Lab Guide

Hands-on-Lab:
Monitor the model with Watson OpenScale

Shivam R Solanki
Data Scientist
Shivam.raj.solanki@ibm.com

IBM

Summary

In this lab tutorial, we'll use Auto Insurance claim data to train, create, and deploy a machine learning model using Watson® Machine Learning on IBM Cloud Pak® for Data. We'll create a data mart for this model with Watson OpenScale™ and configure OpenScale to monitor that deployment. Next, we'll inject data for viewing quality and fairness in the OpenScale Insights dashboard.

Description

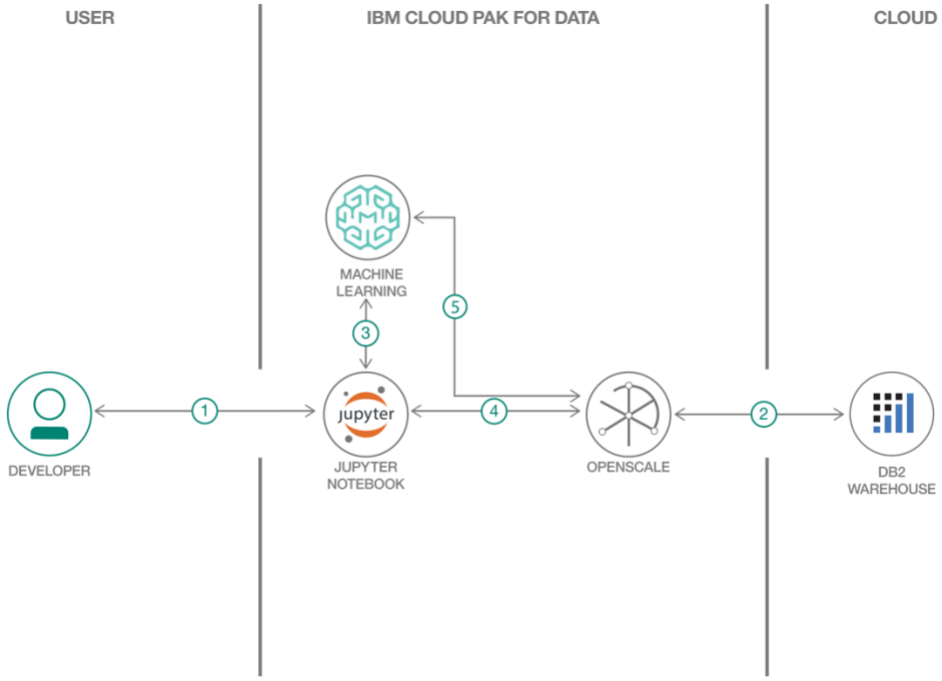
The dataset used for this pattern contains information regarding auto insurance claims from a variety of users. We can use different scikit-learn algorithms to create machine learning model using Watson Machine Learning and deploy this model for use in predicting fraud in auto insurance claims. Because of the sensitive nature of fraud claim prediction, this is an ideal use case for the on-premises solution offered by IBM Cloud Pak for Data.

The deployed ML model can now be monitored by Watson OpenScale. Continued use will generate data that allows administrators to ensure the quality of the model and offer explanations as to what features of the dataset are most influential in creating the fraud claim classification. Bias detection will be configured to allow further insight into the fairness of the model predictions. All of this information is available in the OpenScale dashboard and in great detail.

After completing this code pattern, you'll understand how to:

- Create and deploy a machine learning model using the Watson Machine Learning service on IBM Cloud Pak for Data.
- Set up Watson OpenScale data mart.
- Bind Watson Machine Learning to the Watson OpenScale data mart.
- Add subscriptions to the data mart.
- Enable payload logging and performance monitor for subscribed assets.
- Enable quality (accuracy) monitor.
- Enable fairness monitor.
- Score the Fraud claim prediction model using the Watson Machine Learning.

Flow



1. The developer creates a Jupyter Notebook on IBM Cloud Pak for Data.
2. OpenScale on IBM Cloud Pak for Data is connected to a Db2® database, which is used to store Watson OpenScale data.
3. The notebook is connected to Watson Machine Learning and a model is trained and deployed.
4. Watson OpenScale is used by the notebook to log payload and monitor performance, quality, and fairness.
5. OpenScale will monitor the Watson Machine Learning model for performance, fairness, quality, and explainability.

Instructions

Complete details on how to get started running and using this application are in the README, including how to:

1. Import Jupyter Notebook to IBM Cloud Pak for Data.
2. Run the notebook.
3. Configure OpenScale in a Jupyter Notebook.
4. Utilize the dashboard for OpenScale

Conclusion

This lab tutorial showed you deploy a machine learning model using Watson Machine Learning on IBM Cloud Pak for Data and then configure OpenScale to monitor the model in a Jupyter Notebook.