Machine Learning Model Deployment with IBM Cloud Watson Studio

Phase – 5(Documentation)

Abstract

This project leverages IBM Cloud Watson Studio to develop and deploy predictive models for customer churn analysis in the telecommunications industry. By collecting and analyzing customer data, training machine learning models, and integrating them into operational systems, the project aims to proactively reduce churn rates and enhance customer retention strategies.

Problem Statement

Become a wizard of predictive analytics with IBM Cloud Watson Studio. Train machine learning models to predict outcomes in real-time. Deploy the models as web services and integrate them into your applications. Unlock the magic of data-driven insights and make informed decisions like never before!

Problem Definition

The project involves training a machine learning model using IBM Cloud Watson Studio and deploying it as a web service. The goal is to become proficient in predictive analytics by creating a model that can predict outcomes in real-time. The project encompasses defining the predictive use case, selecting a suitable dataset, training a machine learning model,deploying the model as a web service, and integrating it into applications.

Dataset Link

Project features

1. DATA COLLECTION AND PREPROCESSING.

Data Collection:

The project's foundation is built on a robust dataset that encompasses historical loan application data from a variety of sources, including banks, credit bureaus, and other financial institutions. This dataset is continually updated to reflect changing trends and economic conditions.

Data Preprocessing:

Before the data can be used effectively, it undergoes thorough preprocessing. This includes the identification and handling of missing values, outlier detection, and data standardization. This meticulous data preparation ensures that the machine learning models receive clean, reliable input for decision-making.

2. MODEL TRAINING AND EVALUATION.

Data Splitting:

The dataset is divided into three subsets: training, validation, and test sets. The training set serves as the foundation for model learning, the validation set is used for model selection and hyperparameter tuning, and the test set assesses the final model performance.

Model Training:

Model training is an iterative process where algorithms learn from the training data. Adjustments are made to ensure a balanced trade-off between predictive accuracy and fairness. The training phase involves significant experimentation and refinement to achieve the desired outcomes.

Model Evaluation:

Comprehensive model evaluation is conducted using a battery of metrics, including accuracy, precision, recall, F1-score, and ROC-AUC. The evaluation goes beyond accuracy and extends to fairness metrics to ensure that our automated system is transparent and minimizes discrimination.

3. DEPLOYMENT AND USER INTERFACE

Deployment Process:

Model deployment is executed through IBM Cloud Watson Studio, providing a scalable and reliable platform for serving predictions. The process includes converting trained models into a format suitable for real-time loan approval decisions. Our deployment infrastructure is designed to support increased demand while maintaining efficiency and reliability.

User Interface Design:

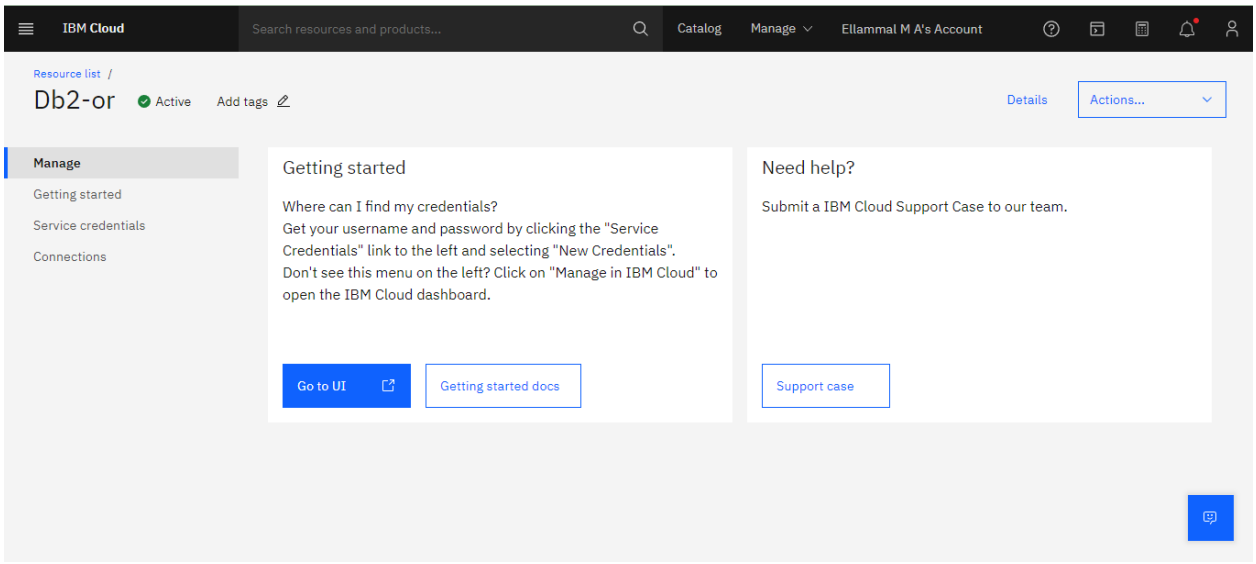
The user interface is meticulously designed to ensure a smooth and transparent application process. It simplifies data submission, reduces documentation requirements, and integrates real-time feedback mechanisms. The interface empowers applicants to make informed decisions about their loan options and improves their overall experience.

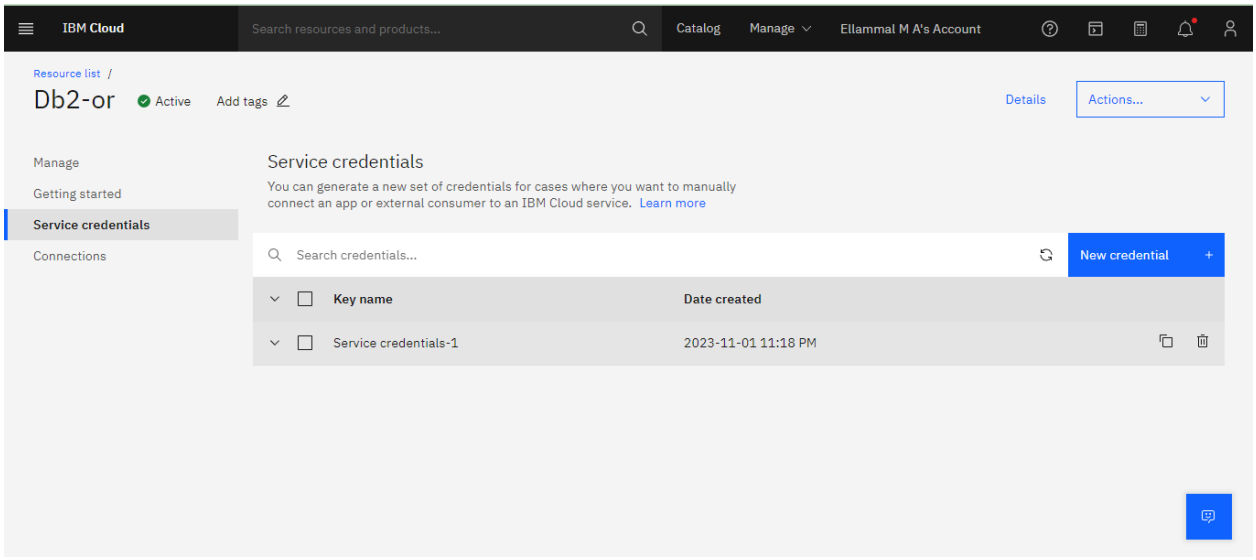
4. TESTING AND VALIDATION

Testing Procedures:

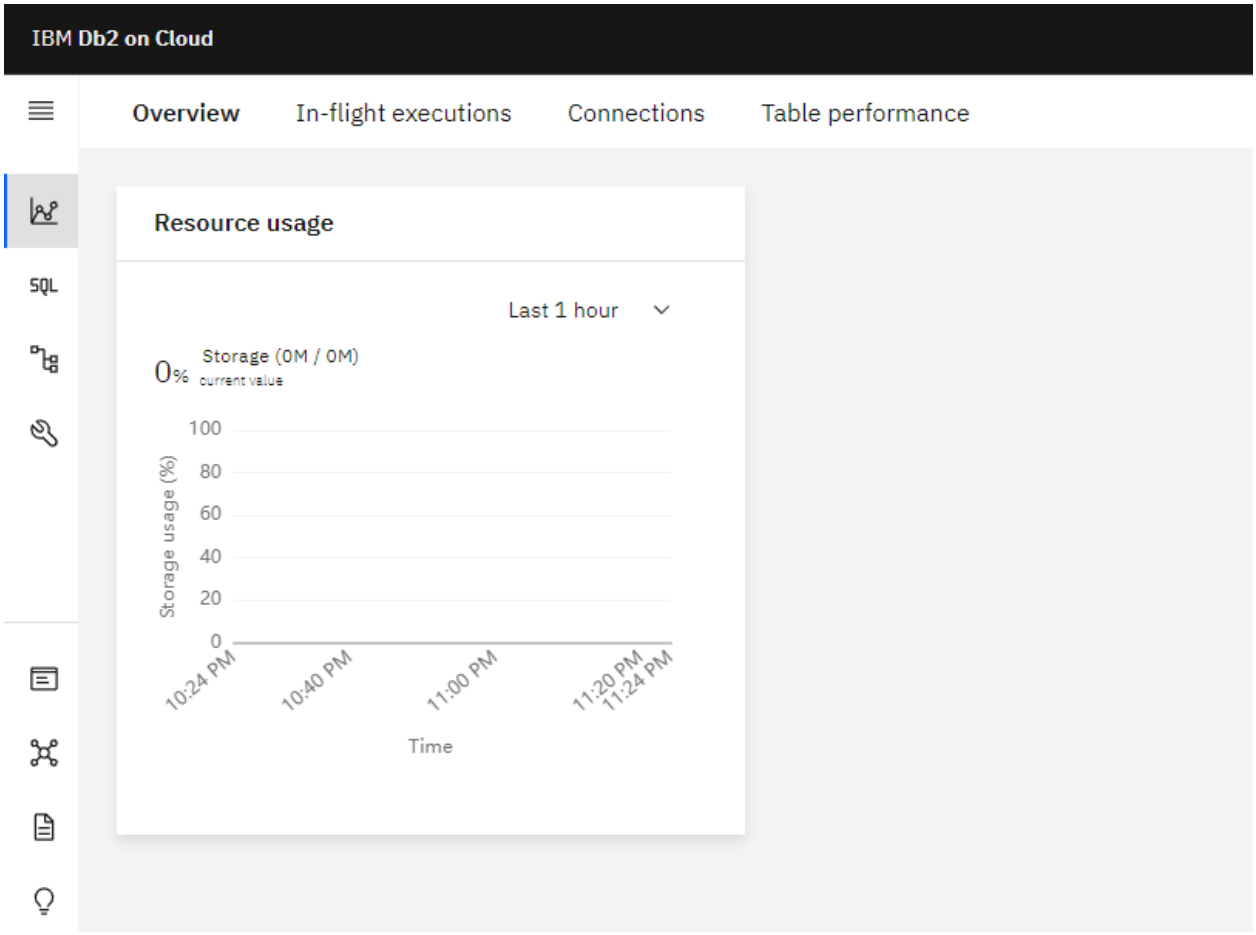
Rigorous testing is conducted to ensure the accuracy and reliability of the system. This includes sensitivity analysis to understand the impact of various factors on loan decisions and stress testing to evaluate the system's performance under high demand. Ongoing testing and validation processes are in place to maintain the system's reliability.

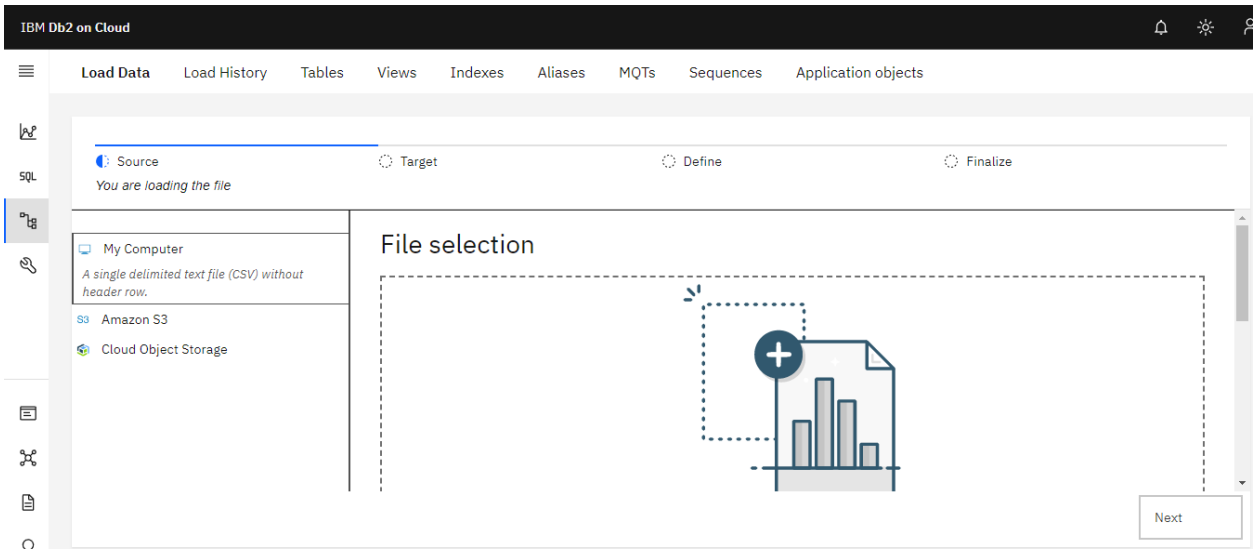
PHASE 3:

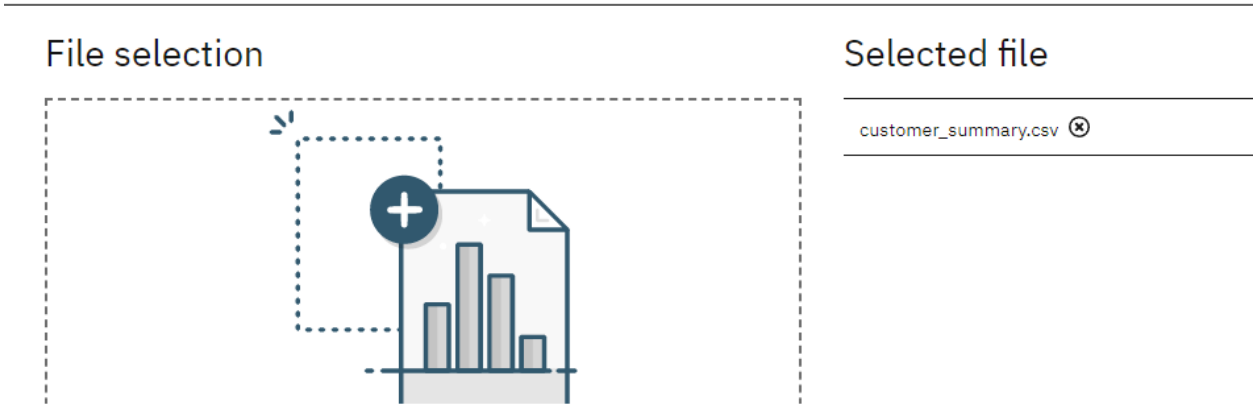


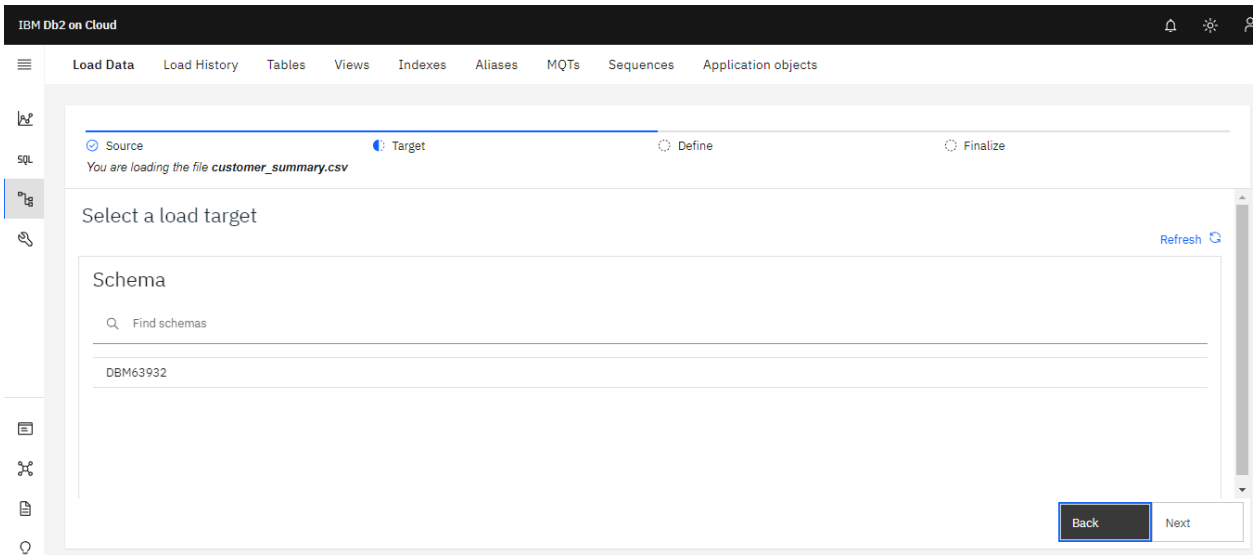


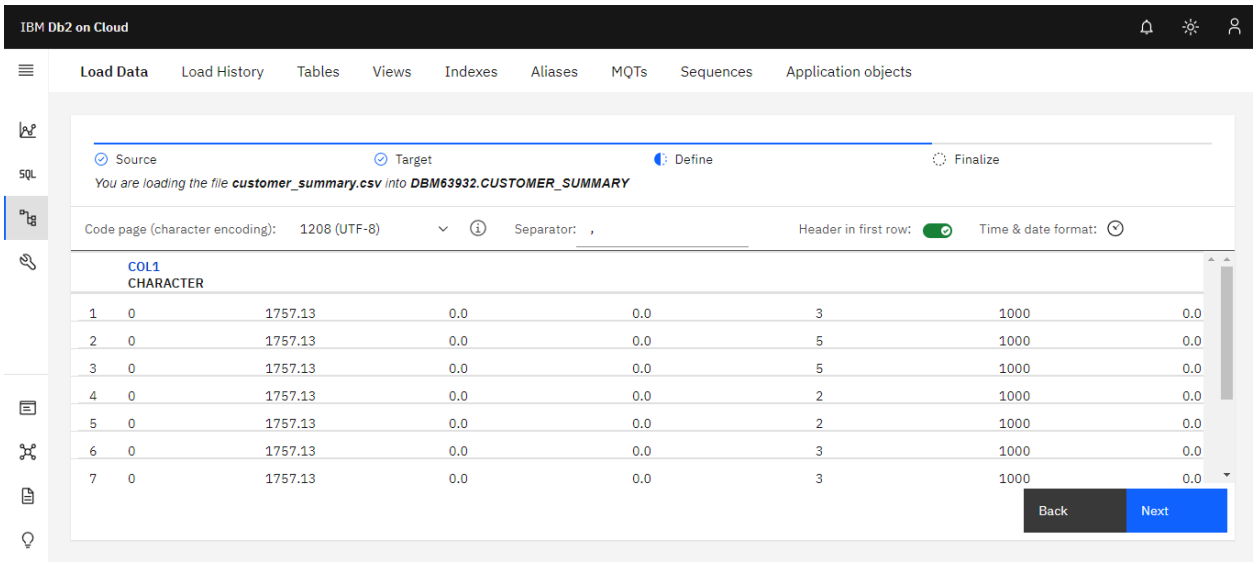

```
"-u",
        "dbm63932",
        "-p",
        "W3KPfJv9swuf699o",
        "--ssl",
        "--sslCAFile",
        "1dd14d0c-1b52-4f63-a606-53ecba28771d",
        "--authenticationDatabase",
        "admin",
        "--host",
```

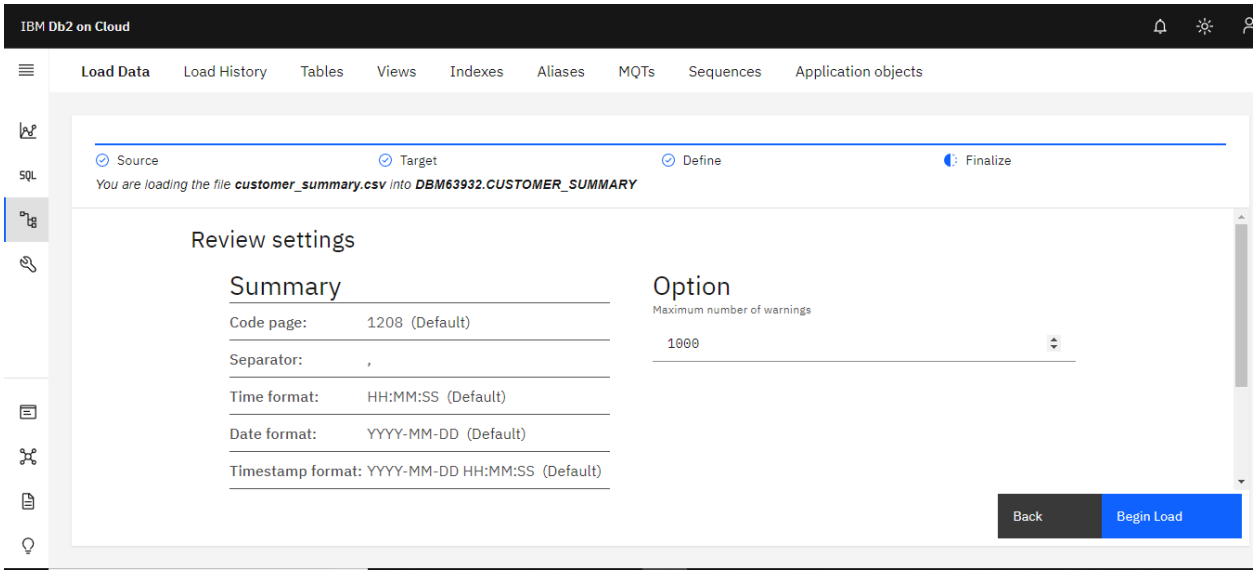

IBM Db2 on Cloud
Overview
In-flight executions
Connections
Table performance
Resource usage
Last 1 hour
0%
Storage (0M / 0M)
current value
100
80
60
40
20
0
Storage usage (%)
10:24 PM
10:40 PM
11:00 PM
11:20 PM
11:24 PM
Time

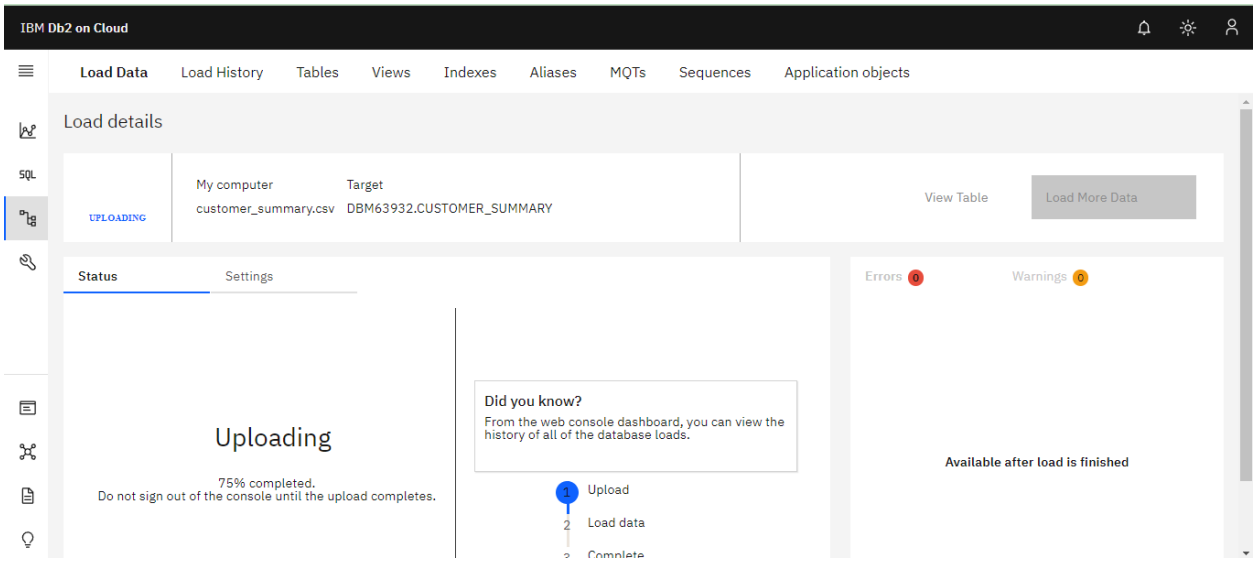

IBM Db2 on Cloud
Load Data
Load History
Tables
Views
Indexes
Aliases
MQTs
Sequences
Application objects
Source
Target
Define
Finalize
You are loading the file
My Computer
A single delimited text file (CSV) without header row.
Amazon S3
Cloud Object Storage
File selection
Next

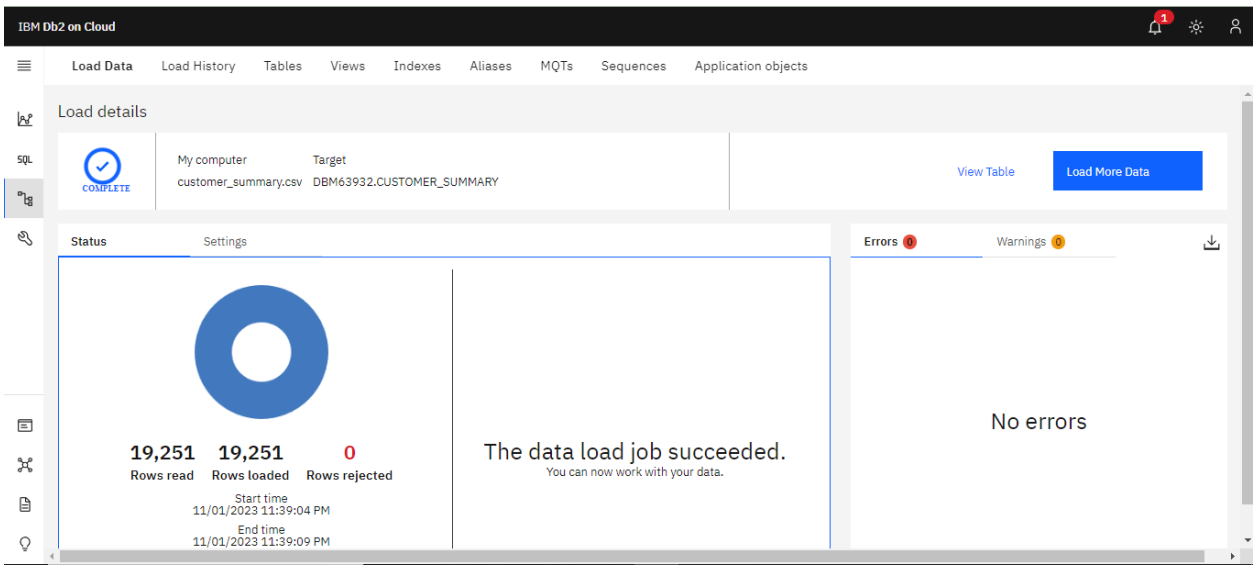

File selection
Selected file
customer_summary.csv


IBM Db2 on Cloud
Load Data
Load History
Tables
Views
Indexes
Aliases
MQTs
Sequences
Application objects
Source
Target
Define
Finalize
You are loading the file customer_summary.csv
Select a load target
Refresh
Schema
Find schemas
DBM63932
Back
Next

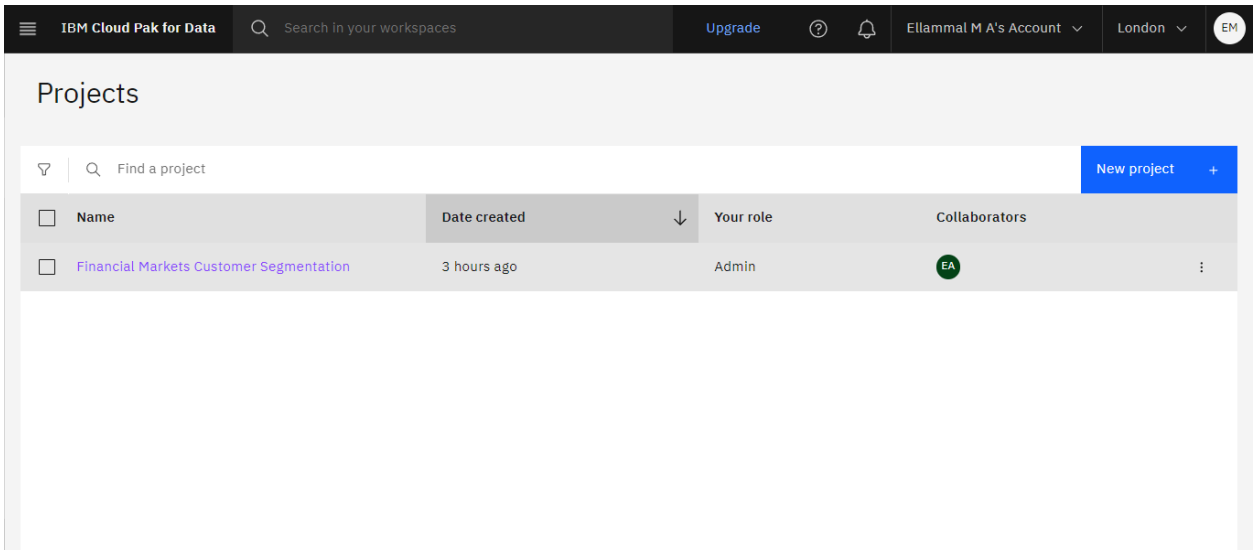

IBM Db2 on Cloud
Load Data
Load History
Tables
Views
Indexes
Aliases
MQTs
Sequences
Application objects
Source
Target
Define
Finalize
You are loading the file customer_summary.csv into DBM63932.CUSTOMER_SUMMARY
Code page (character encoding):
1208 (UTF-8)
Separator:
,
Header in first row:
Time & date format:
COL1
CHARACTER
1 0 1757.13 0.0 0.0 3 1000 0.0
2 0 1757.13 0.0 0.0 5 1000 0.0
3 0 1757.13 0.0 0.0 5 1000 0.0
4 0 1757.13 0.0 0.0 2 1000 0.0
5 0 1757.13 0.0 0.0 2 1000 0.0
6 0 1757.13 0.0 0.0 3 1000 0.0
7 0 1757.13 0.0 0.0 3 1000 0.0
Back
Next

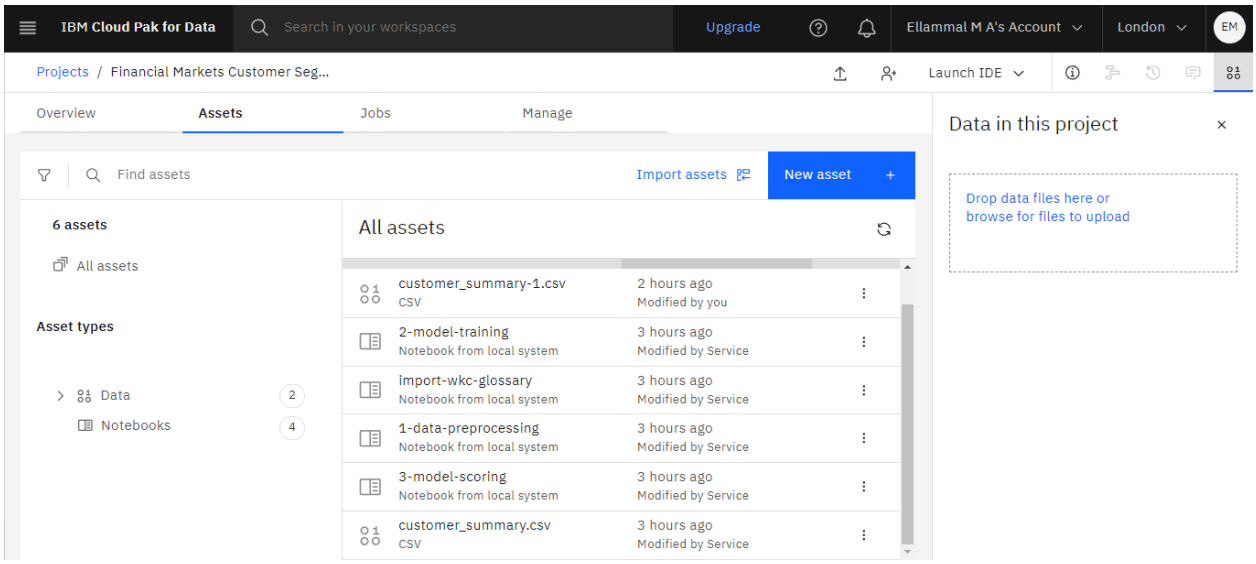

IBM Db2 on Cloud
Load Data
Load History
Tables
Views
Indexes
Aliases
MQTs
Sequences
Application objects
Source
Target
Define
Finalize
You are loading the file customer_summary.csv into DBM63932.CUSTOMER_SUMMARY
Review settings
Summary
Code page: 1208 (Default)
Separator: ,
Time format: HH:MM:SS (Default)
Date format: YYYY-MM-DD (Default)
Timestamp format: YYYY-MM-DD HH:MM:SS (Default)
Option
Maximum number of warnings
1000
Back
Begin Load

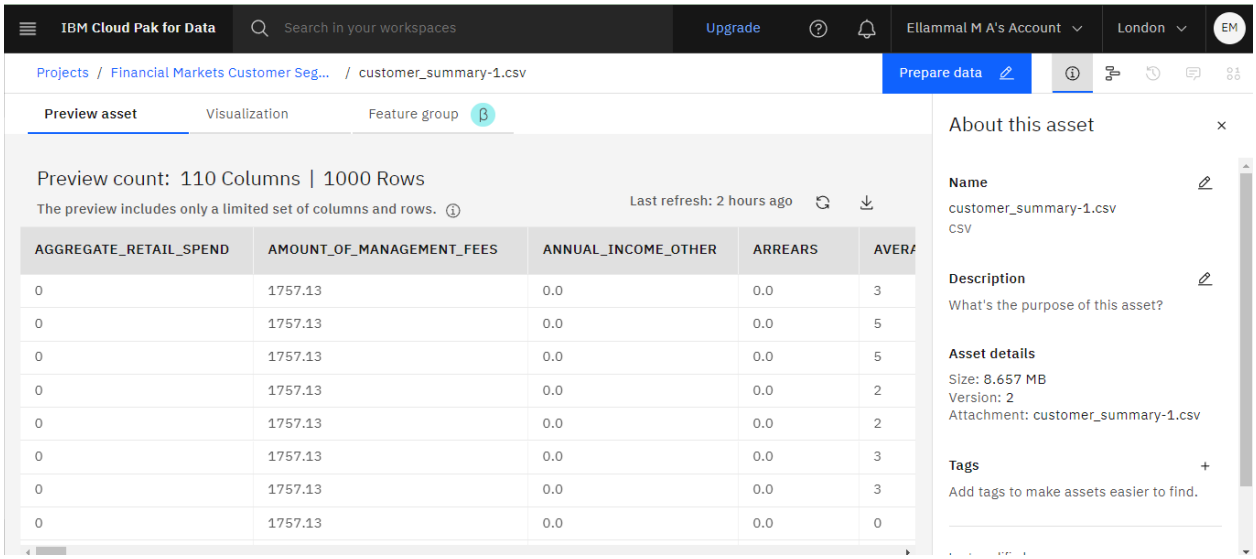

IBM Db2 on Cloud
Load Data
Load History
Tables
Views
Indexes
Aliases
MQTs
Sequences
Application objects
Load details
UPLOADING
My computer
customer_summary.csv
Target
DBM63932.CUSTOMER_SUMMARY
View Table
Load More Data
Status
Settings
Errors 0
Warnings 0
Uploading
75% completed.
Do not sign out of the console until the upload completes.
Did you know?
From the web console dashboard, you can view the history of all of the database loads.
1 Upload
2 Load data
Available after load is finished

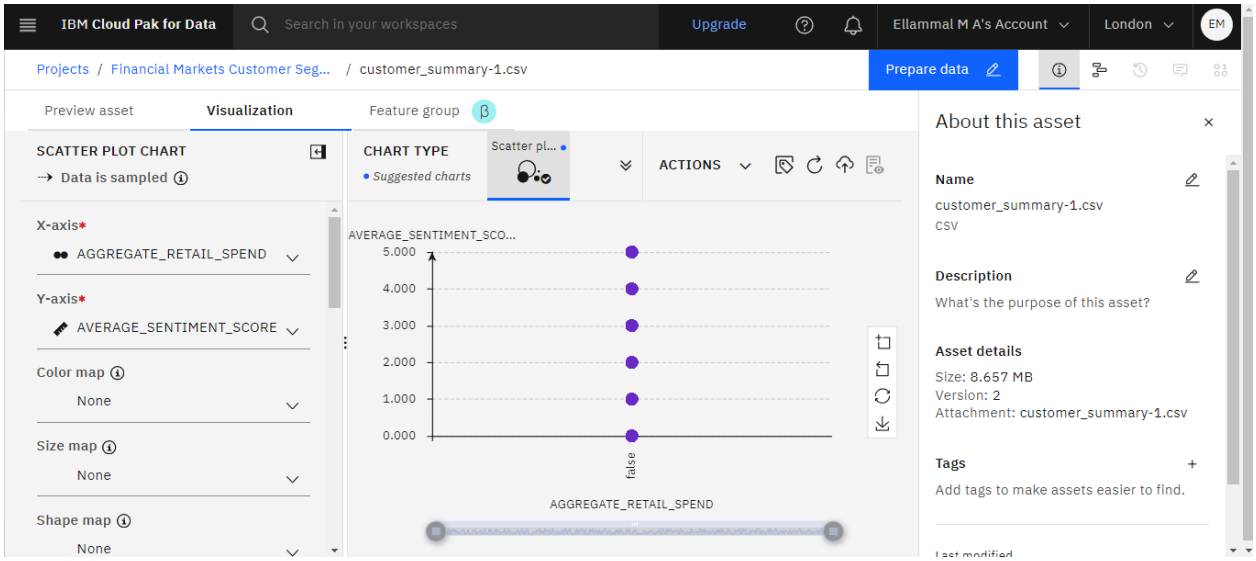

IBM Db2 on Cloud
Load Data
Load History
Tables
Views
Indexes
Aliases
MQTs
Sequences
Application objects
Load details
COMPLETE
My computer
customer_summary.csv
Target
DBM63932.CUSTOMER_SUMMARY
View Table
Load More Data
Status
Settings
Errors 0
Warnings 0
19,251
Rows read
19,251
Rows loaded
0
Rows rejected
Start time
11/01/2023 11:39:04 PM
End time
11/01/2023 11:39:09 PM
The data load job succeeded.
You can now work with your data.
No errors

PHASE 4:


IBM Cloud Pak for Data
Search in your workspaces
Upgrade
Ellammal M A's Account
London
EM
Projects
Find a project
New project
Name
Date created
Your role
Collaborators
Financial Markets Customer Segmentation
3 hours ago
Admin
EA


IBM Cloud Pak for Data
Search in your workspaces
Upgrade
Ellammal M A's Account
London
EM
Projects / Financial Markets Customer Seg...
Launch IDE
Overview
Assets
Jobs
Manage
Find assets
Import assets
New asset
6 assets
All assets
Asset types
Data 2
Notebooks 4
All assets
customer_summary-1.csv
CSV
2 hours ago
Modified by you
2-model-training
Notebook from local system
3 hours ago
Modified by Service
import-wkc-glossary
Notebook from local system
3 hours ago
Modified by Service
1-data-preprocessing
Notebook from local system
3 hours ago
Modified by Service
3-model-scoring
Notebook from local system
3 hours ago
Modified by Service
customer_summary.csv
CSV
3 hours ago
Modified by Service
Data in this project
Drop data files here or
browse for files to upload


IBM Cloud Pak for Data
Search in your workspaces
Upgrade
Ellammal M A's Account
London
EM
Projects / Financial Markets Customer Seg... / customer_summary-1.csv
Prepare data
Preview asset
Visualization
Feature group
Preview count: 110 Columns | 1000 Rows
The preview includes only a limited set of columns and rows.
Last refresh: 2 hours ago
AGGREGATE_RETAIL_SPEND
AMOUNT_OF_MANAGEMENT_FEES
ANNUAL_INCOME_OTHER
ARREARS
AVERA
0 1757.13 0.0 0.0 3
0 1757.13 0.0 0.0 5
0 1757.13 0.0 0.0 5
0 1757.13 0.0 0.0 2
0 1757.13 0.0 0.0 2
0 1757.13 0.0 0.0 3
0 1757.13 0.0 0.0 3
0 1757.13 0.0 0.0 0
About this asset
Name
customer_summary-1.csv
csv
Description
What's the purpose of this asset?
Asset details
Size: 8.657 MB
Version: 2
Attachment: customer_summary-1.csv
Tags
Add tags to make assets easier to find.


IBM Cloud Pak for Data
Search in your workspaces
Upgrade
Ellammal M A's Account
London
EM
Projects / Financial Markets Customer Seg... / customer_summary-1.csv
Prepare data
Preview asset
Visualization
Feature group
SCATTER PLOT CHART
Data is sampled
CHART TYPE
Suggested charts
Scatter pl...
ACTIONS
X-axis*
AGGREGATE_RETAIL_SPEND
Y-axis*
AVERAGE_SENTIMENT_SCORE
Color map
None
Size map
None
Shape map
None
AVERAGE_SENTIMENT_SCO...
5.000
4.000
3.000
2.000
1.000
0.000
false
AGGREGATE_RETAIL_SPEND
About this asset
Name
customer_summary-1.csv
csv
Description
What's the purpose of this asset?
Asset details
Size: 8.657 MB
Version: 2
Attachment: customer_summary-1.csv
Tags
Add tags to make assets easier to find.

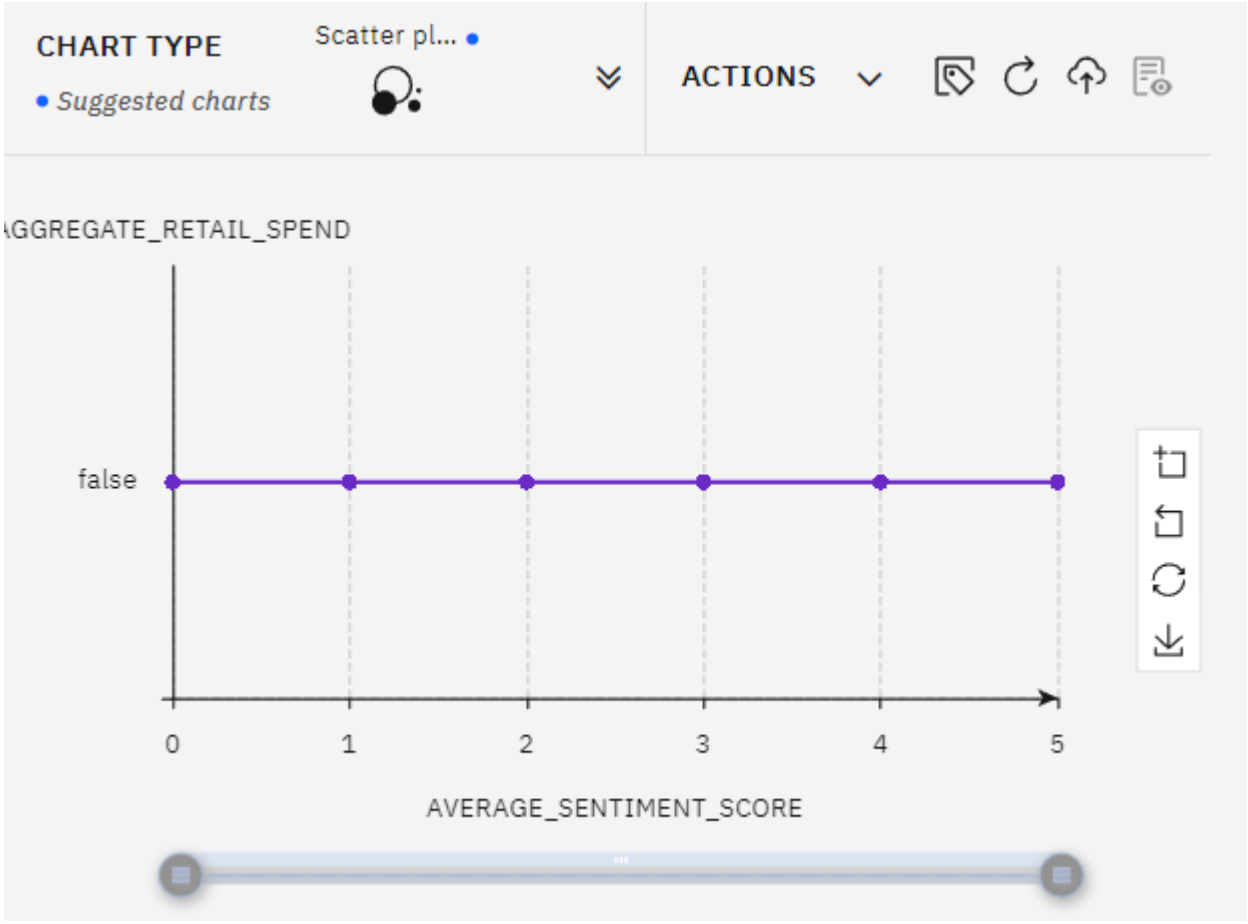
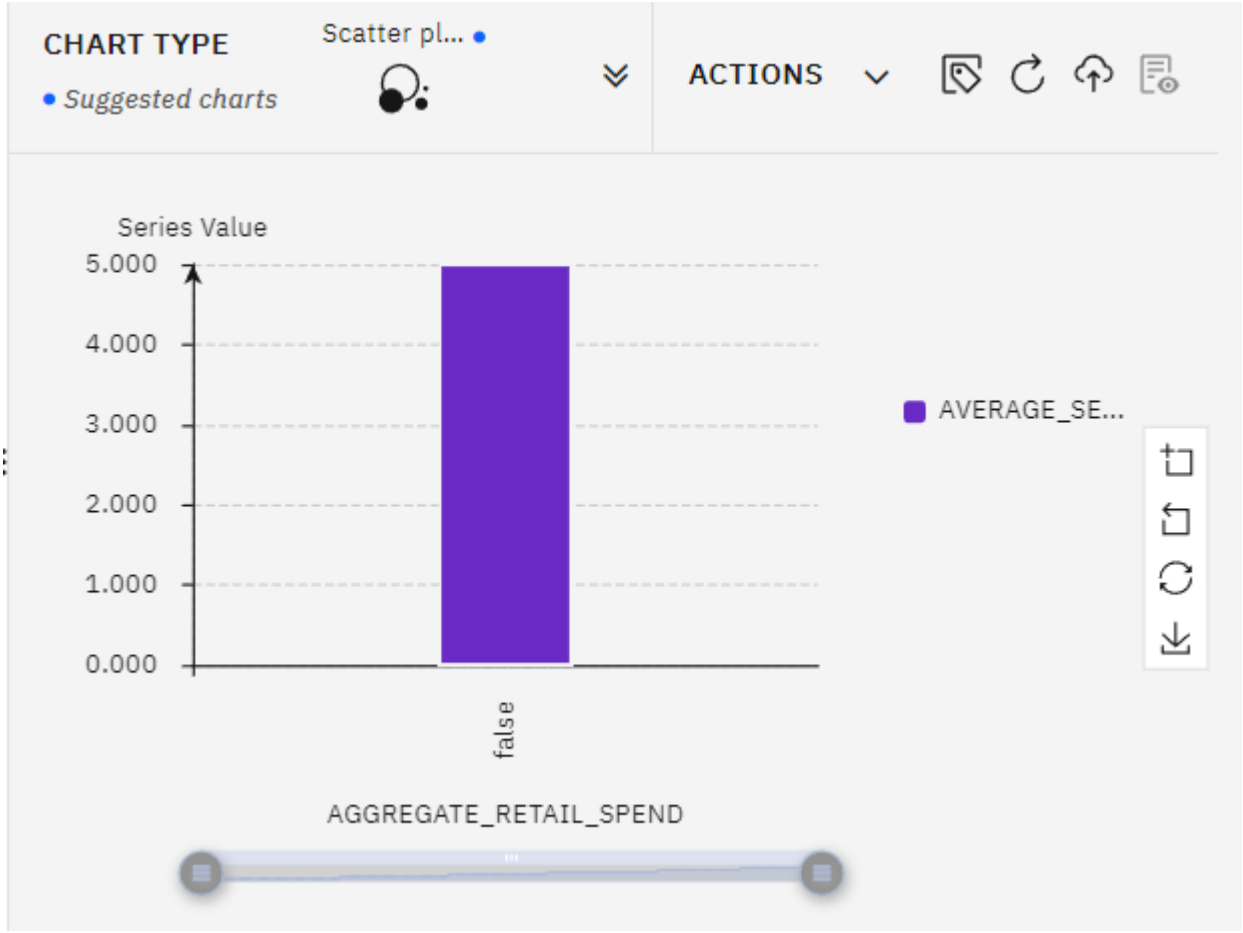


CHART TYPE
Scatter pl...
Suggested charts
ACTIONS
AGGREGATE_RETAIL_SPEND
false
0
1
2
3
4
5
AVERAGE_SENTIMENT_SCORE



Conclusion

In conclusion, deploying a Machine Learning model for a Loan Approval System with IBM Cloud Watson Studio offers enhanced efficiency, accuracy, risk mitigation, cost savings, and a better customer experience. It is a forward-thinking approach for financial institutions seeking to stay competitive in the modern lending landscape.