

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.

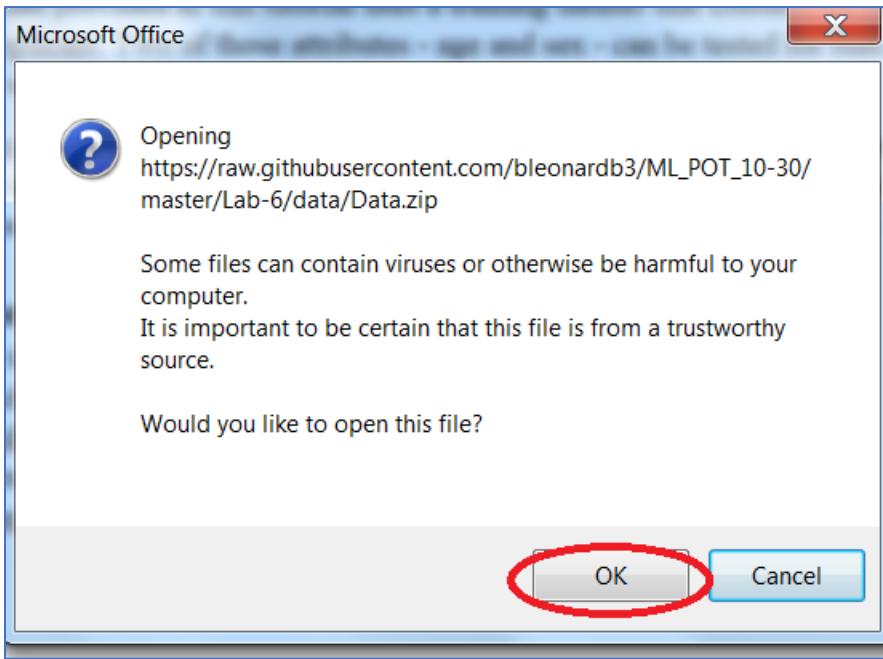
The credit risk model provided in this tutorial uses a training dataset that contains 20 attributes about each loan applicant. Two of those attributes - age and sex - can be tested for bias. For this tutorial, the focus will be on bias against sex and age.

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`.

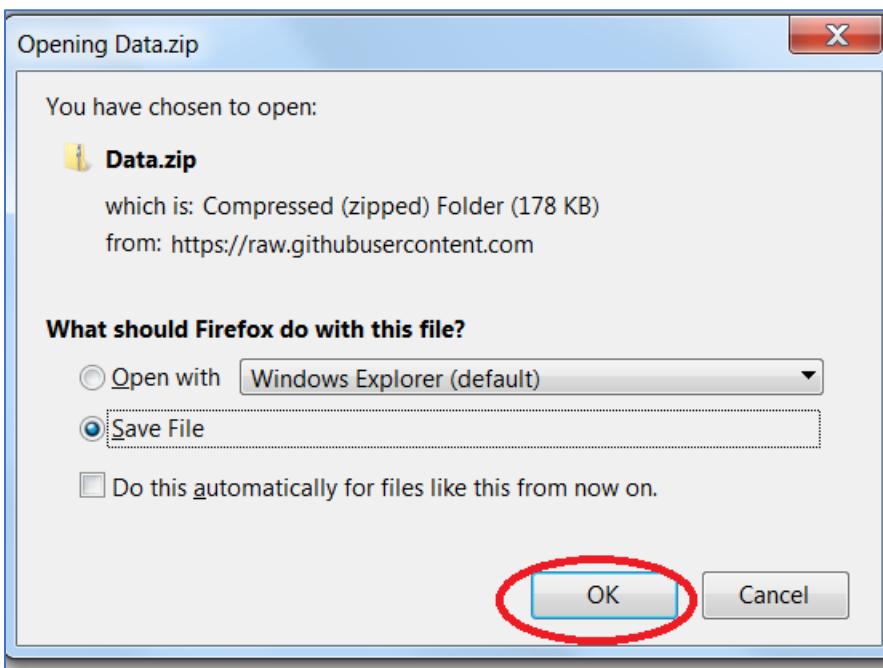
Download the Lab Files

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`

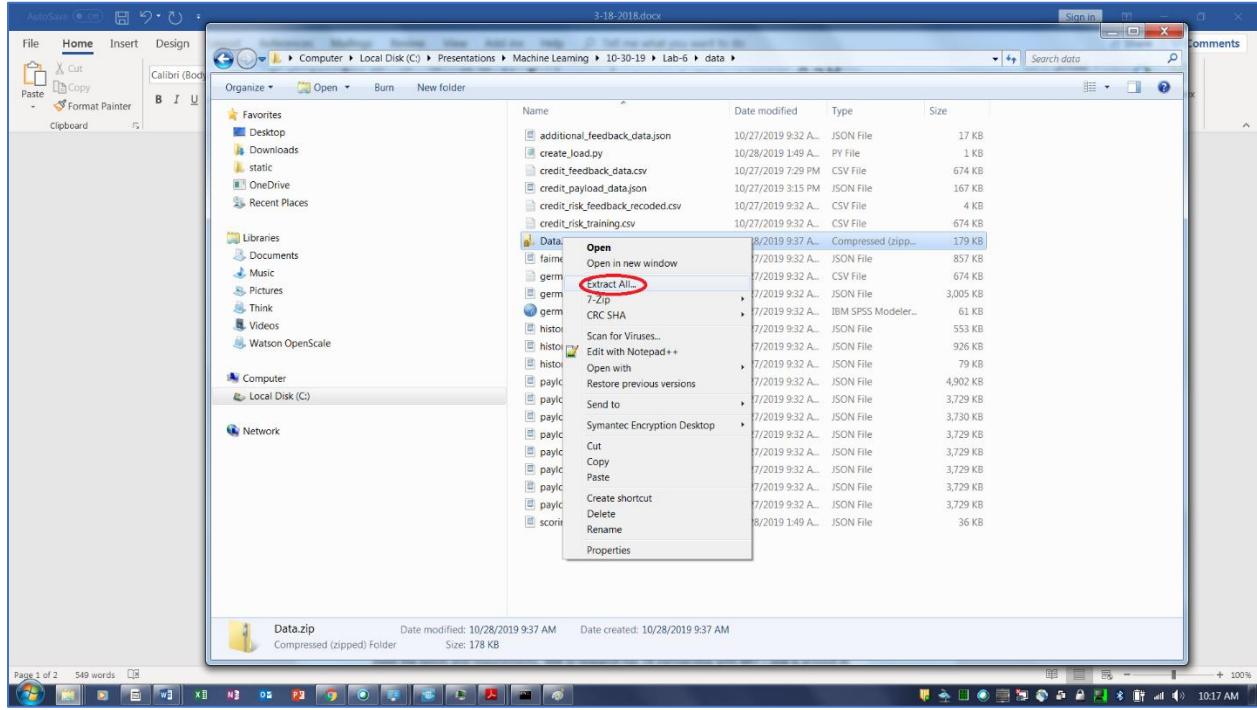
2. Click **OK**. Note your browser may not prompt this message.



3. Click **OK**.



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

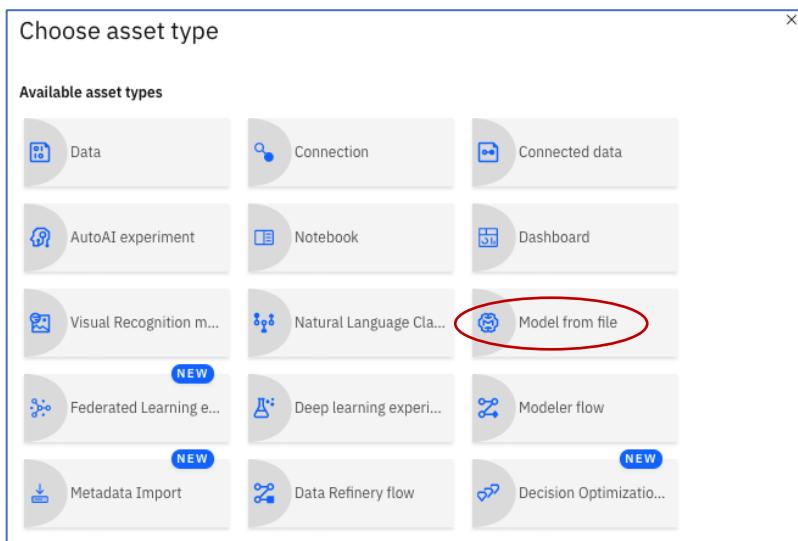


Import the Credit Risk Model

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



2. 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 text as the card. 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 this, there are sections for 'Input Schema' and 'Output' with tables showing column names and types. The 'Input' schema table:

Column	Type
Age	"integer"
CheckingStatus	"string"
CreditHistory	"string"
CurrentResidenceDuration	"integer"
Dependents	"integer"
EmploymentDuration	"string"
ExistingCreditsCount	"integer"
ExistingSavings	"string"

The 'Output' schema table is similar, showing the same columns and types.

- Click on **openscale-express-path** (your path may have long set of numbers and letters after it), for the Target space, check “**Go to the model in the space after promoting it,**” and click **Promote**.

Promote to space

Target space
openscale-express-path-15c8dbf8-a3e6-4da4-ac8d-d7b4180b6a12

Why don't I see all of my spaces? ⓘ Go to the model in the space after promoting it

Selected assets (1)

Asset name	Format
credit-risk	Model

Description (optional)
Description of assets

Promote

- Click Create deployment.

Deployments / openscale-express-path-15c8dbf8-a3e6-4da4-ac8d-d7b4180b6a12 / credit-risk

credit-risk

Create deployment

Deployments Schema

You don't have any deployments yet
Create your first deployment for this model. [Learn more](#)

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

Online <input checked="" type="radio"/>	Run the model on data in real-time, as data is received by a web service.
Batch	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)	credit-risk-deploy	Deployed	May 16, 2021 4:49 PM
Batch	(0)			

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

credit-risk

Create deployment

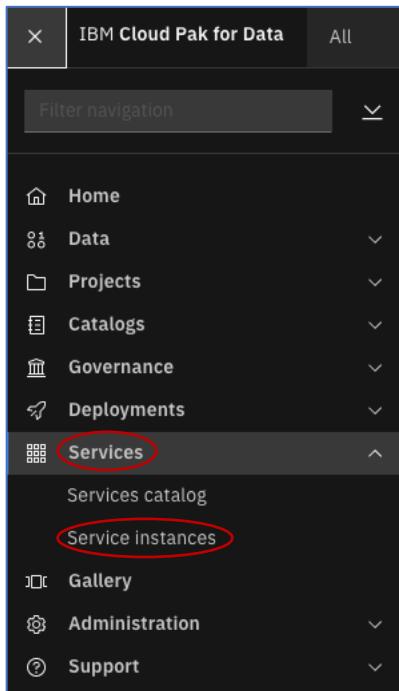
Deployments Schema

DEPLOYMENT TYPES		1 Online Deployment(s)		
		Name	Status	Last modified
Online	(1)	credit-risk-deploy	Deployed	May 16, 2021 4:49 PM
Batch	(0)			

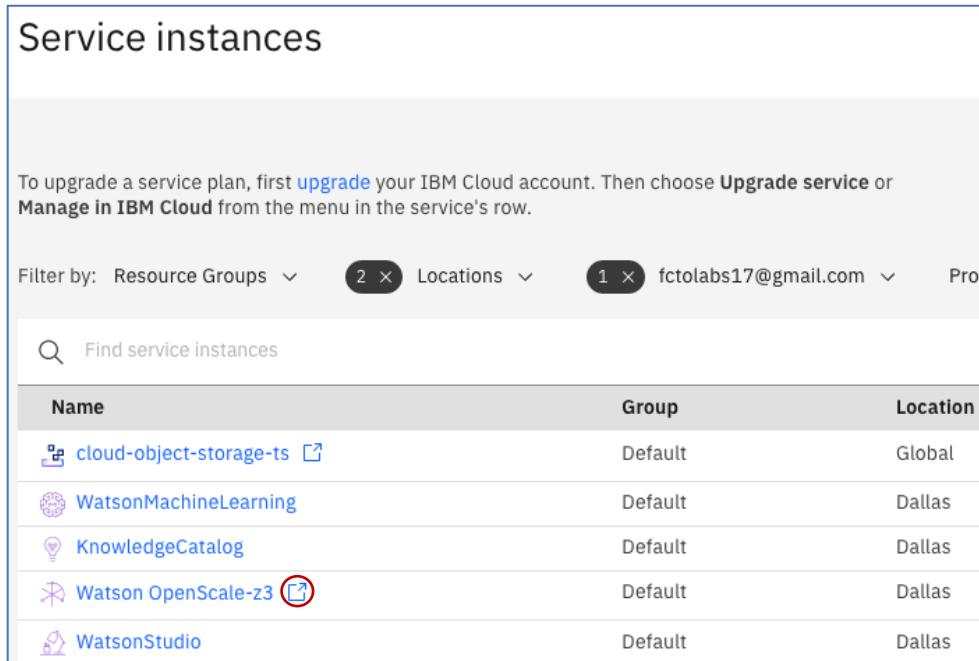
7. Keep this tab open, we will return to this page later.

Begin OpenScale Configuration

1. Click on the hamburger icon  in the top left corner.
2. Click on **Services**, and **Service instances**.



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



The screenshot shows the 'Service instances' page. At the top, there is a message: '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.' Below this, there are filter options: 'Filter by: Resource Groups' (with a dropdown arrow), 'Locations' (set to 'Dallas'), and an email filter 'fctolabs17@gmail.com'. A search bar says 'Find service instances'. The main table lists the following 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

4. Click on **Launch Application.**

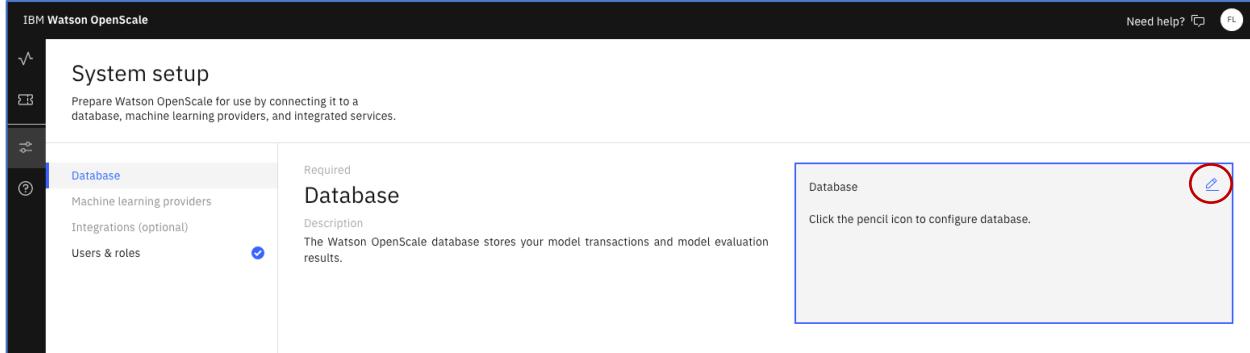
The screenshot shows the Watson OpenScale application interface. At the top, there is a navigation bar with 'Resource list / Watson OpenScale-qe' and status indicators like 'Active'. Below the navigation, there are three tabs: 'Getting started', 'Manage' (which is selected), and 'Plan'. In the center, there is a circular icon with a stylized 'X' or 'scale' symbol. Below the icon, the text 'Watson OpenScale' is displayed. A welcome message says '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, there are links for 'Documentation' and 'Community', and a feedback icon.

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

The screenshot shows the 'Welcome to Watson OpenScale' screen. It features a large circular icon with a blue gradient and some white dots. Below the icon, the text 'Welcome to Watson OpenScale' is displayed. A paragraph explains the purpose of 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.' Below this text, another paragraph 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' (highlighted with a red oval) and 'Auto setup'.

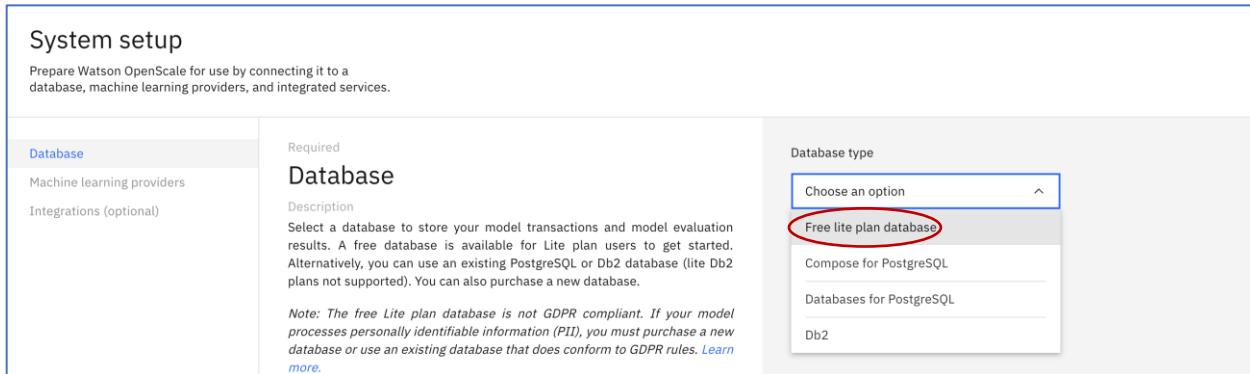
Setup System

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.



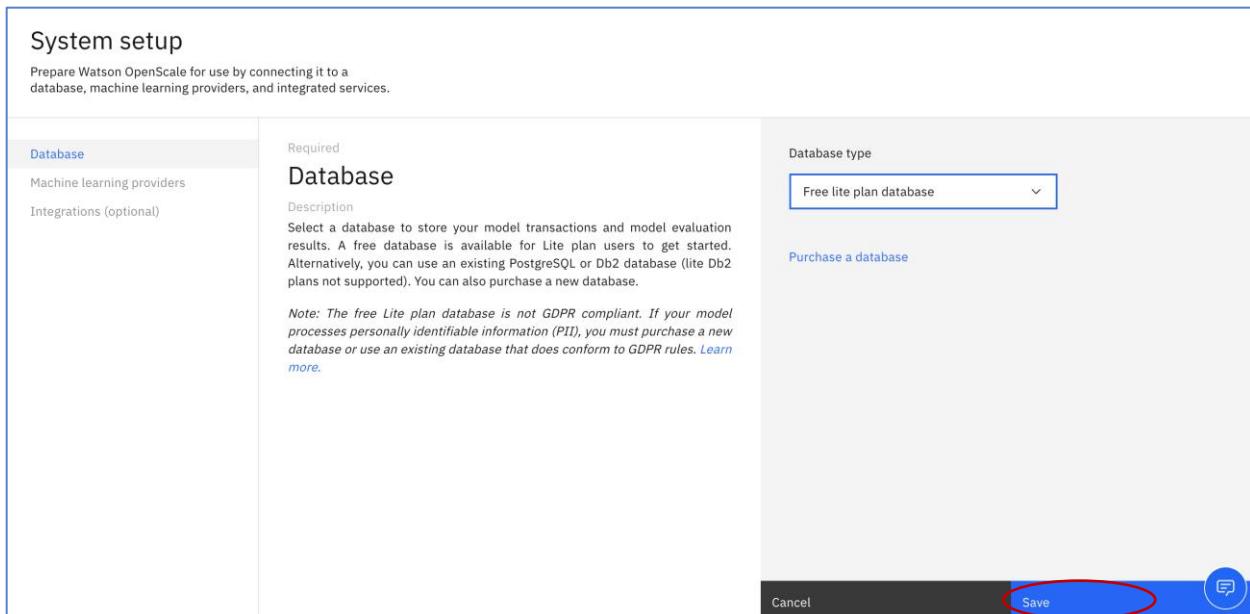
The screenshot shows the 'System setup' page of the IBM Watson OpenScale interface. On the left, there's a sidebar with sections like 'Database', 'Machine learning providers', 'Integrations (optional)', and 'Users & roles'. The 'Database' section is selected and highlighted in blue. The main area has a title 'Database' with a subtitle 'Required'. Below it is a 'Description' section with text about the Watson OpenScale database storing model transactions and evaluation results. To the right, there's a 'Database type' dropdown menu with options: 'Choose an option', 'Free lite plan database' (which is circled in red), 'Compose for PostgreSQL', 'Databases for PostgreSQL', and 'Db2'. A note at the bottom of the description section states: 'Note: The free Lite plan database is not GDPR compliant. If your model processes personally identifiable information (PII), you must purchase a new database or use an existing database that does conform to GDPR rules. [Learn more](#)'.

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



This screenshot shows the same 'System setup' page as the previous one, but the 'Database type' dropdown now has 'Free lite plan database' selected, indicated by a red circle. At the bottom right of the page, there is a large blue 'Save' button, which is also circled in red.

4. Click **Save**.



This screenshot shows the final step where the 'Save' button has been clicked. The 'Database type' dropdown now shows 'Free lite plan database'. The bottom right corner of the page features a blue 'Save' button with a white speech bubble icon, which is circled in red.

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.

System setup

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

Database

Machine learning providers

Integrations (optional)

Users & roles

Machine learning providers

Connection

New provider

Description

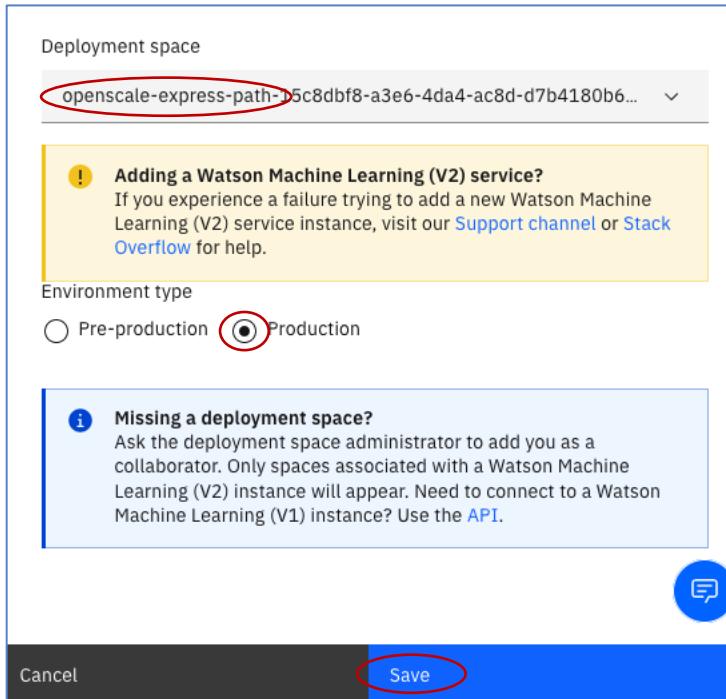
Service provider

Watson Machine Learning (V2)

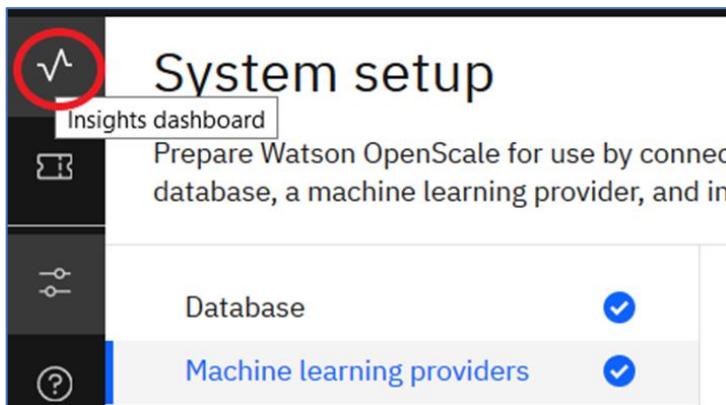
Watson Machine Learning (V2)

Custom Environment

9. Scroll down and click on **openscale-express-path** for the Deployment Space, click on **Production** for the **Environment type**, and click on **Save**.



10. Click on the icon.



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

The screenshot shows the Insights Dashboard interface. On the left is a sidebar with icons for navigation. The main area displays four metrics: Deployments Monitored (0), Quality Alerts (0), Fairness Alerts (0), and Drift Alerts (0). Below these are filter options for Tags, Alert type, Machine learning provider, and Severity, followed by a search bar labeled 'Find models'. In the top right corner, there is a blue button labeled 'Add to dashboard' with a '+' sign, which is circled in red.

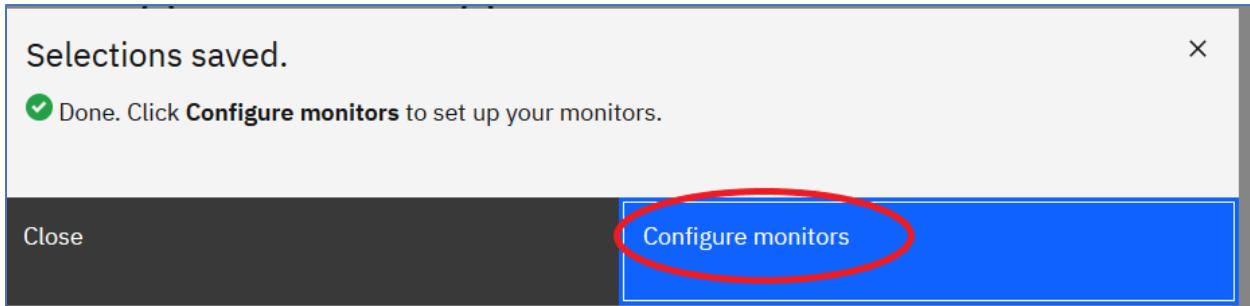
12. Click on **credit-risk-deploy** and then click on **Configure** to configure the model details, and the quality, fairness, and drift monitors.

The 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 the option 'New provider (Production)' selected. Below this is a table listing four model 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
credit-risk-deploy	-	Sun, May 16, 2021, 4:49 PM EDT

At the bottom of the dialog, there are 'Cancel' and 'Configure' buttons. The 'Configure' button is circled in red.

13. Click on **Configure monitors**.



14. 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 ([if you have this tab open skip to step 4](#)), if you do not have this tab open, paste the following link into your browser:
<https://dataplatform.cloud.ibm.com/ml-runtime/spaces?context=cpdaas>
2. Click on your openscale-express deployment space.

The screenshot shows the IBM Cloud Pak for Data interface. At the top, there's a navigation bar with "IBM Cloud Pak for Data", "All", "Search", "Upgrade", and "FCTO Labs's Account". Below the navigation is a "Deployments" section with a sub-header "1 space". There are two tabs: "Activity" and "Spaces", with "Spaces" being the active one. Under "Spaces", there's a search bar and a table. The table has columns: Name, Last modified, Your role, Collaborators, Tags, Online deployments, and Jobs. One row is highlighted with a red circle around the "Name" column, which contains "openscale-express-path".

3. Click on **Deployments** and **credit-risk-deploy**.

The screenshot shows the deployment details for "openscale-express-path". At the top, it says "Deployments / openscale-express-path-15c8db...". There are tabs for "Assets", "Deployments" (which is highlighted with a red circle), "Jobs", and "Manage". Below the tabs is a search bar. The main area shows "Deployments (4)". A table lists four entries, with the first one highlighted by a red circle around its "Name" column, which contains "credit-risk-deploy".

Name	Type	Status	Asset	Tags	Last modified
credit-risk-deploy	Online	Deployed	credit-risk		May 16, 2021 4:49 PM
...
...

4. Click on the **Test** tab.

The screenshot shows the IBM Cloud Pak for Data interface. At the top, there's a navigation bar with 'IBM Cloud Pak for Data' and a search bar. Below it, a breadcrumb trail shows 'Deployments / openscale-express-path-15c8db... / credit-risk / credit-risk-deploy'. The main area displays a deployment named 'credit-risk-deploy' with a green 'Deployed' status and an 'Online' button. Below the deployment name, there are two tabs: 'API reference' and 'Test', with 'Test' being the active tab and circled in red. The background is white with blue accents for the tabs and buttons.

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

This screenshot shows the 'Test' tab of the deployment interface. The 'Enter input data' section has a 'Body' placeholder with a JSON template. The template starts with a brace '{' and includes fields like 'input_data': [{ 'fields': [], 'values': [] }]. To the right of the input area is a small icon with a square and a downward arrow, which is circled in red. Below the input area is a blue 'Predict' button. The overall interface is clean with a white background and blue UI elements.

6. Copy the file contents of **credit_payload_data.json** into the **input data** area and click on **Predict**.

This screenshot shows the 'Enter input data' section with a large red box highlighting the JSON payload. The payload is a list of arrays, each containing a 'Risk' value ('yes' or 'Risk') and a list of numerical features. The 'Predict' button at the bottom is also circled in red. The background is white with blue UI elements.

7. The results should appear as below.

The screenshot shows the Watson OpenScale API Test interface. At the top, it says "credit-risk-deploy" with a green "Deployed" status and an "Online" button. Below that, there are two tabs: "API reference" and "Test", with "Test" being the active tab. On the left, under "Enter input data", there is a large text area containing JSON input data. On the right, under "Result", there is a "Result" section with a code editor showing a JSON response. A red circle highlights the "Predict" button at the bottom center of the interface.

```
[{"id": 1, "fields": ["CheckingStatus", "LoanDuration", "CreditHistory", "LoanPurpose", "LoanAmount", "ExistingSavings", "EmploymentDuration", "InstallmentPercent", "Sex", "OthersOnLoan", "CurrentResidenceDuration", "OwnsProperty", "Age", "InstallmentPlans"]}, {"id": 2, "fields": ["CheckingStatus", "LoanDuration", "CreditHistory", "LoanPurpose", "LoanAmount", "ExistingSavings", "EmploymentDuration", "InstallmentPercent", "Sex", "OthersOnLoan", "CurrentResidenceDuration", "OwnsProperty", "Age", "InstallmentPlans"]}], [{"id": 1, "label": "Predict"}]
```

Configure Model Details

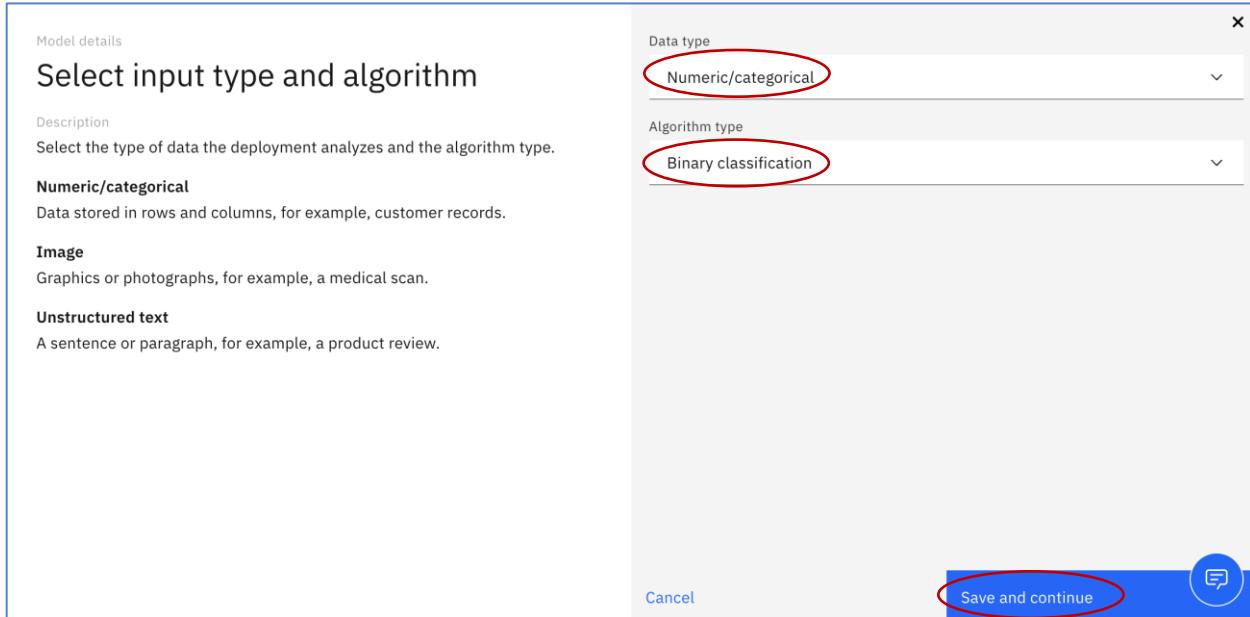
1. Switch back to Watson OpenScale by clicking on the Watson OpenScale browser tab.



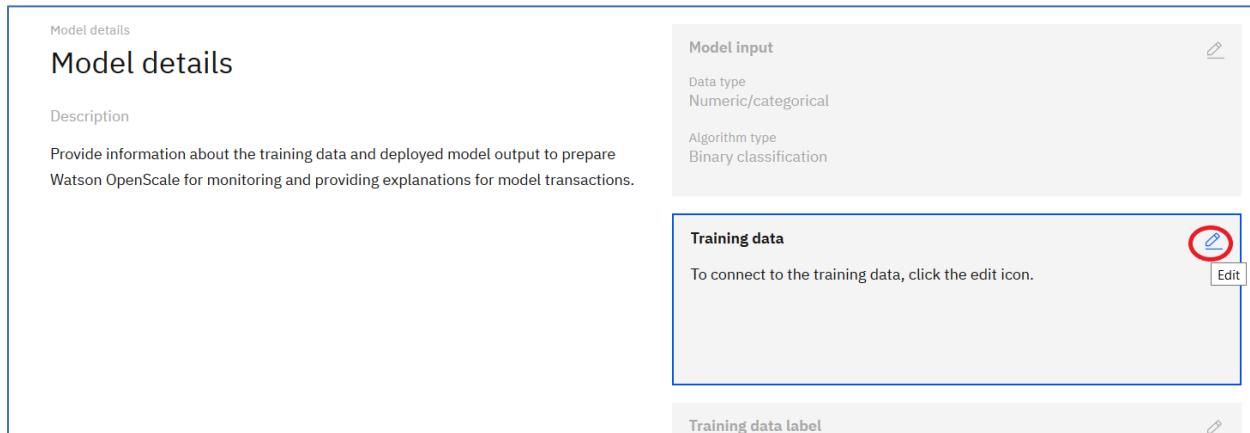
2. Click on pencil icon to edit the **Model input**.

The screenshot shows the "Model details" configuration interface. It has three main sections: "Model details", "Model input", and "Model transaction".
- "Model details" section: Contains a "Description" field with the text: "Provide information about the training data and deployed model output to prepare Watson OpenScale for monitoring and providing explanations for model transactions."
- "Model input" section: Contains a "Model input" field with the text: "To select the data type and algorithm type, click the edit icon." A red circle highlights the edit icon (a pencil symbol) located to the right of the text.
- "Model transaction" section: Contains a "Model transaction" field with the text: "Model transaction Successful". An edit icon is also present here.

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



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



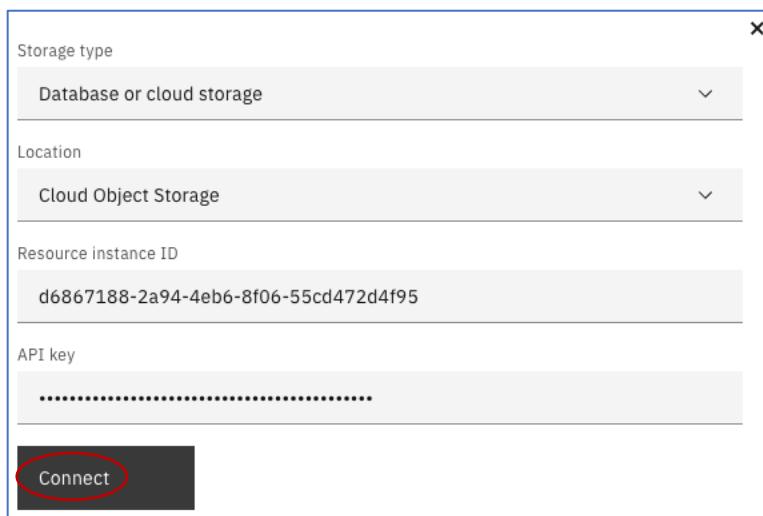
5. Copy and paste the following text as the **resource_id**:

d6867188-2a94-4eb6-8f06-55cd472d4f95

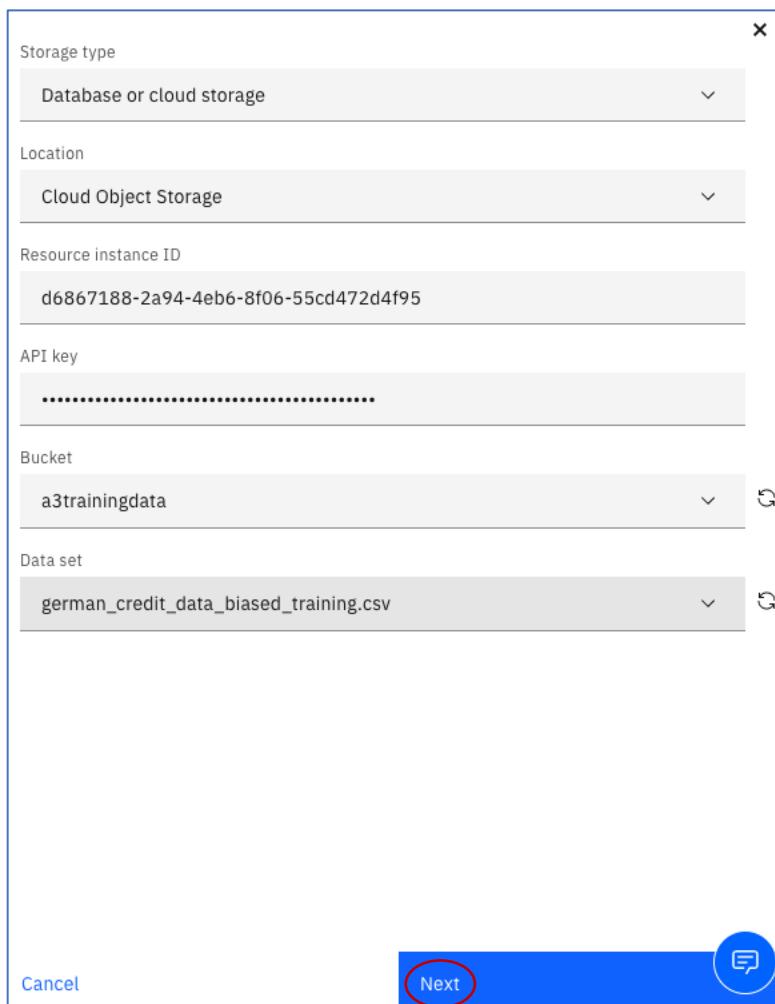
6. Copy and paste the following text as the **apikey**:

JzxKc8k1IZat1W14ZZgY0sKgFBEQGQAi9RmYdPtq1vIE

7. Click Connect



8. Select **a3trainingdata** as the bucket, select **german_credit_data_biased_training.csv** as the Data set, and click **Next**.



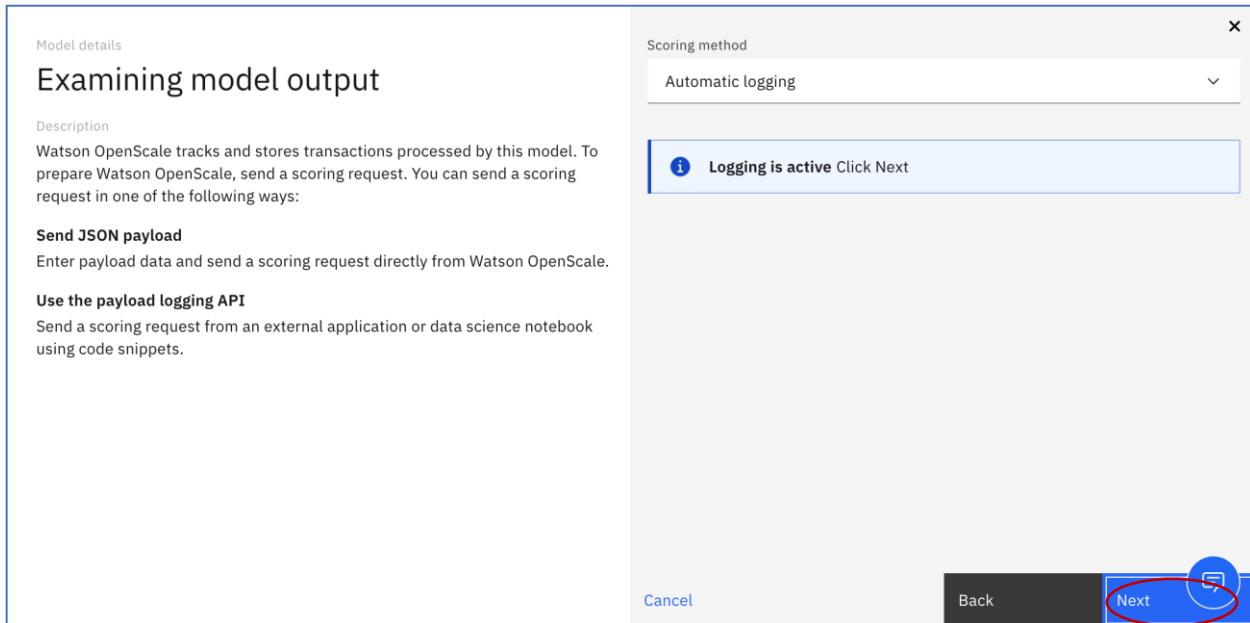
9. Watson OpenScale has determined that Risk is the label (target) column. Click **Next**.

The screenshot shows the 'credit-risk-deploy' interface. On the left, there's a sidebar with 'Model info' sections for 'Model details', 'Endpoints', 'Evaluations' (Fairness, Quality, Drift, Explainability), and 'Import settings'. Below these are links to 'Go to model summary' and 'Go to model details'. The main area is titled 'Select the label column' with a sub-section 'Model details' and 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' is open, showing a table with one row: 'Features (1)' with 'Risk' listed under it, and 'Type' with an arrow icon. At the bottom of the modal are 'Cancel', 'Back', and 'Next' buttons, with 'Next' being circled in red.

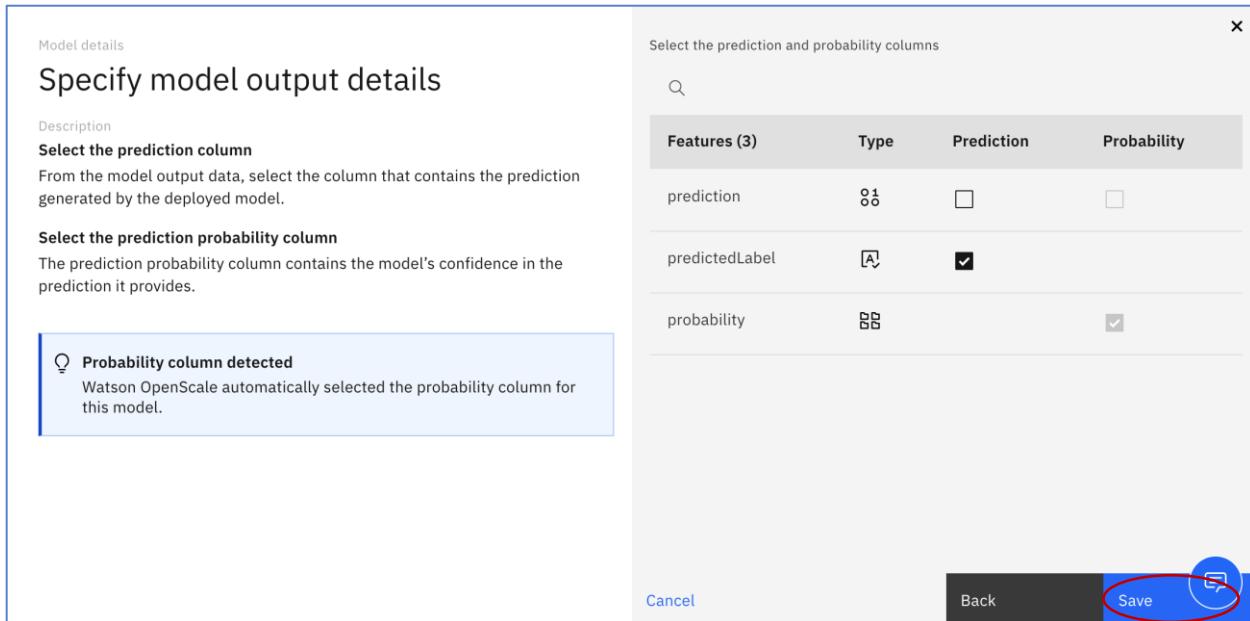
10. All of the features are used to train the model. Watson OpenScale has determined the feature types. Click **Next**.

The screenshot shows the 'credit-risk-deploy' interface. The left sidebar includes 'Model details', 'Endpoints', 'Evaluations' (Fairness, Quality, Drift, Explainability), and 'Import settings'. The main area is titled 'Select the training features' with a sub-section 'Model details' and 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' is open, showing a table with 20 features. The columns are 'Features (20)', 'Type' (with icons for numeric and categorical), and 'Categorical' (checkboxes). Features listed include Age, CheckingStatus (marked as categorical), CreditHistory (marked as categorical), CurrentResidenceDuration, Dependents, and EmploymentDuration. At the bottom of the modal are 'Cancel', 'Back', and 'Next' buttons, with 'Next' being circled in red.

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



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



13. 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 displays a section titled 'Specify the monitored groups for [Sex]'. It includes a 'Description' field, a note about monitoring groups, and a section for setting a 'fairness alert threshold' to 95. A callout box highlights the 'Recommended groups' section, which lists groups identified by Watson OpenScale. On the right, a modal window titled 'Select the groups to monitor [Sex]' is open. It contains a search bar, a table with columns 'Values', 'Monitored', 'Reference', and 'Recommended', and a text input for setting the fairness alert threshold to 95. The 'Save' button at the bottom of the modal is 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, the 'Drift' feature is highlighted with a red circle. The main area shows the 'Fairness' monitor configuration. It includes a 'Description' field stating that the Fairness monitor checks for biases in model outcomes. To the right, there are sections for 'Favorable outcomes' (listing 'No Risk') and 'Unfavorable outcomes' (listing 'Risk'). Below these is a 'Sample size' section set to a minimum of 200. The 'Save' button at the bottom right of the configuration panel is circled in red.

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

Dashboard / Configure

credit-risk-deploy

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

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
Edit icon highlighted with a red circle.

- 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**.

The screenshot shows the 'Drift threshold' configuration screen. On the left, there's descriptive text about tracking model accuracy change. On the right, a input field shows '10 %'. At the bottom, there are 'Cancel', 'Back', and 'Next' buttons, with 'Next' being highlighted with a red circle.

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

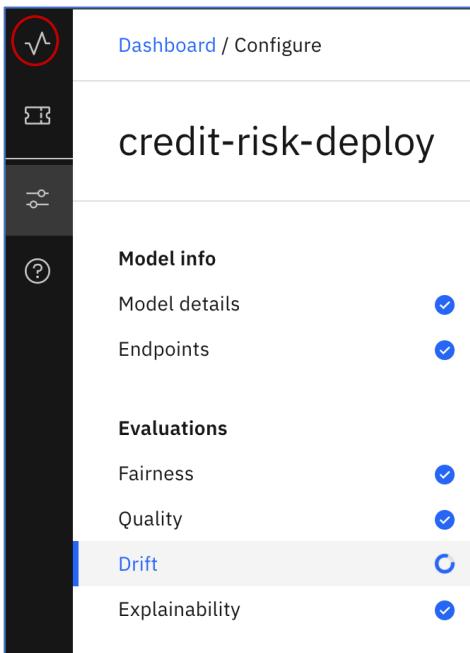
The screenshot shows the 'Sample size' configuration screen. On the left, there's descriptive text about ensuring the sample size is representative. On the right, a input field shows '200'. At the bottom, there are 'Cancel', 'Back', and 'Save' buttons, with 'Save' being highlighted with a red circle.

6. This completes the Drift configuration.

Submit Feedback and View Quality Metrics

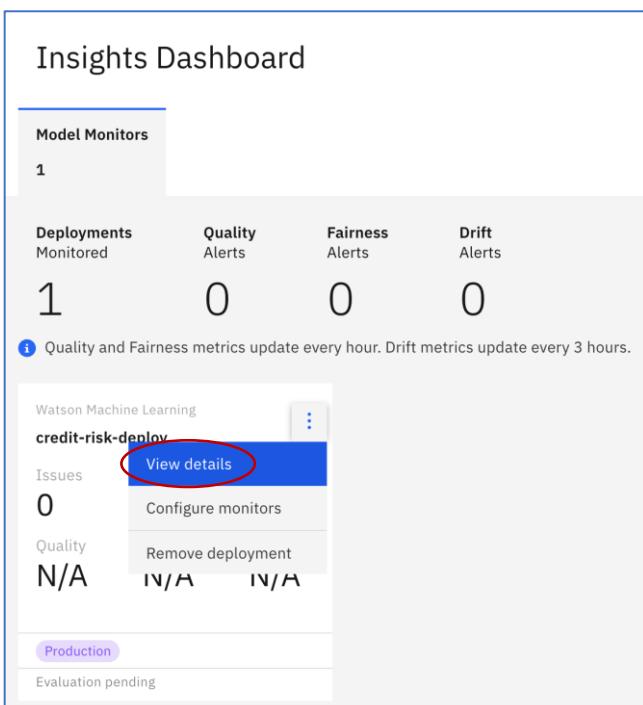
In order to measure quality, scored transactions including human labeled feedback must be provided.

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



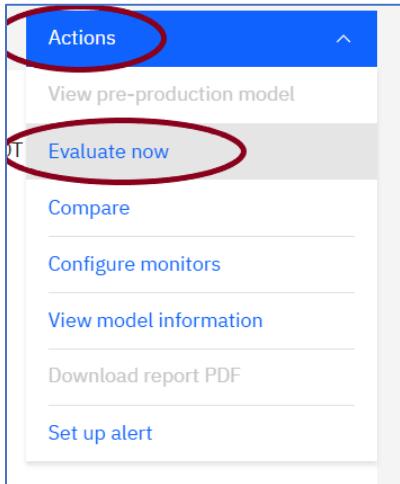
The screenshot shows the 'Model info' section of the Insights Dashboard. On the left is a sidebar with icons for Dashboard, Model info, Evaluations, Drift, and Explainability. The main area displays 'credit-risk-deploy' under 'Model info'. It includes sections for 'Model details' (checked), 'Endpoints'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

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

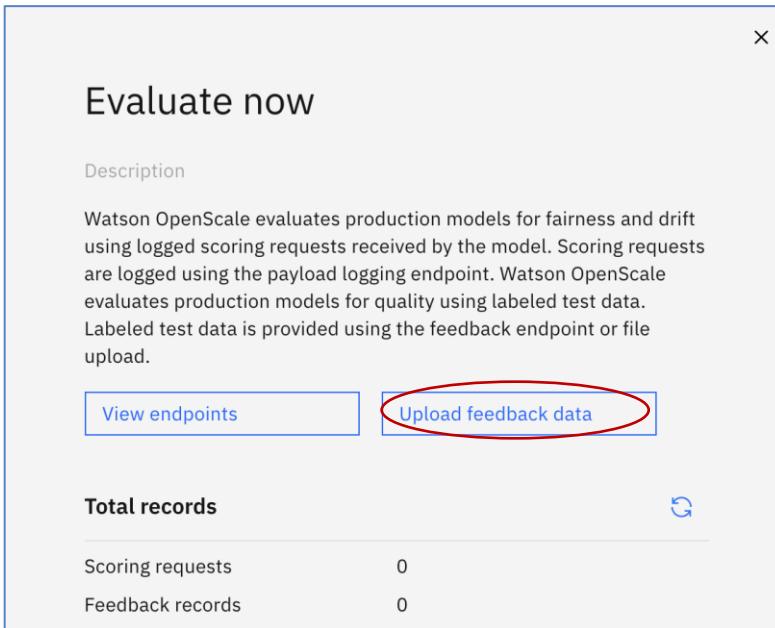


The screenshot shows the 'Insights Dashboard' for 'credit-risk-deploy'. It features a summary table with counts for Deployments Monitored (1), Quality Alerts (0), Fairness Alerts (0), and Drift Alerts (0). A note below states: 'Quality and Fairness metrics update every hour. Drift metrics update every 3 hours.' Below this is a card for 'Watson Machine Learning' with 'credit-risk-deploy' selected. It shows 'Issues' (0) and 'Quality' (N/A). A context menu is open over the 'Issues' field, with the 'View details' option highlighted and circled in red. Other options in the menu include 'Configure monitors' and 'Remove deployment'. At the bottom, there's a 'Production' button and a note 'Evaluation pending'.

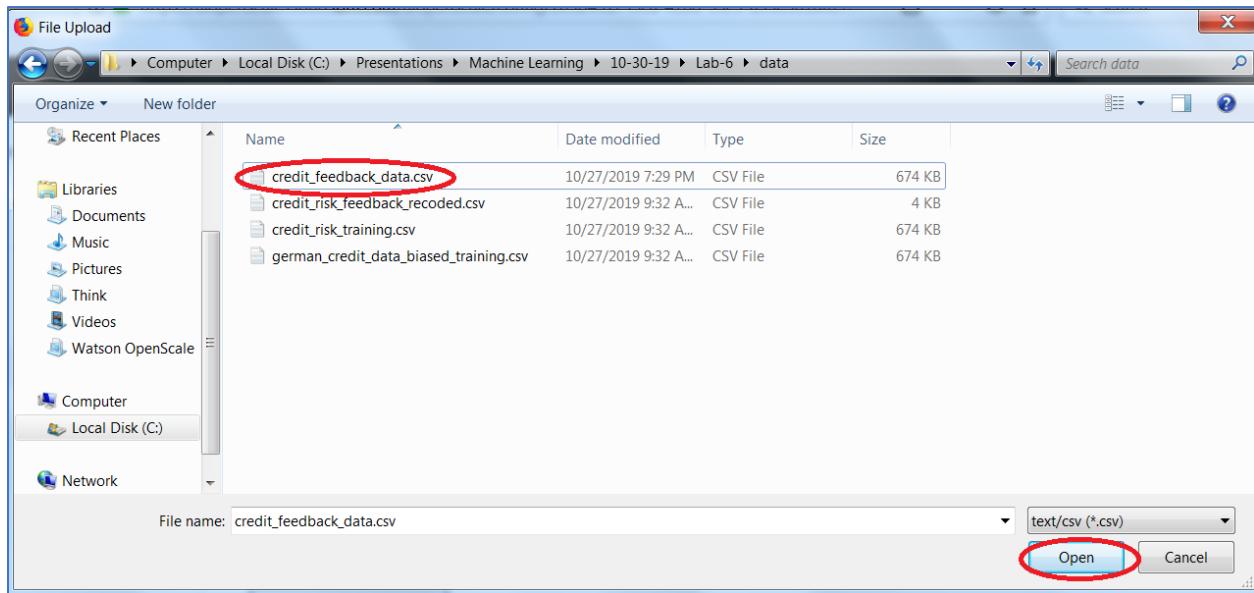
3. It will take a minute or so to bring up the **Evaluations** page. Click on **Actions** in the top right corner of your dashboard and then click **Evaluate now**.



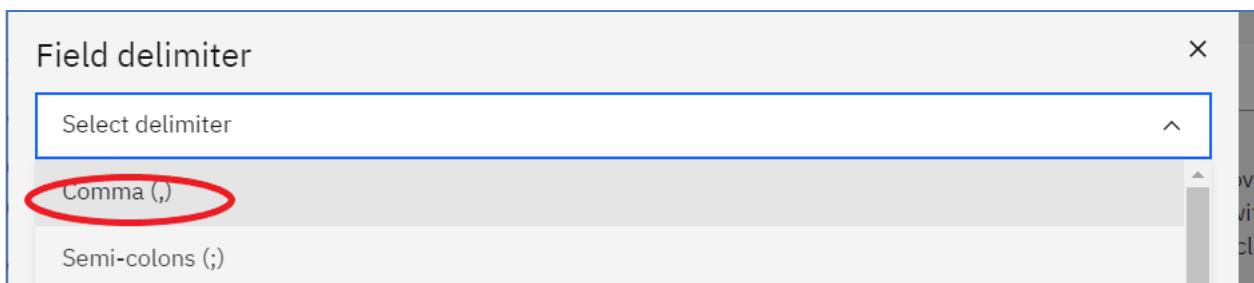
4. Watson OpenScale evaluates models for quality using labeled test data. Labeled test data is provided using the feedback endpoint or by file upload. We will use the file upload options. Click on **Upload feedback data**.



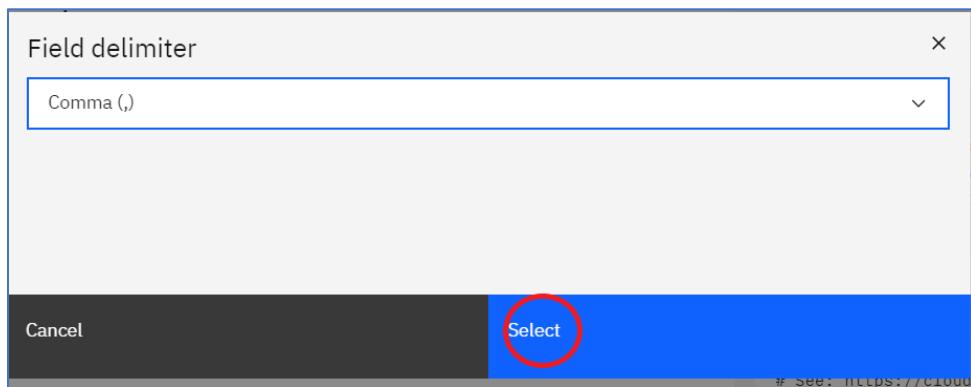
5. Navigate to the feedback file **credit_feedback_data.csv**. Click on **Open**.



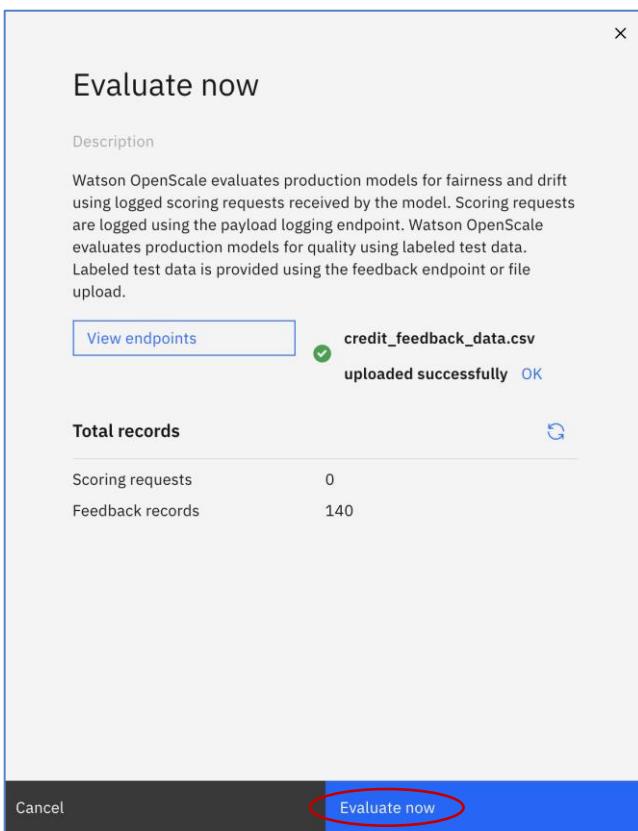
6. Select **Comma** as the **Delimiter**.



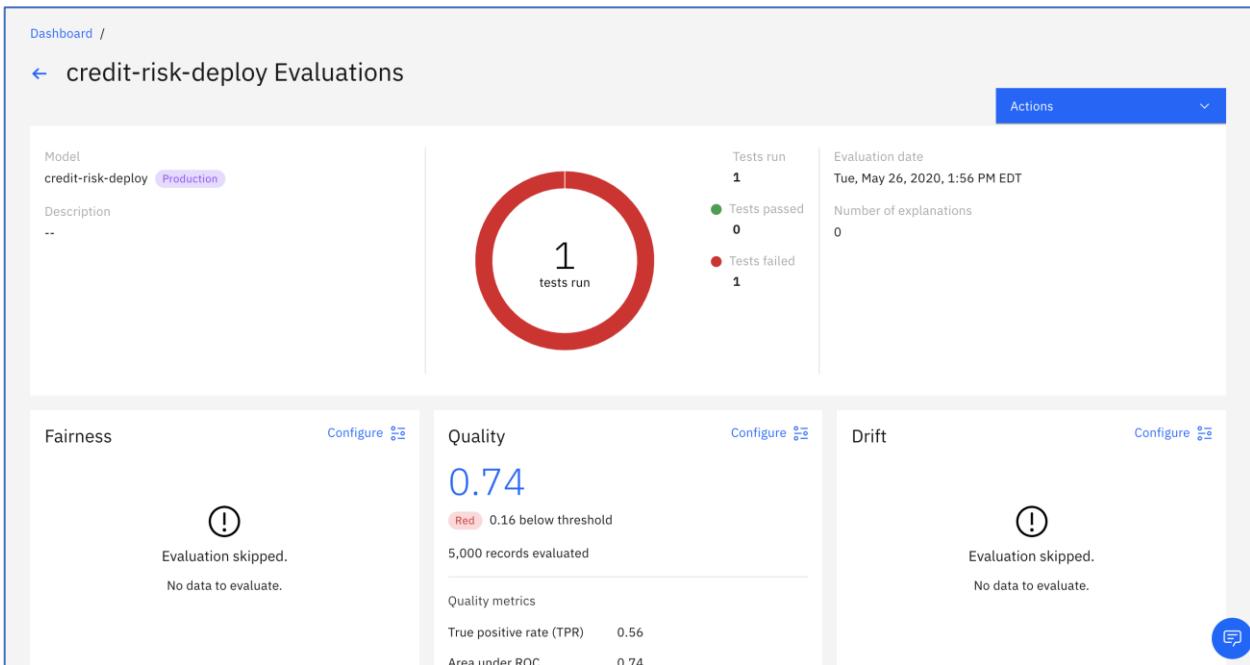
7. Click **Select**.



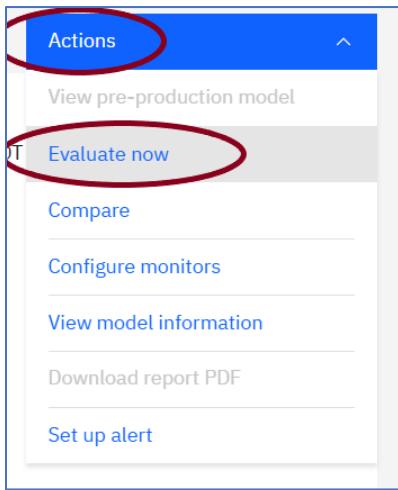
8. After the file is successfully uploaded, click on the **Evaluate now**.



9. Please wait about 3-5 minutes until the evaluation is done. The quality 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 Watson Studio by clicking on the **Watson Studio** browser tab.



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

A screenshot of the Watson Studio interface, specifically the 'credit-risk-deploy' project under the 'Test' tab. On the left, there's a section titled 'Enter input data' with a text area containing JSON data. On the right, there's a large JSON schema for the input fields. The schema includes fields like 'CheckingStatus', 'LoanDuration', 'CreditHistory', 'LoanPurpose', 'LoanAmount', 'ExistingSavings', 'EmploymentDuration', 'InstallmentPercent', 'Sex', 'OthersOnLoan', 'CurrentResidenceDuration', and 'OwnsProperty'. A 'Predict' button is located at the bottom left of the input 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", "CurrentResidenceDuration", "OwnsProperty"]}]
```

5. Click on Predict

The screenshot shows the 'credit-risk-deploy' interface with the 'Test' tab selected. Below it, there's a section titled 'Enter input data' containing a JSON object and a 'Predict' button.

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

Predict

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



7. Click Evaluate now.

The screenshot shows the 'Evaluate now' dialog box. It includes a 'Description' section with text about Watson OpenScale's evaluation process, two buttons ('View endpoints' and 'Upload feedback data'), a summary of 'Total records' (Scoring requests: 200, Feedback records: 5,000), and a 'Cancel' button at the bottom left and an 'Evaluate now' button at the bottom right.

Description

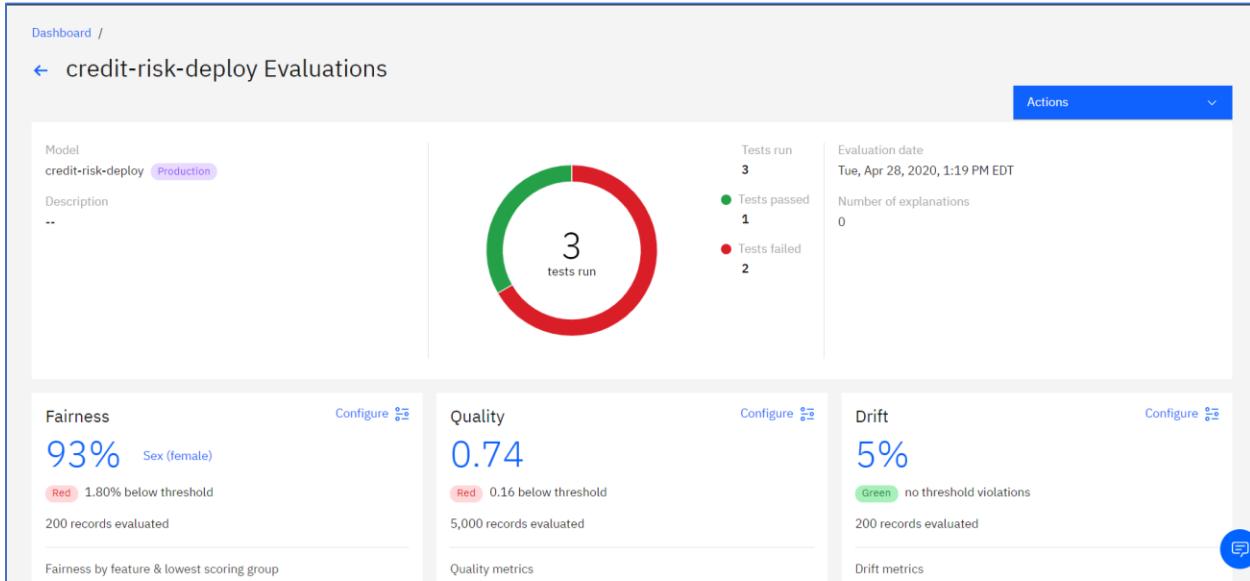
Watson OpenScale evaluates production models for fairness and drift using logged scoring requests received by the model. Scoring requests are logged using the payload logging endpoint. Watson OpenScale evaluates production models for quality using labeled test data. Labeled test data is provided using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

Total records

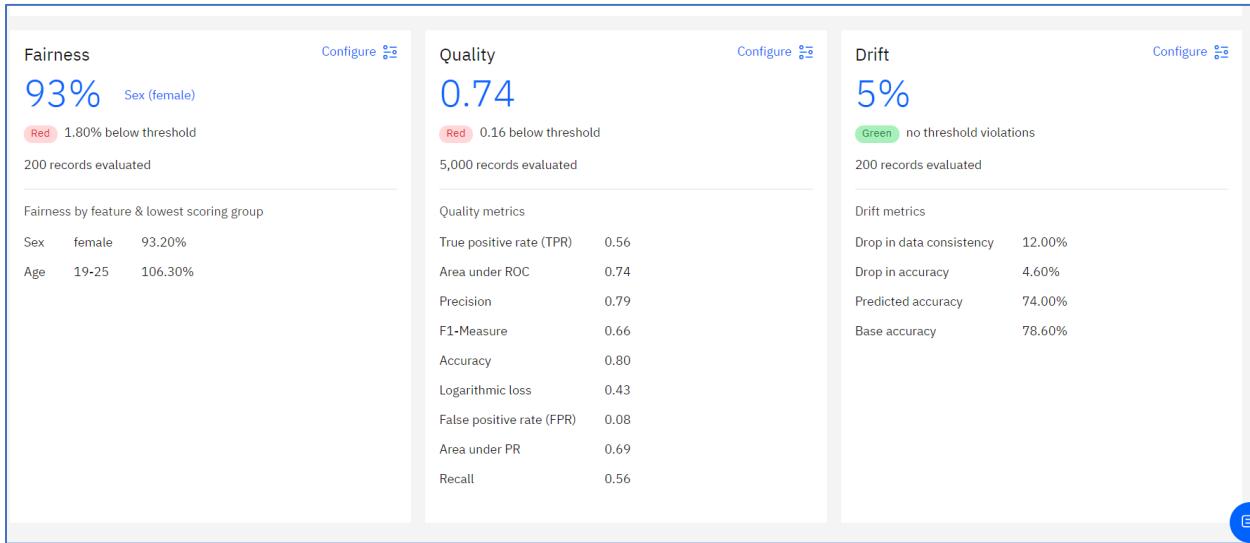
Scoring requests	200
Feedback records	5,000

Evaluate now

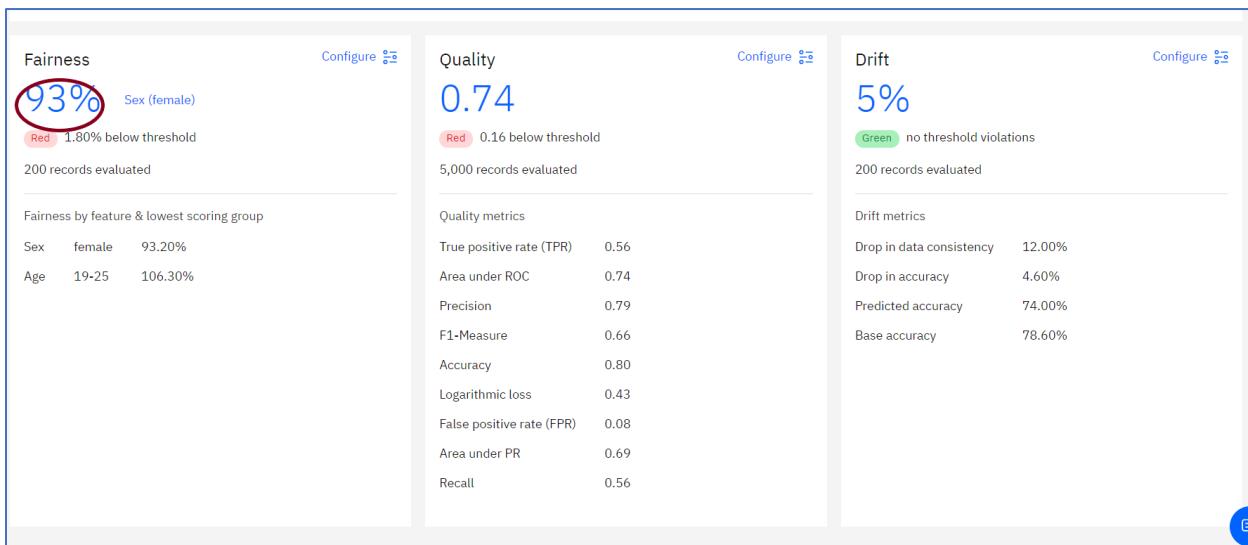
8. Please wait 3-5 minutes for the Evaluations page to display the metrics. Three tests have now been run. The quality test from before, and two new tests. One for fairness, and one for Drift. The fairness test has failed because 93% fairness is lower than our threshold of 95%. The Drift metric passed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



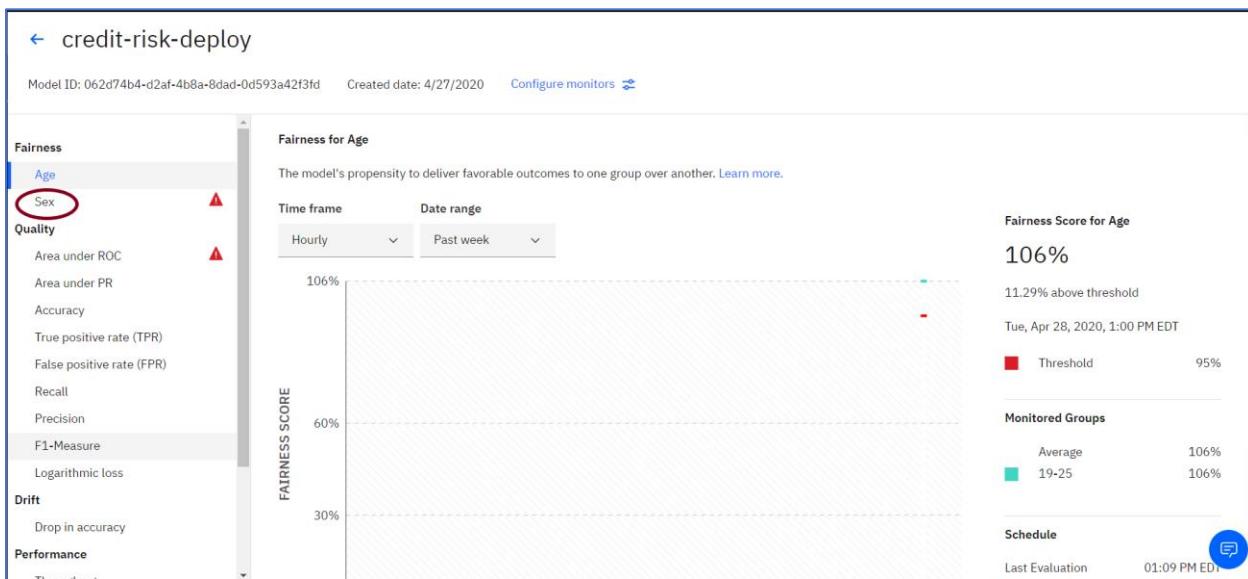
9. Scroll down to see the metric details.



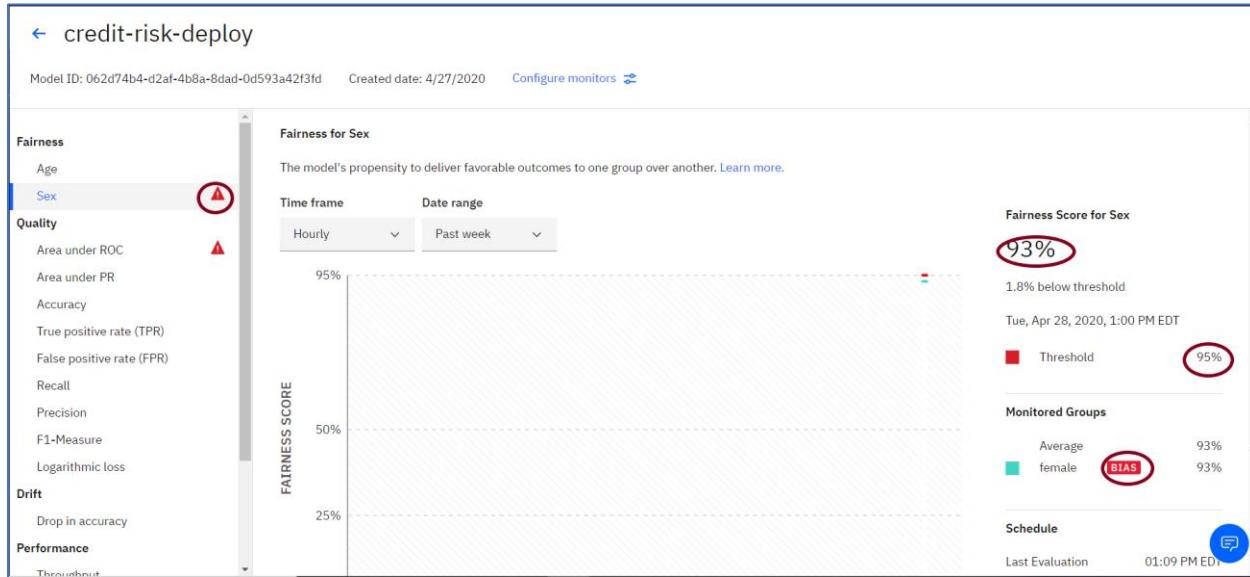
10. Click on the Fairness metric.



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 for Sex

93%

1.8% below threshold

Tue, Apr 28, 2020, 2:00 PM EDT

Threshold: 95%

Monitored Groups

Average	Bias
female	BIAIS

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: 94% **Favorable outcomes:** No Risk **How the fairness score was determined (balanced data set):** 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). [View calculation.](#)

View percentage View count

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 dropdown for 'Monitored attribute' set to 'Sex'. On the right, a dropdown for 'Data Set' is open, showing 'Balanced' (selected) and 'Debiased'. A tooltip for 'Debiased' indicates it's the perfect equality is 81 calculation. Below the dropdowns, the 'Fairness score' is 94% and the outcome is 'No Risk'.

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'. In the top right, there's a 'Date and Time' section and a blue button labeled 'View payload transactions' which is circled in red.

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 the 'Payload Table' section. It has a table of transactions and a bar chart for 'No Risk : Favorable Outcome' and 'Risk : Unfavorable Outcome'. The first transaction in the table has an 'Explain' link next to it, which is circled in red.

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

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

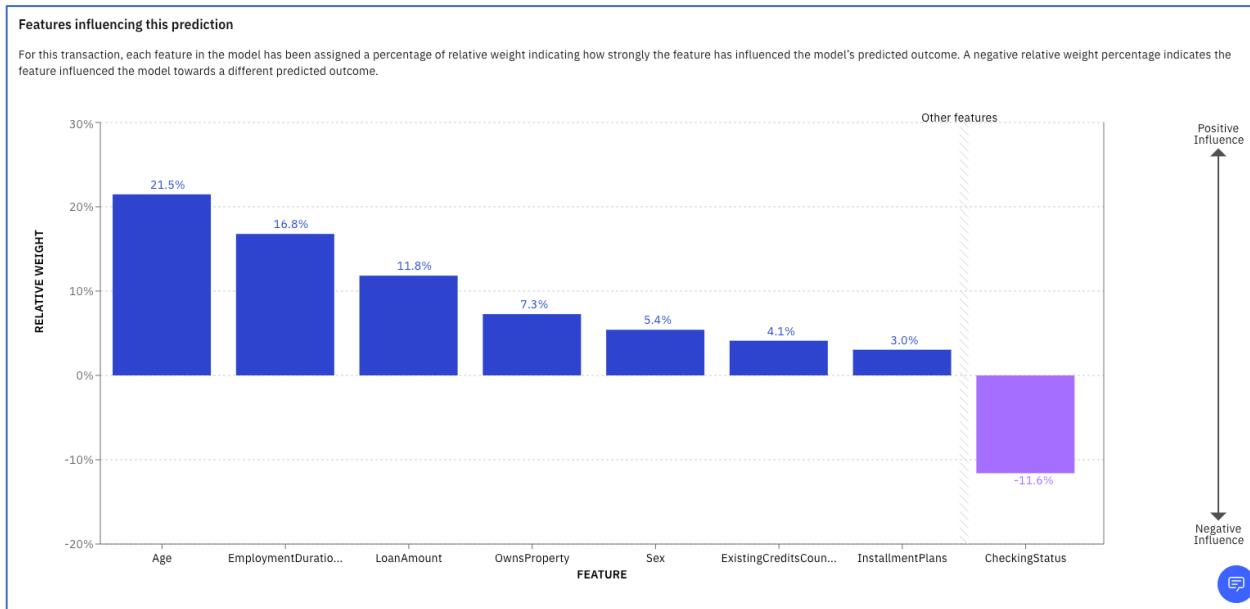
Transaction e6c3f89cd6eeecb79faed2696ea53408e-129	Received on Jul 21, 2020, 03:47:01 PM EDT	Deployed model credit-risk-deploy	Language Not applicable
--	--	--------------------------------------	----------------------------

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 section 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 'Production' status bar at the bottom indicates the data was 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.