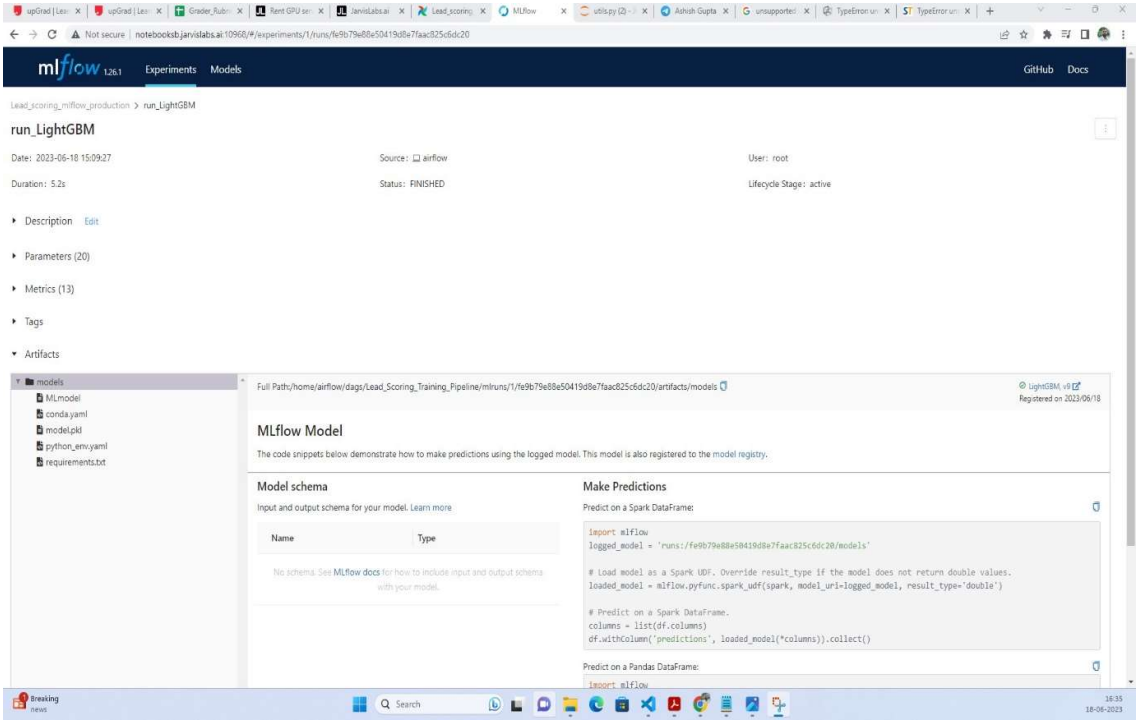
Screenshot of Airflow UI grid



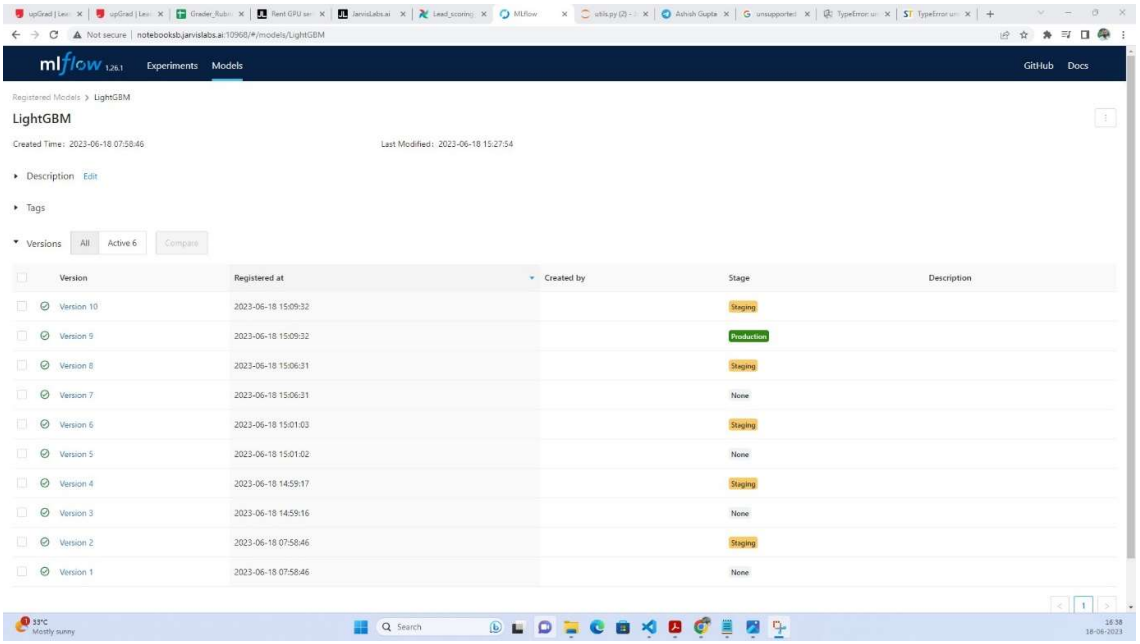
TRAINING PIPELINE

screenshot of experiments with all the artifacts visible and also metric AUC

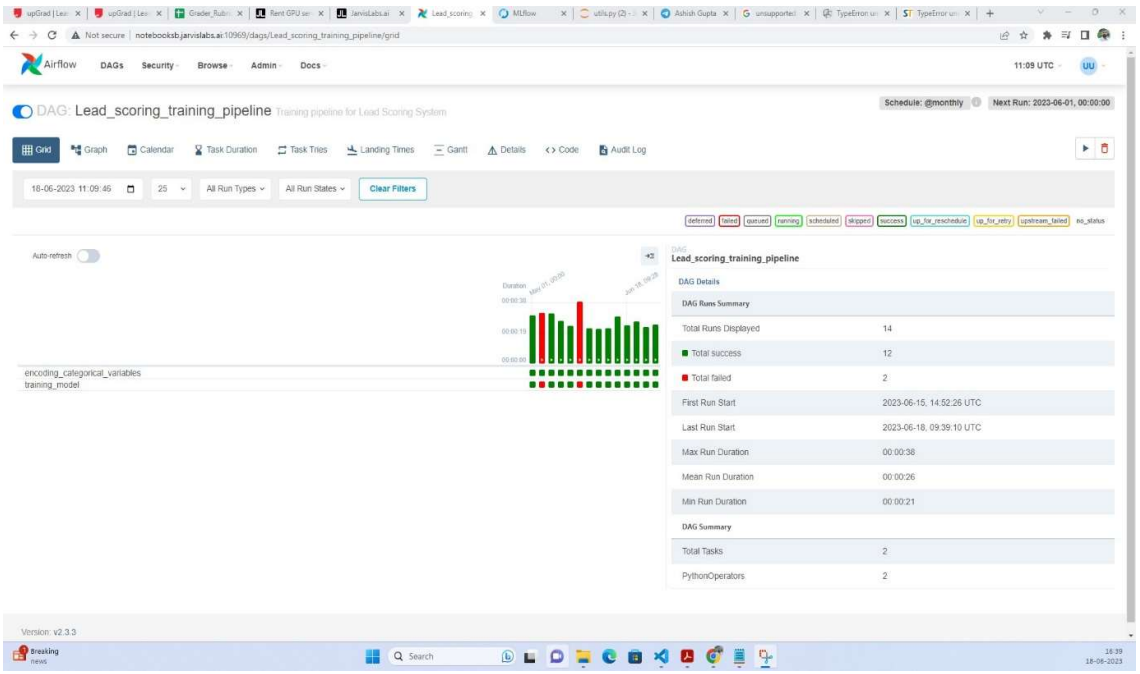




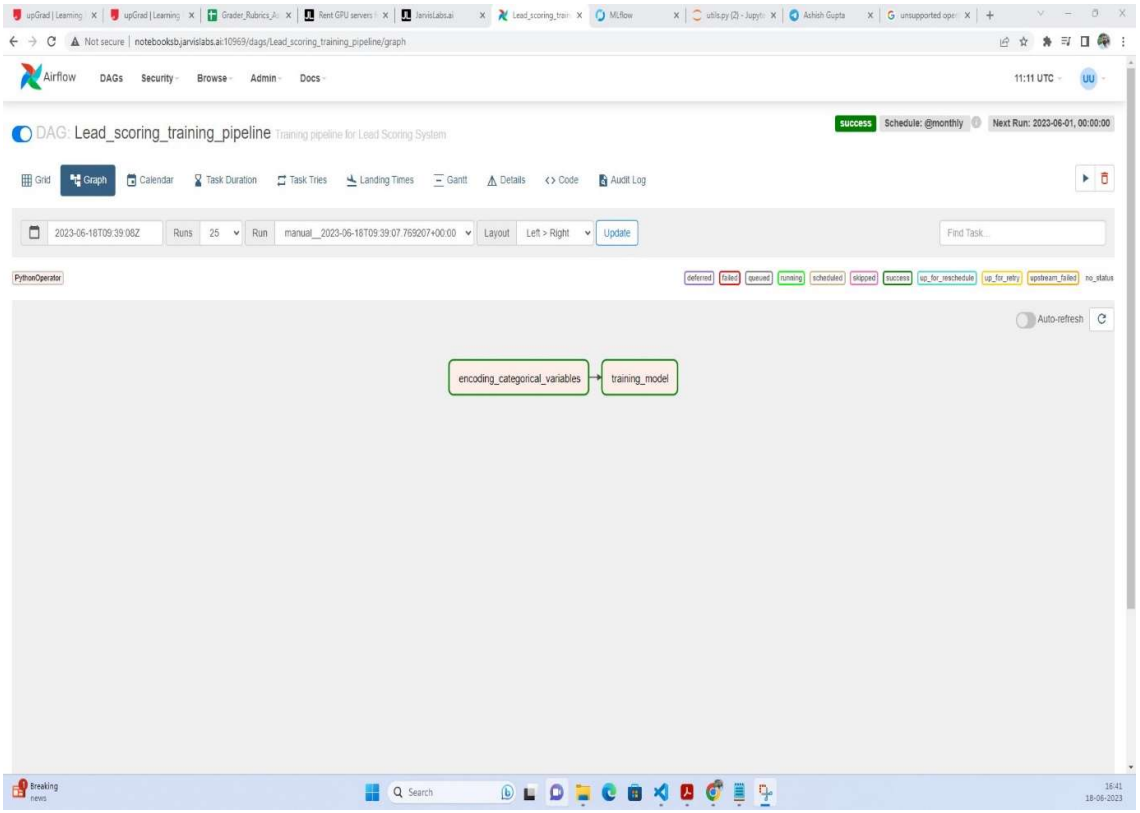
Screenshot of model registry with model name and stage as 'Production'



Screenshot of Airflow UI grid

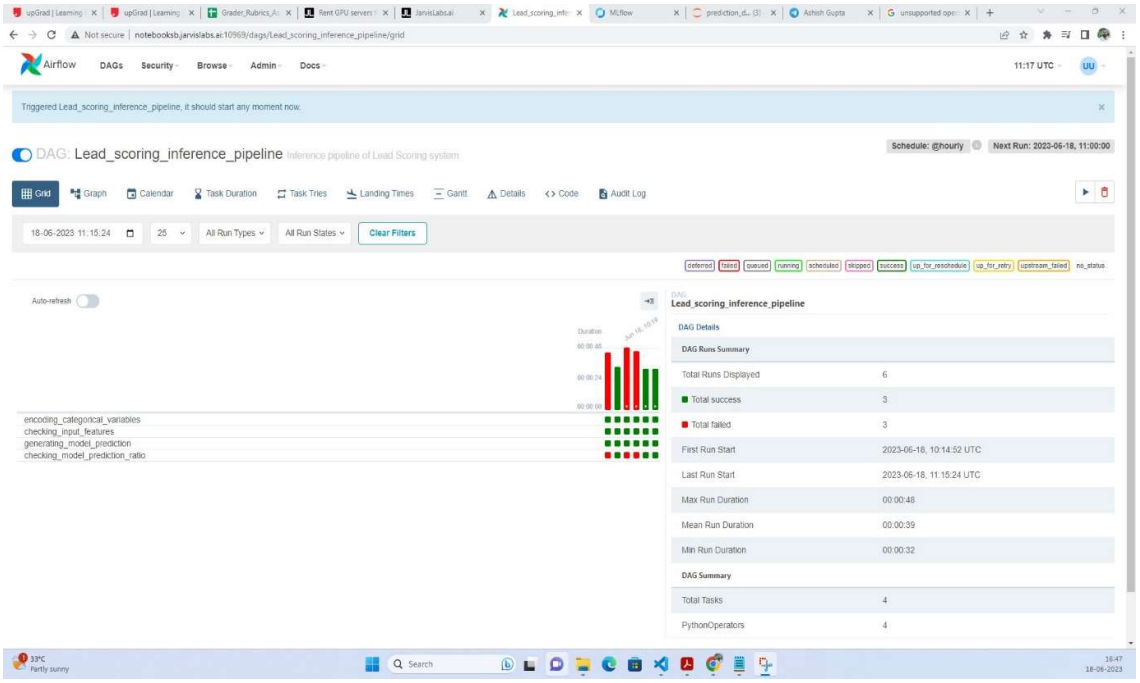


Screenshot of successful execution Airflow DAG in graph

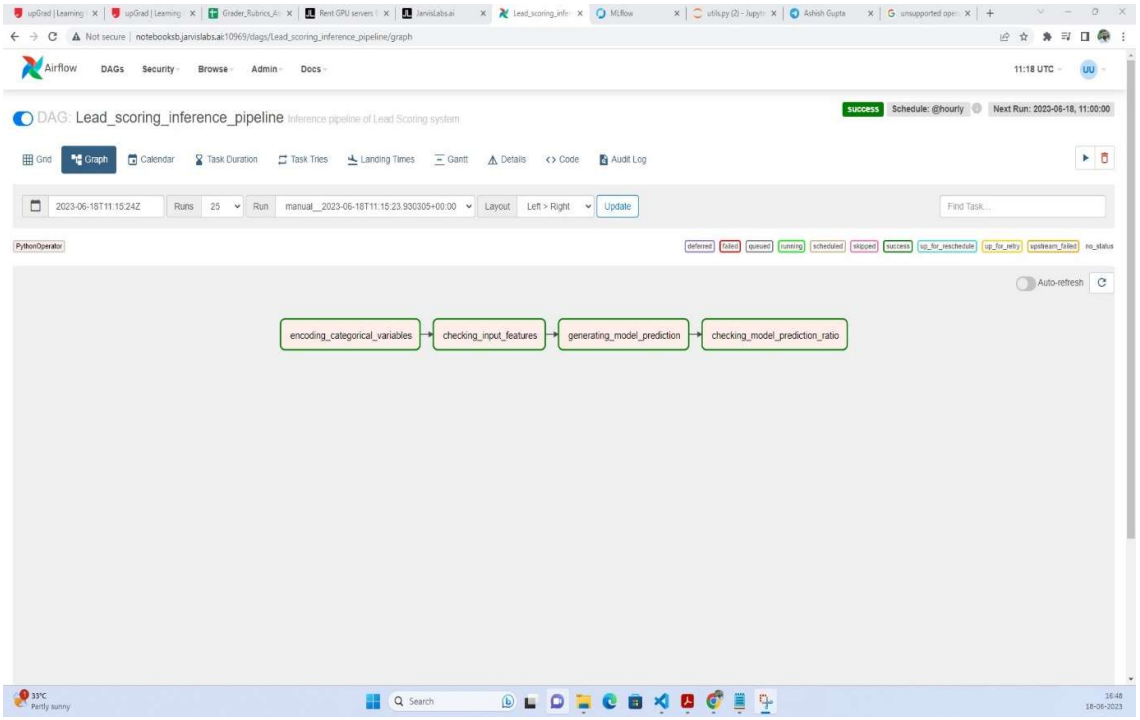


INFERENCE PIPELINE

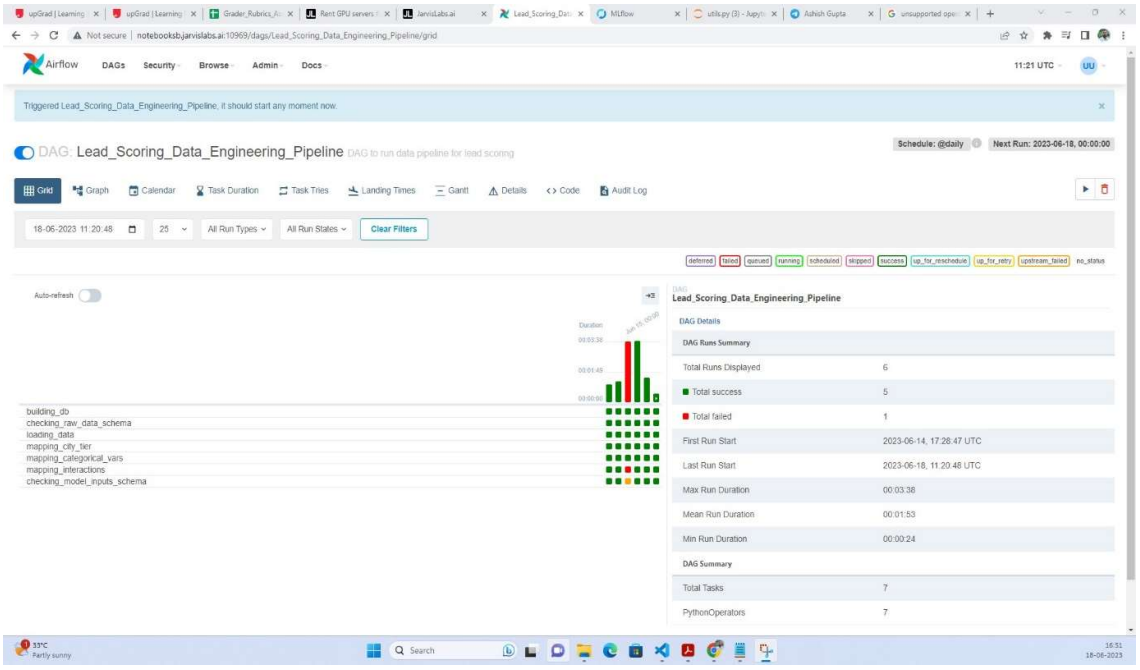
Screenshot of Airflow UI grid

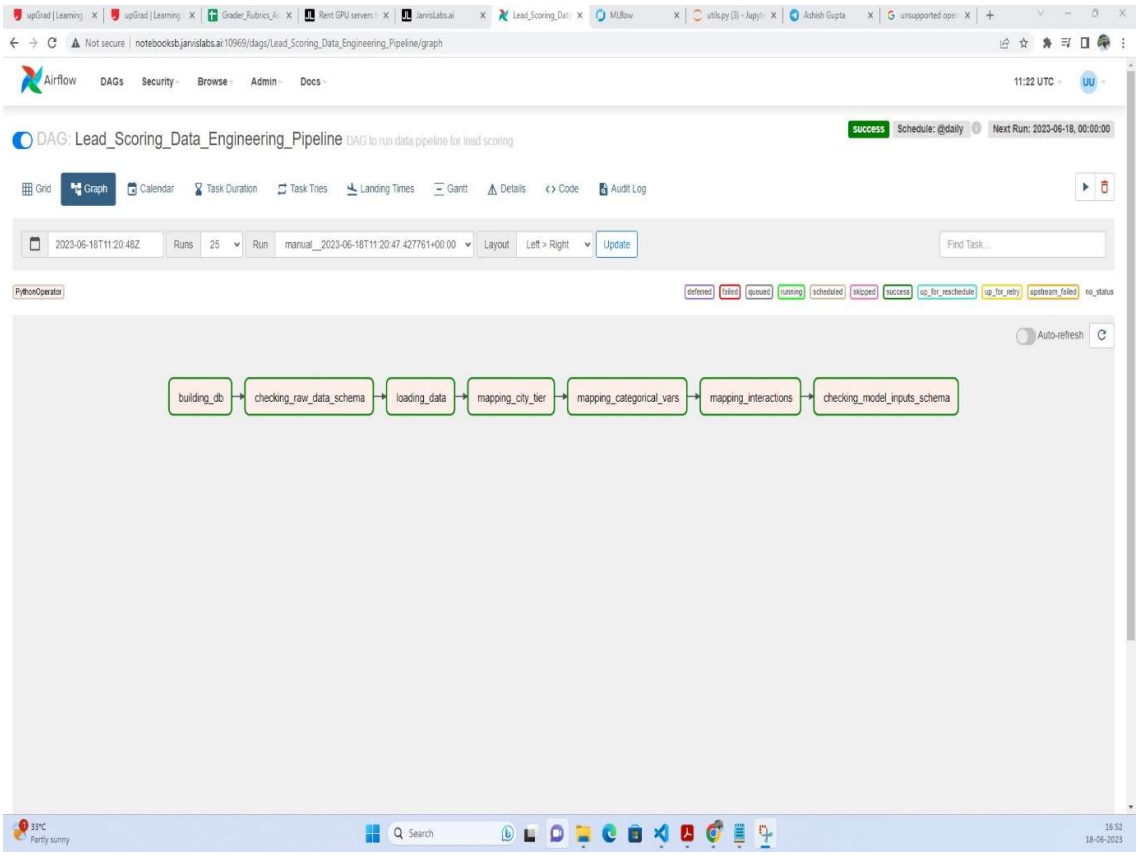


Screenshot of successful execution Airflow DAG in graph

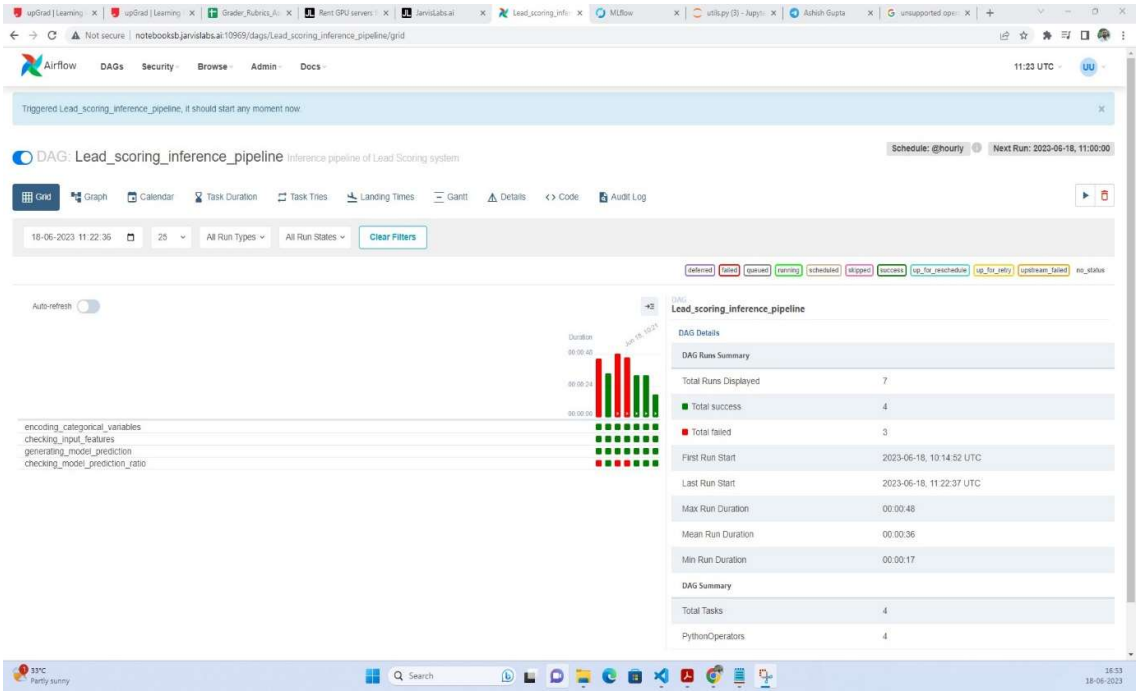


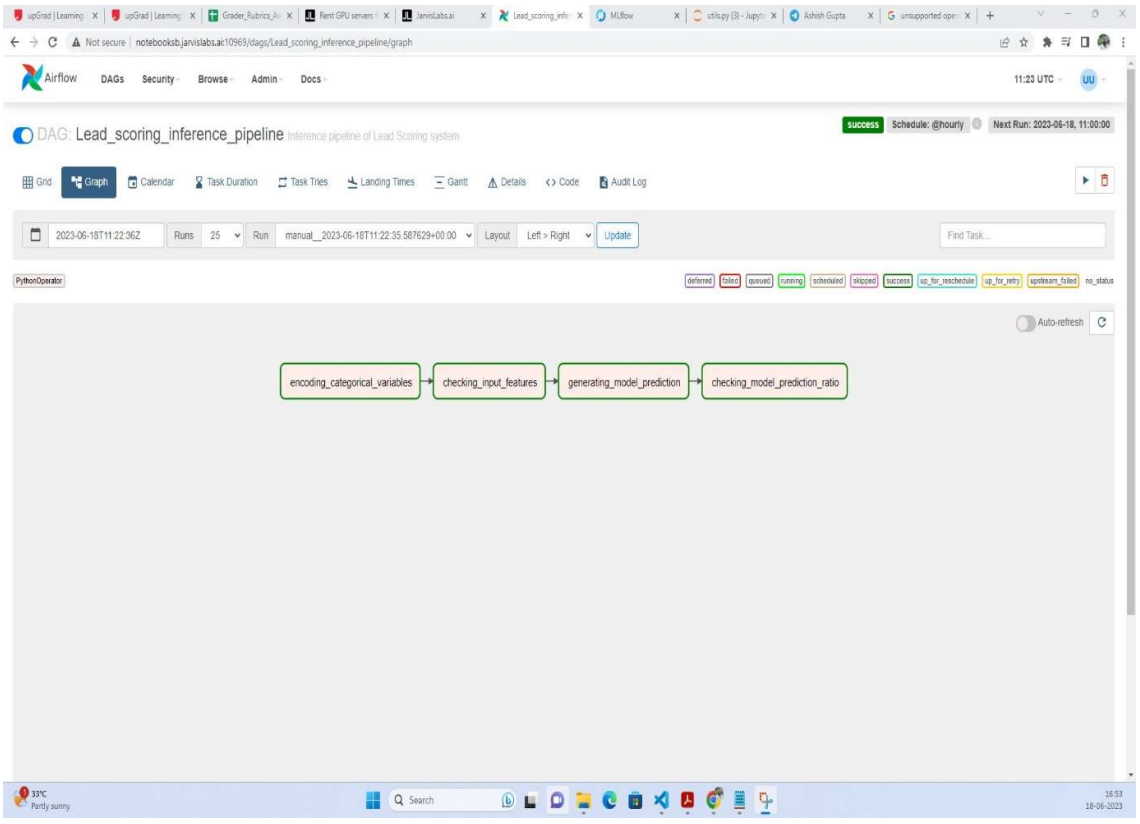
Running data cleaning pipeline on leadscoring_inference.csv



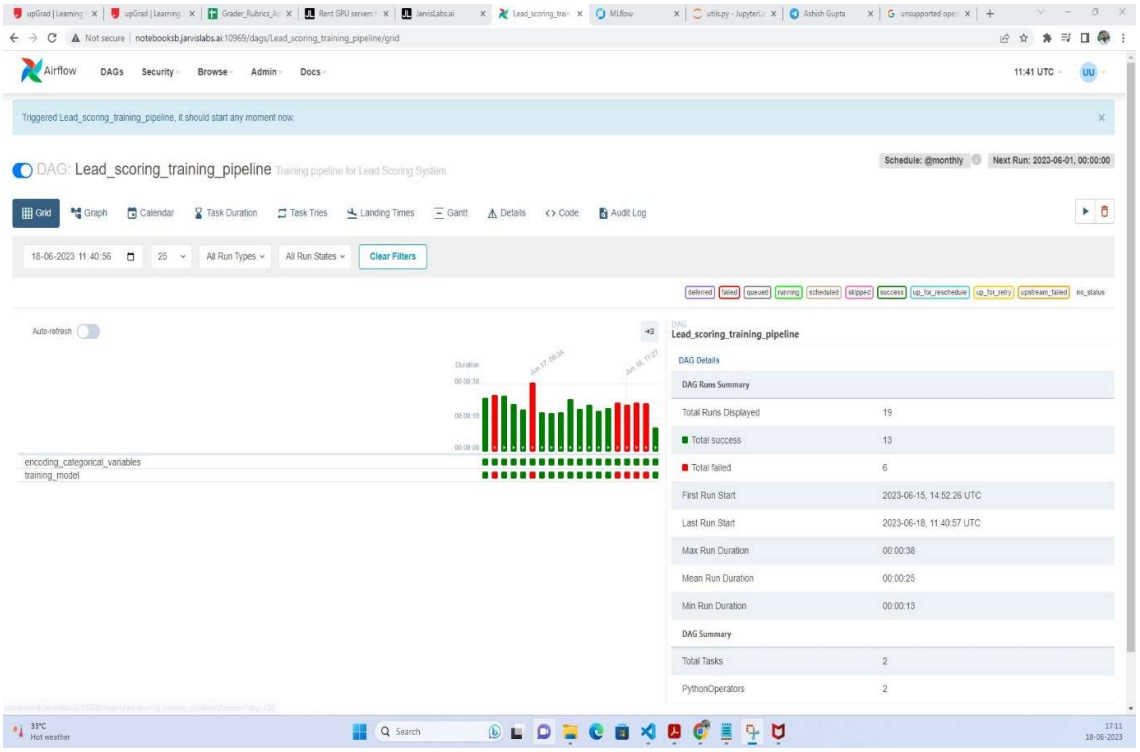


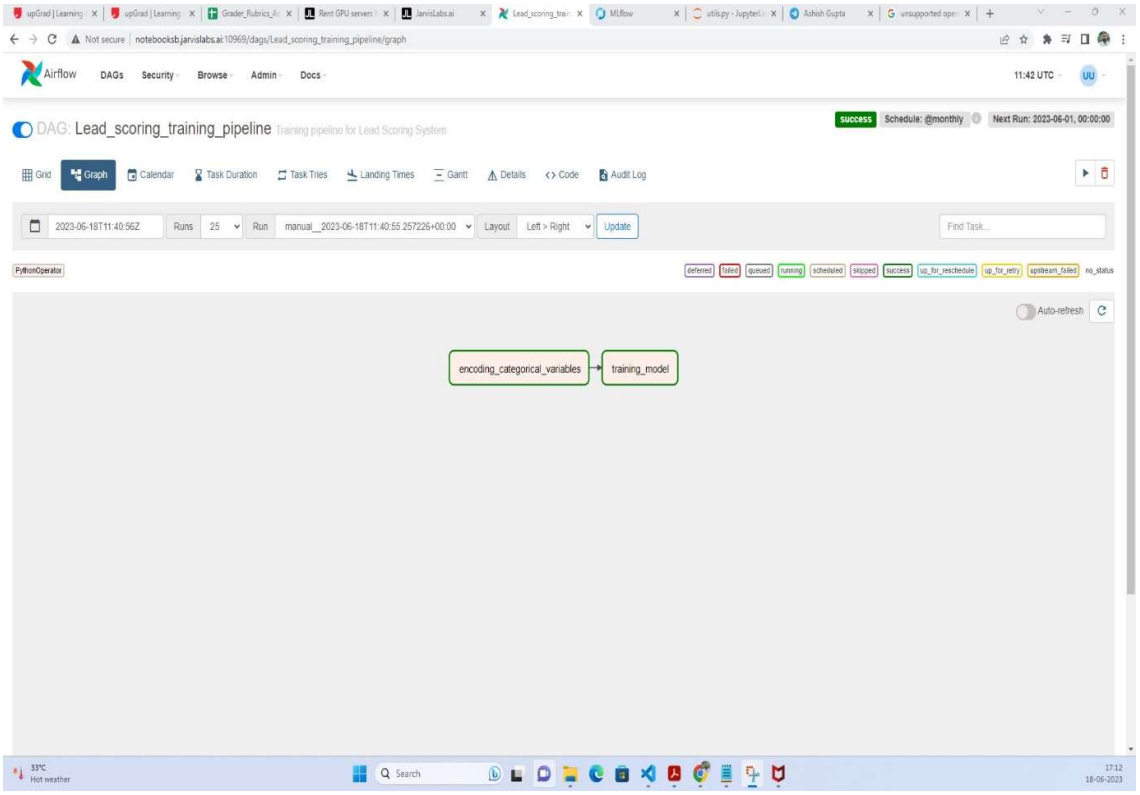
Running data inference pipeline on cleaned leadscoring_inference.csv

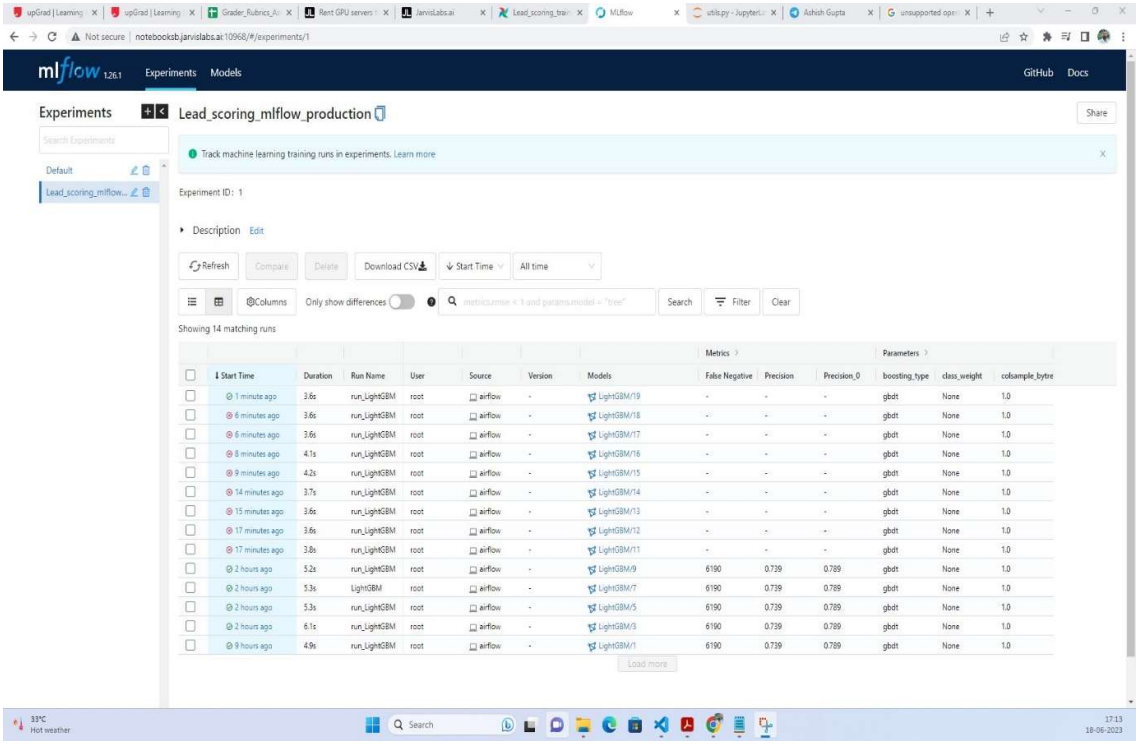


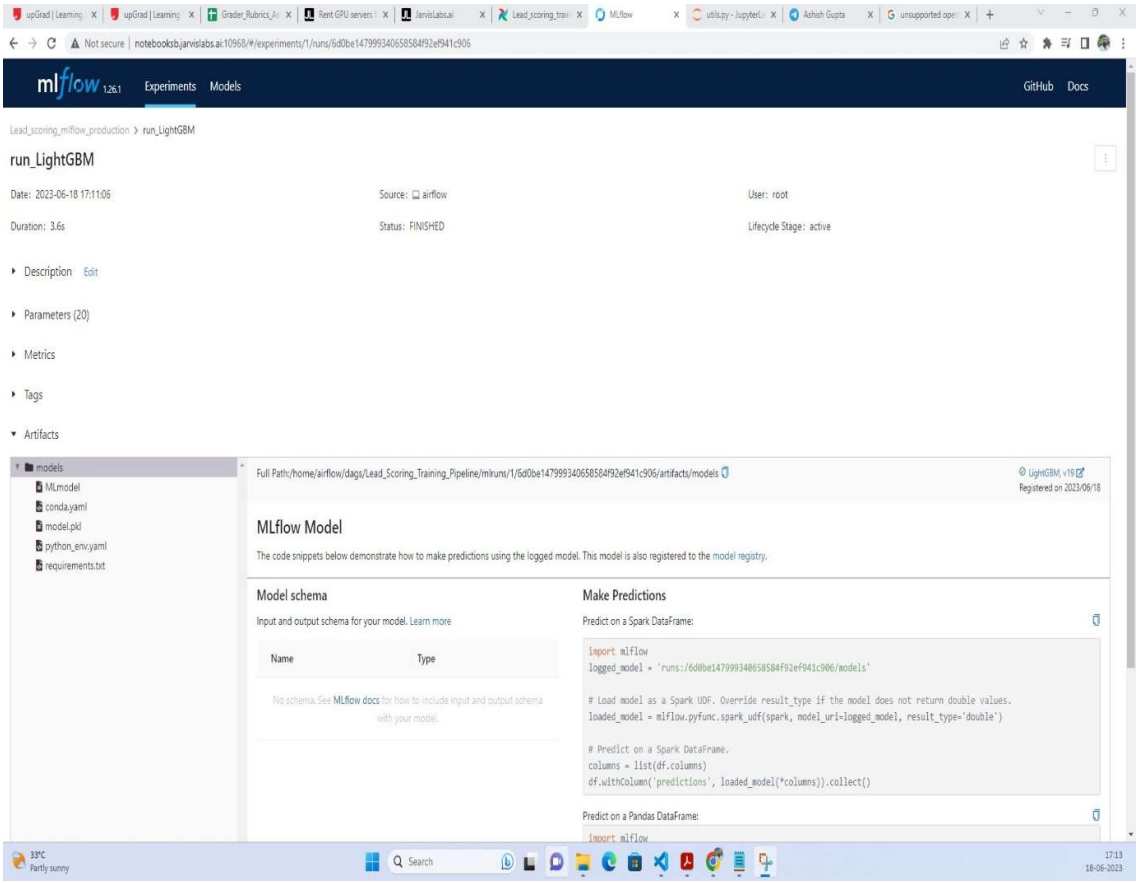


Running training pipeline on cleaned leadscoring_inference.csv






mlflow 1.26.1
Experiments
Models
GitHub
Docs
Experiments
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 14 matching runs
Metrics
Parameters
Start Time
Duration
Run Name
User
Source
Version
Models
False Negative
Precision
Precision_0
boosting_type
class_weight
colsample_bytre
1 minute ago 3.6s run_LightGBM root airflow - LightGBM/19 - - - gbdt None 1.0
6 minutes ago 3.6s run_LightGBM root airflow - LightGBM/18 - - - gbdt None 1.0
6 minutes ago 3.6s run_LightGBM root airflow - LightGBM/17 - - - gbdt None 1.0
8 minutes ago 4.1s run_LightGBM root airflow - LightGBM/16 - - - gbdt None 1.0
9 minutes ago 4.2s run_LightGBM root airflow - LightGBM/15 - - - gbdt None 1.0
14 minutes ago 3.7s run_LightGBM root airflow - LightGBM/14 - - - gbdt None 1.0
15 minutes ago 3.6s run_LightGBM root airflow - LightGBM/13 - - - gbdt None 1.0
17 minutes ago 3.6s run_LightGBM root airflow - LightGBM/12 - - - gbdt None 1.0
17 minutes ago 3.8s run_LightGBM root airflow - LightGBM/11 - - - gbdt None 1.0
2 hours ago 5.2s run_LightGBM root airflow - LightGBM/9 6190 0.739 0.789 gbdt None 1.0
2 hours ago 5.3s LightGBM root airflow - LightGBM/7 6190 0.739 0.789 gbdt None 1.0
2 hours ago 5.3s run_LightGBM root airflow - LightGBM/5 6190 0.739 0.789 gbdt None 1.0
2 hours ago 6.1s run_LightGBM root airflow - LightGBM/3 6190 0.739 0.789 gbdt None 1.0
9 hours ago 4.9s run_LightGBM root airflow - LightGBM/1 6190 0.739 0.789 gbdt None 1.0
Load more
33°C
Hot weather
Search
17:13
18-06-2023


mlflow 1.26.1
Experiments
Models
GitHub
Docs
Lead_scoring_mlflow_production > run_LightGBM
run_LightGBM
Date: 2023-06-18 17:11:06
Source: airflow
User: root
Duration: 3.6s
Status: FINISHED
Lifecycle Stage: active
Description Edit
Parameters (20)
Metrics
Tags
Artifacts
models
MLmodel
conda.yaml
model.pkl
python_env.yaml
requirements.txt
Full Path: /home/airflow/dags/Lead_Scoring_Training_Pipeline/mlruns/1/6d0be14799934065858 4f92ef941c906/artifacts/models
LightGBM, v19
Registered on 2023/06/18
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:/6d0be14799934065858 4f92ef941c906/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
33°C
Partly sunny
Search
17:13
18-06-2023