

# Lab: Introduction to Watson OpenScale

## Introduction

Businesses today are increasingly certain that AI will be a driving force in the evolution of their industries over the next few years. Yet for every successful AI project, there are many that fail to reach widespread adoption in the business and achieve their expected outcomes. This is partly because the mechanics of AI deployment can be complex, and there are still gaps in skills and tooling that can make it difficult for data science, IT operations, and business teams to work in lockstep. But beyond the operational challenges, there are also much more profound issues of trust and transparency that businesses need to address before they can truly turn AI into a business advantage.

Knowledge workers must be able to trust AI and explain the decisions it helps make before they will incorporate it in their business processes. If AI is a black box that simply takes in data and produces obscure, unexplainable outcomes, then there is no way for the business to judge whether these systems are producing fair, accurate outcomes, or have confidence in AI's ability to augment decision-making. Equally, the business will not be able to explain outcomes to customers, auditors, or compliance teams.

IBM Watson OpenScale is an open platform that helps remove barriers to enterprise-scale AI. Watson OpenScale enables the enterprise to:

- Measure performance of production AI and its impact on business goals
- Track actionable metrics in a single console
- Explain AI outcomes
- Detect and mitigate harmful bias to improve outcomes
- Accept feedback to compute accuracy measures
- Accelerate the integration of AI into existing business applications.

## Objectives

The goal of this lab is to familiarize the user with the features of Watson OpenScale. After completing this lab, you will understand how to:

1. Import a machine learning model
2. Deploy the model
3. Provision Watson OpenScale
4. Configure the payload logging database and Machine Learning provider
5. Score Data
6. Prepare Deployed Model for Monitoring
7. Configure Payload Logging
8. Configure Quality
9. Configure Fairness
10. Configure Drift
11. Submit Feedback and View Quality Metrics

12. Score Data and View Fairness Metrics
13. Explain a Transaction.

## Lab Use Case

Traditional lenders are under pressure to expand their digital portfolio of financial services to a larger and more diverse audience, which requires a new approach to credit risk modeling. Their data science teams currently rely on standard modeling techniques - like decision trees and logistic regression - which work well for moderate datasets and make recommendations that can be easily explained. This satisfies regulatory requirements that credit lending decisions must be transparent and explainable.

To provide credit access to a wider and riskier population, applicant credit histories must expand beyond traditional credit, like mortgages and car loans, to alternate credit sources like utility and mobile phone plan payment histories, plus education and job titles. These new data sources offer promise, but also introduce risk by increasing the likelihood of unexpected correlations which introduce bias based on an applicant's age, gender, or other personal traits.

The data science techniques most suited to these diverse datasets, such as gradient boosted trees and neural networks, can generate highly accurate risk models, but at a cost. Such "black box" models generate opaque predictions that must somehow become transparent, to ensure regulatory approval such as Article 22 of the General Data Protection Regulation (GDPR), or the federal Fair Credit Reporting Act (FCRA) managed by the Consumer Financial Protection Bureau.

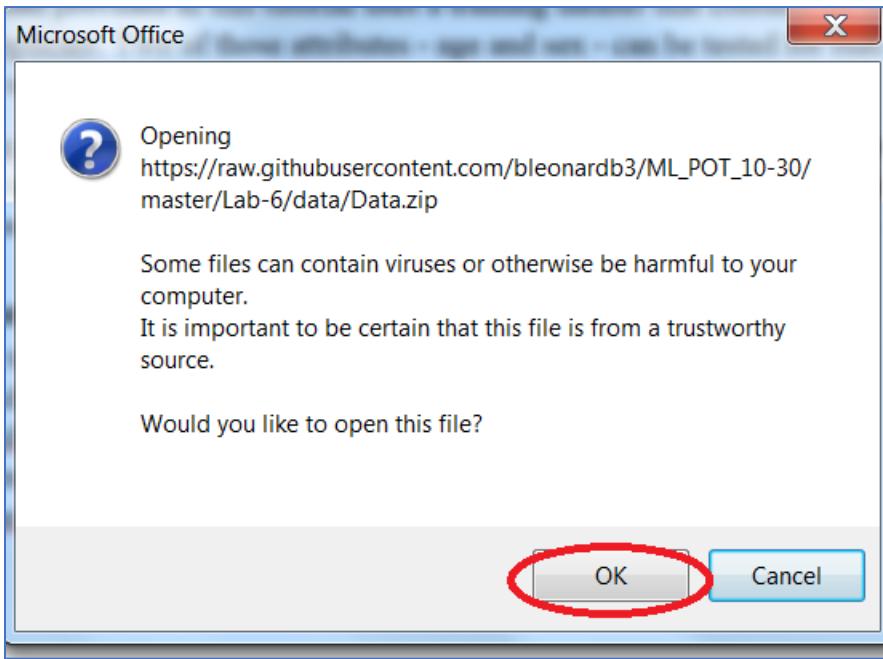
The credit risk model provided in this tutorial uses a training dataset that contains 20 attributes about each loan applicant. Two of those attributes - age and sex - can be tested for bias. For this tutorial, the focus will be on bias against sex and age.

Watson OpenScale will monitor the deployed model's propensity for a favorable outcome ("No Risk") for one group (the Reference Group) over another (the Monitored Group). In this tutorial, the Monitored Group for sex is `female`, while the Monitored Group for age is `19 to 25`.

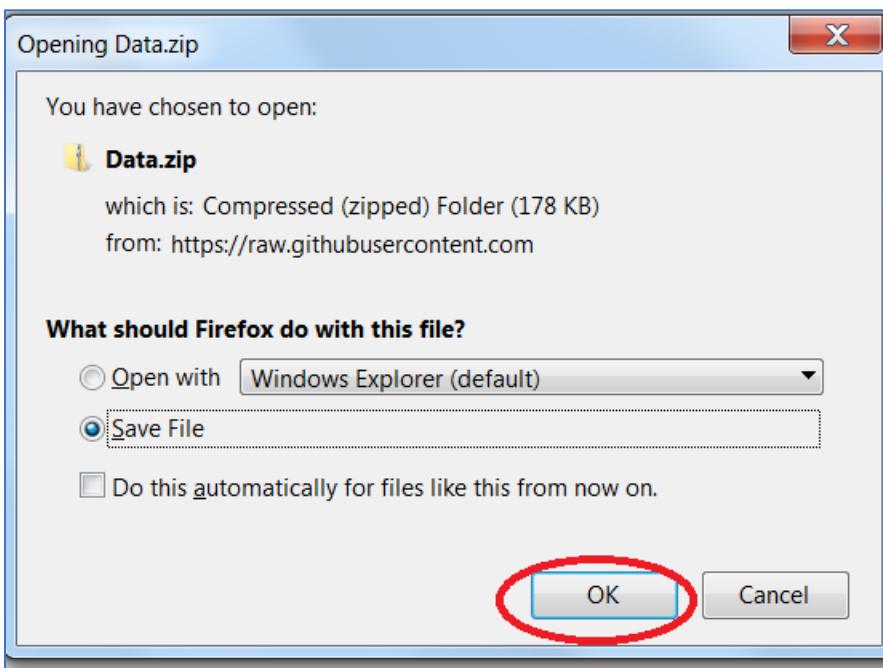
## Download the Lab Files

1. Click [here](#) to download the Data.zip file.
  1. `credit_feedback_data.csv`
  2. `credit_payload_data.json`
  3. `german_credit_data_biased_training.csv`
  4. `scoring.json`

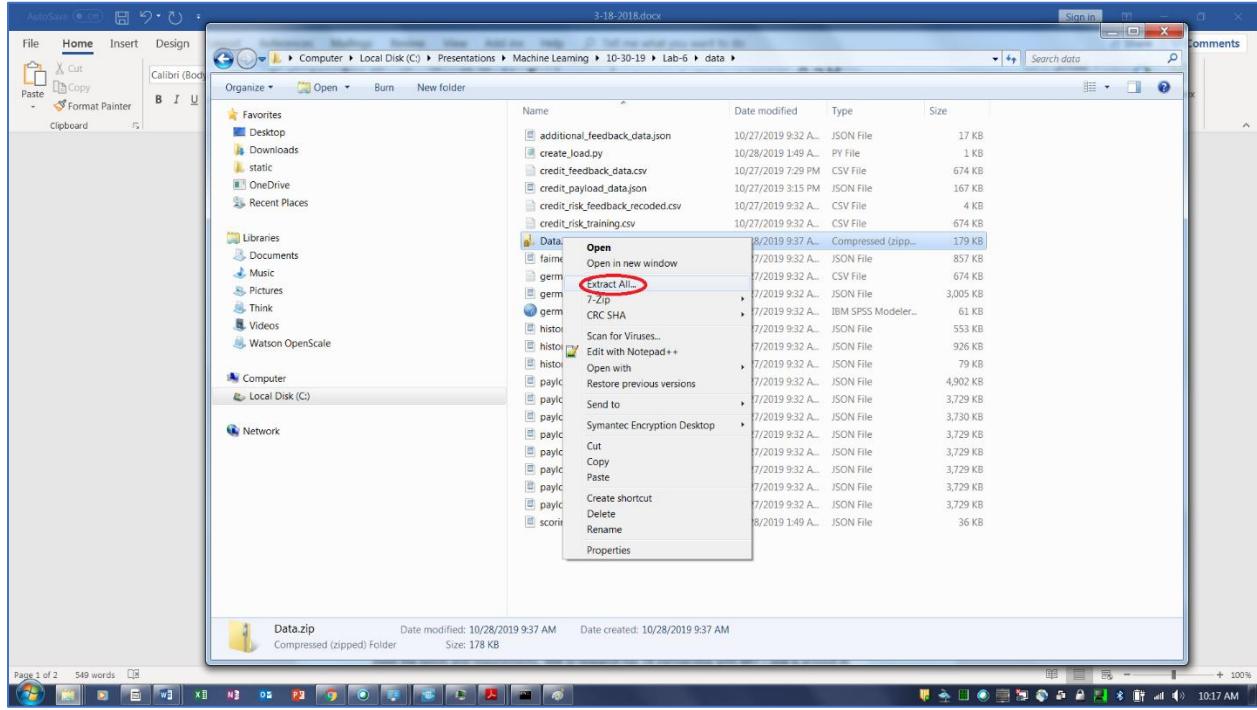
2. Click **OK**. Note your browser may not prompt this message.



3. Click **OK**.



4. Navigate to the folder where the file is saved. Select the Data.zip file, right-click, and click **Extract All**.

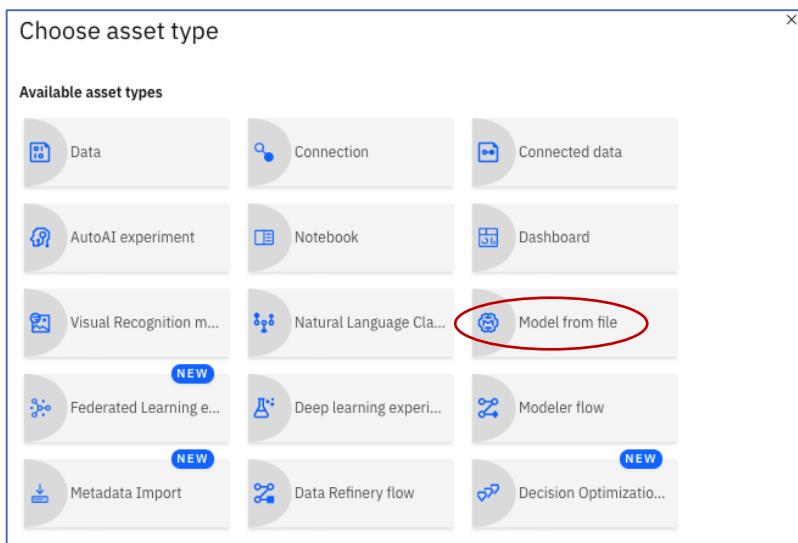


## Import the Credit Risk Model

- From the Watson Studio project, click on **Add to project**.



- Click on **Model from file**.



- From the **Import model** page, click on **Gallery sample**, click on **Credit Risk**, click on **Import**.

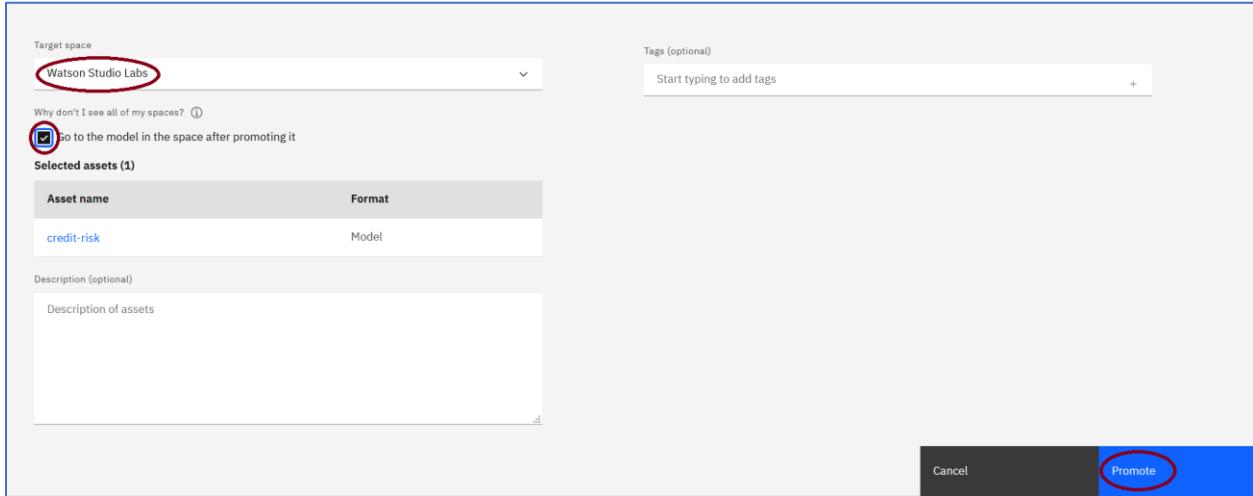
The screenshot shows the 'Import model' interface. On the left, there's a sidebar with a '+ New' button and a 'Gallery sample' button, which is circled in red. In the main area, there are two cards: 'Credit Risk' (circled in red) and 'Customer Satisfaction Prediction'. The 'Credit Risk' card has a detailed description: 'Traditional lenders are under pressure to expand their portfolio to a more diverse audience, requiring a new approach to risk modeling and making transparency and explainability even more important.' Below the cards is a 'Define details' section with a 'Name' field containing 'credit-risk' and a 'Description' field with the same detailed text. At the bottom right of the screen, there are 'Cancel' and 'Import' buttons, with the 'Import' button also circled in red.

## Deploy the Credit Risk Model

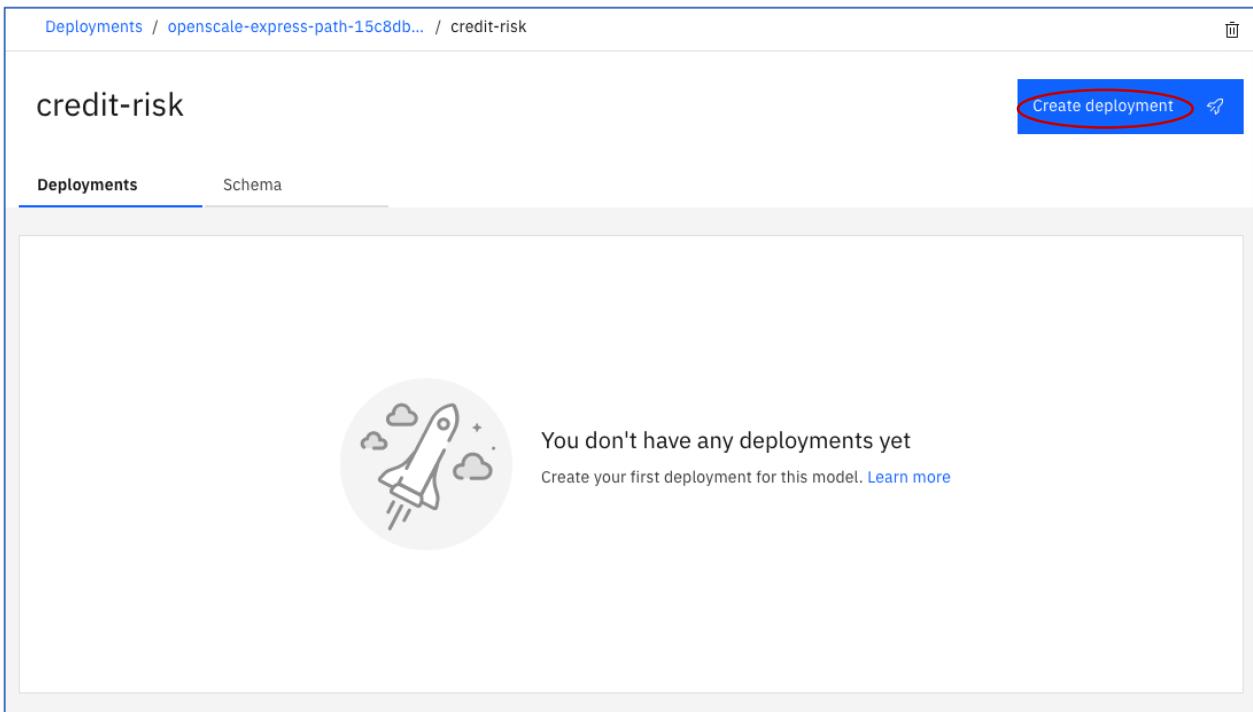
- On the following page, click **Promote to deployment space**.

The screenshot shows the 'credit-risk' project overview. At the top right, there is a 'Promote to deployment space' button, which is circled in red. To the right of the project name, there is a summary box with the following information: 'Last modified at May 16, 2021 4:42 PM', 'Type: mllib\_2.4', 'Model ID: bc250345-ec0b-4579-81a9-ed...', 'Software specification: spark-mllib\_2.4', and 'Tags: Add tags to make assets easier to find.' Below the project name, there are tabs for 'Overview' (which is selected) and 'Activities'.

2. Click on **Watson Studio Labs** for the Target space, check “**Go to the model in the space after promoting it,**” and click **Promote**.



3. Click Create deployment.



4. From the **Create Deployment** page, click **Online**, type **credit-risk-deploy** for the **Name**, and click on **Create**.

Create a deployment

Associated asset  
credit-risk

Deployment type

<b>Online</b> <input checked="" type="radio"/>	Run the model on data in real-time, as data is received by a web service.
<b>Batch</b>	Run the model against data as a batch process.

Name  
**credit-risk-deploy**

Description  
Deployment description

Tags  
Add tags to make assets easier to find.

Cancel **Create**

5. The deployment status should go from **In progress** to **Deployed**. If the status doesn't change after a minute or so, refresh the browser.

credit-risk

**Create deployment**

**Deployments** Schema

DEPLOYMENT TYPES		1 Online Deployment(s)		
		Name	Status	Last modified
Online	(1)	<a href="#">credit-risk-deploy</a>	Deployed	May 16, 2021 4:49 PM
Batch	(0)			

6. Click on **credit-risk-deploy**.

credit-risk

**Create deployment**

**Deployments** Schema

DEPLOYMENT TYPES		1 Online Deployment(s)		
		Name	Status	Last modified
Online	(1)	<a href="#">credit-risk-deploy</a>	Deployed	May 16, 2021 4:49 PM
Batch	(0)			

7. Keep this tab open, we will return to this page later.

The screenshot shows the IBM Cloud Pak for Data interface. At the top, the navigation bar includes 'Deployments / Watson Studio Labs / credit-risk / credit-risk-deploy'. Below this, the deployment name 'credit-risk-deploy' is shown with a green 'Deployed' status and an 'Online' button. A blue-bordered 'API reference' tab is active, while other tabs like 'Test' are inactive. Under the 'API reference' tab, there's a 'Direct link' section with an endpoint URL: <https://us-south.ml.cloud.ibm.com/v4/deployments/f38f4e1c-faf7-46c7-b4b1-2c3a6b502f39/predict>. To the right of the URL is a 'Bearer <token>' field with a help icon, and below it is an 'IAM' button. Below the endpoint URL, there's a 'Code snippets' section with tabs for 'cURL', 'Java', 'JavaScript', 'Python', and 'Scala'.

## Begin OpenScale Configuration

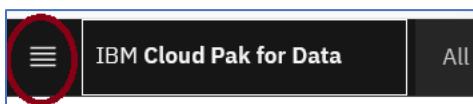
1. Right-click on **IBM Cloud Pak for Data**. Click on **Open Link in New Tab**.

The screenshot shows a browser context menu opened over the 'IBM Cloud Pak for Data' tab. The menu items are: 'Open Link in New Tab' (highlighted with a red circle), 'Open Link in New Window', 'Open Link in New Private Window', 'Bookmark Link', 'Save Link As...', 'Save Link to Pocket', 'Copy Link', 'Search Google for "IBM Cloud Pak f..."', 'Inspect Accessibility Properties', and 'Inspect (Q)'.

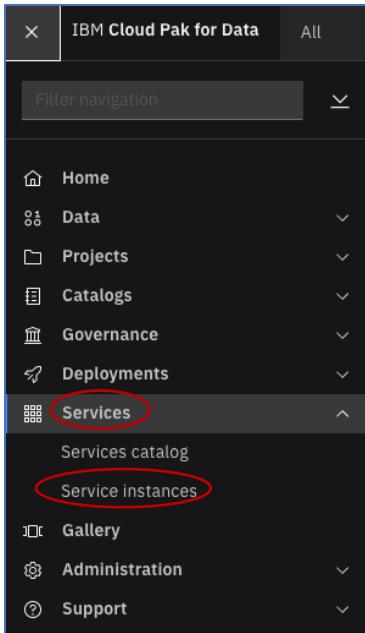
2. Click on the new **IBM Cloud Pak for Data** browser tab.



3. Click on the hamburger icon  in the top left corner.



4. Click on **Services**, and **Service instances**.



5. Click the box to the right of your OpenScale instance.

## Service instances

To upgrade a service plan, first [upgrade](#) your IBM Cloud account. Then choose [Upgrade service](#) or [Manage in IBM Cloud](#) from the menu in the service's row.

Filter by: Resource Groups ▾ 2 × Locations ▾ 1 × fctolabs17@gmail.com ▾ Prod

Find service instances

Name	Group	Location
cloud-object-storage-ts	Default	Global
WatsonMachineLearning	Default	Dallas
KnowledgeCatalog	Default	Dallas
Watson OpenScale-z3	Default	Dallas
WatsonStudio	Default	Dallas

6. Click on **Launch Application**.

Resource list /

Watson OpenScale-qe Active Add tags

Getting started Manage Plan

Watson OpenScale

Welcome to Watson OpenScale, let's get started.

**Launch Application**

Documentation Community

FEEDBACK

7. Make sure to click on **Manual setup**.

Welcome to Watson OpenScale

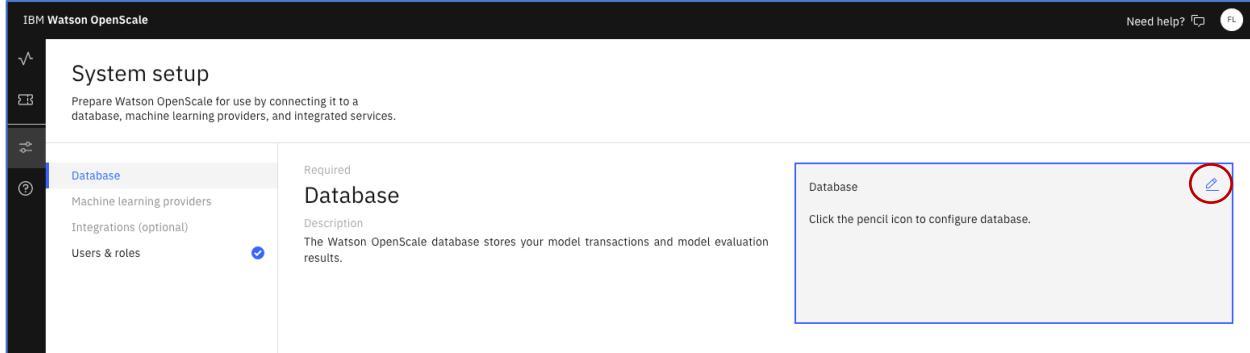
Watson OpenScale maintains the health of AI models in pre-production and production environments by measuring model quality, fairness, and drift in both data and accuracy. It provides AI model transparency by explaining model transactions.

To get up-and-running, we'll set up a machine learning provider, lite database, and sample model for you. The process will take about 10 minutes. Ready to go?

**Manual setup** **Auto setup**

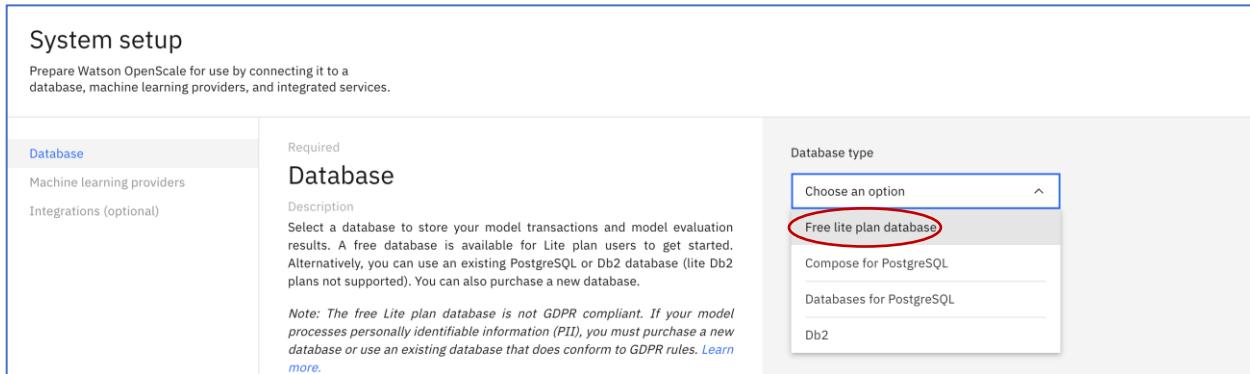
# Setup System

1. In the system setup, we need to set up a database to collect the payload logging data. We also need to specify which deployed model will be monitored.
2. Click on the  icon on the Database tile.



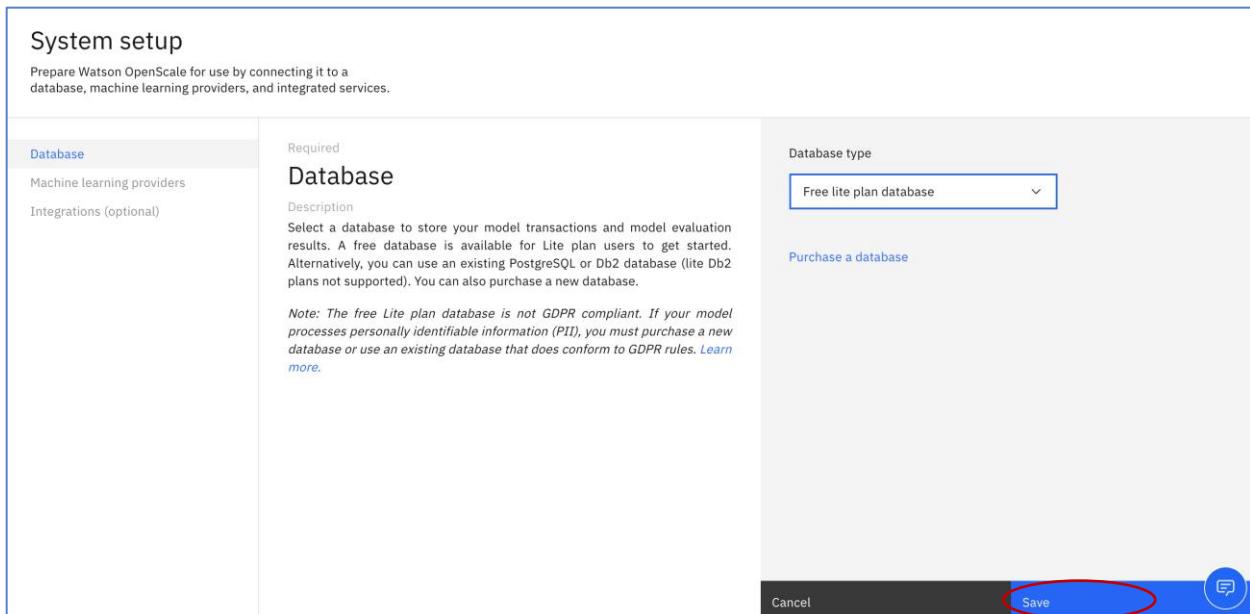
The screenshot shows the 'System setup' page of the IBM Watson OpenScale interface. On the left, there's a sidebar with icons for 'Database', 'Machine learning providers', 'Integrations (optional)', and 'Users & roles'. The 'Database' icon is highlighted with a blue border. The main area has a title 'System setup' and a sub-section 'Database'. It includes a 'Description' field with text about the Watson OpenScale database storing model transactions and evaluation results. To the right, there's a 'Database' section with a 'Description' field containing the text 'Click the pencil icon to configure database.' A red circle highlights the pencil icon at the end of this line.

3. Click on **Free Lite plan database** for **Database type** and click on **Save**.



The screenshot shows the 'System setup' page with the 'Database' tile selected. In the 'Database type' dropdown, the option 'Free lite plan database' is highlighted with a red circle. Other options listed are 'Compose for PostgreSQL', 'Databases for PostgreSQL', and 'Db2'.

4. Click **Save**.



The screenshot shows the 'System setup' page with the 'Database' tile selected. The 'Database type' dropdown now shows 'Free lite plan database'. At the bottom right of the page, there are 'Cancel' and 'Save' buttons. The 'Save' button is highlighted with a red circle.

## 5. Click on Machine learning providers.

System setup

Prepare Watson OpenScale for use by connecting it to a database, machine learning providers, and integrated services.

Database

Required

**Database**

Description

The Watson OpenScale database stores your model transactions and model evaluation results.

Database

Database type  
Free lite plan database

Database  
Internal database

Schema  
public

## 6. Click on Add machine learning provider.

System setup

Prepare Watson OpenScale for use by connecting it to a database, machine learning providers, and integrated services.

Database

Required

**Machine learning providers**

Description

Watson OpenScale connects to deployed models stored in a machine learning environment.

Add machine learning provider +

## 7. Click on pencil icon to edit the connection.

System setup

Prepare Watson OpenScale for use by connecting it to a database, a machine learning provider, and integrated services.

Database

Machine learning providers

Integrations (optional) [beta](#)

Back to all providers

Machine learning providers [edit](#)

New provider

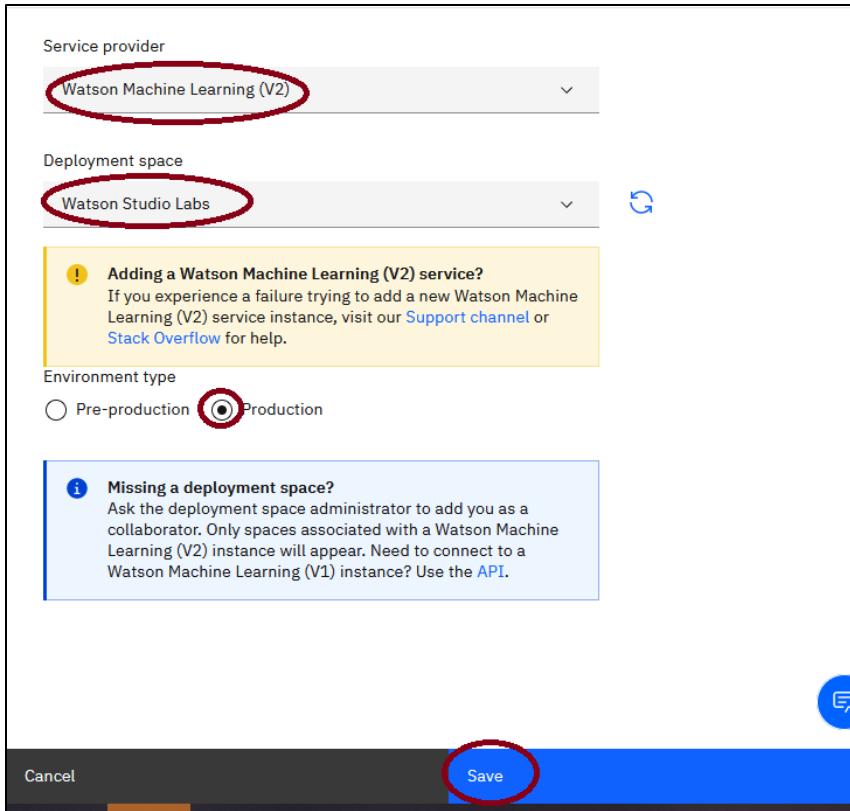
Description [edit](#)

Click edit to enter provider description.

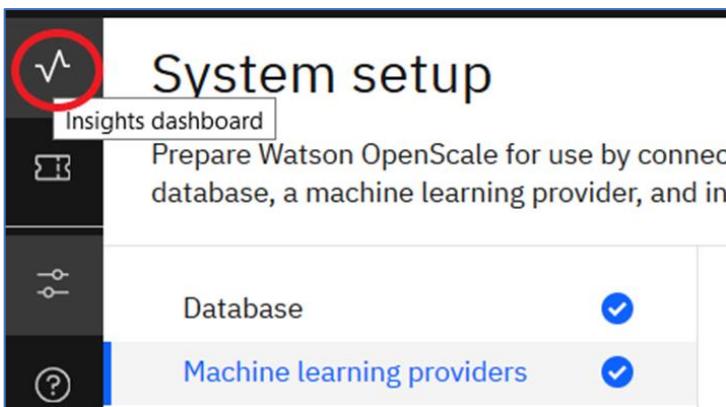
Connection

Click edit to enter the connection information.

## 8. Click on Watson Machine Learning (V2) for the Service provider type. Scroll down and click on Watson Studio Labs for the Deployment Space, click on Production for the Environment type, and click on Save.



9. Click on the icon.



10. Click on **Add** to add a deployment to monitor.

The screenshot shows the IBM Watson OpenScale Insights Dashboard. At the top right, there is a "Need help?" link and a "FL" button. Below the header, the title "Insights Dashboard" is displayed. To the right of the title are "Refresh" and "Add to dashboard" buttons, with the latter being circled in red. The main area contains four cards: "Deployments Monitored" (0), "Quality Alerts" (0), "Fairness Alerts" (0), and "Drift Alerts" (0). Below these cards are filtering options: "Filter by Tags", "Alert type", "Machine learning provider", and "Sort by Severity". A search bar "Find models" is also present.

11. Click on **credit-risk-deploy** and then click on **Configure** to configure the model details, and the quality, fairness, and drift monitors.

This screenshot shows a modal dialog titled "Select a model deployment". It asks to choose a machine learning provider and provides deployment details. A dropdown menu for "Machine learning Provider" shows "New provider (Production)". The main table lists four deployments:

Deployment	Description	Created
GermanCreditRiskModelChallenger	Created by Watson OpenScale Express Path.	Thu, Mar 25, 2021, 10:30 AM EDT
GermanCreditRiskModel	Created by Watson OpenScale Express Path.	Thu, Mar 25, 2021, 10:32 AM EDT
GermanCreditRiskModelPreProd	Created by Watson OpenScale Express Path.	Thu, Mar 25, 2021, 10:31 AM EDT
<input checked="" type="radio"/> credit-risk-deploy	-	Sun, May 16, 2021, 4:49 PM EDT

The "credit-risk-deploy" row has a radio button next to it, which is highlighted with a red circle. At the bottom of the dialog are "Cancel" and "Configure" buttons, with the "Configure" button being circled in red.

12. Click on **Configure monitors**.

This screenshot shows a success message dialog box. It says "Selections saved." and includes a green checkmark icon followed by the text "Done. Click **Configure monitors** to set up your monitors.". At the bottom of the dialog are "Close" and "Configure monitors" buttons, with the "Configure monitors" button being circled in red.

13. Setup is not complete. Do not proceed until you do the scoring step below.

## Score Data

Before proceeding with the monitor configuration, we need to send scoring data to the deployed model in order to generate payload logging data that the monitors can consume. We will use the `credit_risk_data.json` file as sample data that Watson Studio will submit to the deployed model.

1. Return to the tab from earlier that shows your deployed credit-risk model.

The screenshot shows the Watson Studio API reference page for the `credit-risk-deploy` model. At the top, it says "Deployments / Watson Studio Labs / credit-risk / credit-risk-deploy". Below that, it shows the model name "credit-risk-deploy" with a green "Deployed" status and an "Online" button. There are two tabs: "API reference" (which is selected) and "Test". Under "API reference", there's a "Direct link" section with an "Endpoint" field containing the URL `https://us-south.ml.cloud.ibm.com/v4/deployments/f38f4e1c-faf7-46c7-b4b1-2c3a6b502f39/predict`. To the right of the URL are "Bearer <token>" and an "IAM" button. Below the endpoint are "Code snippets" sections for cURL, Java, JavaScript, Python, and Scala. The "cURL" section is currently active.

2. Click on the **Test** tab.

The screenshot shows the same Watson Studio API reference page, but the "Test" tab is now selected, indicated by a red circle around it. The rest of the interface remains the same, showing the "Direct link" section with the same URL and "Code snippets" tabs.

3. Click on the icon to accept input as JSON.

credit-risk-deploy Deployed Online

API reference Test

Enter input data

Body

```
{  
    "input_data": [  
        {  
            "fields": [  
                ],  
            "values": [  
                ]  
        }  
    ]  
}
```

Predict

This screenshot shows the 'Test' tab of the Azure API Management interface for the 'credit-risk-deploy' service. The 'Enter input data' section is active, showing a JSON template for sending data to the API. The 'Body' field contains a JSON object with a single array named 'input\_data'. This array contains one object, which has two properties: 'fields' (an empty array) and 'values' (an array containing a single element). A red circle highlights the 'Predict' button at the bottom right of the input area.

4. Copy the file contents of **credit\_payload\_data.json** into the **input data** area and click on **Predict**.

Enter input data

```
    ", "yes", "Risk"  ],  
  
    ["no_checking",48,"credits_paid_to_date","business",4844,"less_100","unem  
ployed",3,"male","none",2,"car_other",33,"bank","rent",1,"management_self  
-employed",1,"yes","yes","Risk"  ],  
  
    ["less_0",36,"credits_paid_to_date","car_used",8229,"less_100","1_to_4",2  
,"male","none",2,"savings_insurance",26,"none","own",1,"skilled",2,"none"  
,"yes","Risk"  ],  
  
    ["less_0",45,"credits_paid_to_date","radio_tv",1845,"less_100","1_to_4",4  
,"male","none",4,"unknown",23,"none","free",1,"skilled",1,"yes","yes","Ri  
sk"  ]  
}
```

Predict

This screenshot shows the same 'Test' tab interface as the previous one, but with the 'input data' body filled with the contents of 'credit\_payload\_data.json'. The entire JSON payload is enclosed in a red rectangle. The 'Predict' button at the bottom right is also highlighted with a red circle.

5. The results should appear as below.

credit-risk-deploy Deployed Online

API reference Test

**Enter input data**

```
, "less_100", "1_to_4", 4, "male", "none", 2, "car_other", 28, "none", "own", 1, "skilled", 1, "yes", "yes", "Risk" ],
[ "less_0", 24, "credits_paid_to_date", "radio_tv", 2439, "less_100", "less_1", 4, "female", "none", 4, "real_estate", 35, "none", "own", 1, "skilled", 1, "yes", "yes", "Risk" ],
[ "no_checking", 10, "credits_paid_to_date", "furniture", 2210, "less_100", "less_1", 2, "male", "none", 2, "real_estate", 25, "bank", "rent", 1, "unskilled", 1, "none", "yes", "Risk" ],
[ "less_0", 9, "credits_paid_to_date", "car_new", 1422, "less_100", "less_1", 3, "male", "none", 2, "unknown", 27, "none", "free", 1, "management_self-employed", 1, "yes", "yes", "Risk" ],
```

**Result**

```
0 {
  "predictions": [
    {
      "fields": [
        "CheckingStatus",
        "LoanDuration",
        "CreditHistory",
        "LoanPurpose",
        "LoanAmount",
        "ExistingSavings",
        "EmploymentDuration",
        "InstallmentPercent",
        "Sex",
        "OthersOnLoan",
        "CurrentResidenceDuration",
        "OwnsProperty",
        "Age",
        "InstallmentPlans".
      ]
    }
  ]
}
```

Predict

## Configure Model Details

1. Switch back to Watson OpenScale by clicking on the Watson OpenScale browser tab.



2. Click on pencil icon to edit the **Model input**.

**Model details**

Description

Provide information about the training data and deployed model output to prepare Watson OpenScale for monitoring and providing explanations for model transactions.

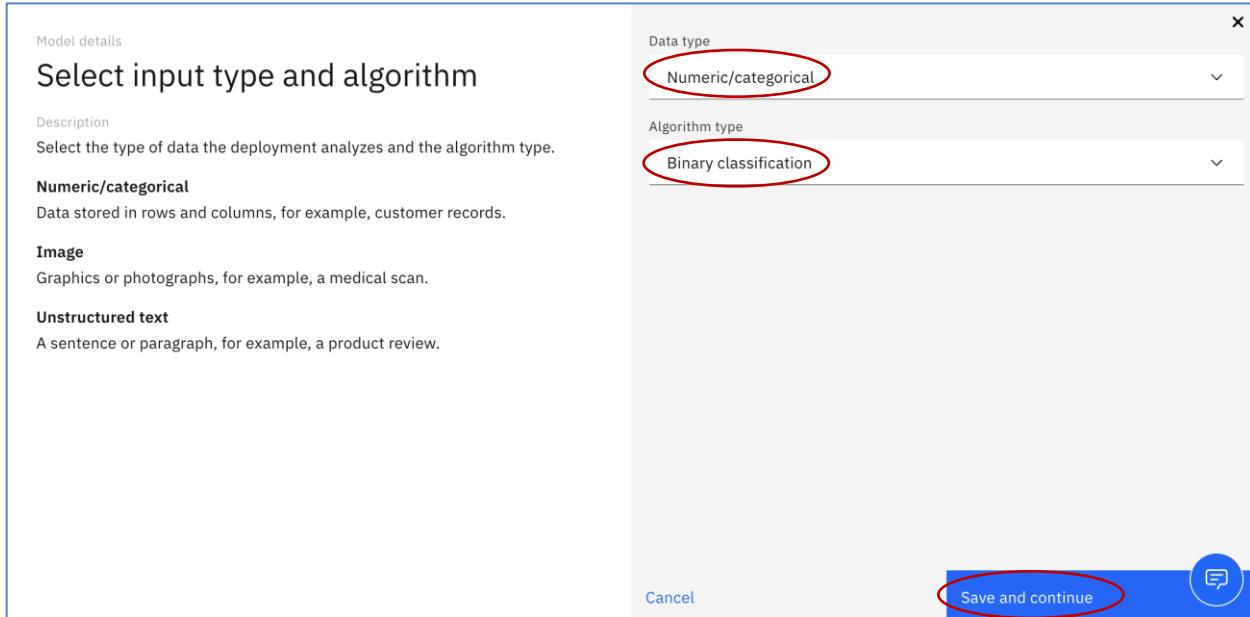
**Model input**

To select the data type and algorithm type, click the edit icon.

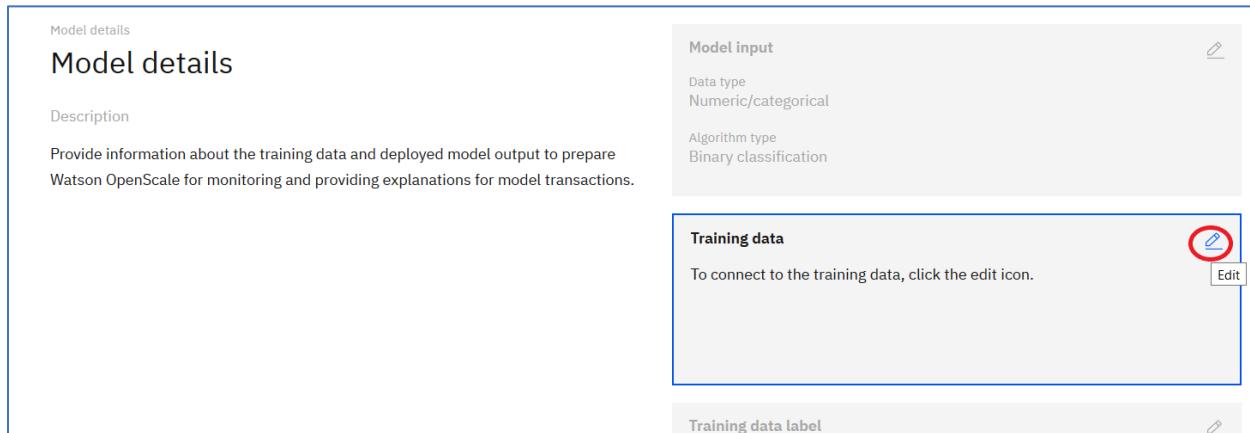
**Model transaction**

Model transaction  
Successful

- Click on **Numeric/categorical** for the **Data type**. Click on **Binary classification** for the **Algorithm type**. Click on **Save and Continue**.



- Click on pencil icon to edit **Training data**.



- Select **Cloud Object Storage** as the **Location**.
- Copy and paste the following text as the **resource\_id**:

1613a7c4-d7b1-4872-a1c2-41db3b7f3d91

- Copy and paste the following text as the **apikey**:

3TMR6f7dKSPuzOWEEGsPoI1I7BCnUPNr6CrmCyL2zGBI

- Click **Connect**

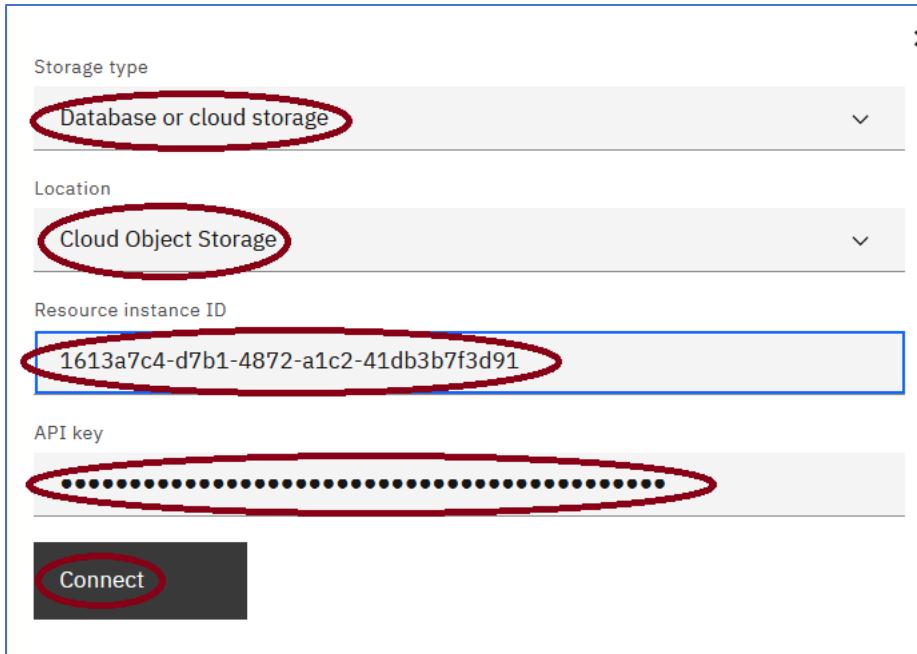
Storage type  
Database or cloud storage

Location  
Cloud Object Storage

Resource instance ID  
1613a7c4-d7b1-4872-a1c2-41db3b7f3d91

API key  
.....

Connect



9. Select **a3trainingdata** as the bucket, select **german\_credit\_data\_biased\_training.csv** as the Data set, and click **Next**.

Storage type  
Database or cloud storage

Location  
Cloud Object Storage

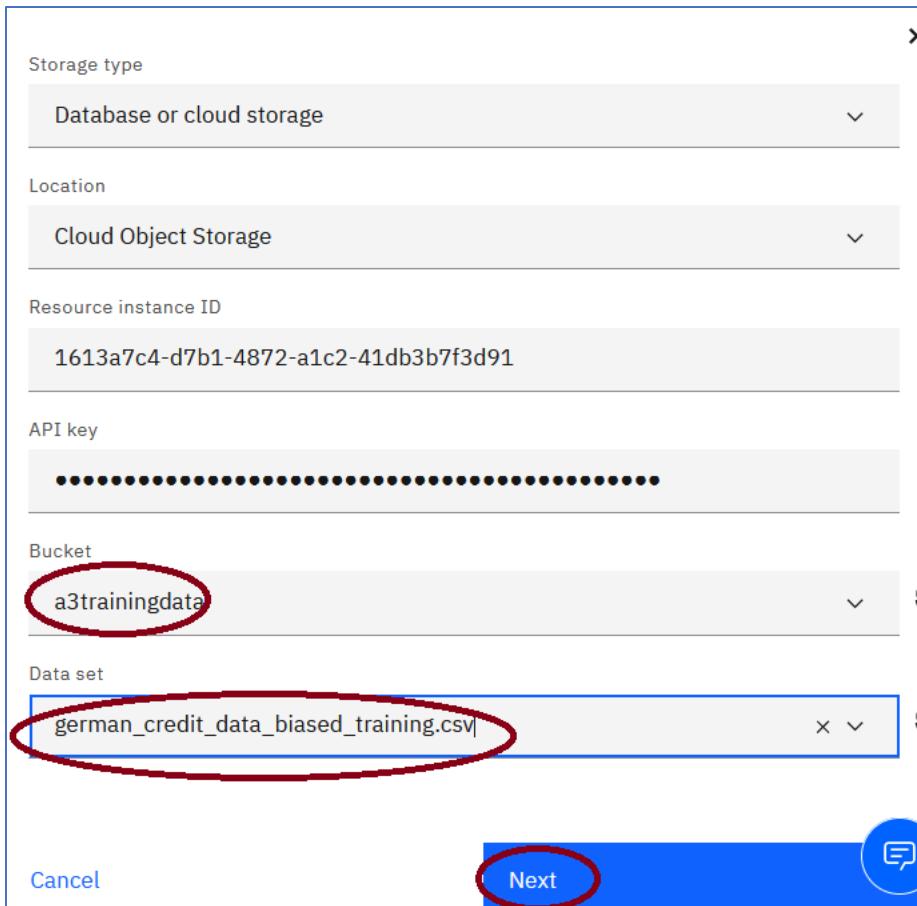
Resource instance ID  
1613a7c4-d7b1-4872-a1c2-41db3b7f3d91

API key  
.....

Bucket  
a3trainingdata

Data set  
german\_credit\_data\_biased\_training.csv

Cancel Next



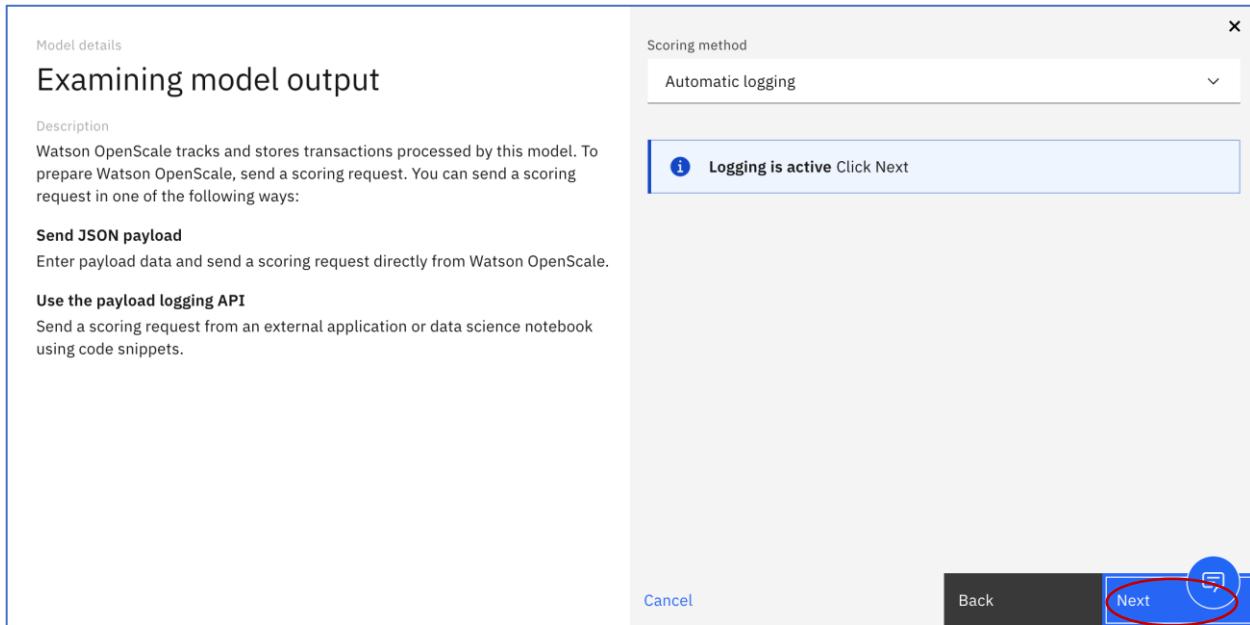
10. Watson OpenScale has determined that Risk is the label (target) column. Click **Next**.

The screenshot shows the 'credit-risk-deploy' interface. On the left, there's a sidebar with 'Model info' sections like 'Model details', 'Endpoints', 'Evaluations' (Fairness, Quality, Drift, Explainability), and 'Import settings'. The main area is titled 'Select the label column' with a description: 'From the selected training features, select a single feature as the label. The label represents the correct prediction (ground-truth) for each record.' A modal window titled 'Select the label column' lists 'Features (1)': 'Risk' (Type: A). At the bottom right of the modal is a 'Next' button, which is circled in red.

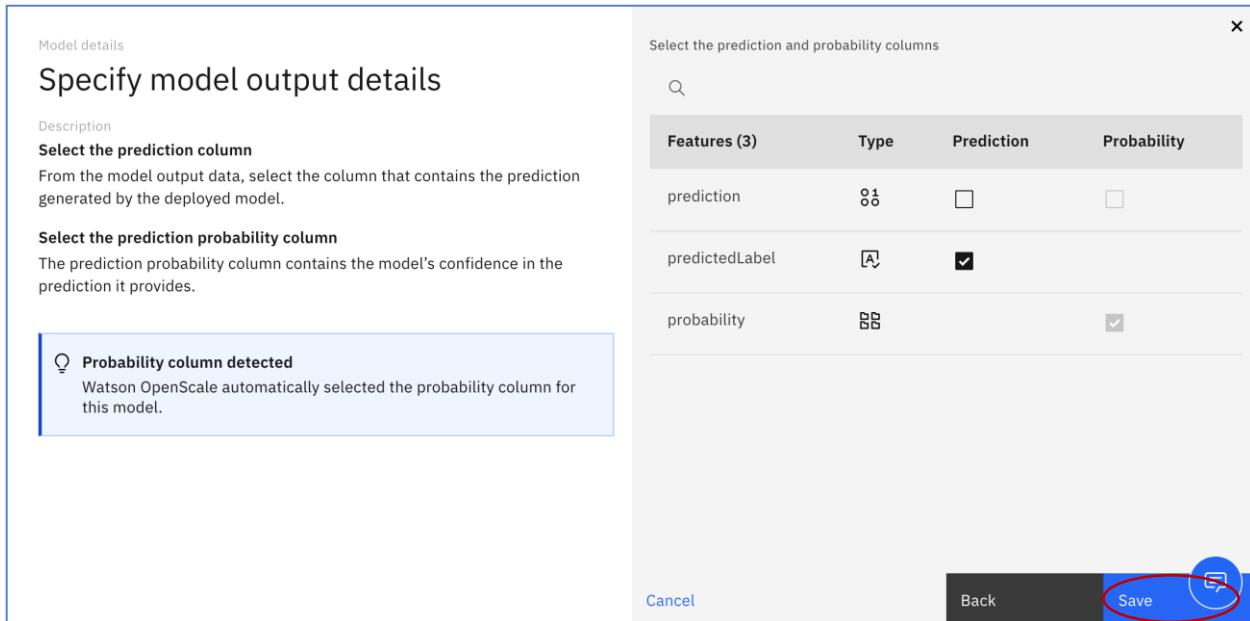
11. All of the features are used to train the model. Watson OpenScale has determined the feature types. Click **Next**.

The screenshot shows the 'credit-risk-deploy' interface. The main area is titled 'Select the training features' with a description: 'From the training data, select the features used to train the model. For each selected feature with a numeric data type, indicate if the numeric feature values represent codes or categories by checking the Categorical checkbox. This instructs OpenScale to process the feature as a set of classifications (ex. ZIP code) rather than continuous numeric values (ex. median household income).' A modal window titled 'Select the training features' lists 'Features (20)': Age, CheckingStatus, CreditHistory, CurrentResidenceDuration, Dependents, EmploymentDuration. The 'Type' and 'Categorical' columns are shown. Several checkboxes are checked, notably for CheckingStatus, CreditHistory, and EmploymentDuration. At the bottom right of the modal is a 'Next' button, which is circled in red.

12. We've already set up the logging (**Logging is active**) so click **Next**.



13. Watson OpenScale has determined the feature that contains the prediction generated by the AI deployment. Click **Save**.



14. Model details are now complete. Note that Explainability input is also complete. The next step is to configure Quality.

# Configure Quality

The Quality Monitor evaluates how well your deployed model predicts accurate outcomes. It identifies when model quality declines so you can retrain appropriately.

1. Click **Quality** to configure the quality inputs.

This screenshot shows the Watson OpenScale configuration interface for a model named "credit-risk-deploy". The left sidebar lists "Model info" (Model details, Endpoints, Evaluations), "Evaluations" (Fairness, Quality, Drift, Explainability), "Import settings", and "Go to model summary". The main panel displays "Model details" with a "Description" section: "Provide information about the training data and deployed model output to prepare Watson OpenScale for monitoring and providing explanations for model transactions." To the right, there are two sections: "Model input" (Data type: Numeric/categorical, Algorithm type: Binary classification) and "Training data" (Storage type: Database or cloud storage, Location of training data: Db2, Hostname or IP address: dashdb-txn-sbox-yp-dal09-03.services.dal.bluemix.net, SSL port). A red circle highlights the "Quality" link under "Evaluations" in the sidebar.

2. Click on the pencil icon to edit the **Quality threshold**.

This screenshot shows the configuration of the "Quality" metric. The left sidebar is identical to the previous screenshot. The central panel shows the "Quality" metric with a "Description": "The Quality monitor evaluates how well your model predicts accurate outcomes. It identifies when model quality declines, so you can retrain your model appropriately." Below this is a note: "Note: The Quality metric measures the model's ability to correctly predict outcomes that match labeled data (ground truth) provided by humans. The quality metrics evaluated are standard data science statistics based on model type. [Learn more](#)". To the right, there are two sections: "Quality threshold" (with a red circle around the edit icon) and "Sample size" (with a red circle around the edit icon). A blue message icon is visible in the bottom right corner of the main panel.

3. Enter 0.9 for the Threshold value and click **Next**.

Quality

## Quality threshold

Description

**Area under ROC**

For binary classification models, it will check the score which measures the model's ability to distinguish two classes. The higher the score, the better the model is at identifying class A as class A and class B as class B. The score of 1 is perfect, while a score of .50 is equal to a random guess.

Additional quality metrics captured for this model include:

- Area under PR
- Accuracy
- True positive rate (TPR)
- False positive rate (FPR)
- Recall
- Precision
- F1-Measure
- Logarithmic loss

Threshold value: Area under ROC

0.9

Cancel Next

4. Set **100** for the minimum the sample size to be analyzed for quality and click **Save**.

Quality

## Sample size

Description

Ensure that your minimum sample size is large enough to be accurately represent the variety of requests the deployment receives.

Minimum sample size (number of transactions)

100

Cancel Back Save

5. This completes the **Quality** configuration. The next step is to configure Fairness.

## Configure Fairness

The Fairness monitor checks your deployment for biases. It tracks when the model shows a propensity to provide a particular outcome more often for one group over another.

1. Click on **Fairness** to configure the Fairness monitor.

Dashboard / Configure

### credit-risk-deploy

**Model info**

- Model details
- Endpoints

**Evaluations**

- Fairness**  (highlighted with a red circle)
- Quality
- Drift
- Explainability

**Import settings**

[Go to model summary](#)

**Quality**

**Description**

The Quality monitor evaluates how well your model predicts accurate outcomes. It identifies when model quality declines, so you can retrain your model appropriately.

**Note:** The Quality metric measures the model's ability to correctly predict outcomes that match labeled data (ground truth) provided by humans. The quality metrics evaluated are standard data science statistics based on model type. [Learn more](#).

**Quality threshold**

Threshold value: Area under ROC  
0.9

**Sample size**

Minimum sample size  
100

Maximum sample size  
10,000

2. Click on the pencil icon to edit the **Favorable outcomes**.

Dashboard / Configure

### credit-risk-deploy

**Model info**

- Model details
- Endpoints

**Evaluations**

- Fairness**  (highlighted with a red circle)
- Quality
- Drift
- Explainability

**Import settings**

[Go to model summary](#)

**Fairness**

**Description**

The Fairness monitor checks your deployments for biases. It tracks when the model shows a tendency to provide a favorable (preferable) outcome more often for one group over another. You will specify which values represent favorable outcomes, select the features to monitor for bias (for example, Age or Sex), and specify the groups to monitor for each selected feature.

**Favorable outcomes**

To select the favorable outcomes, click the edit icon.

**Sample size**

To select the minimum sample size, click the edit icon.

**Features to evaluate (0)**

Add feature +

3. Check the Favorable check box for **No Risk** and check the Unfavorable check box for **Risk**. Then click **Next**.

Fairness

## Select the favorable outcomes

Description

For each group, Watson OpenScale will calculate the percentage of transactions that receive a favorable outcome.

Select the values that represent favorable (preferable) outcomes. You can also add a value manually if it is not included in the list.

Select the favorable outcomes

Enter a value

Add value

Values	Favorable	Unfavorable
No Risk	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Risk	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Cancel

4. Change the **Minimum sample size** to 200 and click **Next**.

Fairness

## Sample size

Description

Ensure that your minimum sample size is large enough to be accurately represent the variety of requests the deployment receives.

Minimum sample size (number of transactions)

200

Cancel

5. Scroll if needed and click on the **Sex** check box and the **Age** check box to monitor those features for bias, and then click on **Next**. Note, these may already be checked.

Fairness

## Select the features to monitor

**Description**

For each feature you select, Watson OpenScale will monitor the deployed model's tendency to provide a favorable (preferred) outcome for one group over another.

Features are monitored individually, but the Watson OpenScale debiasing algorithm will correct bias issues for all monitored features together.

ⓘ With the Lite plan, you can select up to 2 features to monitor. [View upgrade options.](#)

**Recommended features**

Watson OpenScale analyzed your training data to recommend which features should be monitored for fairness. These features are identified in the Recommended column.

Feature	Value	Action
<input checked="" type="checkbox"/> Sex	81	<input checked="" type="checkbox"/>
<input type="checkbox"/> OthersOnLoan		<input type="checkbox"/>
<input type="checkbox"/> CurrentResidenceDuration	81	<input type="checkbox"/>
<input type="checkbox"/> OwnsProperty		<input type="checkbox"/>
<input checked="" type="checkbox"/> Age	81	<input checked="" type="checkbox"/>
<input type="checkbox"/> InstallmentPlans		<input type="checkbox"/>
<input type="checkbox"/> Housing		<input type="checkbox"/>

Cancel Back Next

6. Enter 19 for the **minimum value** and 25 for the maximum value and then click **Add value**. We are defining the age range 19-25 as the age monitoring group.

Fairness

## Specify the monitored groups for [Age]

**Description**

Add value ranges and select the groups to monitor. Minimum and maximum values from the training data are presented for reference.

The percentage of favorable outcomes delivered to the monitored groups will be compared to the percentage of favorable outcomes delivered to the remaining groups (the reference groups) to check for potential bias.

A fairness score of 100% implies that the monitored group and reference group received an equal number of favorable values. Likewise, a fairness score of 50% implies that the monitored group received half as many favorable outcomes as the reference group.

Set the fairness alert threshold to track when the fairness value falls below an acceptable level.

**Recommended groups**

Watson OpenScale analyzed this feature to recommend which groups should be monitored for fairness. These groups are identified in the Recommended column.

Values	Monitored	Reference	Recommended
19-43	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Reference
44-67	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Monitored

Set fairness alert threshold [Age]

80

Cancel Back Next

7. Enter 26 as the **minimum value** and 74 as the **maximum value** and then click **Add value**. We are defining the age range 26-74 as the reference group.

Fairness

## Specify the monitored groups for [Age]

Description

Add value ranges and select the groups to monitor. Minimum and maximum values from the training data are presented for reference.

The percentage of favorable outcomes delivered to the monitored groups will be compared to the percentage of favorable outcomes delivered to the remaining groups (the reference groups) to check for potential bias.

A fairness score of 100% implies that the monitored group and reference group received an equal number of favorable values. Likewise, a fairness score of 50% implies that the monitored group received half as many favorable outcomes as the reference group.

Set the fairness alert threshold to track when the fairness value falls below an acceptable level.

**Recommended groups**

Watson OpenScale analyzed this feature to recommend which groups should be monitored for fairness. These groups are identified in the Recommended column.

Values	Monitored	Reference	Recommended
19-25	<input type="checkbox"/>	<input type="checkbox"/>	
19-43	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Reference
44-67	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Monitored

Select the groups to monitor [Age]

Minimum value: 19      Maximum value: 74

26      74      Add value

Set fairness alert threshold [Age]

80

Cancel      Back      Next

8. Check the **Monitored** check box for the 19-25 group and check the **Reference** check box for the 26-74 group. Set **95** for the **fairness alert threshold**. If OpenScale has already added “recommended” groups (as seen in the screenshots), uncheck those and ensure that the 19-25 group and 26-74 group you added are the only ones checked. Click **Next**.

Fairness

## Specify the monitored groups for [Age]

Description

Add value ranges and select the groups to monitor. Minimum and maximum values from the training data are presented for reference.

The percentage of favorable outcomes delivered to the monitored groups will be compared to the percentage of favorable outcomes delivered to the remaining groups (the reference groups) to check for potential bias.

A fairness score of 100% implies that the monitored group and reference group received an equal number of favorable values. Likewise, a fairness score of 50% implies that the monitored group received half as many favorable outcomes as the reference group.

Set the fairness alert threshold to track when the fairness value falls below an acceptable level.

**Recommended groups**

Watson OpenScale analyzed this feature to recommend which groups should be monitored for fairness. These groups are identified in the Recommended column.

Values	Monitored	Reference	Recommended
19-25	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
19-43	<input type="checkbox"/>	<input type="checkbox"/>	Reference
26-74	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
44-67	<input type="checkbox"/>	<input type="checkbox"/>	Monitored

Select the groups to monitor [Age]

Minimum value: 19      Maximum value: 74

26      74      Add value

Set fairness alert threshold [Age]

95

Cancel      Back      Next

9. Check the **Monitored** check box for female and check the **Reference** check box for male. Set **95** for the **fairness alert threshold**. Click on **Save**.

The screenshot shows the Watson OpenScale interface for configuring Fairness. On the left, the main page has a section titled 'Specify the monitored groups for [Sex]' with a 'Description' and a note about monitoring groups for potential bias. A 'Recommended groups' box is present. On the right, a modal window titled 'Select the groups to monitor [Sex]' lists 'female' and 'male' with checkboxes for 'Monitored' and 'Reference'. The 'female' row has 'Monitored' checked. The 'male' row has 'Reference' checked. Below this is a field 'Set fairness alert threshold [Sex]' containing '95', which is circled in red. At the bottom of the modal are 'Cancel', 'Back', and 'Save' buttons, with 'Save' also circled in red.

10. This completes the Fairness configuration. The next step is to configure Drift.

## Configure Drift

The Drift monitor measures two types of changes. It measures the drop in accuracy of the deployed model during runtime. The model accuracy could drop if there is an increase in the number of transactions similar to ones that the model was not able to accurately evaluate in the training data.

It measures the drop in consistency of the data in runtime as compared to the characteristics of the data at training.

1. Click on **Drift** to configure the Drift monitor.

The screenshot shows the Watson OpenScale interface for configuring Drift. On the left, the 'credit-risk-deploy' project is selected. In the 'Model info' section, 'Fairness' is highlighted with a red circle. In the 'Evaluations' section, 'Drift' is also highlighted with a red circle. The central area shows the 'Fairness' configuration with a description of what it does. To the right, there are sections for 'Favorable outcomes' (listing 'Favorable outcomes' and 'No Risk') and 'Sample size' (listing 'Minimum sample size' and '200').

- Click on the pencil icon to select a **drift model training option**.

Dashboard / Configure

### credit-risk-deploy

<b>Model info</b>	<b>Drift</b>
Model details	<input checked="" type="checkbox"/>
Endpoints	<input checked="" type="checkbox"/>
<b>Evaluations</b>	<b>Drop in accuracy</b>
Fairness	<input checked="" type="checkbox"/>
Quality	<input checked="" type="checkbox"/>
<b>Drift</b>	<input type="radio"/>
Explainability	<input checked="" type="checkbox"/>
<b>Import settings</b>	A drop in model accuracy and data consistency may lead to a negative impact on the business outcomes associated with the model.
<a href="#">Go to model summary</a>	

**Drift model**  
To select a drift model training option, click the edit icon.

**Drift threshold**  
To select drift threshold value, click the edit icon.

**Sample size**  
To select sample size value, click the edit icon.

- Watson OpenScale will detect a drop in accuracy using a custom drift model generated from the training data. Similarly, it detects a drop in data consistency by analyzing your training data. Click on **Train in Watson OpenScale**. Click **Next**.

Drift

### Train a drift model

Description

Watson OpenScale will detect a drop in accuracy using a custom drift model generated from your training data. Similarly, it detects a drop in data consistency by analyzing your training data. Watson OpenScale can analyze the data and train the model for you or you can do it yourself using a custom notebook.

**Train in Watson OpenScale**

If you connected your training data to Watson OpenScale and it is less than 500 MB, use this option.

**Train in a data science notebook**

If you did not connect your training data to Watson OpenScale or it exceeds 500 MB, use this option.

Training option

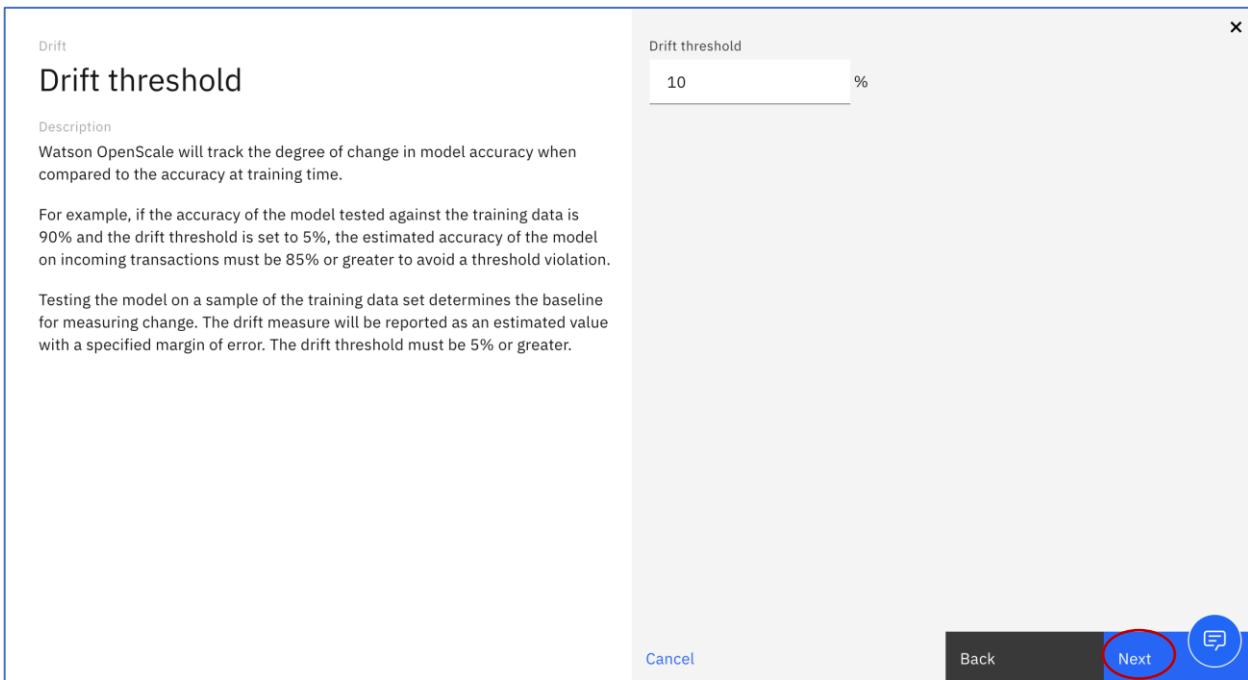
Train in Watson OpenScale

Train in a data science notebook

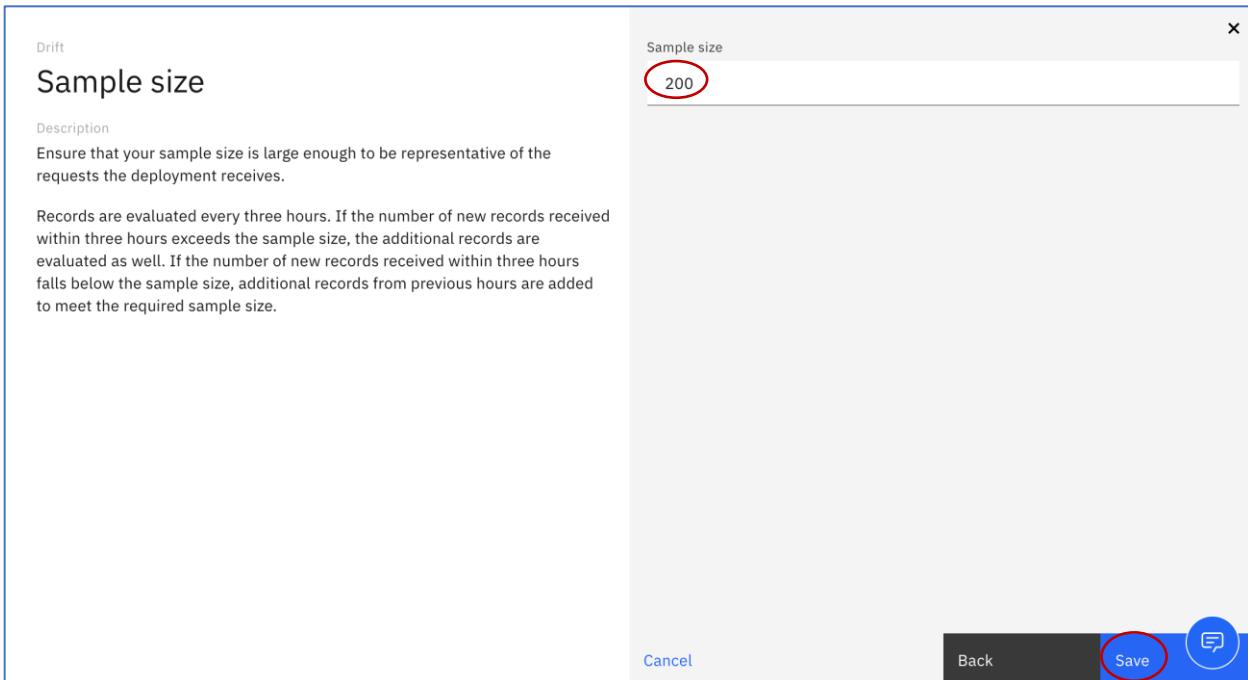
Cancel

Next

4. Set the drift alert threshold. Leave **10%**. Click **Next**.



5. Set the **Sample size** to compute Drift. Select **200**. Click **Save**.

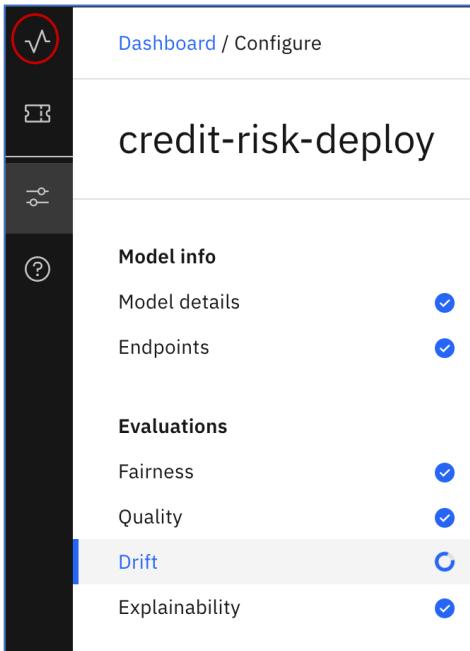


6. This completes the Drift configuration.

## Submit Feedback and View Quality Metrics

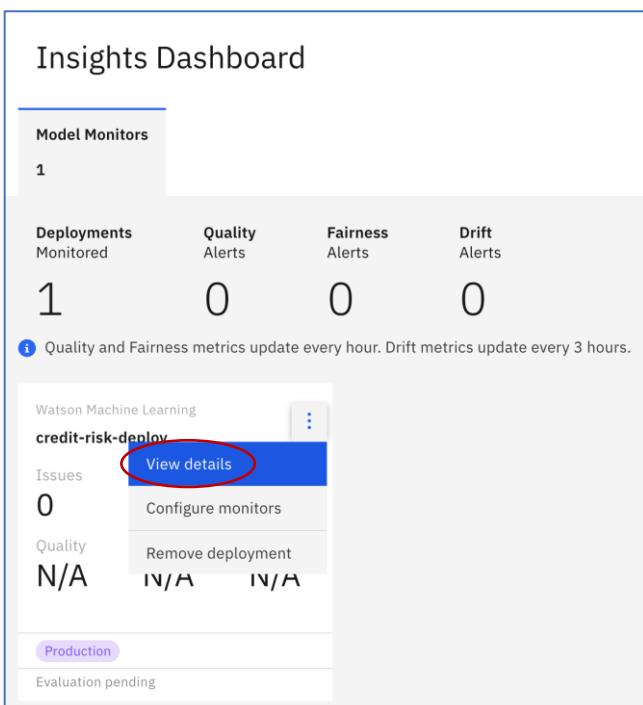
In order to measure quality, scored transactions including human labeled feedback must be provided.

1. Click on the  to display the **Insights Dashboard**.



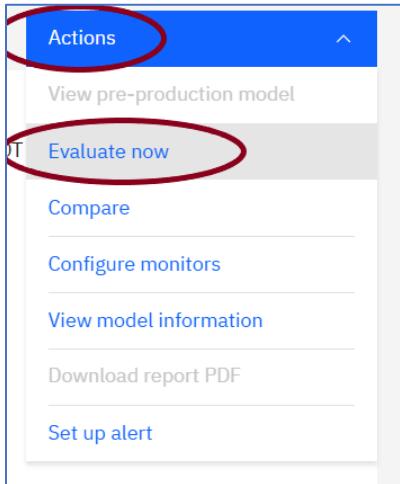
The screenshot shows the 'Model info' section of the Insights Dashboard. On the left is a sidebar with icons for Dashboard, Model info, Evaluations, Drift, and Explainability. The main area displays 'credit-risk-deploy' under 'Model info'. It includes sections for 'Model details' (checked), 'Endpoints' (checked), 'Evaluations' (checked), 'Drift' (highlighted with a blue background), and 'Explainability' (checked).

2. Click on vertical ellipse  and then click on **View Details**.

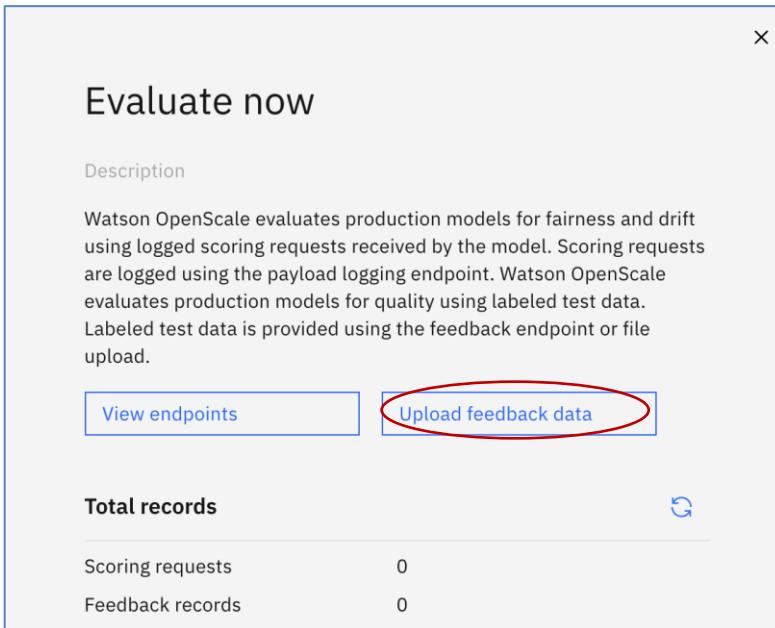


The screenshot shows the 'Insights Dashboard' for 'credit-risk-deploy'. It features a summary of monitored deployments (1), alerts (0 for Quality, 0 for Fairness, 0 for Drift), and a note that Quality and Fairness metrics update every hour while Drift updates every 3 hours. Below this is a detailed view of the deployment, showing 0 issues, 0 Quality alerts, 0 Fairness alerts, and 0 Drift alerts. A context menu is open over the deployment name, with the 'View details' option highlighted and circled in red. Other options in the menu include 'Configure monitors' and 'Remove deployment'. At the bottom, it shows the deployment is in 'Production' status with 'Evaluation pending'.

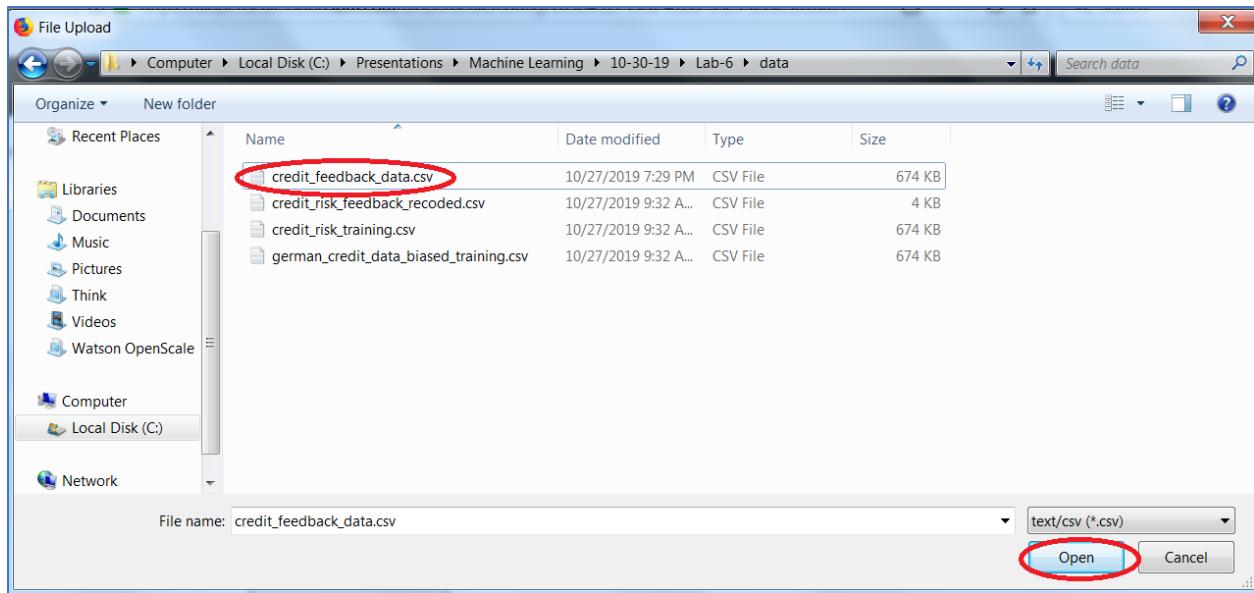
3. It will take a minute or so to bring up the **Evaluations** page. Click on **Actions** in the top right corner of your dashboard and then click **Evaluate now**.



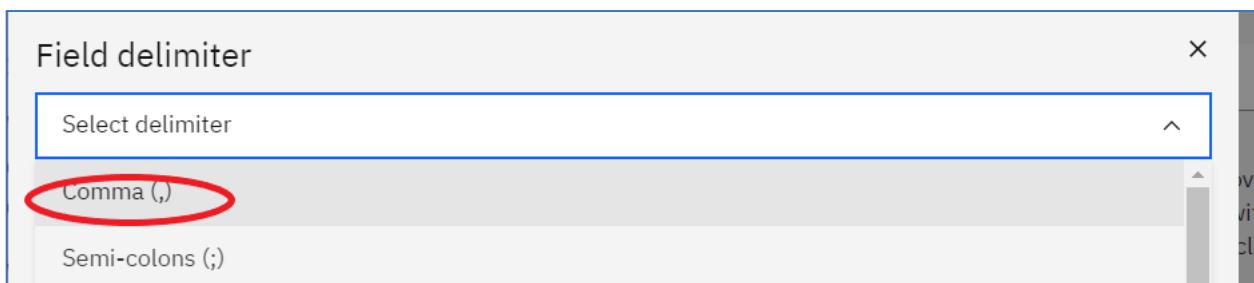
4. Watson OpenScale evaluates models for quality using labeled test data. Labeled test data is provided using the feedback endpoint or by file upload. We will use the file upload options. Click on **Upload feedback data**.



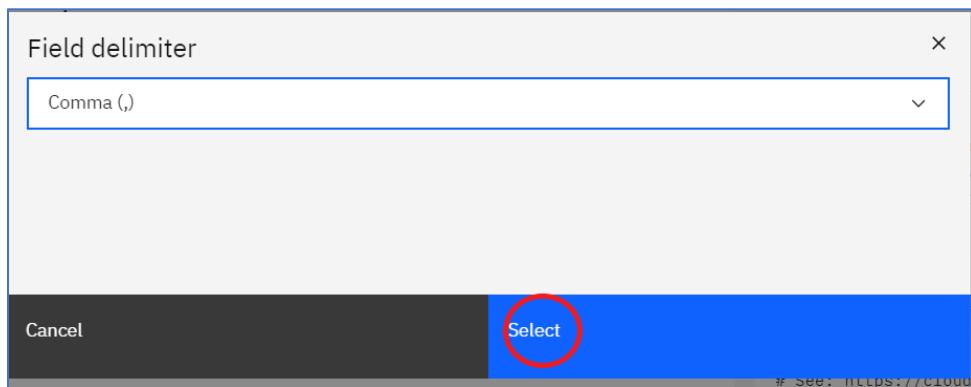
5. Navigate to the feedback file **credit\_feedback\_data.csv**. Click on **Open**.



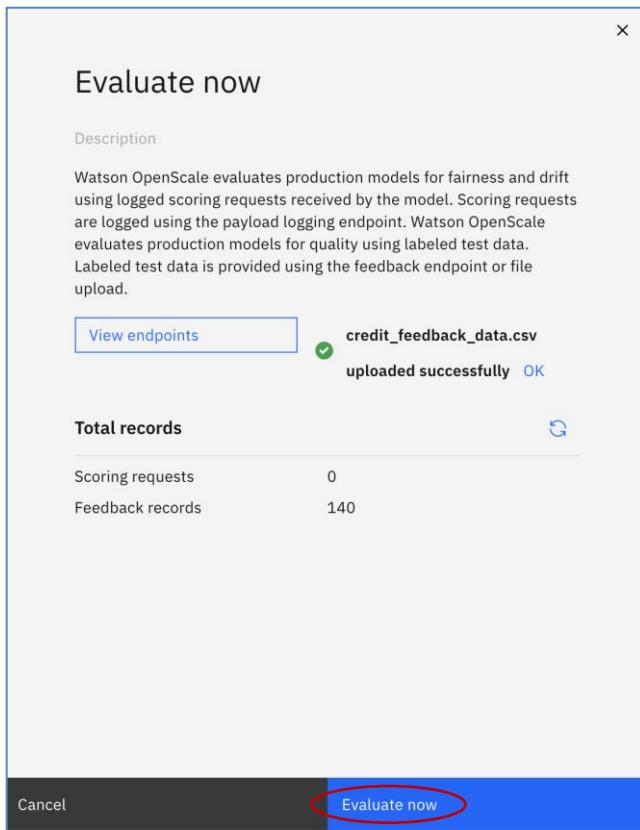
6. Select **Comma** as the **Delimiter**.



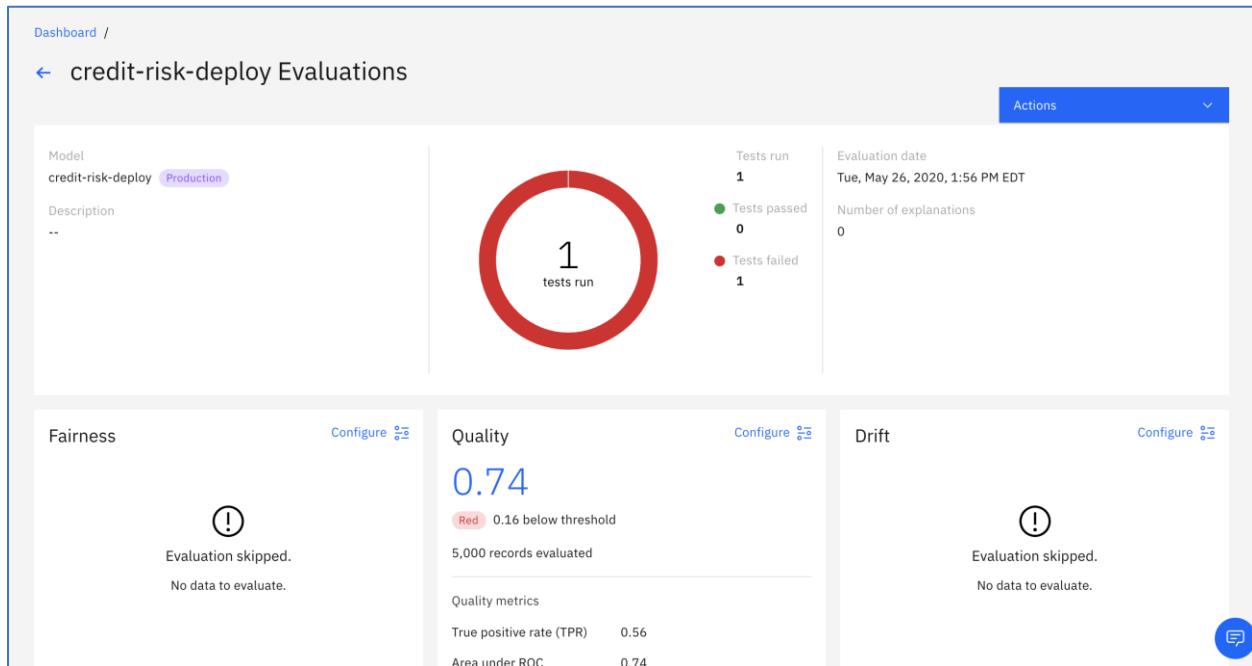
7. Click **Select**.



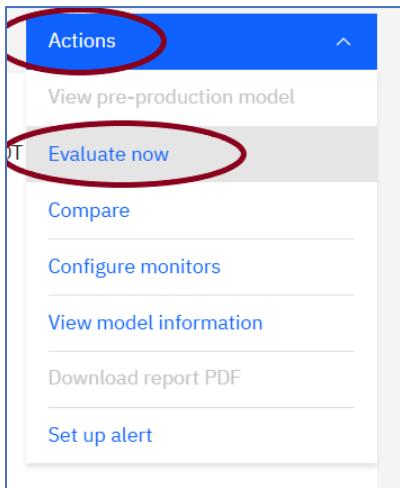
8. After the file is successfully uploaded, click on the **Evaluate now**.



9. Please wait about 3-5 minutes until the evaluation is done. The quality drift results are displayed. The quality test failed because the accuracy is below the quality threshold.



10. Click on **Actions** and click on **Evaluate Now**.

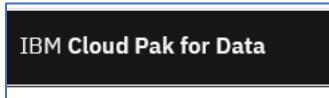


11. Leave Watson OpenScale at the Evaluation page while, you score transactions below.

## Score Transactions and View Fairness and Drift Metrics

In order to display Fairness metrics and Drift metrics, we need to direct transactions to the deployed model. We will use the scoring.json file as sample data that Watson Studio will submit to the deployed model

1. Return to Cloud Pak for Data by clicking on the **Cloud Pak for Data** browser tab.



2. You should be at the **Test** tab of the **credit-risk-deploy** page.

A screenshot of the Watson Studio interface showing the 'credit-risk-deploy' project. The 'Test' tab is selected. On the left, there is an 'Enter input data' section with a JSON editor containing sample data. A 'Predict' button is visible at the bottom left. On the right, there is a large empty workspace area.

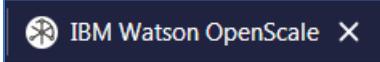
3. CLEAR out the contents of the **input data** area.

4. Navigate to where the scoring.json file and cut and paste the contents of the file into the **input data** area.

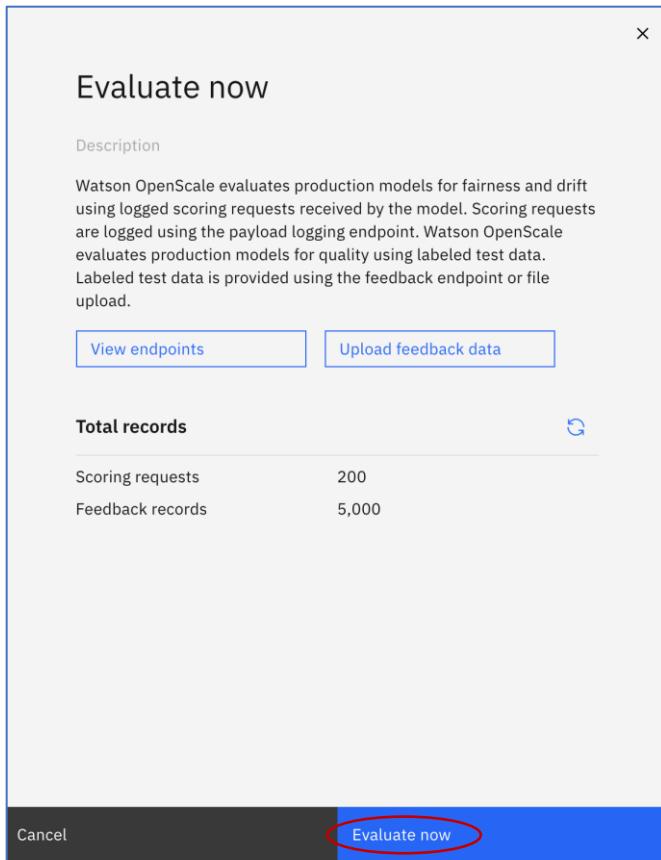
```
  "fields": ["CheckingStatus", "LoanDuration", "CreditHistory", "LoanPurpose", "LoanAmount", "ExistingSavings", "EmploymentDuration", "InstallmentPercent", "Sex", "OthersOnLoan", "Cur
```

## 5. Click on Predict

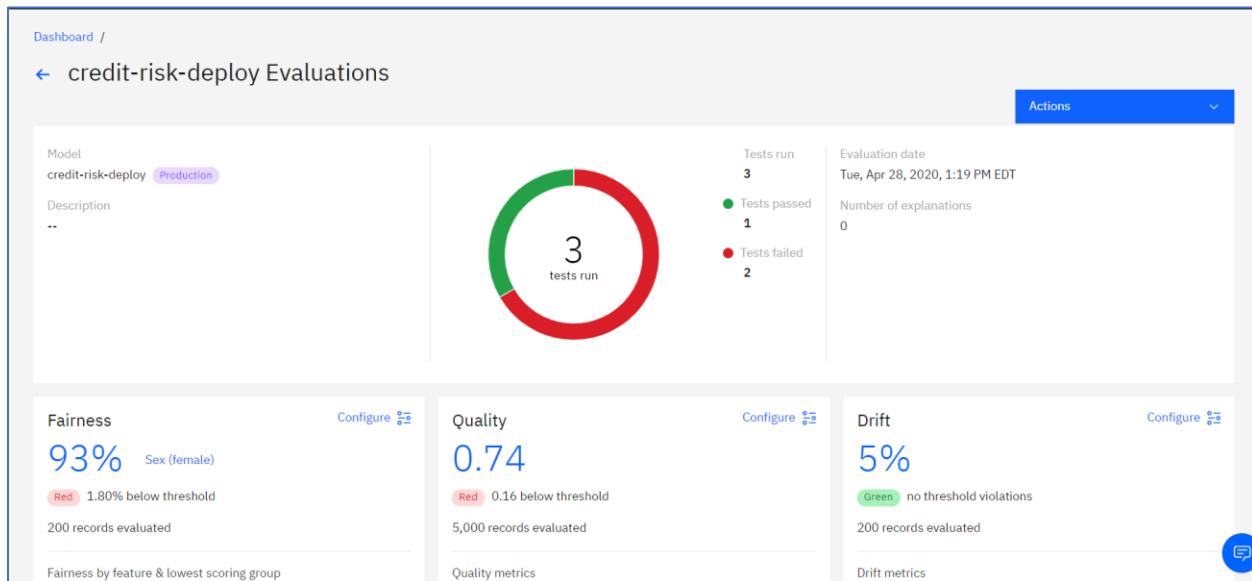
6. Return to Watson OpenScale by clicking on the **Watson OpenScale** browser tab.



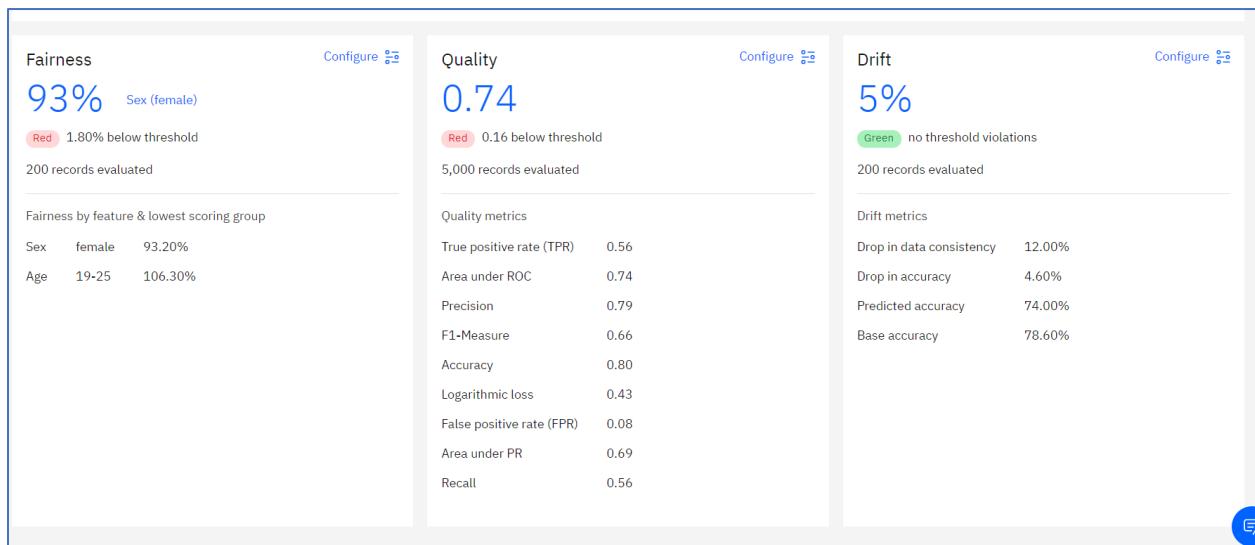
### **7. Click Evaluate now.**



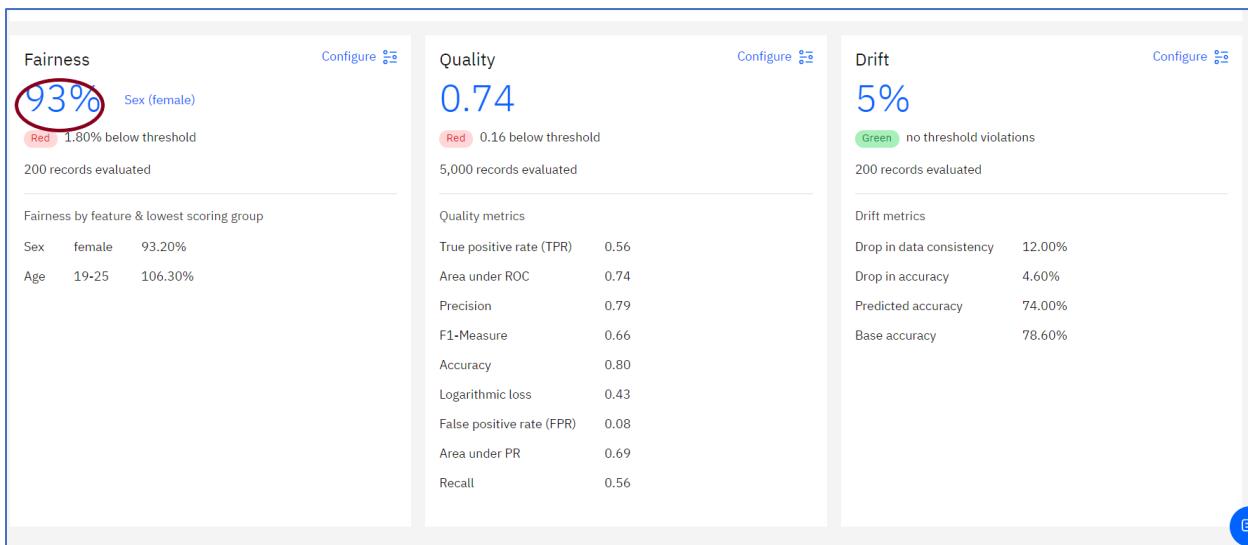
8. Please wait 3-5 minutes for the Evaluations page to display the metrics. Three tests have now been run. The quality test from before, and two new tests. One for fairness, and one for Drift. The fairness test has failed because 93% fairness is lower than our threshold of 95%. The Drift metric passed as the 5% drift is less than our drift threshold of 10%. Note, that lower drift is better, whereas lower fairness is obviously worse. Your evaluation results may be slightly different.



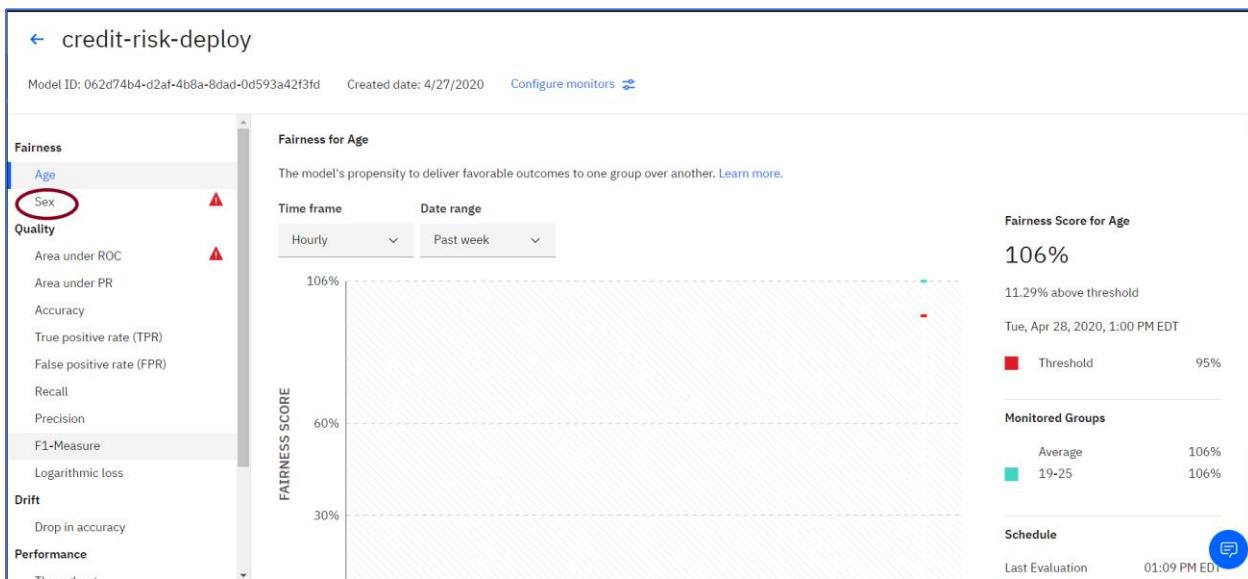
9. Scroll down to see the metric details.



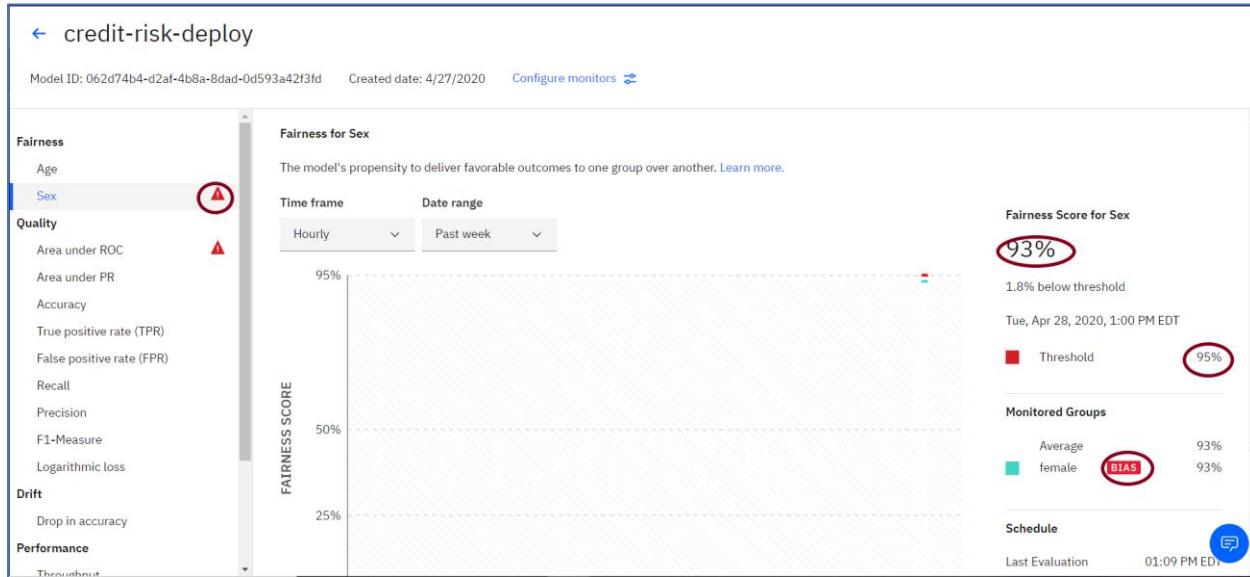
## 10. Click on the Fairness metric.



## 11. The time-series view is displayed. Click on Sex.



12. A time-series display for the Sex metric is shown. Currently, we have only run one evaluation. We can see the red triangle indicating an alert next to Sex. The Fairness score is listed at 93% which is less than the threshold of 95%.



13. Scroll down and click on **Check fairness now**.



14. Click on the timeline to view details.

[credit-risk-deploy](#)

Model ID: 062d74b4-d2af-4b8a-8dad-0d593a42f3fd | Created date: 4/27/2020 | Configure monitors

**Fairness**

- Age
- Sex**
- Quality
- Area under ROC
- Area under PR
- Accuracy
- True positive rate (TPR)
- False positive rate (FPR)
- Recall
- Precision
- F1-Measure
- Logarithmic loss
- Drift
- Drop in accuracy
- Performance
- Throughput

### Fairness for Sex

The model's propensity to deliver favorable outcomes to one group over another. [Learn more.](#)

Time frame: Hourly | Date range: Past week

FAIRNESS SCORE

95%  
50%  
25%

Click to view details

**Fairness Score for Sex**  
93%  
1.8% below threshold  
Tue, Apr 28, 2020, 2:00 PM EDT

**Threshold** 95%

**Monitored Groups**  
Average 93%  
female **BIAIS** 93%

**Schedule**  
Last Evaluation 02:09 PM EDT

Dashboard / credit-risk-deploy / Evaluations / Fairness

### Fairness

Monitored attribute: Sex | Data Set: Balanced | Date and Time: 5/16/2021, 6:00 PM | View payload transactions

No new data added since last fairness computation. Viewing results from May 16, 2021, 06:08 PM

Fairness score	Favorable outcomes	How the fairness score was determined (balanced data set)
94%	No Risk	The monitored group female received favorable outcomes 76.0% of the time. The perfect equality is 81.0%. The fairness score for Sex is 93.8% (76.0/81.0). <a href="#">View calculation.</a>

% FAVORABLE OUTCOMES

View percentage (radio button selected) | View count (radio button)

Sex

Female (Monitored): 76.0% (purple bar)  
Male (Reference): 81.0% (dark green bar)

15. The discrepancy between Male and Female results are sufficient to trigger a bias alert.

Click on **Debiased**. Note that it may take some time for the Debiased endpoint to become available.

The screenshot shows the 'Fairness' section of the dashboard. On the left, there's a 'Monitored attribute' dropdown set to 'Sex'. On the right, a 'Data Set' dropdown is open, showing 'Balanced' as the current selection. A tooltip indicates 'No new data added since last run'. Below the dropdowns, the 'Fairness score' is shown as 94%, labeled 'Favorable outcome' and 'No Risk'.

No new data added since last run

Fairness score  
94%  
Favorable outcome  
No Risk

Balanced

Debiased

16. The results show that applying a debiasing scheme will reduce the bias to close to zero.

The endpoint for invoking the debiasing algorithm can be obtained by clicking on View Debiased Endpoint. For now, click on **View Payload Transaction** to display a list of transactions.

The screenshot shows the 'Fairness' section with 'Monitored attribute' set to 'Sex' and 'Data Set' set to 'Debiased'. The 'Date and Time' is set to 5/16/2021 at 6:00 PM. A red circle highlights the 'View payload transactions' button.

View payload transactions

17. Click **Explain** next to a transaction to get an explanation of the factors that caused the deployed model to make the prediction it did for that transaction. I picked the first transaction that resulted in a Risk prediction. Note that it may take a few minutes for the transactions to show.

The screenshot shows a table of transactions. The first transaction, with ID 61a034fa736dabc57626b399cb139d38-1, has a 'Bias Detected' status of 'male' and an 'Outcome' of 'No Risk'. An 'Explain' link is highlighted with a red circle. To the right, there are sections for 'Payload Table' and 'Corrected Records', and two bar charts comparing 'Current Model' and 'De-biased Model' for 'No Risk : Favorable Outcome' and 'Risk : Unfavorable Outcome'.

Transaction ID	Sex	Bias Detected	Outcome	Action
61a034fa736dabc57626b399cb139d38-1	male	male	No Risk	Explain
61a034fa736dabc57626b399cb139d38-10	female	female	No Risk	Explain
61a034fa736dabc57626b399cb139d38-100	male	male	No Risk	Explain
61a034fa736dabc57626b399cb139d38-101	male	male	No Risk	Explain
61a034fa736dabc57626b399cb139d38-102	male	male	No Risk	Explain
61a034fa736dabc57626b399cb139d38-103	male	male	No Risk	Explain
61a034fa736dabc57626b399cb139d38-104	male	male	No Risk	Explain
61a034fa736dabc57626b399cb139d38-105	female	female	Risk	Explain

Explain

No Risk : Favorable Outcome

Risk : Unfavorable Outcome

18. The results show the confidence level for the result including an explanation of how the result was determined and which features were most important.

Explanations / e6c3f89cd6eeecb79faed2696ea53408e-129

Transaction details

Explain     Inspect

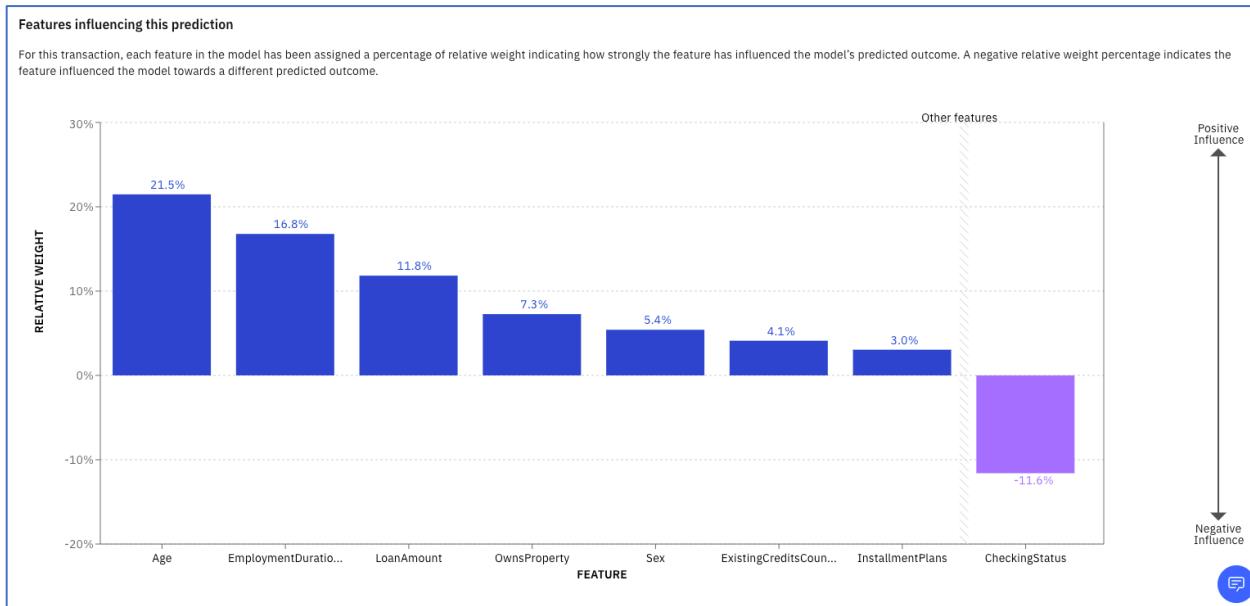
Transaction e6c3f89cd6eeecb79faed2696ea53408e-129	Received on Jul 21, 2020, 03:47:01 PM EDT	Deployed model credit-risk-deploy	Language Not applicable
--	--	--------------------------------------	----------------------------

**Predicted outcome** ⓘ Risk  
The credit-risk model has **50.74% confidence** that the outcome of this transaction would be **Risk**. The top three features influencing the model's predicted outcome are Age, EmploymentDuration, and LoanAmount. The feature CheckingStatus is influencing the model toward a predicted outcome of No Risk.

**Confidence level**  
50.74%

Search by Transaction ID

19. Scroll down to view how features influence the model's predicted outcome towards negative or positive.



20. Click on the Dashboard icon

IBM Watson OpenScale

Explanations / e6c3f89cd6eeecb79faed2696ea53408e-129

Transaction details

Explain     Inspect

Transaction e6c3f89cd6eeecb79faed2696ea53408e-129	Received on Jul 21, 2020, 03:47:01 PM EDT	Deployed model credit-risk-deploy	Language Not applicable
--	--	--------------------------------------	----------------------------

Need help? ⓘ

Search by Transaction ID

21. The Dashboard view provides a snapshot view of how the deployed models are performing. Since we are only monitoring one deployment, there is only 1 monitoring tile displayed. You can drill down from this view to get the details on each of the metrics for a given deployment, as we've shown in this lab.

The screenshot shows the 'Insights Dashboard' interface. At the top left, it says 'Model Monitors' with a value of '1'. On the right, there's a blue button labeled 'Add to dashboard +'. Below this, there are four categories: 'Deployments Monitored' (1), 'Quality Alerts' (1), 'Fairness Alerts' (1), and 'Drift Alerts' (0). A note below states: 'Quality and Fairness metrics update every hour. Drift metrics update every 3 hours.' The main content area is titled 'Watson Machine Learning credit-risk-deploy'. It shows 'Issues' (2) with a 'QUALITY' tab selected. Under 'Quality', the score is 0.74 (1 alert). Under 'Fairness', the score is 93% (1 alert). Under 'Drift', the score is 5%. A small note at the bottom says 'Evaluated 48 seconds ago'.

## Congratulations! You have completed the Lab!!!

- ✓ Imported a machine learning model
- ✓ Deployed the model
- ✓ Provisioned Watson OpenScale
- ✓ Configured the payload logging database and Machine Learning provider
- ✓ Scored Data
- ✓ Prepared the Deployed Model for Monitoring
- ✓ Configured Payload Logging
- ✓ Configured Quality Monitoring
- ✓ Configured Fairness Monitoring
- ✓ Configured Drift Monitoring
- ✓ Submitted Feedback and Viewed Quality Metrics
- ✓ Scored Data and Viewed Fairness and Drift Metrics
- ✓ Explained a Transaction.