**ModelOps - Model Monitoring on Cloud Pak for Data**

Eric Martens, Digital Technical Engagement, Data Science and AI

emartens@us.ibm.com

[emartens@us.ibm.com](mailto:emartens@us.ibm.com)

A close up of a logo

Description automatically generated

**Table of contents**

Contents

[Overview 1](#_Toc59970999)

[Required software, access, and files 1](#_Toc59971000)

[Required skills 4](#_Toc59971001)

[ModelOps Overview 5](#_Toc59971002)

[Feature Parity and Roadmap 6](#_Toc59971003)

[Deployment Options 7](#_Toc59971004)

[IBM Cloud Security and WML API 8](#_Toc59971005)

[API Key 8](#_Toc59971006)

[Cloud Location URL 9](#_Toc59971007)

[Watson Machine Learning Client API 9](#_Toc59971008)

[Manual Deployment to Deployment Spaces 15](#_Toc59971009)

[Programmatic Deployment to Deployment Spaces 27](#_Toc59971010)

[Optional: Additional practice 29](#_Toc59971011)

[Manual Deployment 29](#_Toc59971012)

[Deployment with API 29](#_Toc59971013)

[Deployment of Notebooks as a Job 30](#_Toc59971014)

[Deployment of Notebooks as a URL 34](#_Toc59971015)

[Conclusion and Next Steps 37](#_Toc59971016)

# Overview

In this lab you will learn how to complete the following ModelOps tasks in **Cloud Pak for Data:**

* Deployment of competing models
* Configuring OpenScale monitoring
* Evaluating and comparing models with test data
* Approving models for production

# Required software, access, and files

1. To complete this lab, you will need a **Cloud Pak for Data as a Service** (**CPDaaS**) account: <https://dataplatform.cloud.ibm.com>

* If you don’t have a CPDaaS account, use the same URL to sign up for a free trial. The account will be activated in approximately 5 minutes.

1. If you already have an **IBM Cloud** account, make sure that you provisioned the required services – **Watson Studio, Watson Machine Learning, and Watson OpenScale.**

* Navigate to your *Resource list* in your **IBM Cloud** dashboard: <https://cloud.ibm.com/resources>
* Check if the mentioned services are displayed under **Services**. If not, search for the services in the **Catalog** and add them.

Table

Description automatically generated

1. Download the [project zip file](https://github.com/ericmartens/fs2021/blob/main/fast_start_2021.zip?raw=true) to your machine.

# Required skills

We recommend that users who work through this lab:

* Understand data science lifecycle
* Have at least beginner knowledge of different methods for creating models

# ModelOps Overview

**ModelOps** is a process of developing and deploying data science assets to production. An important focus of ModelOps is automation of deployment, monitoring, and governance.

In this lab we will cover the **Test** phase of ModelOps. We focus on this to show how OpenScale can be used to evaluate and compare candidate models so that users can choose the best one for production.

A picture containing diagram

Description automatically generated

In the interest of time, the models have already been created for you, and are included in the project file. One has been built using IBM’s AutoAI service, and the other has been built using the PySpark library with a Jupyter notebook.

# Project Creation

1. Sign into [Watson Studio](https://dataplatform.cloud.ibm.com/). From the quick navigation on the left, select **Projects**.

Graphical user interface, application

Description automatically generated

1. Click the **New project** link at the top left, then select **Create a project from a sample or file**.

Graphical user interface, text, application

Description automatically generated

1. Drag the project zip file from your machine to the **Upload file** section. Give your project a name, select a Cloud Object Storage service from the dropdown list, and click **Create**.
2. In your project, click the **Settings** tab. Scroll down to the **Associated services** section. Click **Add service** and select **Watson** from the list.

Graphical user interface, application

Description automatically generated

1. Select your Watson Machine learning service from the list and click **Associate service**.

Graphical user interface

Description automatically generated with medium confidence

1. Close the services popup, scroll to the top of the screen, and click on the **Assets** tab.
2. The project contains two models, one built using IBM’s AutoAI service, and one built using a notebook. The training data and notebook used to build the models are included among the assets if you would like to replicate the process yourself.

# Promote and deploy the models

1. Scroll down to the models section. Click the **More** icon for the AutoAI model and select **Promote**.

Graphical user interface, application

Description automatically generated

1. Click **New space** to create a new deployment space for your models. Give your new space a name and select a storage option and WML instance from the dropdowns, and then click **Create**.
2. Once the space has been created, click **Close** to return to the model screen. Ensure that the new space you just created is selected in the **Target space** dropdown and click **Promote**. After your model has been promoted, you will return to the project assets screen. Follow similar steps to promote the PySpark model to the same deployment space.
3. Click the menu button at the top left of the screen, click the **Deployments** section to expand it, and select **View all spaces**.

Graphical user interface, application

Description automatically generated

1. Click the space you just created. From the **Assets** tab, click the **Deploy** icon for the PySpark model.

Graphical user interface, application

Description automatically generated

1. Select the **Online** deployment type and name your model **PySpark Hiring Challenger** to make it easy to identify. Click **Create**.
2. Repeat the two steps above to create a deployment for the AutoAI model, using **AutoAI Hiring Challenger** as the name.

# Configure Watson OpenScale

1. You have successfully deployed two models, using two different technologies. We will now configure OpenScale to monitor those models for fairness, quality, and drift, and compare the results.
2. In a new browser tab, navigate to [Watson OpenScale](https://aiopenscale.cloud.ibm.com/aiopenscale/) and log in. If you have not previously used OpenScale, select the **Manual setup** option.
3. Click the pencil icon on the right of the screen to configure your database. Select the **Free lite plan** database from the **Database type** dropdown, unless you would like to use a paid database service you have already provisioned. Click **Save**.
4. Select **Machine learning providers** from the list at the left of the screen.

Graphical user interface, text, application, chat or text message

Description automatically generated

1. Click **Add machine learning provider**. Click the pencil icon to give your provider a name like **Preproduction models**, then click **Apply**. Then click the pencil icon in the **Connection** tile.
2. Select **Watson Machine Learning (V2)** as the Service provider type. Select the deployment space you have been working with from the **Deployment space** dropdown. Choose **Pre-production** as the environment type, which will simplify running tests on the models. Click **Save**.

Graphical user interface, text, application

Description automatically generated

# Monitor models with OpenScale

You will now configure OpenScale to monitor your deployed models, starting with the AutoAI model. You will then repeat these steps for the second model.

1. Click the dashboard link at the top left to return to the Insights Dashboard, then click **Add to dashboard**. Select the provider you just added from the list, and then select the deployed AutoAI model and click **Configure**.

Graphical user interface, application

Description automatically generated

1. When the model has finished saving, click **Configure monitors**. Click the pencil icon in the **Model input** tile. Set the data type to **Numeric/categorical** and the algorithm type to **Binary classifier**. Click **Save and continue**.
2. Click the pencil icon in the **Training data** tile. Set the location to Cloud Object Storage and enter the following values into the corresponding fields and click **Connect**:

|  |  |
| --- | --- |
| Field | Value |
| Resource instance ID | crn:v1:bluemix:public:cloud-object-storage:global:a/7d8b3c34272c0980d973d3e40be9e9d2:2883ef10-23f1-4592-8582-2f2ef4973639:: |
| API key | yqcPbWZ0AQPHleHVerrR4Wx5e9pymBdMgydbEra5zCif |

1. Select **faststartlab-donotdelete…** as the bucket, and **hr\_training\_data.csv** as the data set. Click **Next**.

Graphical user interface, text, application

Description automatically generated

1. Select **HIRED** as the label column, and click **Next**.
2. Check the box to the left of **Features** to select all columns as training features, then scroll down and uncheck the boxes for **Gender** and **Ethnicity**. These features are included in the training data so OpenScale can find correlations to look for indirect bias but were not used in the actual training of the model. Click **Next**.

Table

Description automatically generated with medium confidence

1. Now, we’ll need to submit data to our model so OpenScale can construct the correct payload log in the datamart. Keep OpenScale open in one tab and return to the browser tab with your deployment space. Click on the **Deployments** tab and select the deployed AutoAI model. From the **Test** tab, click the icon to provide input data as JSON.

Graphical user interface, application

Description automatically generated with medium confidence

1. Delete the contents of the **Enter input data** field. Copy and paste the JSON below into the field, and then click **Predict**:

{"input\_data": [{"fields": ["Age", "BusinessTravel", "Department", "DistanceFromHome", "Education", "EducationField", "RelevantEducationLevel", "JobLevel", "JobRole", "MaritalStatus", "NumCompaniesWorked", "OverTime", "InterviewScore", "ResumeScore", "RequestedBenefits", "TotalWorkingYears", "PreferredSkills", "YearsAtCurrentCompany", "RelevantExperience", "JobType", "SalaryExpectation"],"values": [[24,"Travel\_Frequently","Research & Development",7,"Bachelor","Life Sciences","High","Entry Level","Laboratory Technician","Married",1,"Yes","High","Outstanding","Yes",6,"High",6,3,"contract","High"]],"meta": {"fields": ["Gender","Ethnicity"],"values":[["Female","non-minority"]]}}]}

1. Return to the tab with OpenScale. Verify that the scoring method is set to **Automatic logging** and click **Check now**. You should receive a message that logging is active. If you do not receive this message, wait 30 seconds and check again. Once logging is active, click **Next**, then click **Save**.
2. Click on the **Drift** monitor in the **Evaluations** section on the left side of the screen. Click the pencil icon in the **Drift model** tile and select the **Train in Watson OpenScale** option. Click **Next**.

Graphical user interface, text, application

Description automatically generated

1. Click **Next** to accept the default drift threshold. Set the sample size to 100 and click **Save**. OpenScale will begin training a drift monitor, which will take a few minutes. You may continue configuring the other monitors while this training occurs.
2. Click the **Fairness** monitor in the **Evaluations** section. Click the pencil icon in the **Favorable outcomes** tile. Use the checkboxes to set YES as the favorable value and NO as the unfavorable value, then click **Next**.

Graphical user interface, application

Description automatically generated

1. Set the minimum sample size to 100 records, then click **Next**. For fields to monitor, check the boxes next to **Gender** and **Ethnicity** and click **Next**.

Graphical user interface, application, table

Description automatically generated

1. Set the minority group to **Monitored**, and the non-minority group to **Reference** and click **Save**.

Graphical user interface, application, Teams

Description automatically generated

1. For gender, set Female to **Monitored** and Male to **Reference** and click **Save**.

Graphical user interface, application

Description automatically generated

1. Click **Next** to accept the default values for Monitored and Reference groups for both ethnicity and gender, then click **Save**.
2. In the **Evaluations** section, select the **Quality** monitor. Click on the pencil icon in the Quality threshold tile. Click Next to accept the default value of 0.8. Set the minimum sample size to 100 and click **Save**.
3. Click on the **Dashboard** link at the top left of the screen to return to the Insights Dashboard. Click the **Add to dashboard button**. Once again, select the machine learning provider you set up earlier, but this time, choose the PySpark model and click **Configure**.

Graphical user interface, application

Description automatically generated

1. Return to the beginning of this section and repeat the same steps for the PySpark model, using the exact same values as with the AutoAI model. However, during the **Specify model ouput details** step, set the **Prediction** value to **predictedLabel**.

Graphical user interface, application

Description automatically generated

# Evaluate and Compare Models

1. You will now compare the two pre-production models to see how they perform on a set of test data. Download this file to your machine:

<https://raw.githubusercontent.com/ericmartens/fs2021/main/payload_100.csv>

1. From the [OpenScale Insights dashboard](https://aiopenscale.cloud.ibm.com/aiopenscale/insights), click on the tile for the AutoAI model. Then, from the **Actions** menu, select **Evaluate now**.

Graphical user interface, text, application, chat or text message

Description automatically generated

1. Set the **Import** option to **from CSV file**. Drag the file you downloaded from your machine to the upload window, and click **Upload and evaluate**.

Graphical user interface, text, application

Description automatically generated

1. The evaluation will take a few minutes to run. When it is finished, return to the Insights dashboard and repeat the steps above for the PySpark model.
2. When the PySpark model has finished, you can compare the two models.

There’s currently a blocking issue with OpenScale that’s preventing me from finishing the lab instructions; I will update as soon as it’s fixed:

https://github.ibm.com/aiopenscale/tracker/issues/20099

# Conclusion and Next Steps

In this lab you learned how to complete the following ModelOps tasks in Cloud Pak for Data:

* Promote and deploy models
* Configure OpenScale datamarts and machine learning providers
* Configure model monitors with OpenScale
* Evaluate and compare models using test data