

# 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.

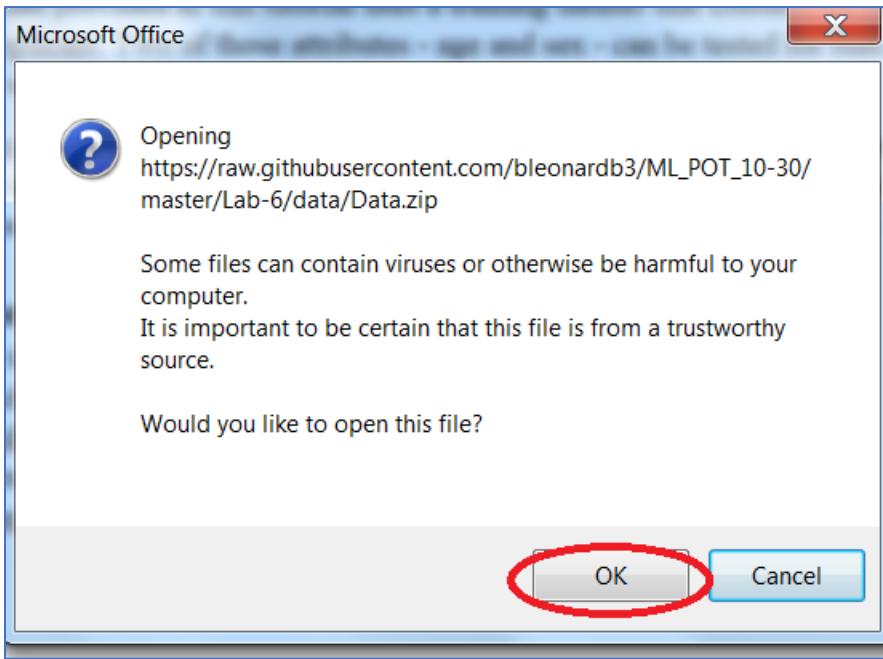
The credit risk model provided in this tutorial uses a training dataset that contains 20 attributes about each loan applicant. Two of those attributes - age and sex - can be tested for bias. For this tutorial, the focus will be on bias against sex and age.

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](#).

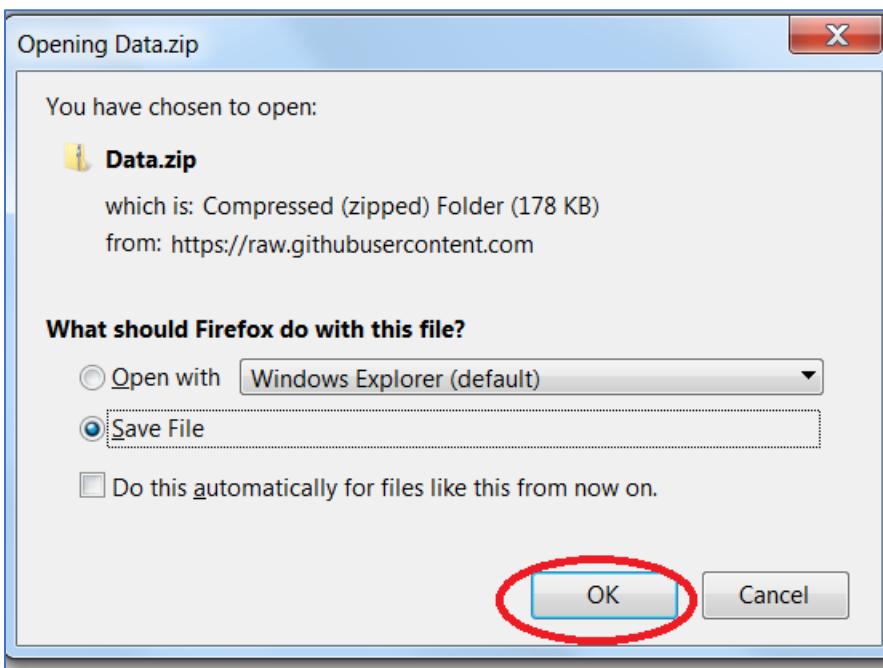
## Download the Lab Files

1. Click [here](#) to download the Data.zip file containing the following files.
  1. credit\_feedback\_data.csv
  2. credit\_payload\_data.json
  3. german\_credit\_data\_biased\_training.csv
  4. scoring.json

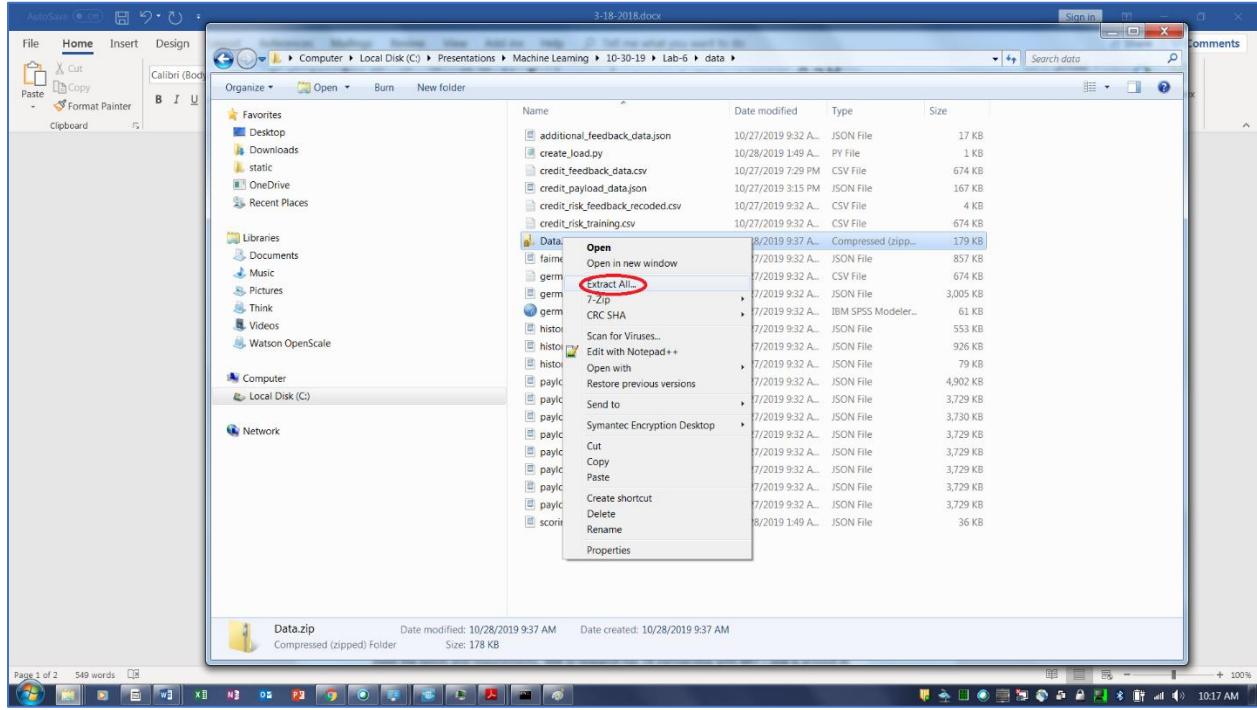
2. Click **OK**. Note your browser may not prompt this message.



3. Click **OK**.



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

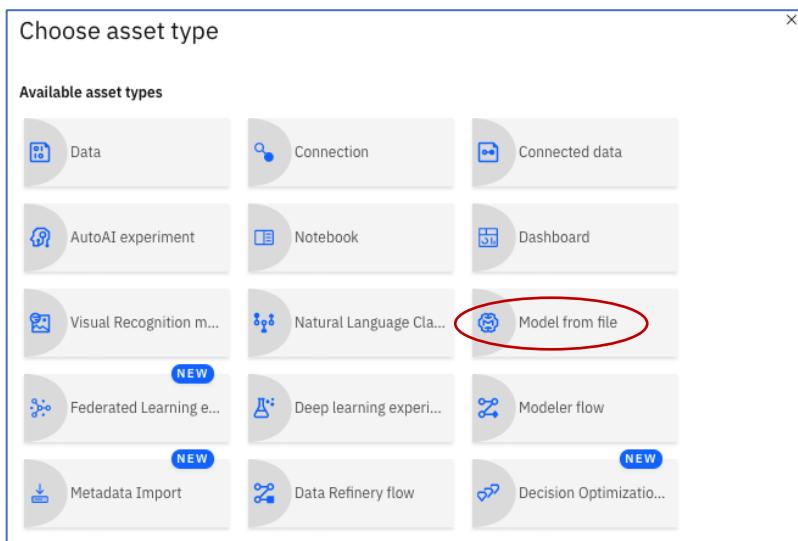


## Import the Credit Risk Model

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



2. Click on **Model from file**.



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

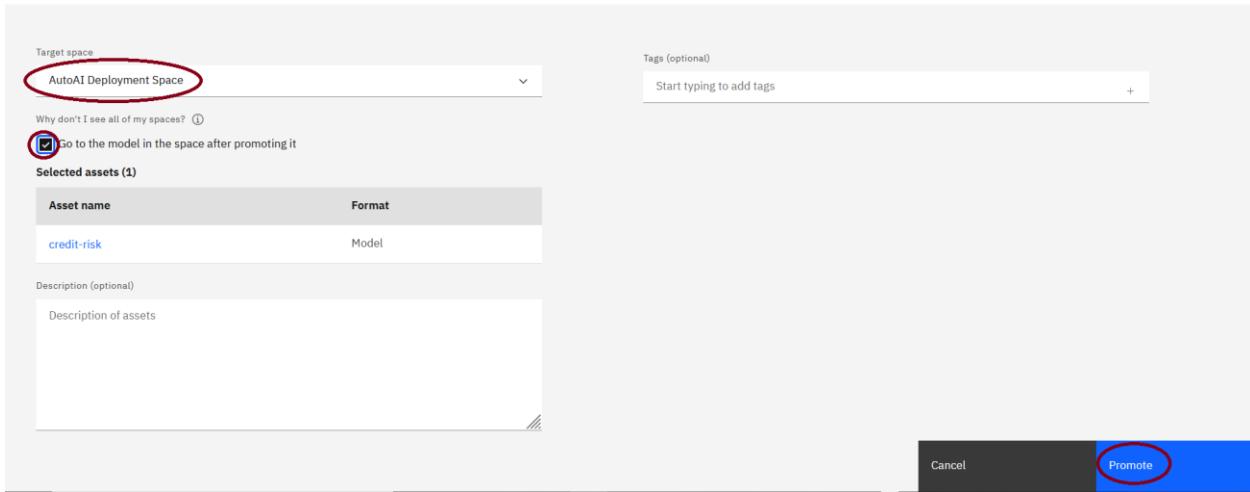
The screenshot shows the 'Import model' interface. On the left, there's a sidebar with a '+ New' button and a 'Gallery sample' button, which is circled in red. In the main area, there are two cards: 'Credit Risk' (circled in red) and 'Customer Satisfaction Prediction'. The 'Credit Risk' card has a detailed 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.' At the bottom right of the screen, the 'Import' button is also circled in red.

## Deploy the Credit Risk Model

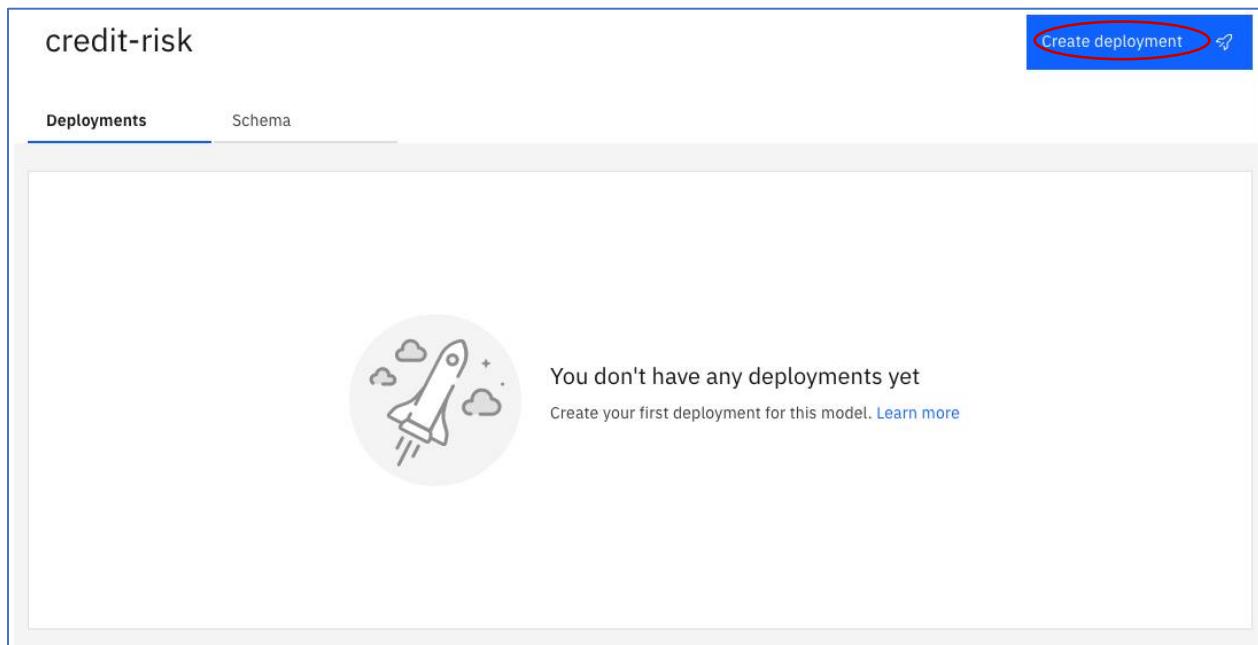
1. Click **Promote to deployment space**.

The screenshot shows the 'credit-risk' project overview. At the top right, there's a 'Promote to deployment space' button, which is circled in red. To the right of the project name, there's a summary box with the following details: '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 a 'Tags' section with a note to 'Add tags to make assets easier to find.'

2. Click on **AutoAI Deployment Space** 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 Watson Studio API reference page for a deployment named 'credit-risk-deploy'. The deployment status is 'Deployed' and 'Online'. The 'API reference' tab is selected. Below it, there's a 'Direct link' section with an endpoint URL: <https://us-south.ml.cloud.ibm.com/ml/v4/deployments/f38f4e1c-faf7-46c7-b4b1-2c3a6b502f39/predict>. A 'Bearer <token>' header is listed. The 'Code snippets' section includes tabs for 'cURL', 'Java', 'JavaScript', 'Python', and 'Scala'. The 'Test' tab is also visible at the top.

## 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 for the 'IBM Cloud Pak for Data' link. The menu items are: All, Open Link in New Tab (circled in red), 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). The 'Open Link in New Tab' option is highlighted with a red circle.

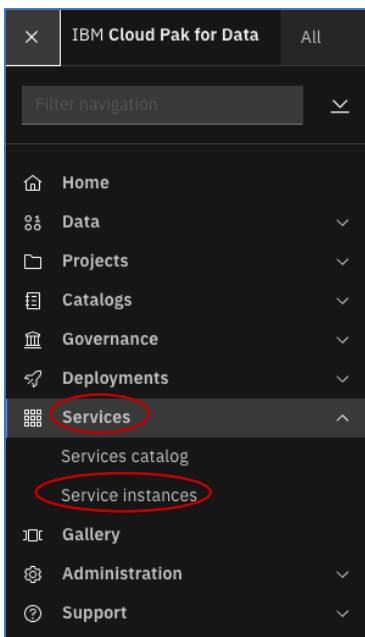
2. Click on the new **IBM Cloud Pak for Data** browser tab.

The screenshot shows a browser window with two tabs. Both tabs have the title 'IBM Cloud Pak for Data'. The second tab is active and has a red circle around its title bar.

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

The screenshot shows the IBM Cloud Pak for Data interface. In the top-left corner, there is a red circle around the hamburger menu icon (three horizontal lines). The main title 'IBM Cloud Pak for Data' is also visible.

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



5. Click the box to the right of your OpenScale instance. Login again if required.

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**.

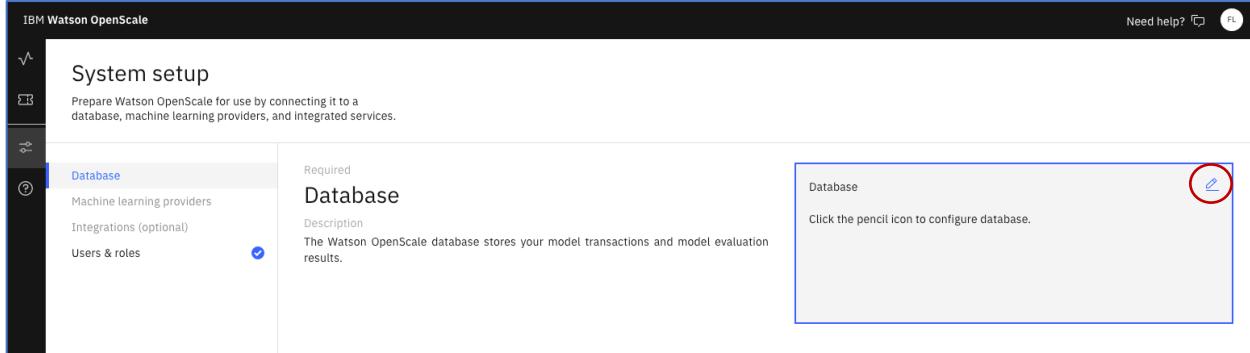
The screenshot shows the Watson OpenScale application interface. At the top, there is a navigation bar with 'Resource list / Watson OpenScale-qe' and status indicators like 'Active'. Below the navigation, there are three tabs: 'Getting started', 'Manage' (which is selected), and 'Plan'. In the center, there is a circular icon with a stylized purple and white design. Below the icon, the text 'Watson OpenScale' is displayed. A welcome message reads 'Welcome to Watson OpenScale, let's get started.' Below the message is a blue button labeled 'Launch Application', which is circled in red. At the bottom of the screen, there are links for 'Documentation' and 'Community', and a feedback icon.

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

The screenshot shows the Watson OpenScale manual setup interface. It features a large circular logo with a blue and white abstract design. Below the logo, the text 'Welcome to Watson OpenScale' is displayed. A detailed description follows: '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.' Below this text, a message says: '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?' At the bottom, there are two buttons: 'Manual setup' (circled in red) and 'Auto setup'.

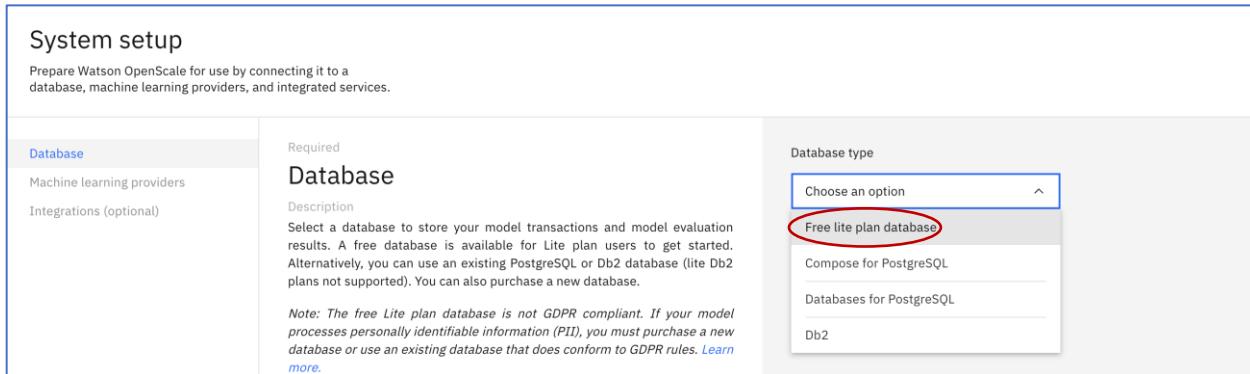
# Setup System

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.



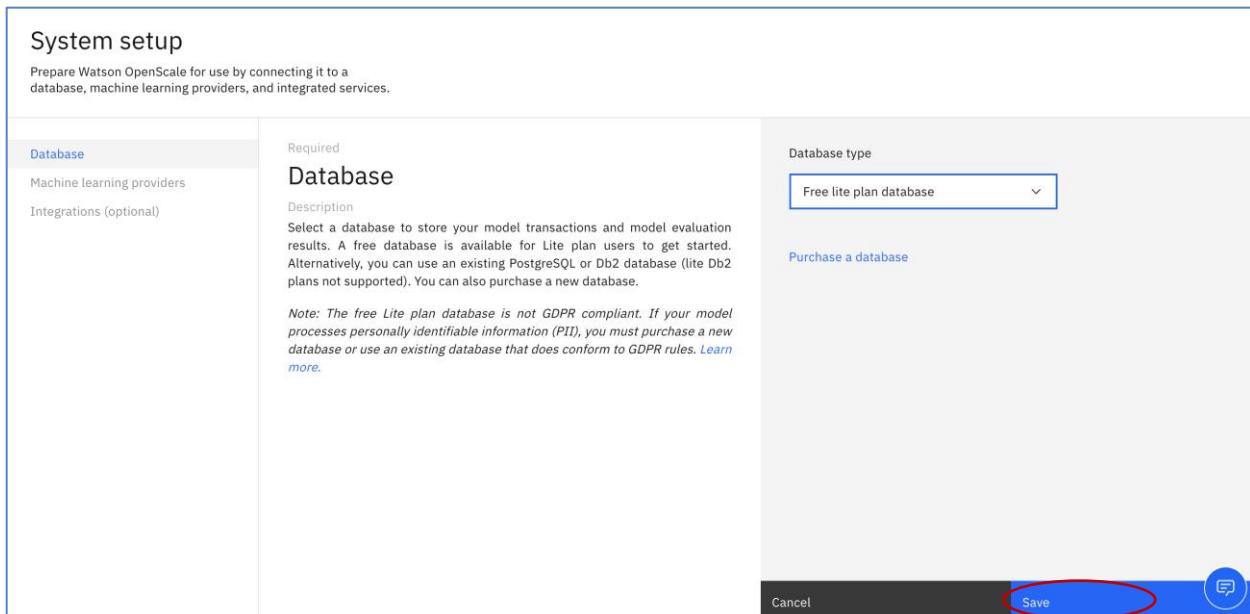
The screenshot shows the 'System setup' page of the IBM Watson OpenScale interface. On the left, there's a sidebar with icons for 'Database', 'Machine learning providers', 'Integrations (optional)', and 'Users & roles'. The 'Database' icon is highlighted with a blue border. The main area has a title '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 note: 'Click the pencil icon to configure database.' A red circle highlights the pencil icon in the top right corner of this section.

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



This screenshot shows the 'System setup' page with the 'Database' tile selected. In the 'Database type' dropdown menu, 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 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

New provider

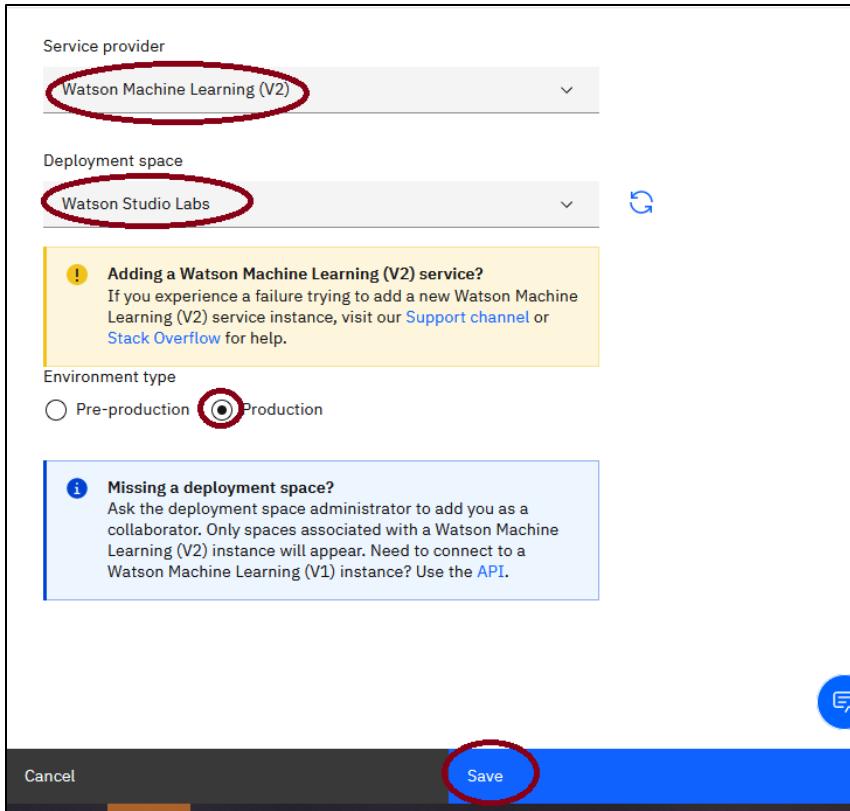
Description

Click edit to enter provider description.

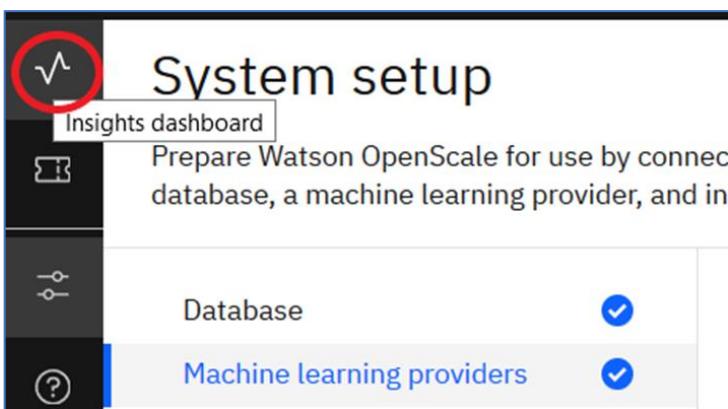
Connection

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 Insights Dashboard interface. At the top, there are four status indicators: Deployments Monitored (0), Quality Alerts (N/A), Fairness Alerts (N/A), and Drift Alerts (N/A). Below these are filter options for Tags, Alert type, and Machine learning provider, along with a search bar. In the top right corner, there is a 'Refresh' button and an 'Add to dashboard' button, which is highlighted with a red circle. The overall background is light gray with dark blue header elements.

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 instructs the user to choose a machine learning provider and provides a dropdown menu for 'Machine learning Provider' with 'New provider (Production)' selected. Below this is a table listing two deployments:

Deployment	Description	Created
FHT AutoAI Deployment	-	Tue, Aug 24, 2021, 11:52 AM EDT
credit-risk-deploy	-	Tue, Aug 24, 2021, 8:43 PM EDT

At the bottom of the dialog, there are 'Cancel' and 'Configure' buttons, with 'Configure' being circled in red.

12. Click on **Configure monitors**.

This screenshot shows a confirmation dialog box with the message 'Selections saved.' and a checkmark icon. Below the message, it says 'Done. Click **Configure monitors** to set up your monitors.' At the bottom of the dialog, there are 'Close' and 'Configure monitors' buttons, with 'Configure monitors' 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 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 interface for a deployed model named "credit-risk-deploy". The "API reference" tab is selected. Below it, the "Test" tab is visible. The "Direct link" section displays the endpoint URL: `https://us-south.ml.cloud.ibm.com/v4/deployments/f38f4e1c-faf7-46c7-b4b1-2c3a6b502f39/predict`. To the right of the URL are two buttons: "Bearer <token>" with a help icon and "IAM". A red circle highlights the "Test" tab. At the bottom, there are tabs for "cURL", "Java", "JavaScript", "Python", and "Scala".

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

This screenshot is identical to the previous one, showing the "Test" tab highlighted with a red circle. The rest of the interface, including the endpoint URL and code snippet tabs, remains the same.

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

credit-risk-deploy Deployed Online

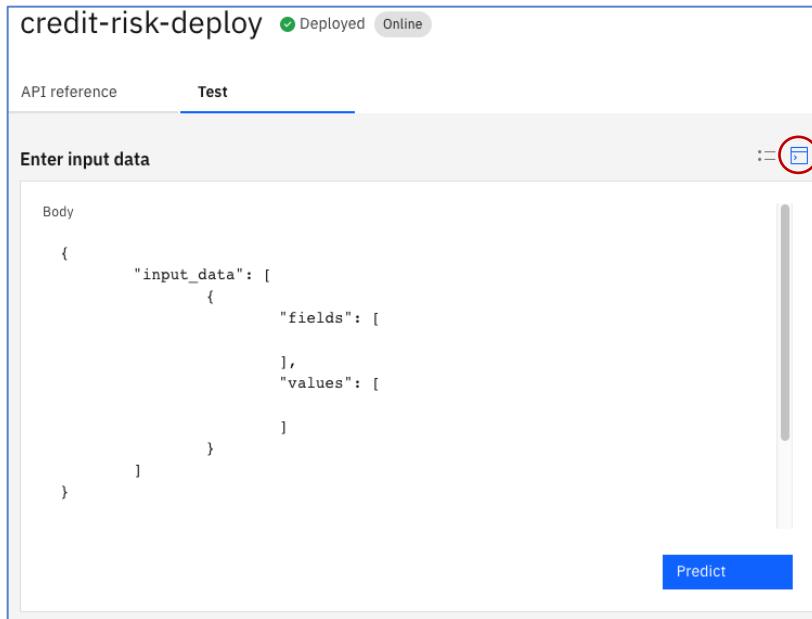
API reference Test

Enter input data

Body

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

Predict



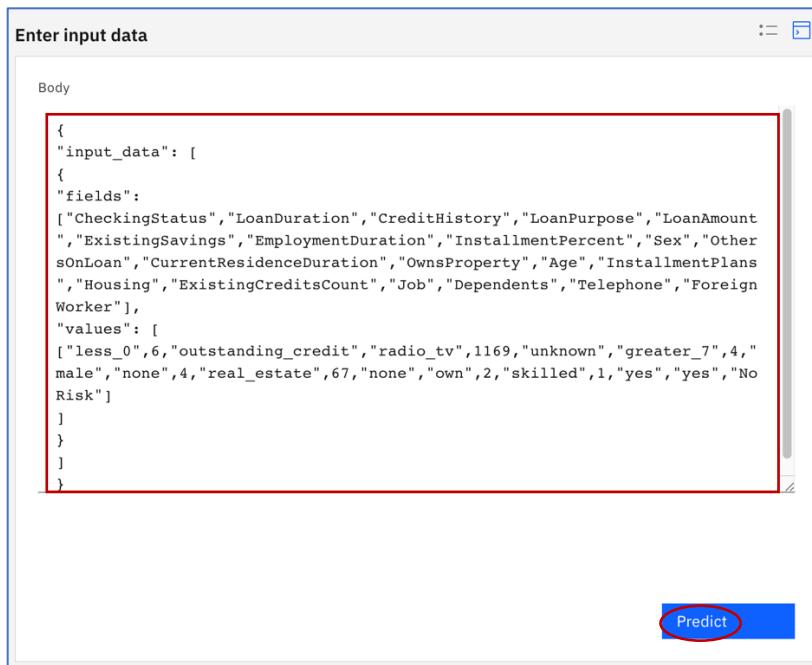
4. Copy the file contents of **credit\_payload\_data.json** into the **input data** area and click on **Predict**.

Enter input data

Body

```
{  
    "input_data": [  
        {  
            "fields": [  
                "CheckingStatus", "LoanDuration", "CreditHistory", "LoanPurpose", "LoanAmount",  
                "ExistingSavings", "EmploymentDuration", "InstallmentPercent", "Sex", "Other  
                sOnLoan", "CurrentResidenceDuration", "OwnsProperty", "Age", "InstallmentPlans",  
                "Housing", "ExistingCreditsCount", "Job", "Dependents", "Telephone", "Foreign  
                Worker"],  
            "values": [  
                ["less_0", 6, "outstanding_credit", "radio_tv", 1169, "unknown", "greater_7", 4, "male",  
                "none", 4, "real_estate", 67, "none", "own", 2, "skilled", 1, "yes", "yes", "No  
                Risk"]  
            ]  
        }  
    ]  
}
```

Predict



5. The results should appear as below.

The screenshot shows the Watson OpenScale API Test interface. The 'Test' tab is selected. In the 'Enter input data' section, there is a 'Body' field containing a JSON object representing input data for a credit risk prediction. The 'Result' section displays the JSON response from the API, which includes a 'predictions' array with a single object containing various fields like 'CheckingStatus', 'LoanDuration', etc.

```

{
  "input_data": [
    {
      "fields": [
        "CheckingStatus", "LoanDuration", "CreditHistory", "LoanPurpose", "LoanAmount",
        "ExistingSavings", "EmploymentDuration", "InstallmentPercent", "Sex", "OthersOnLoan",
        "CurrentResidenceDuration", "OwnsProperty", "Age", "InstallmentPlans",
        "Housing", "ExistingCreditsCount", "Job", "Dependents", "Telephone", "ForeignWorker"
      ],
      "values": [
        ["less_0", 6, "outstanding_credit", "radio_tv", 1169, "unknown", "greater_7", 4,
         "male", "none", 4, "real_estate", 67, "none", "own", 2, "skilled", 1, "yes", "yes", "NoRisk"]
      ]
    }
  ]
}

```

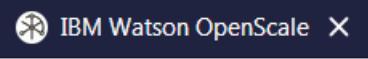
```

0 {
  "predictions": [
    {
      "fields": [
        "CheckingStatus",
        "LoanDuration",
        "CreditHistory",
        "LoanPurpose",
        "LoanAmount",
        "ExistingSavings",
        "EmploymentDuration",
        "InstallmentPercent",
        "Sex",
        "OthersOnLoan",
        "CurrentResidenceDuration",
        "OwnsProperty",
        "Age",
        "InstallmentPlans",
        "Housing",
        "ExistingCreditsCount",
        "Job",
        "Dependents",
        "Telephone",
        "ForeignWorker",
        "CheckingStatus_IX",
        "CreditHistory_IX"
      ]
    }
  ]
}

```

## 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**.

The screenshot shows the 'Model details' configuration page. It includes sections for 'Model details', 'Model input', and 'Model transaction'. The 'Model input' section is highlighted with a red circle around its edit icon. The 'Model transaction' section shows a successful transaction record.

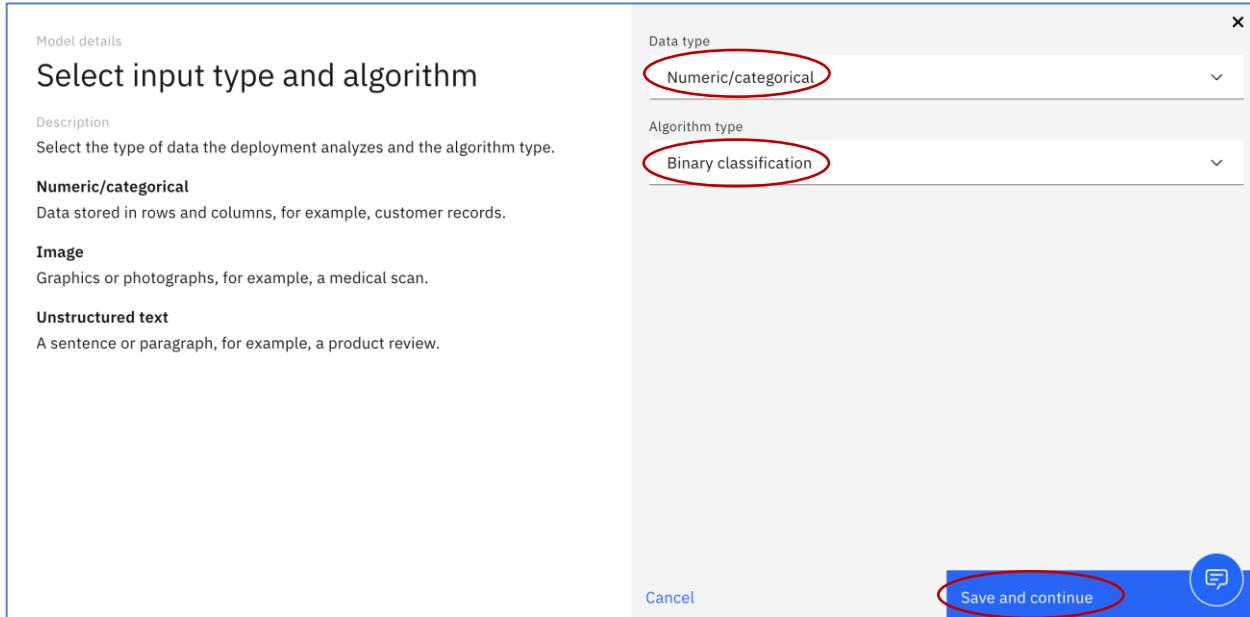
**Model details**

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

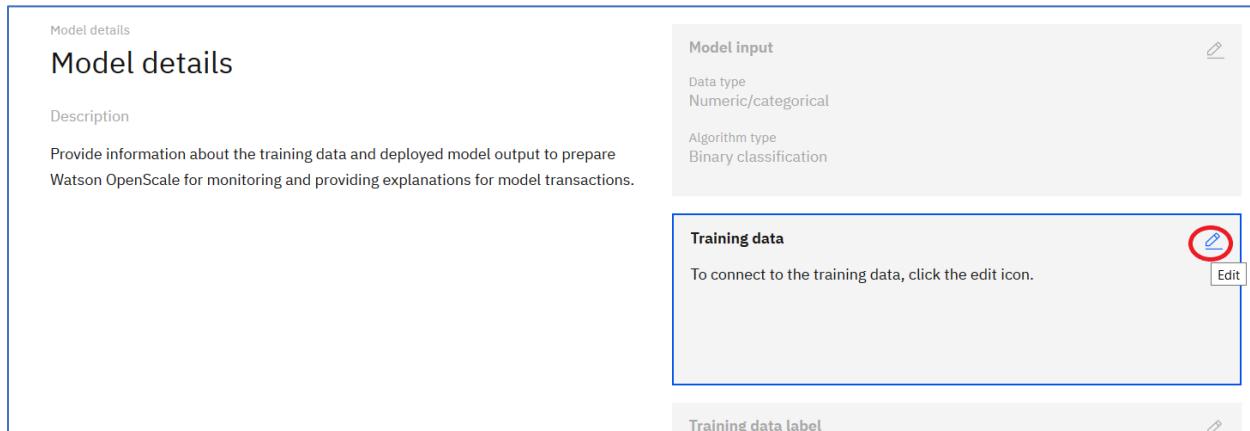
**Model input**  
To select the data type and algorithm type, click the edit icon.

**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**:

43bf01a8-03e1-4a1e-8c1a-1cc77e31657f

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

U79XrAvJ6HUBBfmmFo\_0\_CjdhfRW3Qr5D9-ygfZ-kq5n

## 8. Click Connect

The screenshot shows a configuration dialog for connecting to cloud storage. It includes fields for 'Storage type' (set to 'Database or cloud storage'), 'Location' (set to 'Cloud Object Storage'), 'Resource instance ID' (containing the value '1613a7c4-d7b1-4872-a1c2-41db3b7f3d91'), 'API key' (redacted), and a 'Connect' button at the bottom.

## 9. Select **a3-trainingdata** as the bucket, select **german\_credit\_data\_biased\_training.csv** as the Data set, and click **Next**.

The screenshot shows a configuration dialog for selecting a bucket. It includes fields for 'Storage type' (set to 'Database or cloud storage'), 'Location' (set to 'Cloud Object Storage'), 'Resource instance ID' (containing the value '43bf01a8-03e1-4a1e-8c1a-1cc77e31657f'), 'API key' (redacted), 'Bucket' (set to 'a3-trainingdata'), and 'Data set' (set to 'german\_credit\_data\_biased\_training.csv'). At the bottom, there are 'Cancel' and 'Next' buttons, with 'Next' being redacted.

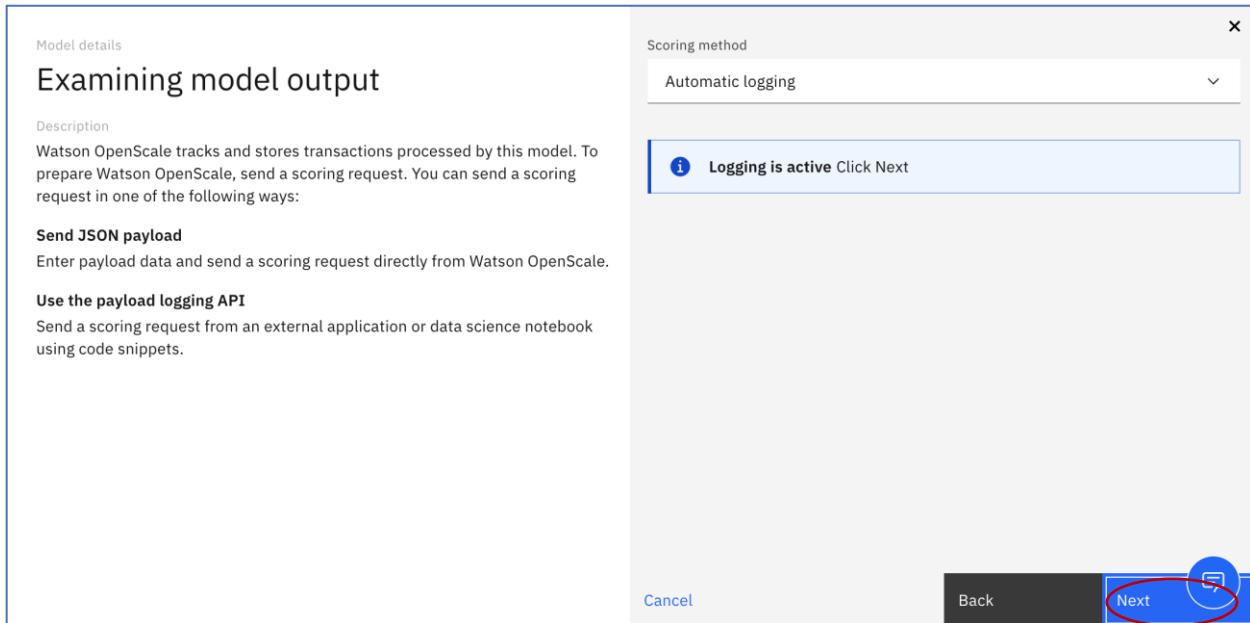
10. Watson OpenScale has determined that Risk is the label (target) column. Click **Next**.

The screenshot shows the 'credit-risk-deploy' interface. On the left, there's a sidebar with 'Model info' sections like 'Model details' (selected), '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.

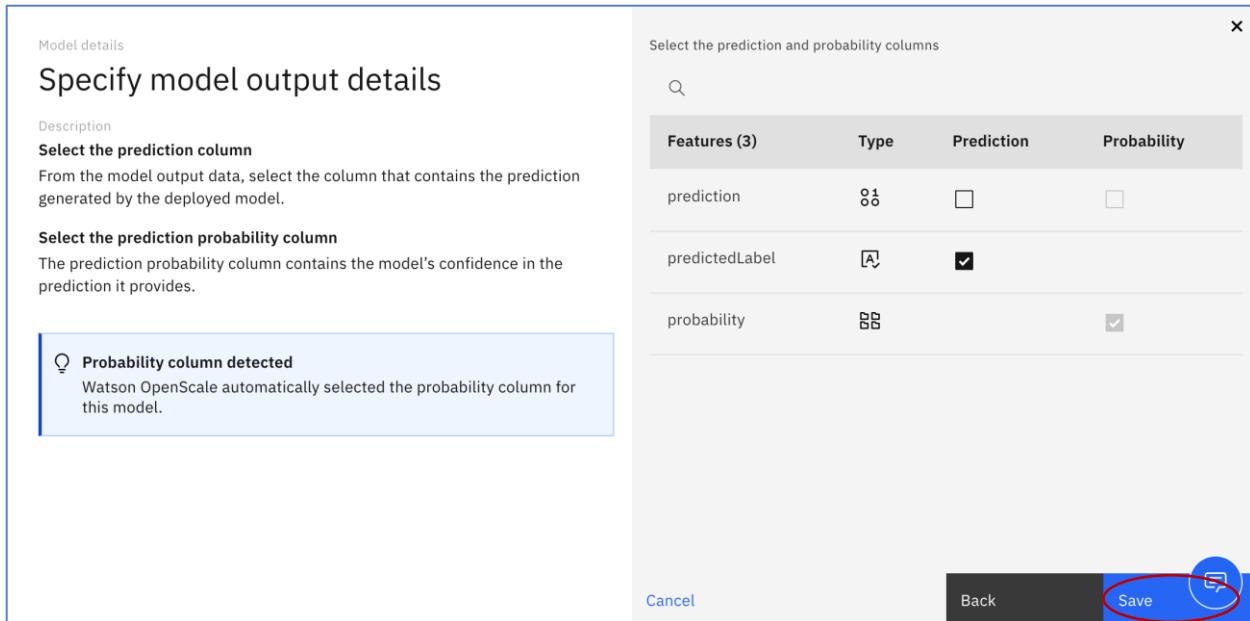
11. All of the features are used to train the model. Watson OpenScale has determined the feature types. Click **Next**.

The screenshot shows the 'credit-risk-deploy' interface. The main area is titled 'Select the training features' with a 'Description' section explaining how to select features for training. A modal window titled 'Select the training features' lists 'Features (20)': 'Age', 'CheckingStatus', 'CreditHistory', 'CurrentResidenceDuration', 'Dependents', and 'EmploymentDuration'. For 'CheckingStatus', 'CreditHistory', and 'EmploymentDuration', the 'Categorical' checkbox is checked. 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" (Fairness, Quality, Drift, Explainability), and "Import settings". 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 panels: "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".

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 panels: "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 at the bottom right.

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  Next

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  Back  Next

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	Monitored
Sex	81	<input checked="" type="checkbox"/>
OthersOnLoan	80	<input type="checkbox"/>
CurrentResidenceDuration	80	<input type="checkbox"/>
OwnsProperty	80	<input type="checkbox"/>
Age	80	<input checked="" type="checkbox"/>
InstallmentPlans	80	<input type="checkbox"/>
Housing	80	<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 "Recommended groups" callout. 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 the modal is a field "Set fairness alert threshold [Sex]" containing the value "95", which is circled in red. At the bottom right of the modal are buttons for "Cancel", "Back", and "Save", 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" and "Drift" are listed under "Evaluations", with "Drift" circled in red. The "Fairness" section contains a "Description" of the Fairness monitor. To the right, there are three panels: "Favorable outcomes" (listing "Favorable outcomes" and "No Risk"), "Unfavorable outcomes" (listing "Risk"), and "Sample size" (listing "Minimum sample size" and "200").

2. 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>	
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 href="#">Go to model summary</a>	

**Drift**

Description  
The drift test measures two types of changes.

**Drop in accuracy**  
*structured binary and multi-class classification models only*  
Watson OpenScale estimates the drop in accuracy of the model at runtime. The model accuracy could drop if there is an increase in transactions similar to those which the model was unable to evaluate correctly in the training data.

**Drop in data consistency**  
Watson OpenScale estimates the drop in consistency of the data at runtime as compared to the characteristics of the data at training time.

A drop in model accuracy and data consistency may lead to a negative impact on the business outcomes associated with the model.

**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.

3. 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.

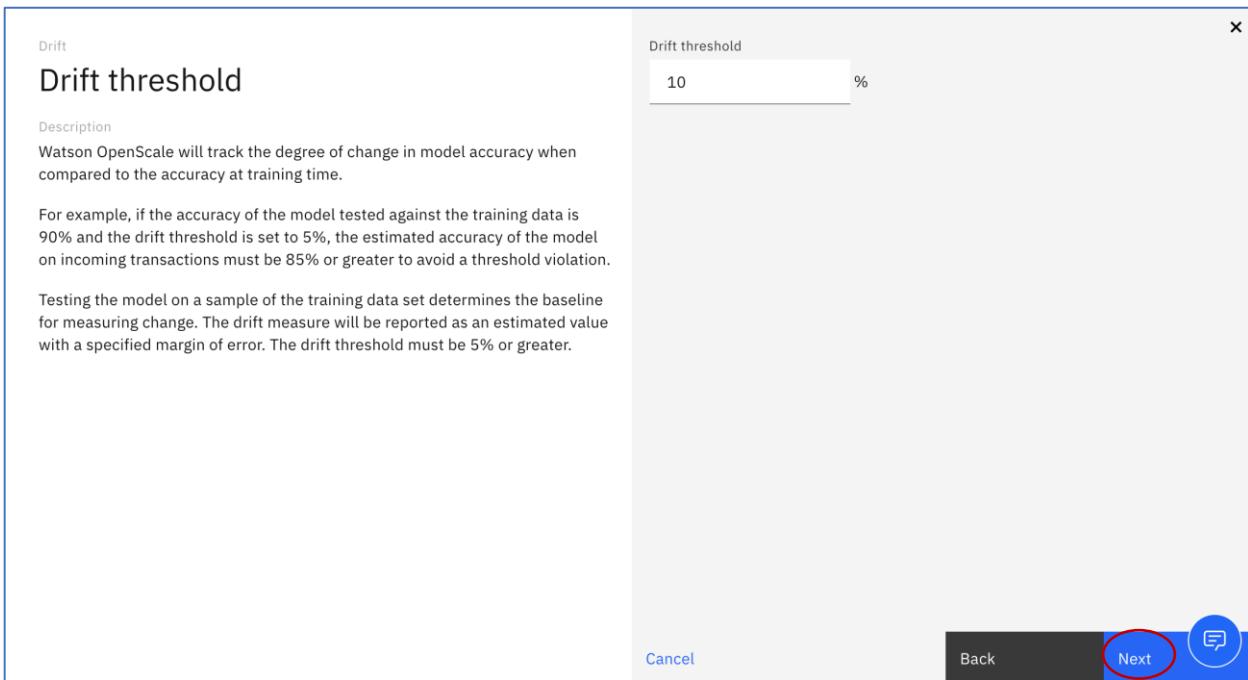
Training option

Train in Watson OpenScale

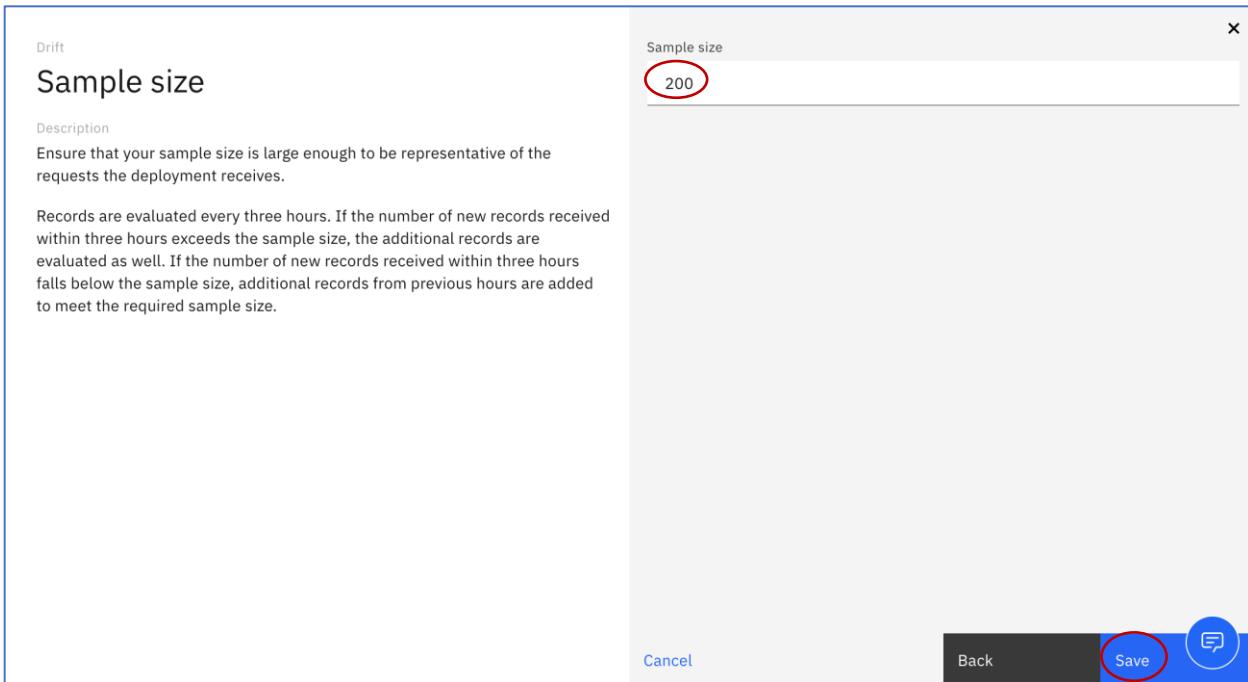
Train in a data science notebook

Cancel **Next**

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



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

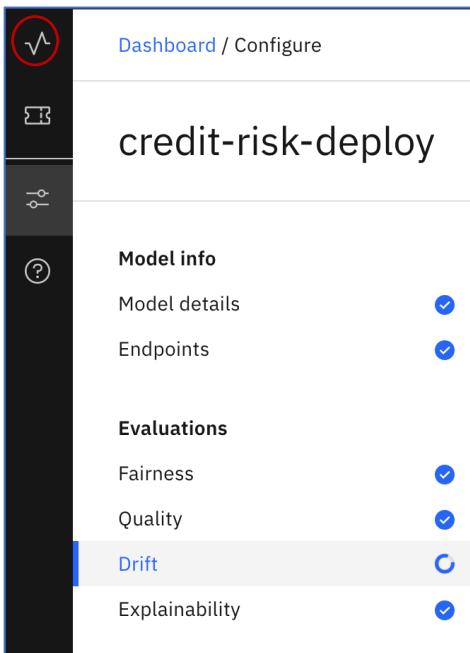


6. This completes the Drift configuration. A drift model will be created. You can move onto the next section while the drift model is being created.

## Submit Feedback and View Quality Metrics

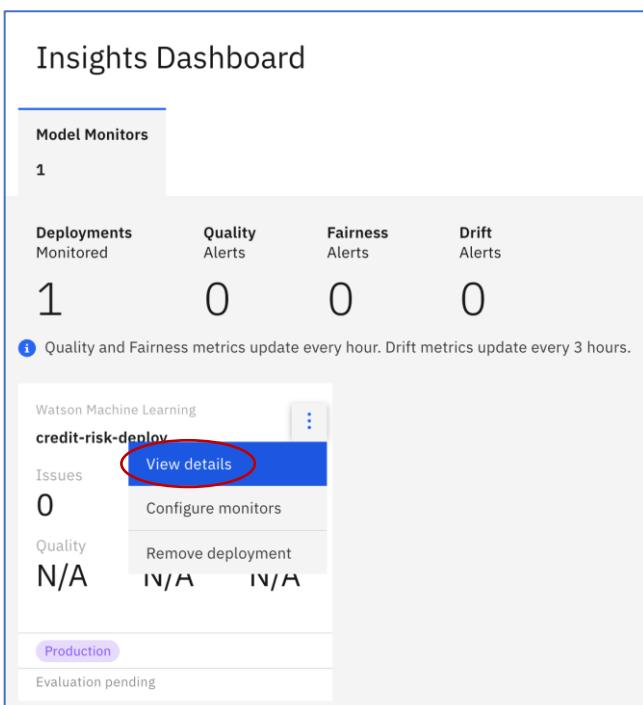
In order to measure quality, scored transactions including human labeled feedback must be provided.

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



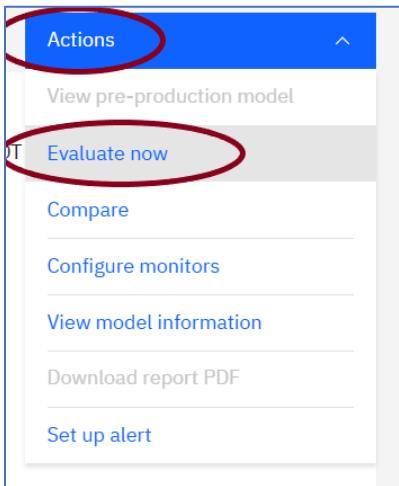
The screenshot shows the 'Model info' section of the Insights Dashboard. On the left is a sidebar with icons for Dashboard, Model info, Evaluations, Drift, and Explainability. The main area displays 'credit-risk-deploy' under 'Model info'. It includes sections for 'Model details' (checked), 'Endpoints'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

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

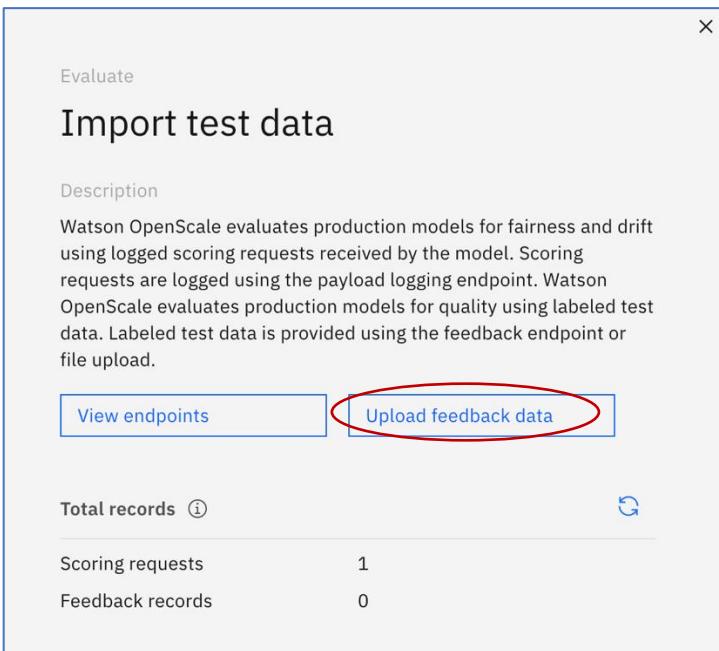


The screenshot shows the 'Insights Dashboard' for 'credit-risk-deploy'. It features a summary of monitored deployments (1), alerts (0 for Quality, 0 for Fairness, 0 for Drift), and a note that Quality and Fairness metrics update every hour while Drift updates every 3 hours. Below this is a detailed view of the deployment, showing 0 issues, 0 Quality alerts (N/A), 0 Fairness alerts (N/A), and 0 Drift alerts (N/A). 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'.

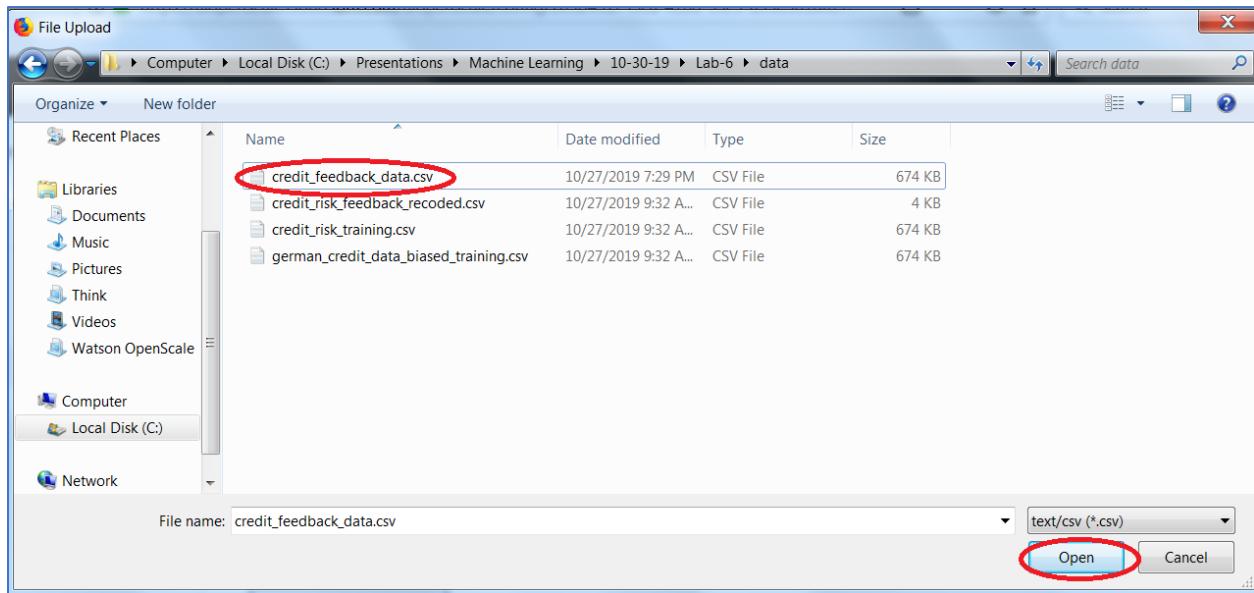
3. It will take a minute or so to bring up the **Evaluations** page. Click on **Actions** in the top right corner of your dashboard and then click **Evaluate now**.



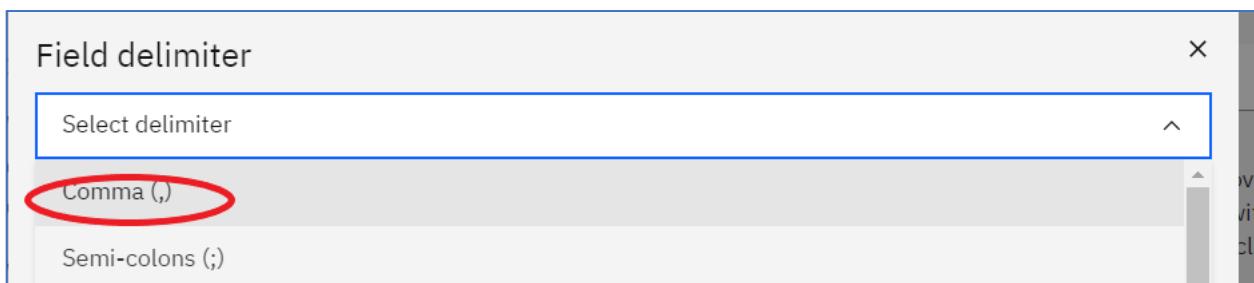
4. Watson OpenScale evaluates models for quality using labeled test data. Labeled test data is provided using the feedback endpoint or by file upload. We will use the file upload options. Click on **Upload feedback data**.



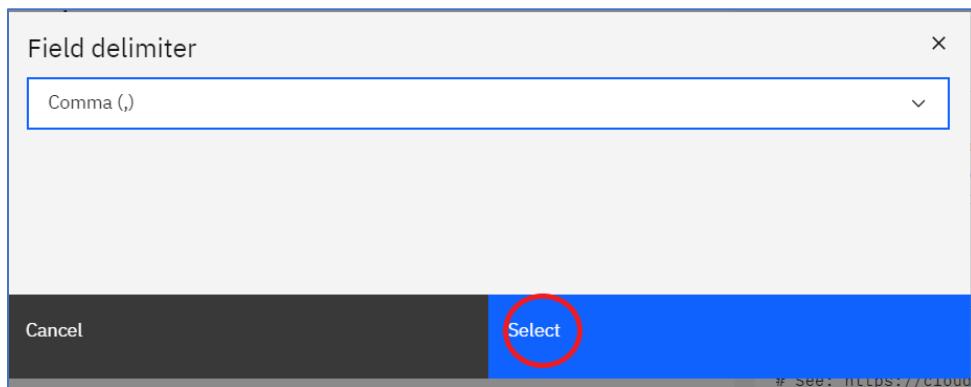
5. Navigate to the feedback file **credit\_feedback\_data.csv**. Click on **Open**.



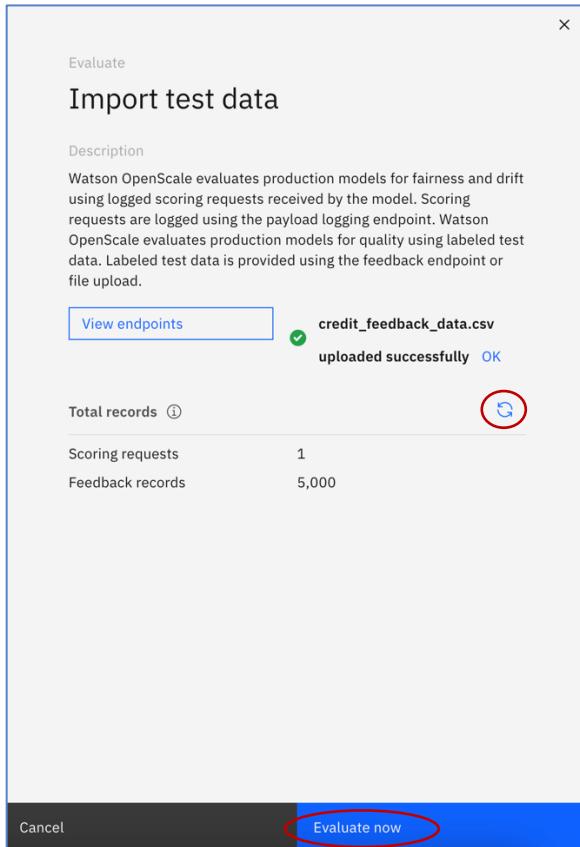
6. Select **Comma** as the **Delimiter**.



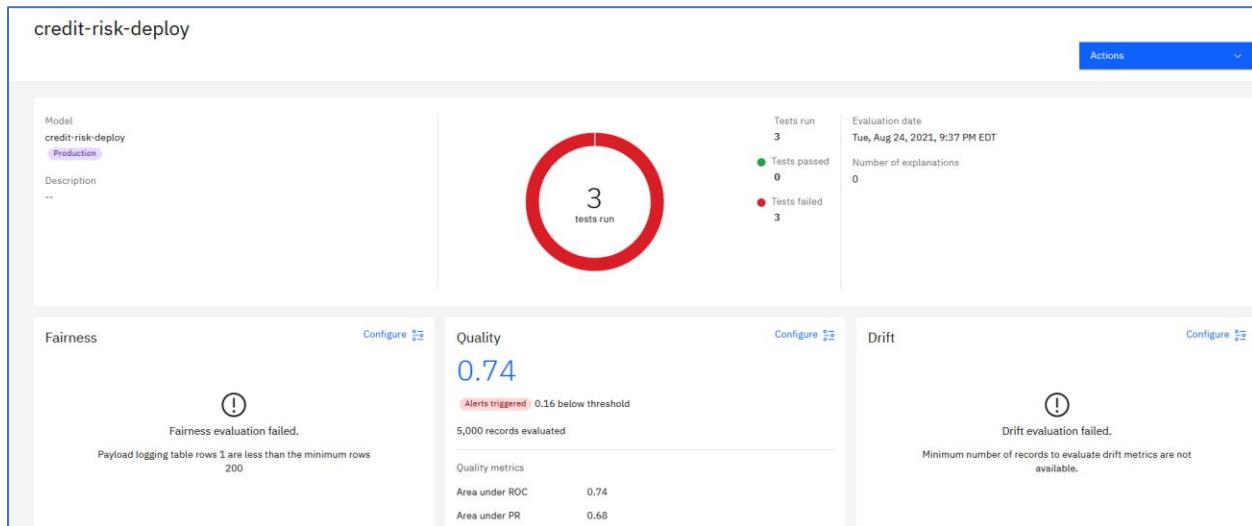
7. Click **Select**.



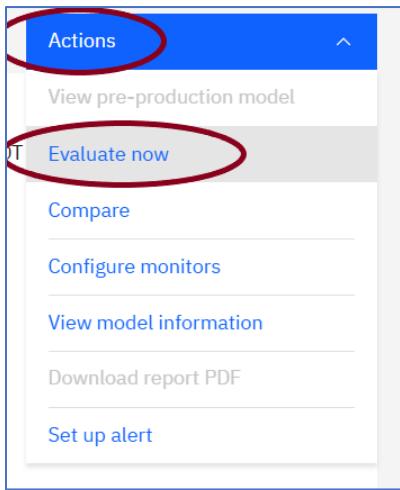
8. After the file is successfully uploaded, click on the refresh button to see the Feedback records increment, then click on the **Evaluate now**.



9. Please wait about 3-5 minutes until the evaluation is done. The quality results are displayed. The quality test failed because the accuracy is below the quality threshold. The Fairness and drift are not evaluated yet because we only had 1 entry in the credit\_payload\_data.json file.



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

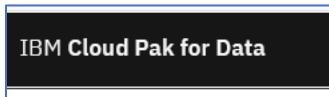


11. Leave Watson OpenScale at the Evaluation page while, you score transactions below.

## 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.

The screenshot shows the 'credit-risk-deploy' page in Watson Studio. The top navigation bar indicates the model is 'Deployed' and 'Online'. The main area has two tabs: 'API reference' and 'Test', with 'Test' being the active tab. On the left, under 'Enter input data', there is a 'Body' section containing a JSON object:

```
Body
{
  "input_data": [
    {
      "fields": [
        "CheckingStatus", "LoanDuration", "CreditHistory", "LoanPurpose", "LoanAmount",
        "ExistingSavings", "EmploymentDuration", "InstallmentPercent", "Sex", "Other
        sOnLoan", "CurrentResidenceDuration", "OwnsProperty", "Age", "InstallmentPlans",
        "Housing", "ExistingCreditsCount", "Job", "Dependents", "Telephone", "Foreign
        Worker"
      ],
      "values": [
        ["less_0", 6, "outstanding_credit", "radio_tv", 1169, "unknown", "greater_7", 4,
        "male", "none", 4, "real_estate", 67, "none", "own", 2, "skilled", 1, "yes", "yes", "No
        Risk"]
      ]
    }
  ]
}
```

Below this is a 'Predict' button. On the right, under 'Result', the response is shown as a JSON object:

```
0 {
  "predictions": [
    {
      "fields": [
        "CheckingStatus",
        "LoanDuration",
        "CreditHistory",
        "LoanPurpose",
        "LoanAmount",
        "ExistingSavings",
        "EmploymentDuration",
        "InstallmentPercent",
        "Sex",
        "OthersOnLoan",
        "CurrentResidenceDuration",
        "OwnsProperty",
        "Age",
        "InstallmentPlans",
        "Housing",
        "ExistingCreditsCount",
        "Job",
        "Dependents",
        "Telephone",
        "ForeignWorker",
        "CheckingStatus_IX",
        "CreditHistory_IX"
      ]
    }
  ]
}
```

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

credit-risk-deploy

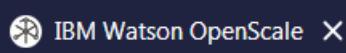
Overview Implementation Test

Enter input data

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

Predict

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



## 7. Click Evaluate now.

Evaluate

Import test data

Description

Watson OpenScale evaluates production models for fairness and drift using logged scoring requests received by the model. Scoring requests are logged using the payload logging endpoint. Watson OpenScale evaluates production models for quality using labeled test data. Labeled test data is provided using the feedback endpoint or file upload.

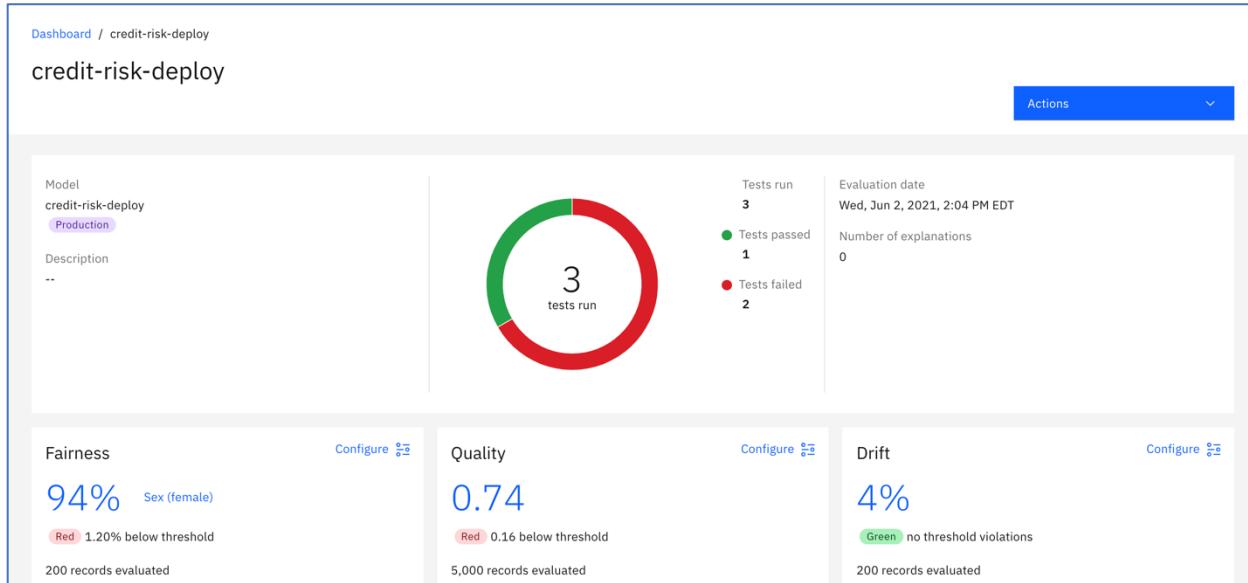
[View endpoints](#) [Upload feedback data](#)

Total records ⓘ

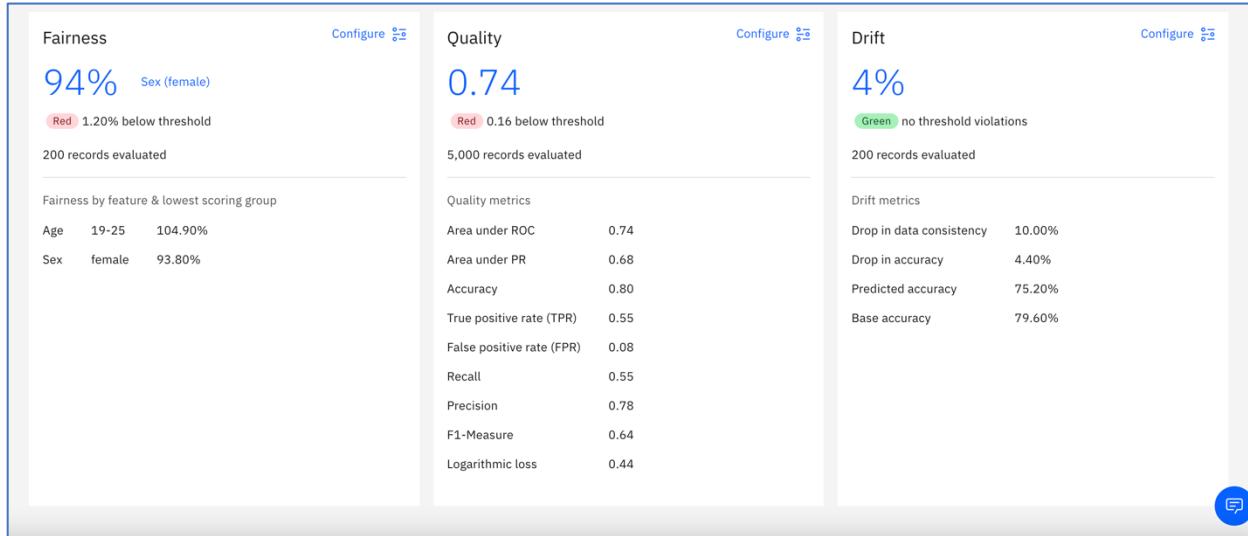
Scoring requests	201
Feedback records	5,000

Cancel **Evaluate now**

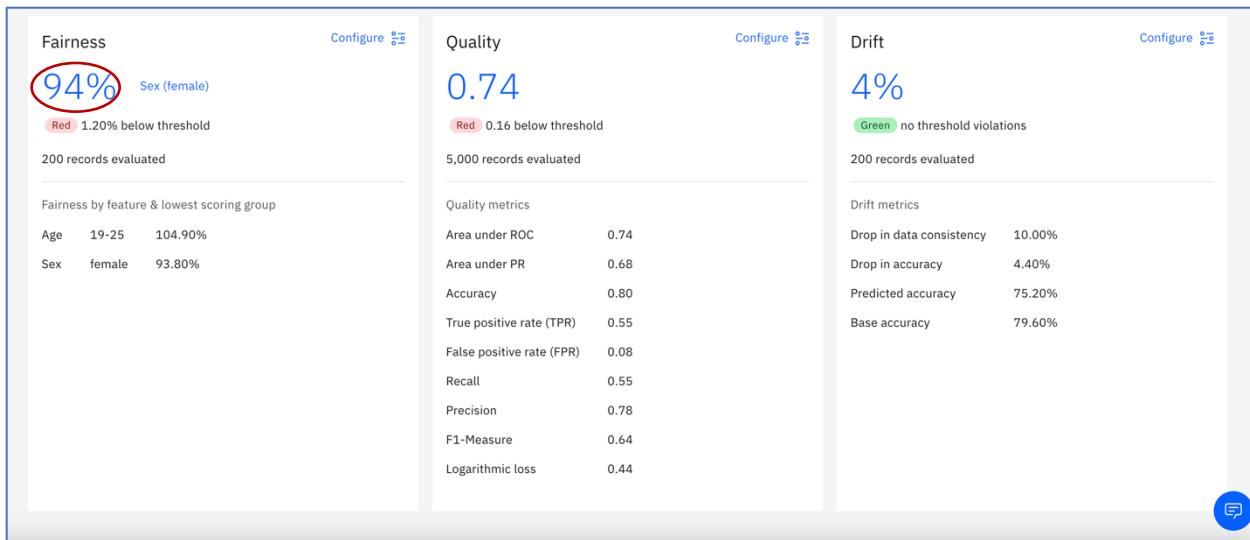
8. Please wait 3-5 minutes for the Evaluations page to display the metrics. Three tests have now been run. The quality test from before, and two new tests. One for fairness, and one for Drift. The fairness test has failed because 93% fairness is lower than our threshold of 95%. The Drift metric passed 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.



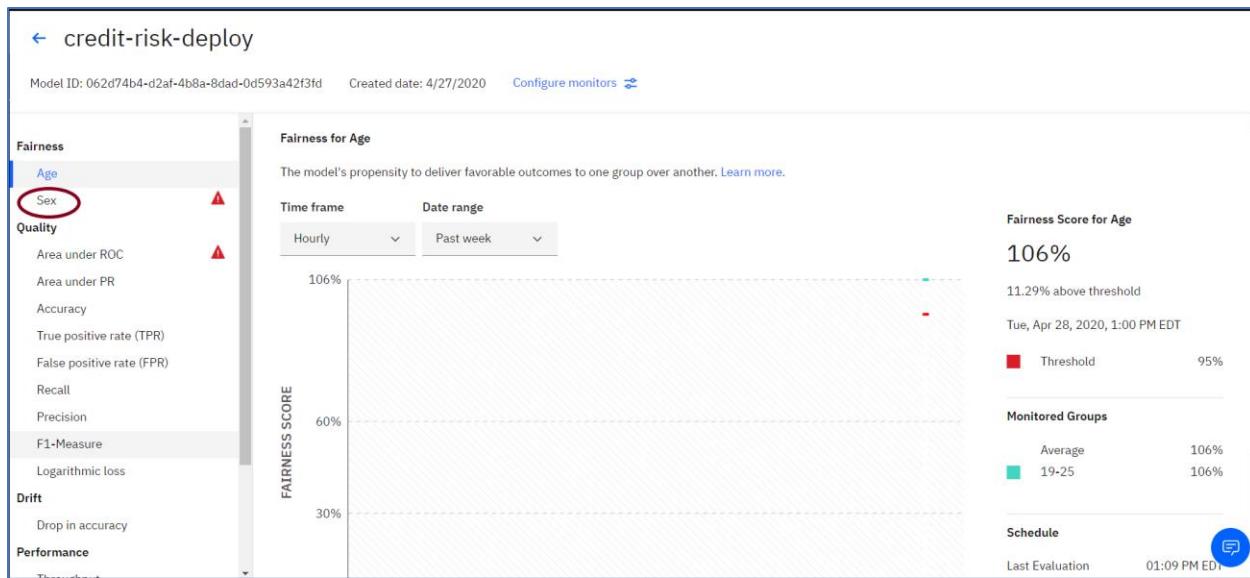
9. Scroll down to see the metric details.



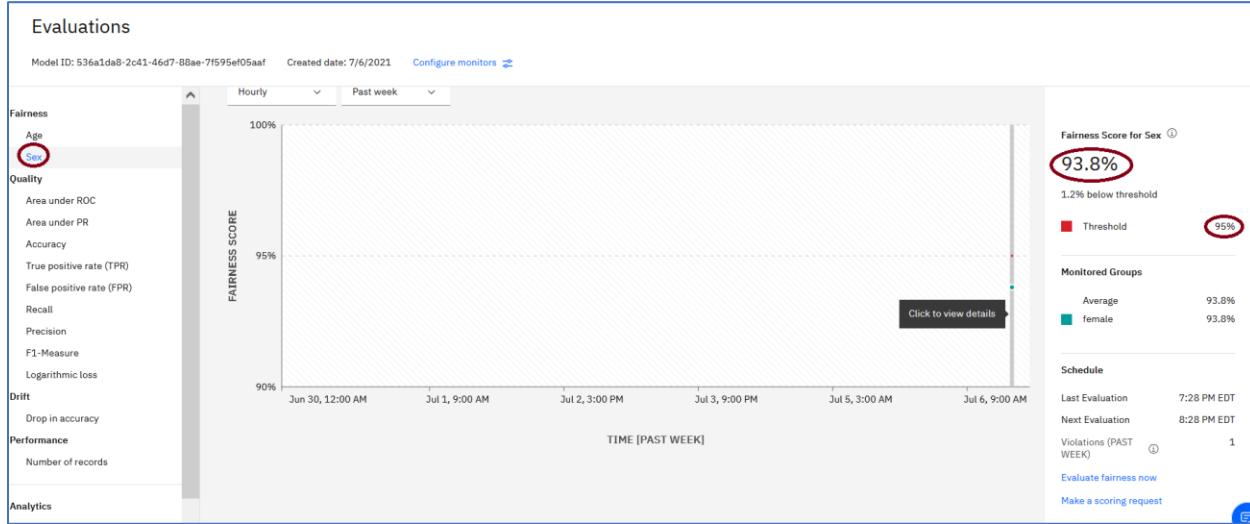
## 10. Click on the Fairness metric.



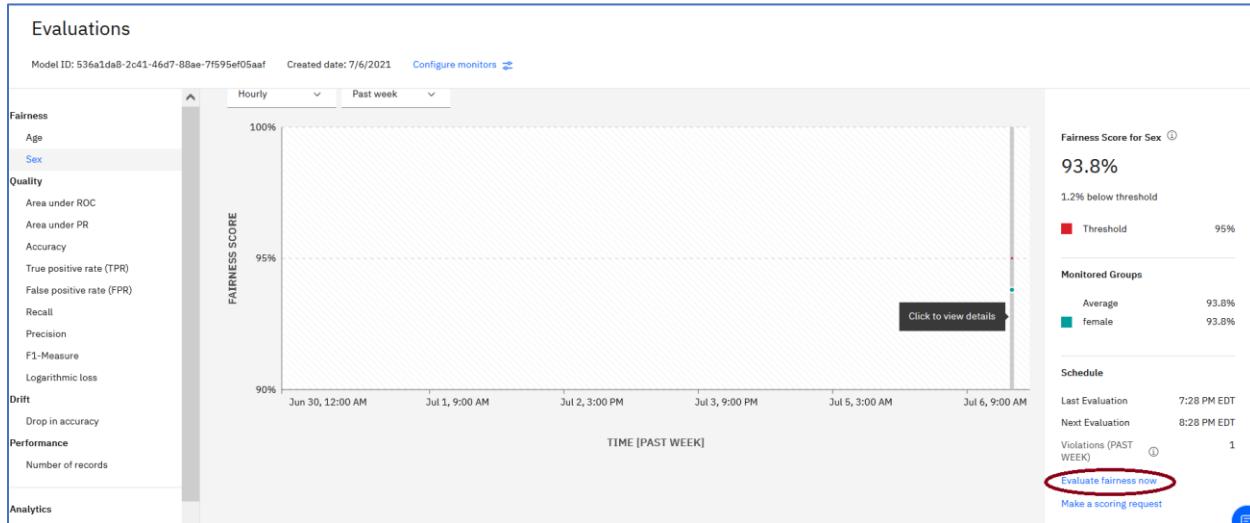
## 11. The time-series view is displayed. Click on Sex.



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. Move the cursor so that it is positioned on the one point. The Fairness score is listed at 93% which is less than the threshold of 95%.



13. Scroll down and click on Evaluate 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

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

Time frame: Hourly    Date range: Past week

FAIRNESS SCORE

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 ■ BIAS 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

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>

View percentage     View count

% FAVORABLE OUTCOMES

female Monitored    male Reference

▲ View payload transactions

This screen allows you to view the fairness score broken up by % of favorable outcomes for female (monitored group) vs male (reference group).

15. Click on **View Payload Transaction** to display a list of transactions.

Dashboard / credit-risk-deploy / Evaluations / Fairness

### Fairness

Monitored attribute: Sex    Data Set: Debaised    Date and Time: 5/16/2021 6:00 PM

View payload transactions

16. 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.

View  
All transactions Biased transactions

Review the transactions that were received for the selected time period. Click an Explain link to determine the features that contributed to each outcome.

Transaction ID	Sex	Bias Detected	Outcome	Action
61a034fa736dabc57626b399cb139d38-1	male		No Risk	<a href="#">Explain</a>
61a034fa736dabc57626b399cb139d38-10	female		No Risk	<a href="#">Explain</a>
61a034fa736dabc57626b399cb139d38-100	male		No Risk	<a href="#">Explain</a>
61a034fa736dabc57626b399cb139d38-101	male		No Risk	<a href="#">Explain</a>
61a034fa736dabc57626b399cb139d38-102	male		No Risk	<a href="#">Explain</a>
61a034fa736dabc57626b399cb139d38-103	male		No Risk	<a href="#">Explain</a>
61a034fa736dabc57626b399cb139d38-104	male		No Risk	<a href="#">Explain</a>
61a034fa736dabc57626b399cb139d38-105	female		Risk	<a href="#">Explain</a>

Payload Table  
Payload\_a2fe6bb3-9439-4ef9-b8e8-4372cf05f...

Corrected Records

No Risk : Favorable Outcome

Current Model  
De-biased Model

Risk : Unfavorable Outcome

Current Model  
De-biased Model

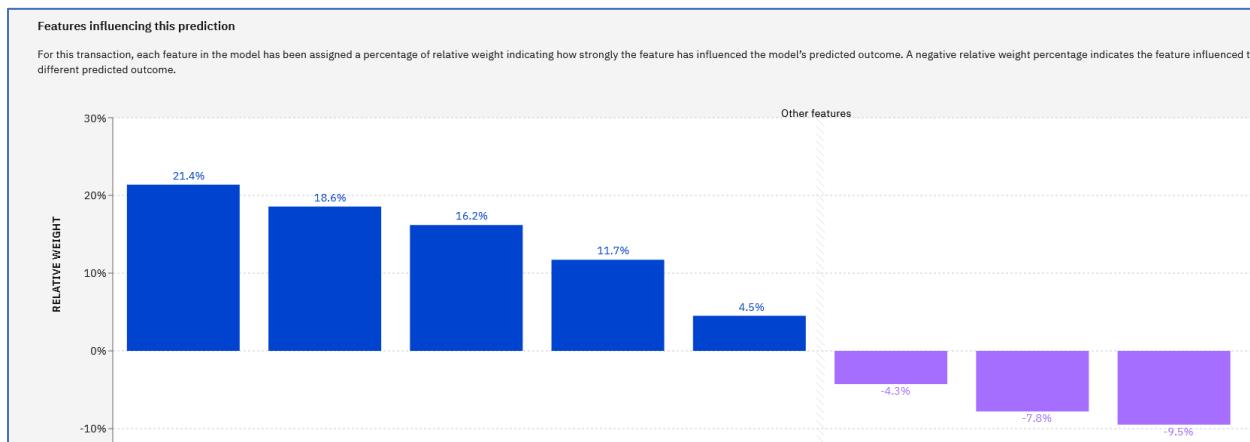
17. The results show the confidence level for the result including an explanation of how the result was determined and which features were most important.

f724856002cd4d1a475837e1799ef987-105 Aug 24, 2021, 09:54:58 PM EDT credit-risk-deploy Production Not applicable

Predicted outcome Risk How this prediction was determined  
The credit-risk model has 56.49% confidence that the outcome of this transaction would be Risk. The top three features influencing the model's predicted outcome are EmploymentDuration, Age, and LoanAmount. The top three features CheckingStatus, OthersOnLoan, and Telephone are influencing the model toward a predicted outcome of No Risk.

Confidence level 56.49%

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



19. Click on the Dashboard icon

IBM Watson OpenScale

f724856002cd4d1a475837e1799ef987-105 Aug 24, 2021, 09:54:58 PM EDT

Predicted outcome Risk

How this prediction was determined  
The credit-risk model has 56.49% confidence in predicting Risk based on the features: LoanAmount. The top three features are:

20. 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.

IBM Watson OpenScale

Need help? FL

Insights Dashboard

Refresh Add to dashboard +

Deployments Monitored	Quality Alerts	Fairness Alerts	Drift Alerts
1	1	1	0

Filter by Tags ▼ Alert type ▼ Machine learning provider ▼

Sort by Severity ▼

Q Find models

New provider credit-risk-deploy ⋮

Issues 2 QUALITY BIAS

Quality	Fairness	Drift
0.74	94%	4%
1 alert	1 alert	

Production Evaluates 4 minutes 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.