

Part 0 : Preparation

Starting by entering the deployment space to the overview page

The screenshot shows the IBM WatsonX interface. At the top, there's a navigation bar with tabs for Overview, Assets, Deployments, Jobs, and Manage. The Overview tab is selected. Below the navigation bar, there's a section titled "Jump back in" with a link to "final-P2_hpo_d_output" (Deployed 6 minutes ago). To the right, there are two summary boxes: "Deployments" (1 Deployed, 0 Failed) and "Space history" (No notifications). Further down, there are sections for "Job runs" (0 Active, 0 Failed last 24 hours) and "AI governance".

Then, go to the deployments tab and click on the deployment which you have deployed earlier

The screenshot shows the "Deployments" tab in the IBM WatsonX interface. The URL in the address bar is "Deployments / lab 4 final 3 / final-P2_hpo_d_output". The main area displays a table of deployments:

Name	Type	Status	Tags	Last modified
Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8	Online	Deployed		1 minute ago Chananop Wannatipaporn (You)

To the right of the table, there are sections for "About this asset" (Name: final-P2_hpo_d_output, Description: No description provided), "Asset Details" (Type: wml-hybrid_01, Model ID: 822a356c-baea-4f..., Software specification: hybrid_0.1), and "Tags" (parent_run_0b6cad9b-f2f5-41fe-9070-2c137c18). At the bottom, there are pagination controls: "Items per page: 20" and "1 of 1 pages".

Once you clicked on the deployment, you should be entering this page

The screenshot shows the deployment details for a model named 'Deployment for model 822a356c-baea-4fcf-a098-12d32ba84ef8'. The status is 'Deployed' and 'Online'. The 'API reference' tab is selected, showing endpoints for inferencing:

- Private endpoint: <https://private.us-south.ml.cloud.ibm.com/v4/deployments/d1b9903c-9284-4c3d-ae19-37e8b9f63bfa/predictions?version=2021-05-01>
- Public endpoint: <https://us-south.ml.cloud.ibm.com/v4/deployments/d1b9903c-9284-4c3d-ae19-37e8b9f63bfa/predictions?version=2021-05-01>

Code snippets section:

- cURL
- Java
- JavaScript
- Python
- Scala

```
# NOTE: you must set $API_KEY below using information retrieved from your IBM Cloud account (https://dataplatform.cloud.ibm.com/docs/content/wsj/analyze-data/ml-authentication.html#api-key)
curl --insecure -X POST -H "Content-Type: application/x-www-form-urlencoded" -H "Accept: application/json" -H "data-urlencode grant_type=urn:ibm:params:oauth:grant-type:apikey" -H "data-urlencode apikey=$API_KEY" "https://iam.cloud.ibm.com/identity/token"

# The above CURL request will return an auth token that you will use as $IAM_TOKEN in the scoring request below
# TODO: manually define and pass values to be scored below
curl -X POST -H "Content-Type: application/json" -H "Accept: application/json" -H "Authorization: Bearer $IAM_TOKEN" -d '[{"input_data": [{"fields": [{"array": "INPUT_FIELDS"}, {"values": [{"array": "VALUES_TO_BE_SCORED"}, {"array": "$ANOTHER_ARRAY_OF_VALUES_TO_BE_SCORED"}]}]}, {"url": "https://private.us-south.ml.cloud.ibm.com/v4/deployments/d1b9903c-9284-4c3d-ae19-37e8b9f63bfa/predictions?version=2021-05-01"}]]'
```

About this deployment:

- Name: Deployment for model 822a356c-baea-4fcf-a098-12d32ba84ef8
- Description: No description provided.
- Deployment Details:
 - Deployment ID: d1b9903c-9284-4c3d-ae19-37e8b9f63bfa
 - Serving name: No serving name.
 - Software specification: hybrid_0.1
 - Hybrid pipeline software specifications: autoai-kb_r24.1-py3.11
 - Copies: 1
- Tags: Add tags to make assets easier to find.
- Associated asset: final-P2_hpo_d_output
- Last modified: 1 minute ago
- Created on: Oct 7, 2024

Part 1 : Configuring model evaluations

Click on Evaluations tab then click on Configure OpenScale evaluation settings

The screenshot shows the deployment details for the same model. The 'Evaluations' tab is selected. The page displays the following content:

Evaluation results
View Watson OpenScale evaluation results to gain insights about your model performance

Start configuring model evaluations
To view evaluation results, configure evaluations to monitor the deployment for fairness or performance drift.

Configure OpenScale evaluation settings

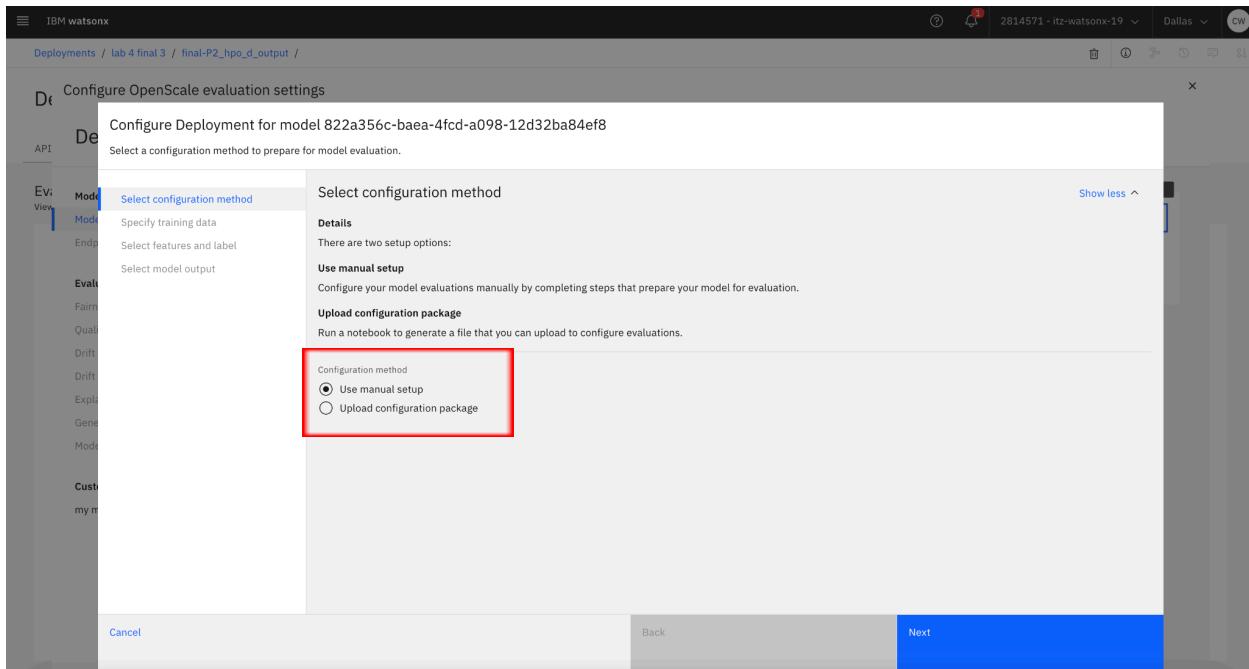
Once you click on the Configure OpenScale evaluation settings, you should be entering this page. Set the data type and algorithm type according to your data asset. When done click view summary.

The screenshot shows the 'Prepare for evaluation' configuration page. It includes sections for 'Provide model information' and 'Details'. A red box highlights the 'Data type' and 'Algorithm type' dropdowns. The 'Data type' dropdown is set to 'Numeric/categorical' and the 'Algorithm type' dropdown is set to 'Binary classification'. At the bottom right, there is a blue button labeled 'View summary'.

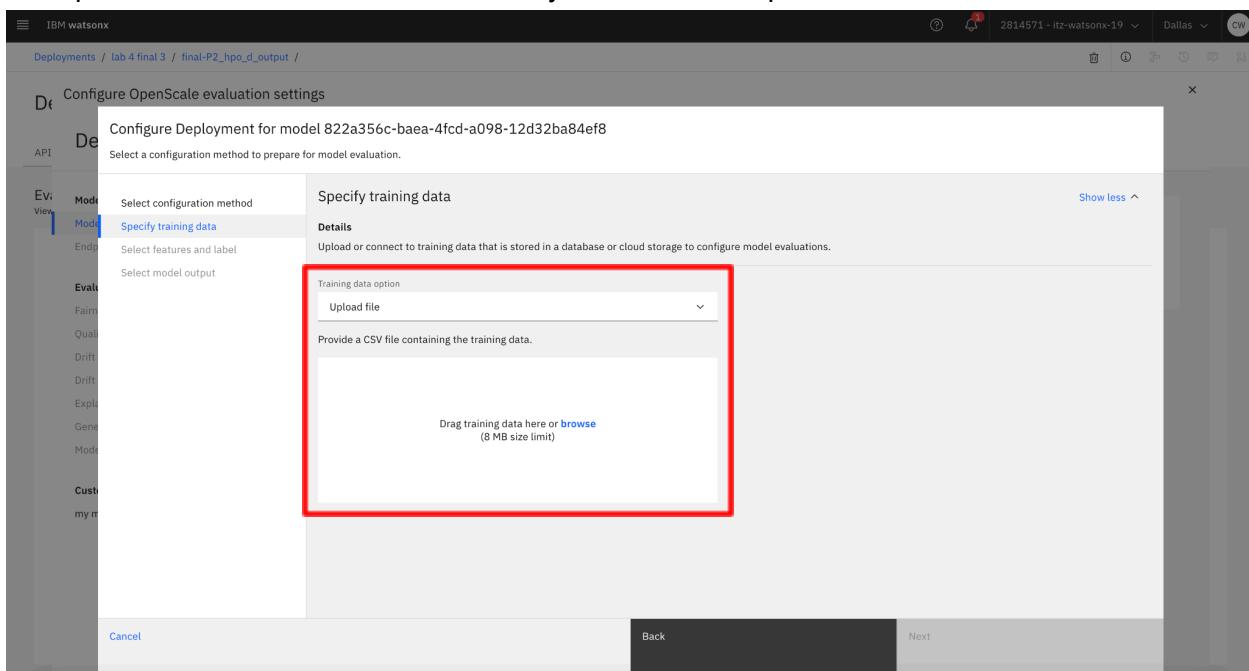
After clicking view summary, check your deployed model, model input data type, algorithm type and feature columns whether it's correct or not. Once it's all correct click save and continue.

The screenshot shows the 'Review setup for Deployment' summary page. It lists four items with radio buttons: 'Deployed model' (selected), 'Model input data type' (set to 'Numeric/categorical'), 'Algorithm type' (set to 'Binary classification'), and 'Feature columns' (list of columns: ApplicantIncome, CoapplicantIncome, Credit_History, Dependents, Education, Gender, Loan_Amount_Term, LoanAmount, LoanAmount_Term, Married, Property_Area, Self_Employed). A red box highlights the 'Save and continue' button at the bottom right.

Once you click save and continue, you should now be at this page. Select your configuration method as use manual setup then click next.



After that, upload your training data. However, the label column of the data file must not include a unique value and need to be named as 'y' as in the example data format below



Data format should have no unique value and required the label column to be named as 'y'

dataset_loan_example_no_loanID_with_Ylabel_withNA												
Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoaapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	y	
Male	No	0	Graduate	No	5849	0		360	1	Urban	Y	
Male	Yes	1	Graduate	No	4583	1508	128	360	1	Rural	N	
Male	Yes	0	Graduate	Yes	3000	0	66	360	1	Urban	Y	
Male	Yes	0	Not Graduate	No	2583	2358	120	360	1	Urban	Y	
Male	No	0	Graduate	No	6000	0	141	360	1	Urban	Y	
Male	Yes	2	Graduate	Yes	5417	4196	267	360	1	Urban	Y	
Male	Yes	0	Not Graduate	No	2333	1516	95	360	1	Urban	Y	
Male	Yes	3	Graduate	No	3036	2504	158	360	0	Semiurban	N	
Male	Yes	2	Graduate	No	4006	1526	168	360	1	Urban	Y	
Male	Yes	1	Graduate	No	12841	10968	349	360	1	Semiurban	N	
Male	Yes	2	Graduate	No	3200	700	70	360	1	Urban	Y	
Male	Yes	2	Graduate		2500	1840	109	360	1	Urban	Y	
Male	Yes	2	Graduate	No	3073	8106	200	360	1	Urban	Y	
Male	No	0	Graduate	No	1853	2840	114	360	1	Rural	N	
Male	Yes	2	Graduate	No	1299	1086	17	120	1	Urban	Y	
Male	No	0	Graduate	No	4950	0	125	360	1	Urban	Y	
Male	No	1	Not Graduate	No	3596	0	100	240		Urban	Y	
Female	No	0	Graduate	No	3510	0	76	360	0	Urban	N	
Male	Yes	0	Not Graduate	No	4887	0	133	360	1	Rural	N	
Male	Yes	0	Graduate		2600	3500	115		1	Urban	Y	
Male	Yes	0	Not Graduate	No	7660	0	104	360	0	Urban	N	
Male	Yes	1	Graduate	No	5955	5625	315	360	1	Urban	Y	
Male	Yes	0	Not Graduate	No	2600	1911	116	360	0	Semiurban	N	
Yes	2	Not Graduate	No		3365	1917	112	360	0	Rural	N	
Male	Yes	1	Graduate		3717	2925	151	360		Semiurban	N	
Male	Yes	0	Graduate	Yes	9560	0	191	360	1	Semiurban	Y	
Male	Yes	0	Graduate	No	2799	2253	122	360	1	Semiurban	Y	

After you are done uploading the data file, select the delimiter according to your file format. Which for .csv is Comma (,). Then click next to continue onto the next page.

W Configure OpenScale evaluation settings

Configure with loan id y label

Select a configuration method to prepare for model evaluation.

Select configuration method

Specify training data

Details

Upload or connect to training data that is stored in a database or cloud storage to configure model evaluations.

Training data option

Upload file

Training data file

dataset_loan_example_with_loanID_Ylabel.csv

Select delimiter

Comma (,)

Cancel Back Next

Once you are done with the previous step, now you should be landing on this page. Select the column type for feature column (categorical/non-categorical)

Configure Deployment for model 94b076c2-b087-4769-9b2c-b2016ff56cb3

Select a configuration method to prepare for model evaluation.

Select features and label

From the training data, select the input features. Indicate if numeric feature values represent categories by checking the Categorical checkbox. Next, select the label column. The label represents the correct prediction (ground-truth) for each record. A column can be marked as a feature or label but not both.

	Features (13)	Type	Categorical	Label / Target
<input checked="" type="checkbox"/>	ApplicantIncome	0.0	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	CoapplicantIncome	0.0	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	Credit_History	0.0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	Dependents	0.0	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	Education	A	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	Gender	A	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Items per page: 25 1 - 13 of 13 items

Cancel Back Next

Then select label column which in this place is 'y' by ticking in this area

Configure Deployment for model 822a356c-baea-4fcda098-12d32ba84ef8

Select a configuration method to prepare for model evaluation.

Select features and label

From the training data, select the input features. Indicate if numeric feature values represent categories by checking the Categorical checkbox. Next, select the label column. The label represents the correct prediction (ground-truth) for each record. A column can be marked as a feature or label but not both.

	Features (13)	Type	Categorical	Label / Target
<input checked="" type="checkbox"/>	ApplicantIncome	0.0	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	CoapplicantIncome	0.0	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	LoanAmount	0.0	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	Loan_Amount_Term	0.0	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	Credit_History	0.0	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	Property_Area	A	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	LoanAmount_Term	0.0	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	y	A	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Items per page: 25 1 - 13 of 13 items

Cancel Back Next

Once you done marking all the features and label column to its type, click next.

Configure Deployment for model 822a356c-baea-4fcf-a098-12d32ba84ef8

Select a configuration method to prepare for model evaluation.

Details

From the training data, select the input features. Indicate if numeric feature values represent categories by checking the Categorical checkbox. Next, select the label column. The label represents the correct prediction (ground-truth) for each record. A column can be marked as a feature or label but not both.

Select the feature columns

Column	Type	Categorical	Label
ApplicantIncome	0.0	<input type="checkbox"/>	<input type="checkbox"/>
CoapplicantIncome	0.0	<input type="checkbox"/>	<input type="checkbox"/>
LoanAmount	0.0	<input type="checkbox"/>	<input type="checkbox"/>
Loan_Amount_Term	0.0	<input type="checkbox"/>	<input type="checkbox"/>
Credit_History	0.0	<input type="checkbox"/>	<input type="checkbox"/>
Property_Area	A	<input checked="" type="checkbox"/>	<input type="checkbox"/>
LoanAmount_Term	0.0	<input type="checkbox"/>	<input type="checkbox"/>
y	A	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Items per page: 25 | 1 - 13 of 13 items | 1 of 1 page | < >

Cancel Back Next

Then continue with selecting the model output by marking prediction as 'Prediction' and probability as 'Probability'.

Configure Deployment for model 822a356c-baea-4fcf-a098-12d32ba84ef8

Select a configuration method to prepare for model evaluation.

Details

From the model output data, select the column that contains the prediction generated by the deployed model. Select the prediction probability column which contains the model's confidence in the prediction.

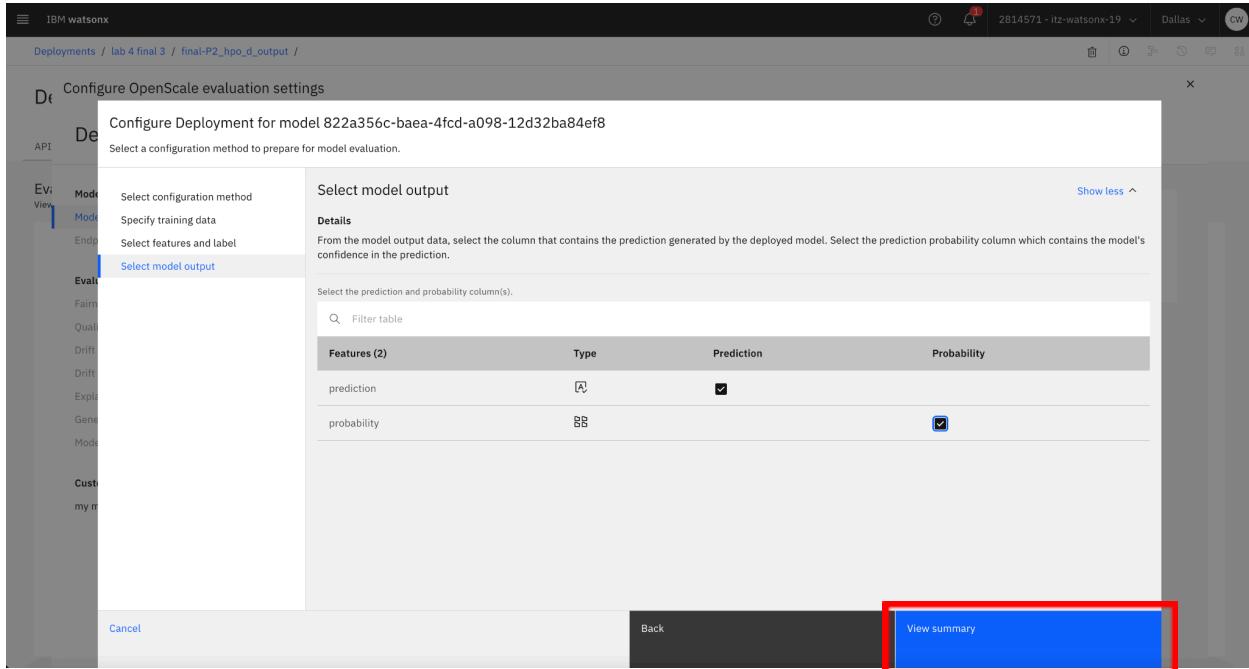
Select the prediction and probability column(s).

Features (2)	Type	Prediction	Probability
prediction	A	<input type="checkbox"/>	<input type="checkbox"/>
probability	B	<input type="checkbox"/>	<input type="checkbox"/>

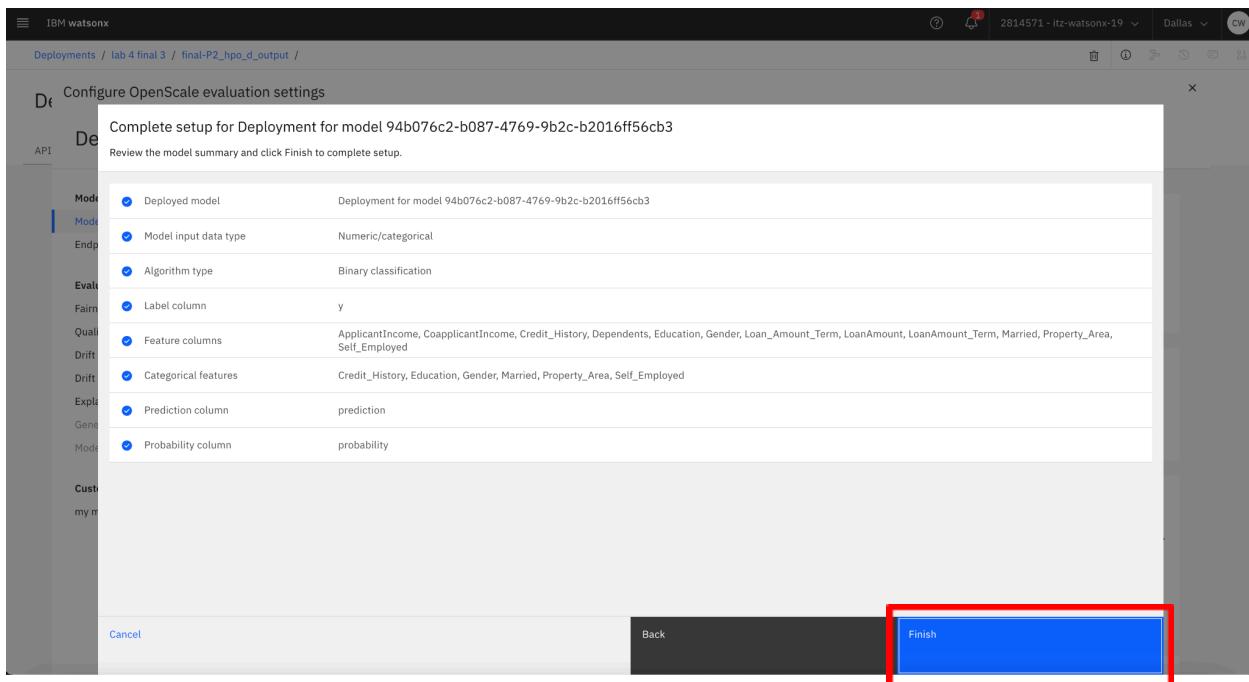
No probability column selected
To support explainability, drift detection, and debiasing for binary classification you must select one or two prediction probability columns

Cancel Back View summary

Once done marking, click ‘View Summary’.



Check all your setup and make sure that what you choose aligns with your dataset. Once done, click finish.



Now you have finished your model setup and should be able to see this page.

The screenshot shows the IBM Watsonx interface with the URL [IBM Watsonx](#) at the top. The main title is "Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8". On the left, there's a sidebar with tabs: "Deployments", "API", "Events", "View", and "Logs". The "Events" tab is selected. The main content area is titled "Model details" under "Model info". It includes sections for "Model details", "Endpoints", "Evaluations" (with options like Fairness, Quality, Drift v2, Drift, Explainability, Generative AI Quality, and Model health), and "Custom" (with my model performance). The "Model details" section has a "Description" field containing placeholder text about training data and deployed model output. To the right, there are three panels: "Training data" (Storage type: Offline (CSV file), Training data file name: dataset_loan_example_with_loanID_Ylabel4 copy.csv), "Training data label" (Label column: y), and "Training data features" (Feature columns: ApplicantIncome, CoapplicantIncome, Credit_History, Dependents, Education, Gender, LoanAmount, LoanAmount_Term, Loan_Amount_Term, Married, Property_Area, Self_Employed; Categorical features: Education, Gender, Married, Property_Area, Self_Employed).

Part 2 : Configuring model explainability

Go to general settings under the explainability topic and you should be able to see this page.
After that click on the edit icon of the explanation method to start configuring the model explanation method.

The screenshot shows the IBM Watsonx interface with the same deployment URL. The "Events" tab is still selected in the sidebar. The main content area is titled "Explainability" under "Evaluations". It includes sections for "Local" (Factors that influence a model outcome of a specific transaction) and "Global" (Holistic factors that influence model outcomes in general). A red box labeled "1." highlights the "Explainability" option in the sidebar. To the right, there are three panels: "Parameters" (To select an explanation method, click the edit icon, with a red box labeled "2." over the edit icon), "Controllable features" (To select controllable features, click the edit icon), and "Language support" (To enable word segmentation for Chinese, Japanese, and Korean languages, click the edit icon).

Then choose ‘LIME (enhanced)’ which is a local explanation as an explanation method and click next to continue onto the next page.

Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8

Model info

- Model details
- Endpoints

Evaluations

- Fairness
- Quality
- Drift v2
- Drift
- Explainability**
- Generative AI Quality
- Model health

Custom

- my model performance

Explainability

Description

OpenScale uses SHAP or LIME (enhanced) to calculate the feature influence and reveals which features contributed to the model's predicted outcome.

Parameters

Description

SHAP assigns each model feature an importance value called a SHAP value. The SHAP values are the estimate of the average marginal contribution of a feature value across all possible groups of features. The sum of input features SHAP values is the difference between average model prediction and the prediction of the transaction being explained. SHAP requires access to the model scoring endpoint and scores the perturbations for each local explanation.

LIME (enhanced) generates explanations by modifying model input and observing the model's response around the vicinity of the transaction and creating a local linear model using the perturbations. LIME (enhanced) explanations for structured data could be generated using pre scored perturbations, which allows scoring the model once per subscription and generate the explanations for all the transactions.

Learn more

Global explanation

Off

Local explanation method

- SHAP (Kernel Explainer)
- LIME (enhanced)

Cancel Next

Set number of perturbations per record, it is a measure used to indicate how frequently each data record has been subject to modifications during the perturbation-based analysis.

Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8

Model info

- Model details
- Endpoints

Evaluations

- Fairness
- Quality
- Drift v2
- Drift
- Explainability**
- Generative AI Quality
- Model health

Custom

- my model performance

Explainability

Description

OpenScale uses SHAP or LIME (enhanced) to calculate the feature influence and reveals which features contributed to the model's predicted outcome.

Parameters

Description

SHAP assigns each model feature an importance value called a SHAP value. The SHAP values are the estimate of the average marginal contribution of a feature value across all possible groups of features. The sum of input features SHAP values is the difference between average model prediction and the prediction of the transaction being explained. SHAP requires access to the model scoring endpoint and scores the perturbations for each local explanation.

LIME (enhanced) generates explanations by modifying model input and observing the model's response around the vicinity of the transaction and creating a local linear model using the perturbations. LIME (enhanced) explanations for structured data could be generated using pre scored perturbations, which allows scoring the model once per subscription and generate the explanations for all the transactions.

Learn more

LIME settings

Number of perturbations per record

1000

Set this to 0 to automatically choose a suitable count

Back Next

Select the features that are controllable.

The screenshot shows the 'Configure OpenScale evaluation settings' dialog box. On the left, there's a sidebar with sections like Model info, Evaluations, Explainability, and Custom. The 'Explainability' section is currently selected. In the main area, there's a 'Controllable features' section with a table. The table has columns for Feature, Type, and Controllable. All rows in the 'Controllable' column have their toggle switches set to 'On'. A red box highlights this column.

Feature	Type	Controllable
Gender	A	On
Married	A	On
Dependents	SS	On
Education	A	On
Self_Employed	A	On
ApplicantIncome	SS	On
CoplicantIncome	SS	On
LoanAmount	SS	On
Loan_Amount_Term	SS	On

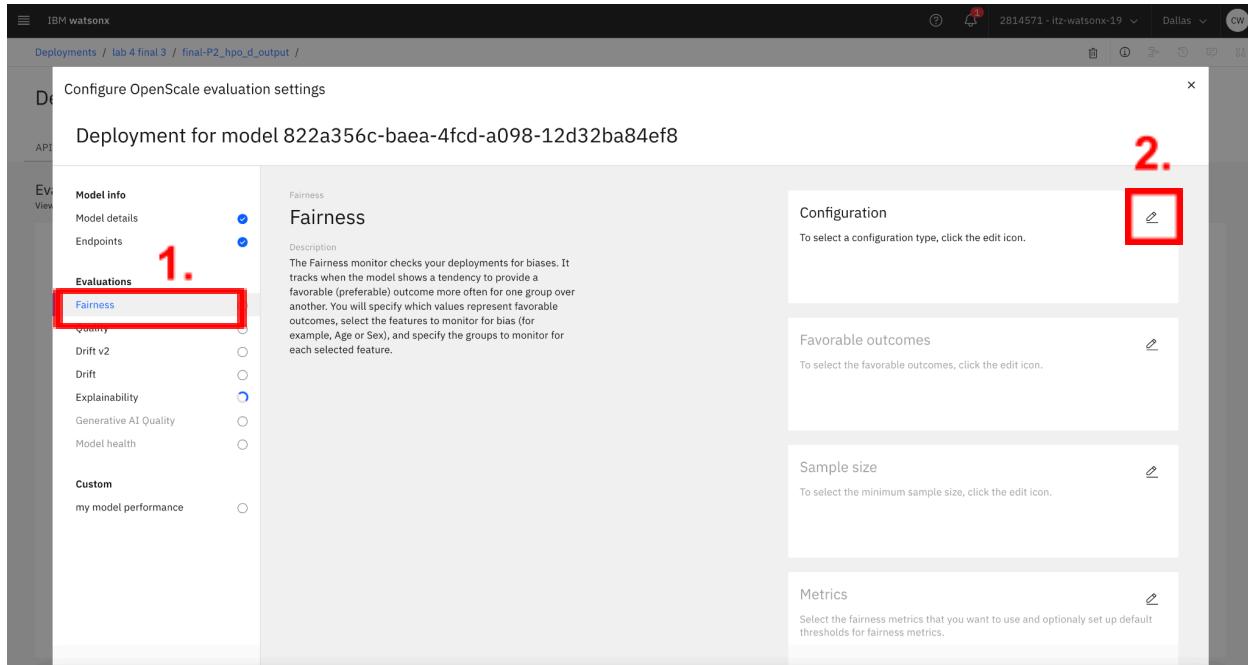
Once done choosing click save

The screenshot shows the same 'Configure OpenScale evaluation settings' dialog box after changes have been saved. The 'Controllable' column in the table now shows all switches set to 'Off'. A red box highlights the 'Save' button at the bottom right of the dialog.

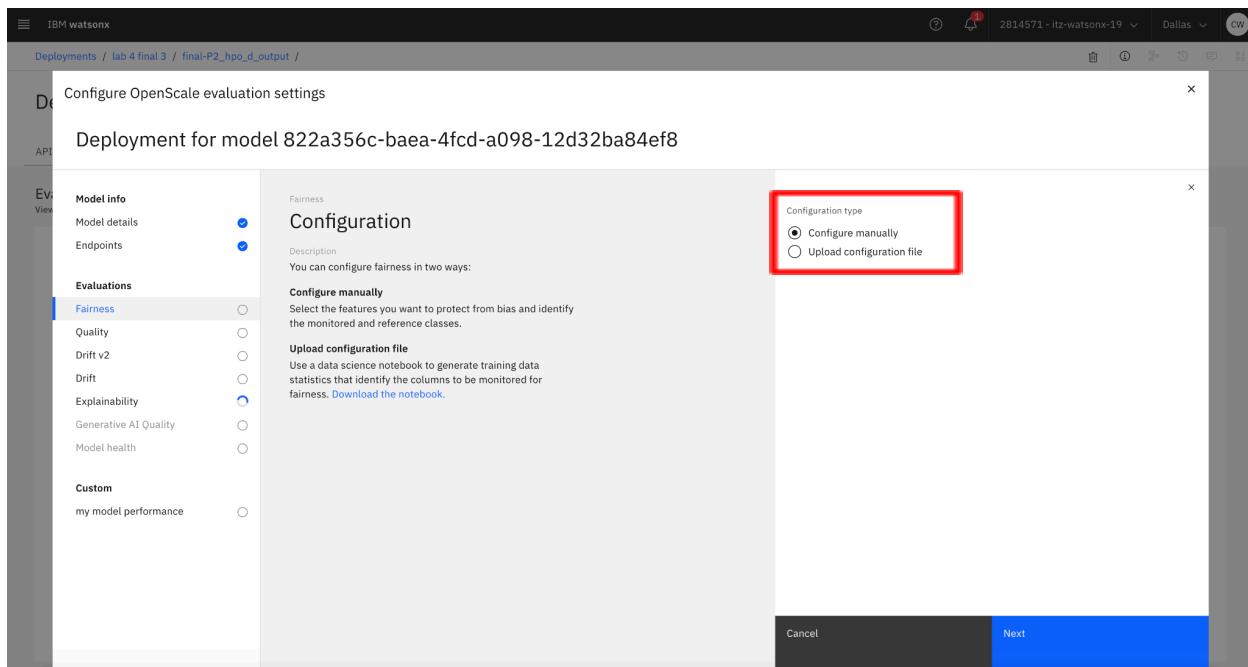
Feature	Type	Controllable
Gender	A	Off
Married	A	Off
Dependents	SS	Off
Education	A	Off
Self_Employed	A	Off
ApplicantIncome	SS	Off
CoplicantIncome	SS	Off
LoanAmount	SS	On
Loan_Amount_Term	SS	On
Credit_History	SS	Off
Property_Area	A	Off

Part 3 : Configuring model fairness monitor

Go to Fairness under the Evaluations topic and you should be able to see this page. After that click on the edit icon of the configuration to start configuring the model fairness monitor.



Then choose configure manually under configuration type and click next to continue onto the next page.



Once you land on this page, choose the value of the outcomes according to the dataset ('Y' is favorable and 'N' is unfavorable)

The screenshot shows the IBM Watsonx interface for configuring OpenScale evaluation settings. On the left, a sidebar lists various evaluation categories: Model info, Endpoints, Evaluations (Fairness, Quality, Drift v2, Drift, Explainability, Generative AI Quality, Model health), and Custom (my model performance). The 'Fairness' tab is selected. The main panel displays a 'Fairness' configuration step titled 'Select the favorable outcomes'. It includes a description: 'For each group, Watson OpenScale will calculate the percentage of transactions that receive a favorable outcome.' Below this is a table for selecting favorable outcomes:

Values	Favorable	Unfavorable
N	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Y	<input checked="" type="checkbox"/>	<input type="checkbox"/>

A red box highlights this table. At the bottom right of the configuration panel are 'Back' and 'Next' buttons.

Once you are done selecting, click next to continue.

This screenshot is identical to the one above, showing the 'Select the favorable outcomes' configuration step. The table for selecting outcomes is shown again, with the 'Y' row having checked boxes in both the 'Favorable' and 'Unfavorable' columns. A red box highlights the 'Next' button at the bottom right of the configuration panel.

Set the minimum sample size to be whatever number which is smaller than the prepared dataset. Once done, click next to continue onto the next page.

The screenshot shows the 'Configure OpenScale evaluation settings' dialog for a deployment. On the left, there's a sidebar with 'Model info' (selected), 'Endpoints', and 'Evaluations' (selected). Under 'Evaluations', 'Fairness' is selected. The main panel shows the 'Fairness' section with a 'Sample size' field. A red box highlights this field, containing 'Minimum sample size' set to '100' and 'Maximum sample size (optional)' set to '10000'. At the bottom right are 'Back' and 'Next' buttons.

Select the metrics which you want to monitor which in this case we only select Disparate impact. Once done, click next to continue.

The screenshot shows the 'Configure OpenScale evaluation settings' dialog for a deployment. The 'Metrics' section is highlighted with a red box. It lists several metrics under 'Metrics generated with payload data, training data, and balanced data': Disparate impact (selected), Statistical parity difference, and Impact score. Below this, under 'Metrics generated with feedback data', are other options: Average odds difference, Average absolute odds difference, False negative rate difference, False positive rate difference, False discovery rate difference, False omission rate difference, and Error rate difference. At the bottom right are 'Back' and 'Next' buttons.

Set the threshold for the metrics we selected previously which in this place is Disparate impact. Once done, click next to continue.

Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8

Fairness

Description

The Fairness monitor tracks multiple fairness metrics. You can select the metrics that will be monitored across the features in the future, and set up some default thresholds. You can always overwrite the default thresholds when setting up the monitored features. Disparate impact is pre-selected as the default fairness metric.

Metric thresholds track when a metric value is outside of an acceptable range. There are two types of thresholds.

Lower threshold - High metric value is better
Upper threshold - Low metric value is better

At least one metric must have a threshold. Metrics without a threshold will not trigger alerts.

Disparate impact

Lower threshold for fairness value: 80
Upper threshold for fairness value: 120

Other metrics

(Difference metrics are not selected.)

Back Next

Select field for tendency monitoring to make sure that model provides output which have acceptable fairness level. Once done, click next to continue.

Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8

Select the fields to monitor

Description

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

If you select a field that is not a feature column, called a added field, Watson OpenScale will look for indirect bias by finding associated values in the feature columns. For example, the profession "student" may imply a younger individual even though the Age field was excluded from model training.

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.

Fields	Recommended	Type
<input checked="" type="checkbox"/> Gender	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Married	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Dependents	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Education	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Self_Employed	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> ApplicantIncome	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> CoapplicantIncome	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> LoanAmount	<input type="checkbox"/>	<input type="checkbox"/>

Items per page: 25 1 - 12 of 12 items 1 of 1 page

Back Next

Select value from the selected field to be monitored or to be a reference which the output value from those two groups will be compared in order to check for potential bias. Once done, click next to continue.

Configure OpenScale evaluation settings

Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8

Evaluations

- Fairness
- Quality
- Drift v2
- Drift
- Explainability
- Generative AI Quality
- Model health

Custom

- my model performance

Fairness

Description

Select the groups to monitor.

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 provide recommendations for group assignments. Less frequently occurring values in the training data may be suggested as monitored groups. Always review these recommended selections before proceeding.

Specify the monitored groups for [Gender]

Select the groups to monitor [Gender]

Add custom value

Enter a value

Add value

Values	Monitored	Reference	Recommended
Female	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Monitored
Male	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Reference

Back Next

Select threshold for the monitored field which in this place we will use the default threshold. Once done, click next to continue.

Configure OpenScale evaluation settings

Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8

Evaluations

- Fairness
- Quality
- Drift v2
- Drift
- Explainability
- Generative AI Quality
- Model health

Custom

- my model performance

Fairness

Description

Select the groups to monitor.

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 provide recommendations for group assignments. Less frequently occurring values in the training data may be suggested as monitored groups. Always review these recommended selections before proceeding.

Set up threshold for feature [Gender]

Use default thresholds

Set up different thresholds

Disparate impact 80 - default

Back Save

Now you have finished your model fairness monitor setup and should be able to see this page.

The screenshot shows the IBM Watsonx interface with the URL `Deployments / lab 4 final 3 / final-P2_hpo_d_output /`. A modal window titled "Configure OpenScale evaluation settings" is open, specifically for the "Fairness" section. The left sidebar lists "Model info", "Endpoints", and "Evaluations" (with "Fairness" selected). The main content area shows a "Fairness" section with a "Description" and a "Configuration" panel containing fields for "Favorable outcomes" (Y) and "Unfavorable outcomes" (N). Other sections like "Sample size" and "Metrics" are also visible.

Part 4 : Configuring model quality monitor

Go to Quality under the Evaluations topic and you should be able to see this page. After that click on the edit icon of the configuration to start configuring the model quality monitor.

The screenshot shows the IBM Watsonx interface with the same deployment URL. A modal window titled "Configure OpenScale evaluation settings" is open, specifically for the "Quality" section. The left sidebar shows "Model info", "Endpoints", and "Evaluations" (with "Quality" selected). The main content area shows a "Quality" section with a "Description" and a "Quality thresholds" panel. The "Quality thresholds" panel includes sections for "Lower thresholds" (with an edit icon highlighted by a red box), "Area under ROC", "Area under PR", "Accuracy", "True positive rate (TPR)", "Recall", "Precision", "F1-Measure", and "Matthews correlation coefficient". A red number "1." is placed over the "Quality" section in the sidebar, and a red number "2." is placed over the edit icon in the "Quality thresholds" panel.

Once you click on the configuration icon, you should enter this page. In this page you can adjust the thresholds for model quality as you prefer but for now, we will leave it as it is. Click next to continue.

Configure OpenScale evaluation settings

Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8

Model info

- Model details
- Endpoints

Evaluations

- Fairness
- Quality**
- Drift v2
- Drift
- Explainability
- Generative AI Quality
- Model health

Custom

- my model performance

Quality

Quality thresholds

Description

Area under ROC

Quality metrics are based on model type. Metric thresholds track when a metric value is outside of an acceptable range. There are two types of thresholds.

- Lower threshold - High metric value is better
- Upper threshold - Low metric value is better

At least one metric must have a threshold. Metrics without a threshold will not trigger alerts.

Lower thresholds

Area under ROC	0.8
Area under PR	0.8
Accuracy	0.8
True positive rate (TPR)	0.8
Recall	0.8
Precision	0.8
F1-Measure	0.8

Cancel Next

Set the minimum sample size to be whatever number which is smaller than the prepared dataset. Once done, click next to continue onto the next page.

Configure OpenScale evaluation settings

Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8

Model info

- Model details
- Endpoints

Evaluations

- Fairness
- Quality**
- Drift v2
- Drift
- Explainability
- Generative AI Quality
- Model health

Custom

- my model performance

Quality

Sample size

Description

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

Minimum sample size

100

Maximum sample size (optional)

1000

Back Save

Now you have finished your model quality monitor setup and should be able to see this page.

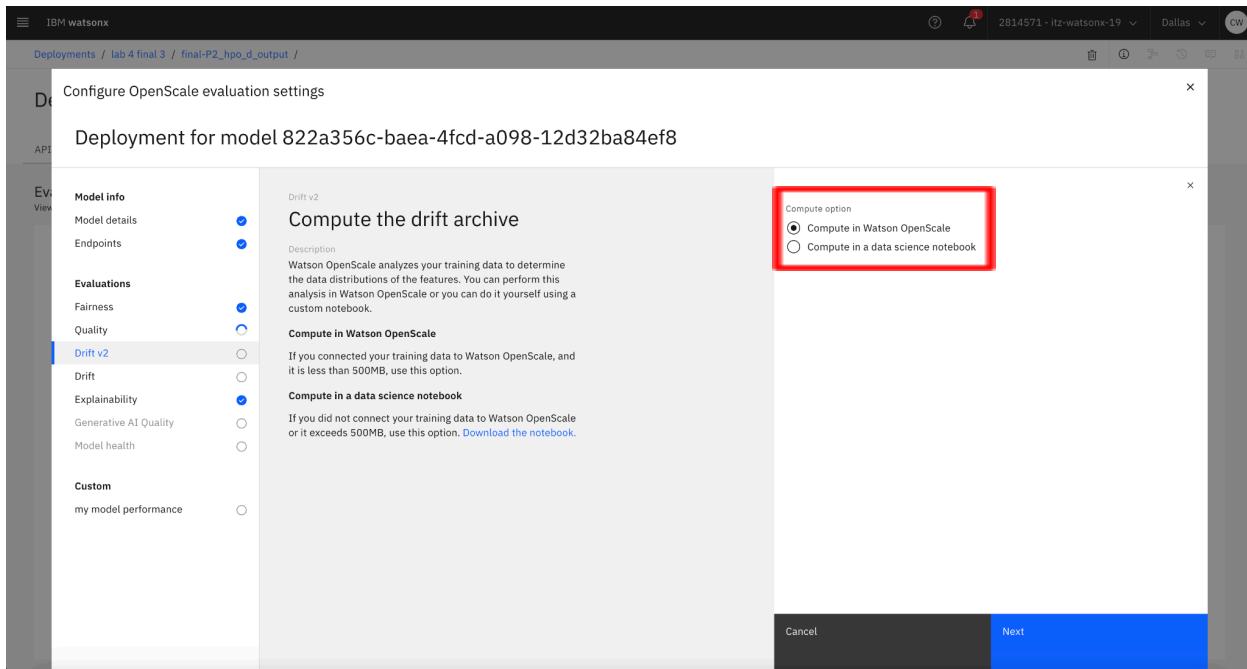
The screenshot shows the IBM Watsonx interface with the URL `Deployments / lab 4 final 3 / final-P2_hpo_d_output /`. The main title is "Configure OpenScale evaluation settings" for "Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8". On the left, there's a sidebar with "Evaluations" selected. Under "Evaluations", "Quality" is highlighted with a blue background. The central panel displays the "Quality" configuration with a detailed description and various quality thresholds like Area under ROC (0.8), Area under PR (0.8), Accuracy (0.8), True positive rate (TPR) (0.8), Recall (0.8), Precision (0.8), F1-Measure (0.8), Matthews correlation coefficient (0.8), and Label skew (-0.5).

Part 5 : Configuring model Drift v2 monitor

Go to Drift v2 under the Evaluations topic and you should be able to see this page. After that click on the edit icon of the configuration to start configuring the model drift v2 monitor.

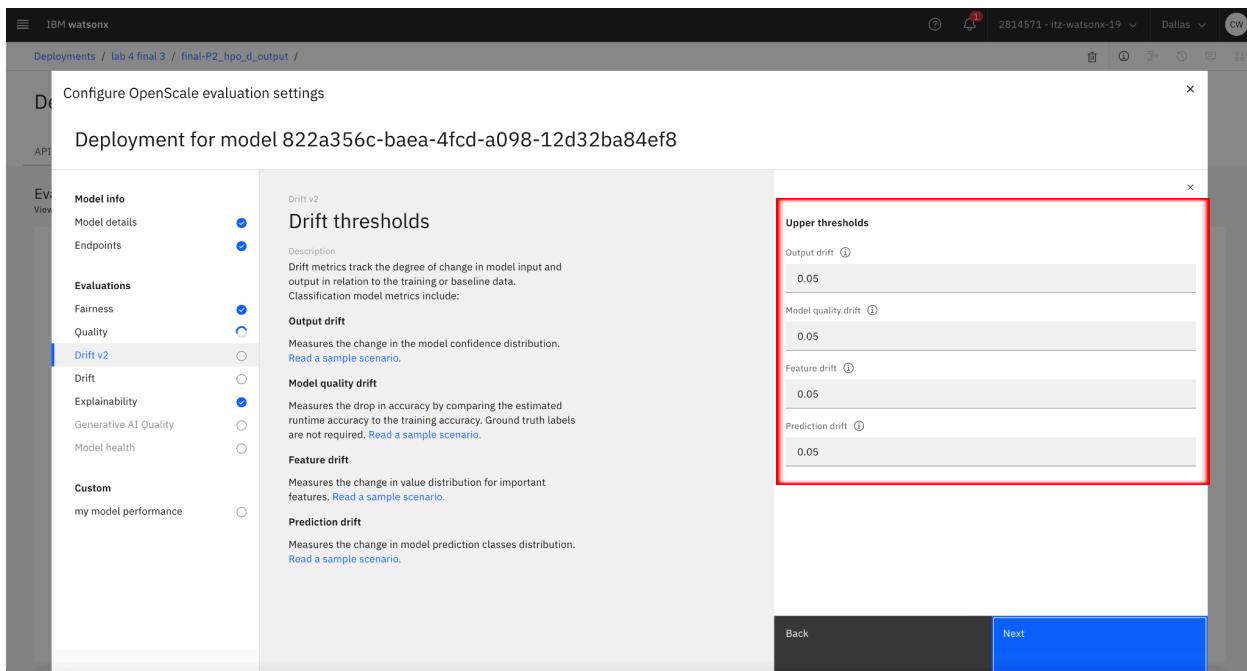
The screenshot shows the IBM Watsonx interface with the same deployment URL. The "Evaluations" section now has "Drift v2" highlighted with a red box and the number "1." above it. The right panel shows the "Drift v2" configuration with sections for "Compute the drift archive" (with an edit icon highlighted with a red box and the number "2."), "Drift thresholds", "Important features", and "Most important features".

Select the compute option to be within Watson OpenScale. Once done, click next to continue.



The screenshot shows the 'Configure OpenScale evaluation settings' dialog for a deployment. On the left, a sidebar lists various evaluations: Model info, Endpoints, Fairness, Quality, Drift v2 (selected), Drift, Explainability, Generative AI Quality, Model health, and Custom. The 'Drift v2' section is expanded, showing 'Compute the drift archive' and two options for computation: 'Compute in Watson OpenScale' (selected) and 'Compute in a data science notebook'. A red box highlights the 'Compute in Watson OpenScale' radio button. At the bottom right are 'Cancel' and 'Next' buttons.

Drift upper thresholds track the degree of change of model input and output from its training data. Set it as you prefer but in this case we will leave it as it is. Click next to continue.



The screenshot shows the 'Configure OpenScale evaluation settings' dialog for a deployment. The sidebar shows 'Model info', 'Endpoints', 'Fairness', 'Quality', 'Drift v2' (selected), 'Drift', 'Explainability', 'Generative AI Quality', 'Model health', and 'Custom'. The 'Drift v2' section is expanded, showing 'Drift thresholds' and four types of drift: 'Output drift' (0.05), 'Model quality drift' (0.05), 'Feature drift' (0.05), and 'Prediction drift' (0.05). A red box highlights the 'Upper thresholds' section. At the bottom right are 'Back' and 'Next' buttons.

Select the important features which impact the model from the list.

The screenshot shows the 'Configure OpenScale evaluation settings' step for a deployment. The 'Drift v2' section is selected. On the right, there's a 'Important features' panel with a 'Select from list' option selected. A table lists 12 features: ApplicantIncome, CoapplicantIncome, Credit_History, Dependents, Education, Gender, LoanAmount, and LoanAmount_Term. Each feature has a checkbox next to it, and the entire list is highlighted with a red box.

Once done click next to continue.

The screenshot shows the same configuration screen after some features have been selected. The 'Credit_History' and 'LoanAmount' checkboxes are now checked. A red box highlights the 'Next' button at the bottom right of the screen.

Select the most important features which have high impact on the model from the list.

The screenshot shows the IBM WatsonX interface for configuring OpenScale evaluation settings. On the left, there's a sidebar with various deployment and evaluation options. The main area is titled 'Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8'. A modal window titled 'Select most important features' is open, listing four features: 'Features (4)' (selected), 'ApplicantIncome', 'Credit_History', 'LoanAmount', and 'LoanAmount_Term'. Each feature has a small icon next to it. At the bottom of the modal are 'Back' and 'Next' buttons. The entire 'Select most important features' dialog is highlighted with a red box.

Set the minimum sample size to be whatever number which is smaller than the prepared dataset. Once done, click save to continue onto the next page.

The screenshot shows the IBM WatsonX interface for configuring OpenScale evaluation settings. The sidebar and main deployment title are identical to the previous screenshot. A modal window titled 'Sample size' is open, containing a description of what the sample size means for drift evaluation. It includes input fields for 'Minimum sample size' (set to 100) and 'Maximum sample size (optional)' (set to 10000). The 'Minimum sample size' field is highlighted with a red box. At the bottom of the modal are 'Back' and 'Save' buttons.

Now you have finished your model drift v2 monitor setup and should be able to see this page.

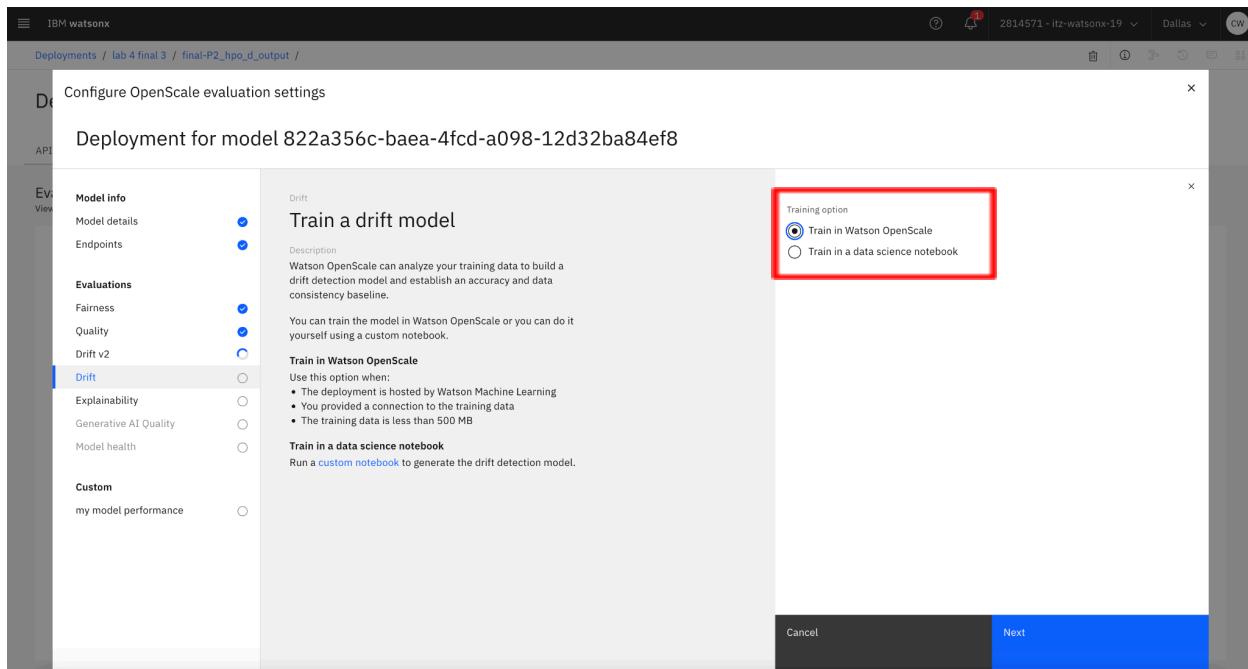
The screenshot shows the IBM Watsonx interface with the URL [Deployments / lab 4 final 3 / final-P2_hpo_d_output /](#). On the left, a sidebar lists various evaluation topics: Model info, Endpoints, Evaluations (Fairness, Quality, Drift v2), Drift, Explainability, Generative AI Quality, Model health, and Custom (my model performance). The 'Drift v2' topic is selected and highlighted with a blue border. The main content area is titled 'Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8'. It contains sections for 'Drift v2' (Description: 'The Drift monitor checks if your deployments are up-to-date and behaving consistently. Model input/output data is analyzed in relation to the training/baseline data.'), 'Compute the drift archive' (Status: 'Monitor is being prepared'), 'Drift thresholds' (Upper thresholds: Output drift 0.05, Model quality drift 0.05, Feature drift 0.05, Prediction drift 0.05), and 'Important features' (Selected features: 4).

Part 6 : Configuring model Drift monitor

Go to Drift under the Evaluations topic and you should be able to see this page. After that click on the edit icon of the configuration to start configuring the model drift monitor.

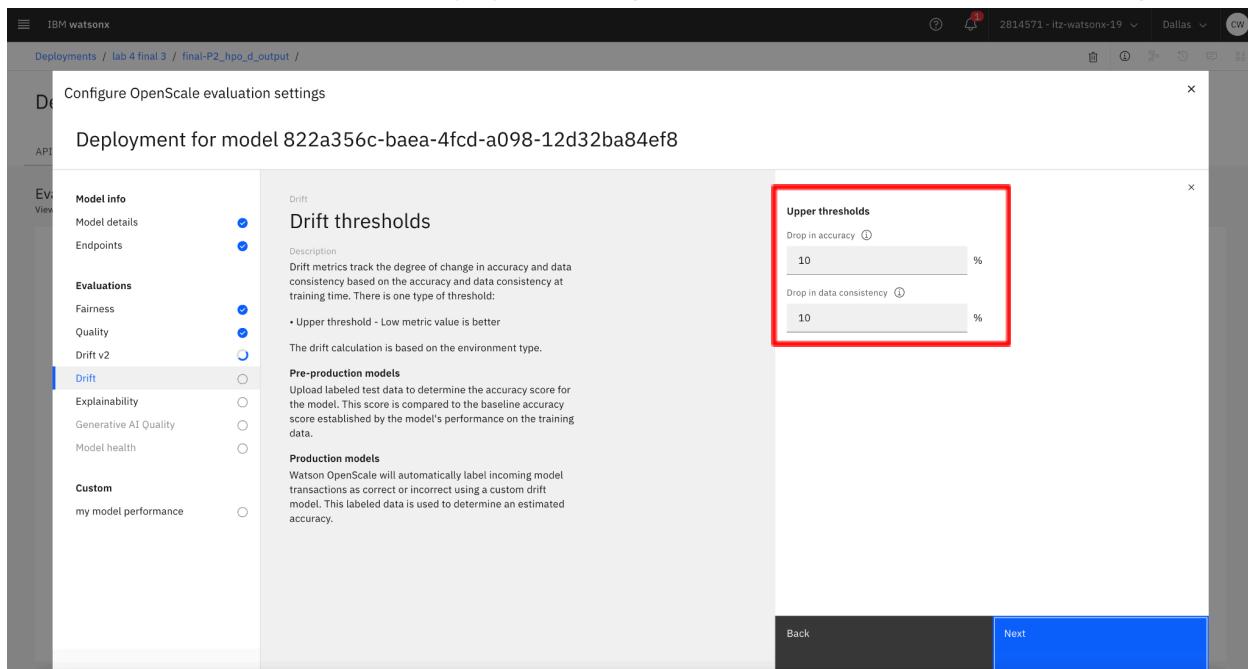
The screenshot shows the IBM Watsonx interface with the same URL as the previous screenshot. The 'Drift v2' topic is selected (marked with a red box labeled '1.'). The main content area shows the 'Drift' configuration page. It includes sections for 'Drift' (Description: 'The drift test measures two types of changes. Drop in accuracy (structured binary and multi-class classification models only) Measures the drop in accuracy from a base accuracy score determined by the training data. Drop in data consistency Measures the drop in data consistency by comparing recent model transactions to the training data.'), 'Drift model' (To select a drift model training option, click the edit icon, marked with a red box labeled '2.'), 'Drift thresholds' (Upper thresholds: To select drift threshold value, click the edit icon.), and 'Sample size' (To select the sample size, click the edit icon.).

Select the training option to be within Watson OpenScale. Once done, click next to continue.



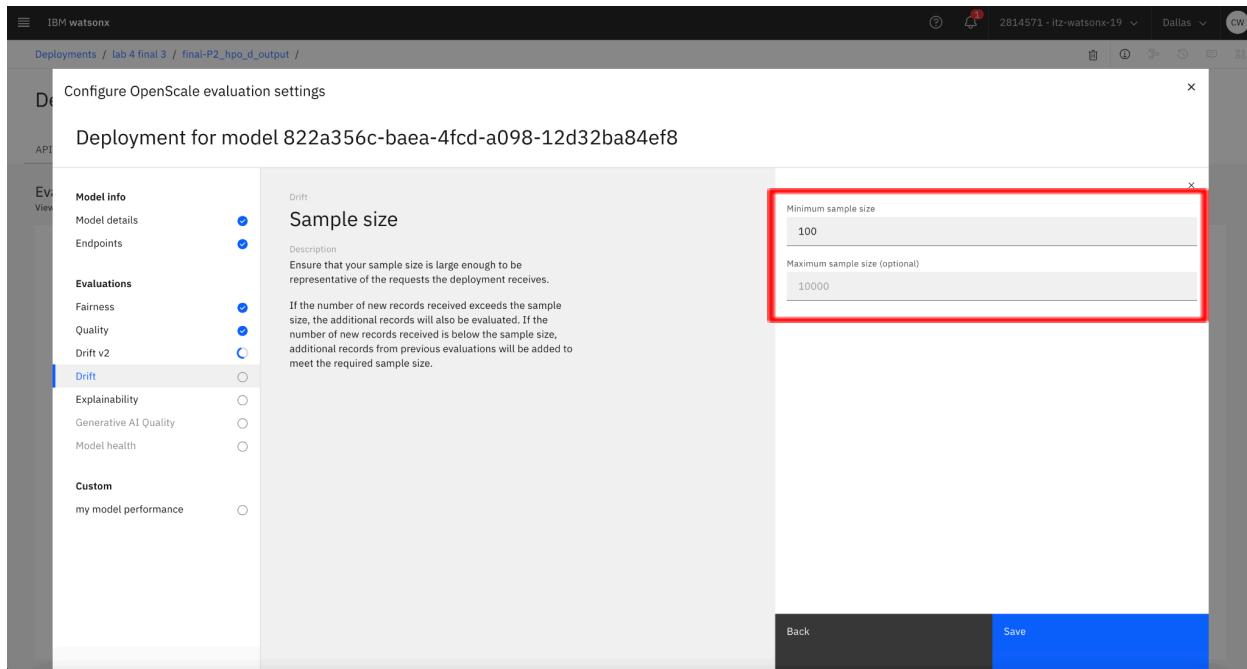
The screenshot shows the 'Configure OpenScale evaluation settings' dialog for a deployment. On the left, a sidebar lists various evaluation types: Model info, Endpoints, Evaluations (Fairness, Quality, Drift v2, Drift, Explainability, Generative AI Quality, Model health), and Custom (my model performance). The 'Drift' tab is selected. The main panel shows the 'Train a drift model' configuration. Under 'Description', it says: 'Watson OpenScale can analyze your training data to build a drift detection model and establish an accuracy and data consistency baseline.' Below this, there are two options: 'Train in Watson OpenScale' (selected) and 'Train in a data science notebook'. At the bottom right are 'Cancel' and 'Next' buttons. A red box highlights the 'Training option' section.

Setup drift upper thresholds which “drop in accuracy” means to measure the drop in accuracy from a base accuracy score determined by the training data and “drop in data consistency” means to measure the drop in data consistency by comparing recent model transactions to the training data.

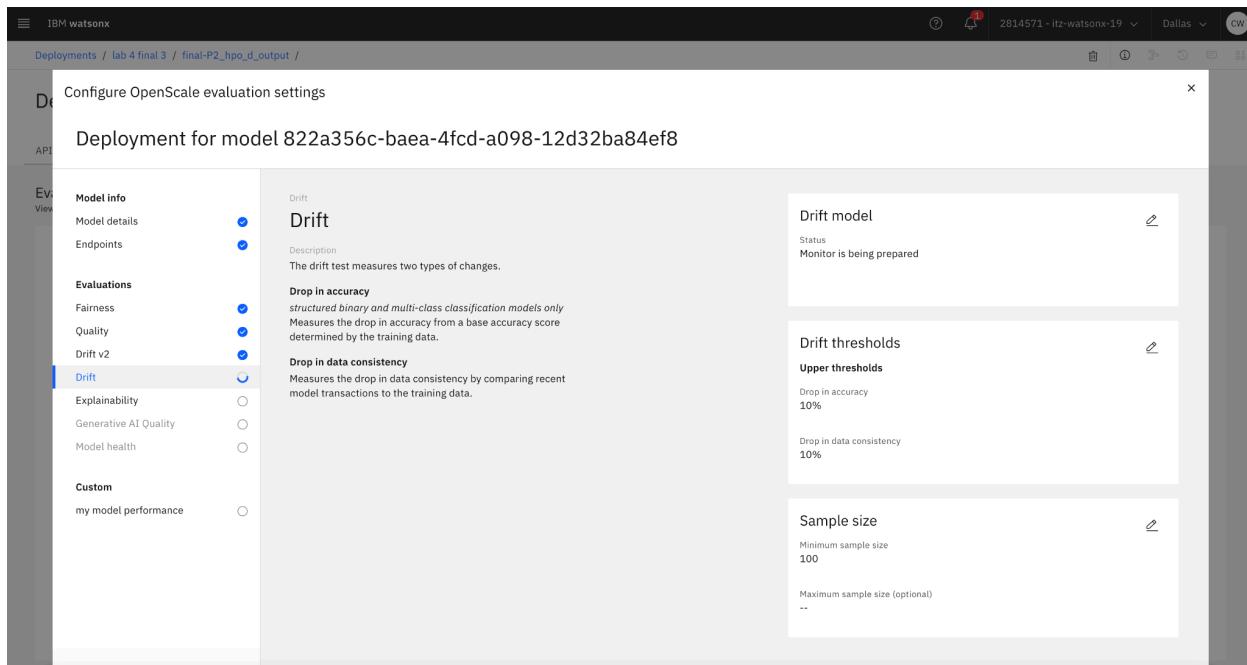


The screenshot shows the 'Configure OpenScale evaluation settings' dialog for a deployment. The 'Drift' tab is selected in the sidebar. The main panel shows the 'Drift thresholds' configuration. Under 'Description', it says: 'Drift metrics track the degree of change in accuracy and data consistency based on the accuracy and data consistency at training time. There is one type of threshold: Upper threshold - Low metric value is better'. Below this, there are two input fields: 'Drop in accuracy' (set to 10%) and 'Drop in data consistency' (set to 10%). At the bottom right are 'Back' and 'Next' buttons. A red box highlights the 'Upper thresholds' section.

Set the minimum sample size to be whatever number which is smaller than the prepared dataset. Once done, click save to continue onto the next page.



Now you have finished your model drift monitor setup and should be able to see this page.



Part 7 : Start evaluation

Once you are done configuring the evaluation monitor you should be on this page. Click on evaluate now under actions dropdown.

Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8 Deployed Online

Last evaluation: --

Test data set
--

Number of explanations
0

0 Tests run

Tests passed
0

Tests failed
0

Actions

(i) Awaiting test data to evaluate.
Select the Actions button to evaluate now.

Once you click evaluate now, click import from CSV file to import the test data using .csv file.

Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8 Deployed

Last evaluation: --

Test data set
--

Number of explanations
0

0 Tests run

Tests passed
0

Tests failed
0

(i) Awaiting test data to evaluate.
Select the Actions button to evaluate now.

Import test data

Description

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

Import test data by uploading a CSV file containing labeled test data or connecting to a database table. watsonx.governance will evaluate the test data for fairness, quality, and drift according to the configuration settings.

Import

Choose an option

from database or cloud storage

from CSV file

Close Upload and evaluate

upload the test data into the red area.

The screenshot shows the IBM Watsonx interface. On the left, the 'Evaluations' tab is selected. The main panel displays 'Last evaluation: --'. Below it, there's a section for 'Test data set' which is currently empty. To the right, two circular progress indicators show 'Tests run' at 0 (green) and 'Tests' at 0 (red). A tooltip indicates 'Awaiting test data to evaluate. Select the Actions button to evaluate now.' On the right side, a modal window titled 'Import test data' is open. It contains a 'Description' section with details about watsonx.governance. The 'Import' section has a dropdown menu set to 'from CSV file'. Below it is a large red-bordered area containing a 'Drag test data file here or browse' input field and a checkbox for 'Test data includes model output'. At the bottom of the modal are 'Close' and 'Upload and evaluate' buttons.

Once you upload the data, your file name should appear on the test data set. Click upload and evaluate to evaluate.

This screenshot is similar to the previous one, showing the IBM Watsonx interface with the 'Evaluations' tab selected. The 'Test data set' section now lists a file named 'dataset_loan_example_with_loanID_Ylabel4 copy.csv'. The circular progress indicators remain at 0. The 'Import test data' modal is still open, showing the same 'from CSV file' dropdown and the file name in the input field. The 'Upload and evaluate' button at the bottom of the modal is highlighted with a blue background.

Part 8 : Evaluation result

Once done with evaluate, you should be able to get the evaluation result as shown below. You can click on the arrow button inside each metrics monitor to find further information for that monitor type.

The screenshot shows the IBM Watsonx AI Model Evaluation interface. The top navigation bar includes 'IBM watsonx', deployment details ('2814571 - itz-watsonx-19'), location ('Dallas'), and a 'CW' button. The main page title is 'Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8' with status 'Deployed' and 'Online'.

The 'Evaluations' tab is selected. The interface displays the following sections:

- Last evaluation:** Mon, Oct 7, 2024, 1:10 PM +07
- Test data set:** dataset_loan_example_with_loanID_Ylabel4 copy.csv
- Number of explanations:** 2
- Tests run:** 4 (1 passed, 3 failed)
- Fairness:** No alerts (0 total, 0 alerts for Gender).
- Quality:** Alerts triggered (1). Metrics table:

Metric	Score	Violation
Disparate impact	100.00%	none
True positive rate (TPR)	0.98	none
Area under ROC	0.94	none
- Drift v2:** Drift computation failed. Argument 'value' has incorrect type (expected int, got numpy.int64)
- Associated group:** Female (84 records evaluated)
- Drift:** Alerts triggered (1). Metrics table:

Metric	Score	Violation
Drop in data consistency	100.00%	90.00%
Drop in accuracy	1.63%	none
Evaluated accuracy	96.29%	--
Base accuracy	97.92%	--
- AI Factsheet:** Area under PR (0.95), Recall (0.98), Brier score (0.21). (566 records evaluated)

Here is the example of further information for each evaluation metrics.

IBM Watsonx

Deployments / lab 4 final 3 / final-P2_hpo_d_output /

Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8 Deployed Online

Evaluations Fairness Test Transactions AI Factsheet

Fairness Quality Drift Back to summary

Fairness

Monitored attribute: Gender, Fairness metric: Disparate impact, Data set: Balanced, Date and time: 10/7/2024 1:11 PM

Disparate impact: 100% **Favorable outcomes:** Y

How the disparate impact score was determined (balanced data set):
The monitored group Female received favorable outcomes 67.857% of the time. The perfect equality is 67.857%. The fairness score for Gender is 100.0% (67.857/67.857). [View calculation](#).

View percentage View count

Gender	% favorable outcomes
Female	~67.857%
Male	~67.857%

IBM Watsonx

Deployments / lab 4 final 3 / final-P2_hpo_d_output /

Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8 Deployed Online

Evaluations Quality Test Drift Back to summary

Quality

Quality monitor calculates a list of metrics to compare an automatically produced classification against a reference (human-produced) classification.

Mon, Oct 7, 2024, 1:10 PM +07

Area under ROC	Area under PR	Accuracy	True positive rate (TPR)	False positive rate (FPR)	Recall	Precision	F1-Measure	Logarithmic loss	Brier score	Matthews correlation coefficient	Label skew
0.94	0.95	0.96	0.98	0.11	0.98	0.96	0.97	0.32	0.21	0.9	-0.88 <small>0.38 lower threshold violation</small>

Confusion matrix

Prediction		Predicted			
Actual		N		Y	
N		151		18	
Y		6		391	
				Percent correct	
				89.3% 98.5%	

IBM Watsonx

Deployments / lab 4 final 3 / final-P2_hpo_d_output /

Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8 Deployed Online

API reference Test Evaluations Transactions AI Factsheet

Fairness Quality Drift Back to summary

Drift

View the transactions with incorrect predictions, inconsistent data, or both.

10/7/2024 1:10 PM

Incorrect predictions (21) **Inconsistent data (566)** **Both (21)**

Transactions with incorrect predictions

Drop in accuracy: 1.6%

Grouped by	Dominant values	Grouped by	Dominant values	Grouped by	Dominant values
Credit_History	1	Dependents	0	Dependents	3
Loan_Amount_Term	378.46	CoapplicantIncome	0	Loan_Amount_Term	360
LoanAmount_Term	378.46	Loan_Amount_Term	360	LoanAmount_Term	360
Self_Employed	No	Education	Graduate	Education	Graduate
Gender	Male	Gender	Male	Gender	Male

For model explainability, click on transactions tab in the red area

IBM Watsonx

Deployments / lab 4 final 3 / final-P2_hpo_d_output /

Deployment for model 822a356c-baea-4fcd-a098-12d32ba84ef8 Deployed Online

API reference Test Evaluations Transactions AI Factsheet

Last evaluation: Mon, Oct 7, 2024, 1:10 PM +07

Test data set: dataset_loan_example_with_loanID_Ylabel4 copy.csv

Number of explanations: 2

Actions

Tests passed: 1

Tests failed: 3

Fairness No alerts

Total alerts: 0 Alerts for Gender: 0

Metric	Score	Violation
Disparate impact	100.00%	none

Quality Alerts triggered

Alerts: 1

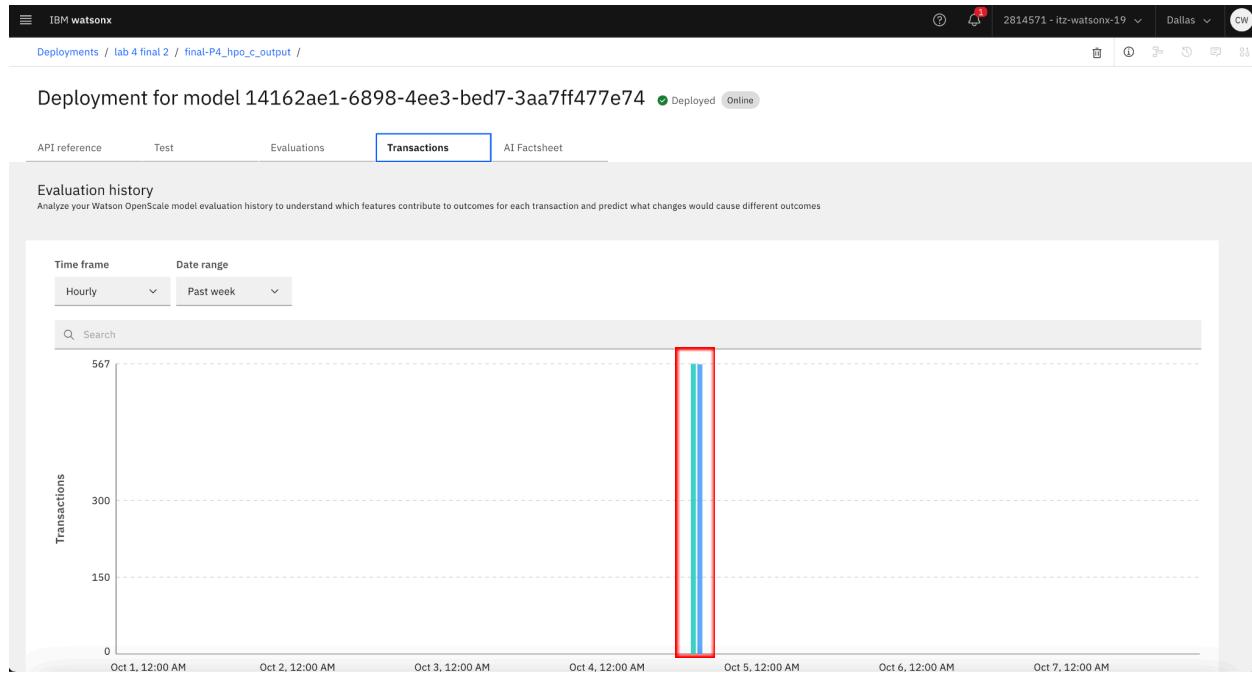
Metric	Score	Violation
True positive rate (TPR)	0.98	none
Area under ROC	0.94	none

Drift v2

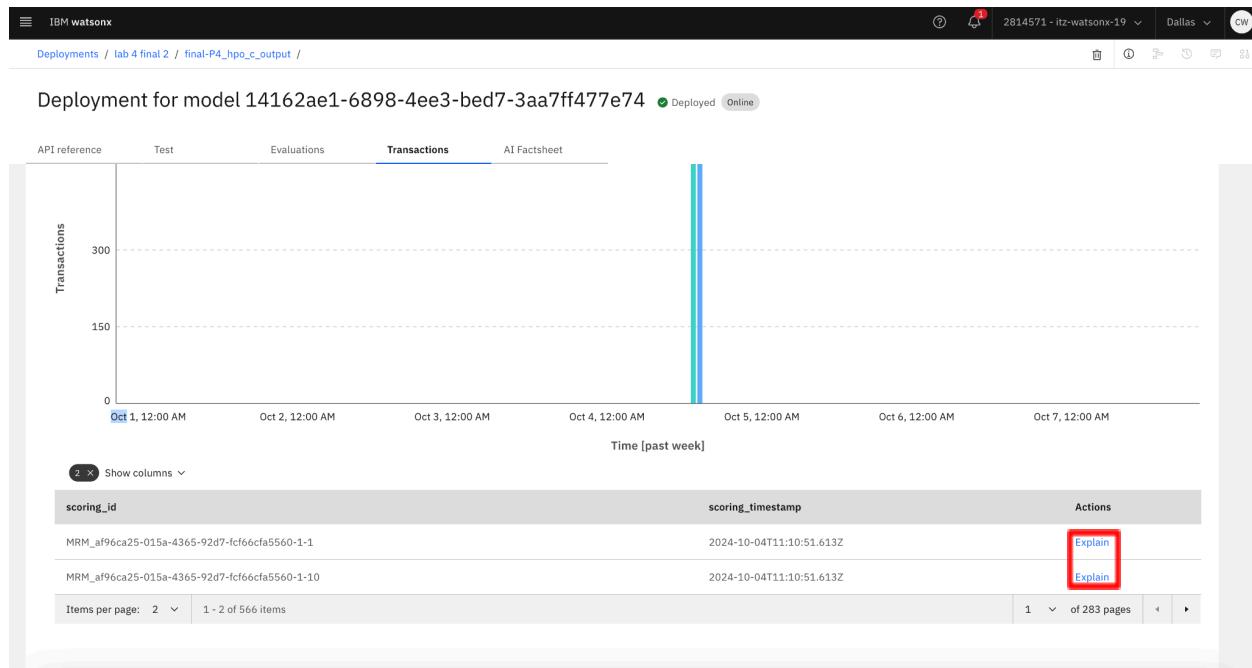
Drift v2 evaluation failed.

Drift computation failed. Argument 'value' has incorrect type (expected int, got numpy.int64)

Once you click on transactions you should see this page. Then, click on the bar chart to view its detail



After you click on the bar chart the list underneath should appear. Click explain to view model explainability.



Now you are able to observe the model explainability through feature importance. If you want to see how the change of the feature affects the model you can use the 'inspect' tab to observe.

TRANSACTION
Transaction details Online

Explain **Inspect** Inspect

Feature influence analyzed with LIME

LIME (enhanced) generates explanations by modifying model input and observing the model's response. A relative weight indicates the feature's influence on the model's outcome. Features with a negative relative weight influence the model towards a different prediction. [Learn more](#)

The top three features influencing the model's predicted outcome are **ApplicantIncome**, **Property_Area**, and **Gender**. The top three features **LoanAmount**, **CoapplicantIncome**, and **Loan_Amount_Term** are influencing the model toward a predicted outcome of Y.

Relative weight

Positive influence ↑

Feature	Relative weight
ApplicantIncome	28.1%
Property_Area	15.5%
Gender	2.9%
CoapplicantIncome	2.5%
LoanAmount	1.1%
Other features	
Dependents	-6.5%
Education	-8.5%
Self_Employed	-12.0%

Predicted outcome: N
Confidence level: 76.08%

How this prediction was determined:
The final-P4_hpo_c_output model has 76.08% confidence that the outcome of this transaction would be N. See the LIME analysis on the left to analyze the top features influencing the model's prediction.

Deployed model:
Deployment for model 14162ae1-6898-4ee3-bed7-3aa7ff477e74
Pre-production

Transaction:
MRM_af96ca25-015a-4365-92d7-fc166cfa5560-1-1

Received on:
Oct 07, 2024, 09:54:17 AM +07

Language:
Not applicable

Once you click inspect, you should be entering this page. In this page you can config the feature values to predict a different outcome by adjusting the feature value in the red area.

Explain **Inspect** Inspect

Reaching a different predicted outcome

For the model to have predicted a different outcome for this transaction, the value of all listed features would need to change to the indicated minimum value. Note that changing a feature value by more than the minimum value may affect the minimum change of other features for the model to predict a different outcome. Higher feature importance numbers indicate a greater likelihood of changing the prediction.

Analyze controllable features only

Feature	Original value	New value	Value for a different outcome	Importance
Gender	Male	Male		
Married	Yes	Yes		
Dependents	1	1	①	
Education	Graduate	Graduate		
Self_Employed	No	No	▼	
ApplicantIncome	4583	4583	①	
CoapplicantIncome	1508	1508	①	
LoanAmount	128	128	①	
Predicted outcome	N	N	Confidence 79.34%	
Confidence	79.34%		Predicted outcome	N
Confidence	79.34%		Confidence	79.34%

Want to know what feature values will result in a different outcome? Analyze the model to get a suggestion. This analysis will take a few minutes.

Run analysis

Once you are done configuring the value, click “Score new values” and then click “Run analysis”.

For the model to have predicted a different outcome for this transaction, the value of all listed features would need to change to the indicated minimum value. Note that changing a feature value by more than the minimum value may affect the minimum change of other features for the model to predict a different outcome. Higher feature importance numbers indicate a greater likelihood of changing the prediction.

Analyze controllable features only ⓘ

Feature	Original value	New value	Value for a different outcome ⓘ	Importance
Gender	Male	Male		
Married	Yes	Yes		
Dependents	1	1 ⓘ		
Education	Graduate	Graduate		
Self_Employed	No	Yes ⓘ		
ApplicantIncome	4583	4583 ⓘ		
CoapplicantIncome	1508	1508 ⓘ		
LoanAmount	128	128 ⓘ		
Predicted outcome N	Confidence 79.34%	Score new values	Run analysis	

After you click run analysis, you should be able to see the outcome in the red area.

Want to know what feature values will result in a different outcome? Analyze the model to get a suggestion. This analysis will take a few minutes.

Value for a different outcome ⓘ	Importance
4666.5	1.00
Male	0.00
Yes	0.00
1	0.00
Graduate	0.00
No	0.00
1508	0.00
128	0.00
Predicted outcome Y	Confidence 80.24%
N	Confidence 79.34%

Furthermore, you can reconfigure the evaluation metrics monitor by clicking on the configure button in the red area.

Last evaluation: Wed, Oct 2, 2024, 9:51 AM +07

Test data set
dataset_loan_example_with_loanID_Ylabel.csv

Number of explanations
0

Fairness: 4 Tests run (No alerts)

Quality: 5 Alerts triggered (Alerts triggered)

Drift v2: Drift v2 evaluation failed. Drift computation failed. Argument 'value' has incorrect type (expected int, got numpy.int64)

Metric	Score	Violation
Disparate impact	98.15%	none

Which once you clicked on the reconfigure button in the red area. Now you can reconfigure OpenScale evaluation settings.

Configure OpenScale evaluation settings

with loan id y label 3

Model info

- Model details
- Endpoints

Explainability

- General settings
- SHAP
- LIME (enhanced)

Evaluations

- Fairness** (selected)
- Quality
- Drift v2
- Drift
- Generative AI Quality
- Model health

Custom

- my model performance

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.

Configuration

Description: Configuration file

Favorable outcomes

Favorable outcomes: Y

Unfavorable outcomes: N

Sample size

Minimum sample size: 100

Maximum sample size (optional): --

Metrics