

Monitoring Vertex AI Model

2 hours 30 minutesFree

Overview

Model monitoring is the close tracking of the performance of ML models in production so that production and AI teams can identify potential issues before they affect the business.

If production traffic differs from training data or varies substantially over time, the quality of the answers your model produces is probably affected. When that happens, you will want to be alerted automatically and responsively so that you can anticipate problems before they affect your customer experiences or your revenue streams.

Objectives

- **Deploy a pre-trained model.**
- **Configure model monitoring.**
- **Generate some artificial traffic.**

- Interpret the data reported by the model monitoring feature.

Setup and requirements

For each lab, you get a new Google Cloud project and set of resources for a fixed time at no cost.

1. Sign in to Qwiklabs using an incognito window.
2. Note the lab's access time (for example, 1:15:00), and make sure you can finish within that time. There is no pause feature. You can restart if needed, but you have to start at the beginning.
3. When ready, click Start lab.
4. Note your lab credentials (Username and Password). You will use them to sign in to the Google Cloud Console.
5. Click Open Google Console.
6. Click Use another account and copy/paste credentials for this lab into the prompts. If you use other credentials, you'll receive errors or incur charges.
7. Accept the terms and skip the recovery resource page.

Note: Do not click End Lab unless you have finished the lab or want to restart it. This clears your work and removes the project.

Task 1. Set up your environment

Enable the Notebooks API

1. In the Google Cloud Console, on the Navigation menu, click APIs & Services > Library.
2. Search for Notebooks API and press enter.
3. Click on the `Notebooks API` result.
4. If the API is not enabled, you'll see the Enable button. Click Enable to enable the API.

Enable the Vertex AI API

- In the Google Cloud Console, on the Navigation menu, click Vertex AI > Dashboard, and then click Enable Vertex AI API.

Task 2. Launch Vertex AI Notebooks

1. In the Google Cloud Console, on the Navigation Menu, click Vertex AI > Workbench. Select User-Managed Notebooks.
2. On the Notebook instances page, click New Notebook > TensorFlow Enterprise > TensorFlow Enterprise 2.6 (with LTS) > Without GPUs.
3. In the New notebook instance dialog, confirm the name of the deep learning VM, if you don't want to change the region and zone, leave all settings as they are and then click Create. The new VM will take 2-3 minutes to start.
4. Click Open JupyterLab. A JupyterLab window will open in a new tab.
5. You will see "Build recommended" pop up, click Build. If you see the build failed, ignore it.

Task 3. Clone a course repo within your Vertex AI Notebooks instance

To clone the training-data-analyst notebook in your JupyterLab instance:

1. In JupyterLab, to open a new terminal, click the Terminal icon.
2. At the command-line prompt, run the following command:

```
git clone https://github.com/GoogleCloudPlatform/training-data-analyst
```

Copied!

content_copy

3. To confirm that you have cloned the repository, double-click on the training-data-analyst directory and ensure that you can see its contents. The files for all the Jupyter notebook-based labs throughout this course are available in this directory.

Task 4. Monitor your Vertex AI model

1. In the notebook interface, navigate to training-data-analyst > courses > machine_learning > deepdive2 > machine_learning_in_the_enterprise > labs, and open model_monitoring.ipynb.
2. Click Edit > Clear All Outputs.
3. Carefully read through the notebook instructions and fill in lines marked with #TODO where you need to complete the code.

Tip: To run the current cell, click the cell and press Shift+Enter. Other cell commands are listed in the notebook UI under Run.

- Hints may also be provided for the tasks to guide you. Highlight the text to read the hints, which are in white text.
- If you need more help, to view the complete solution, navigate to training-data-analyst > courses > machine_learning > deepdive2 > machine_learning_in_the_enterprise > solutions, and open model_monitoring.ipynb.

Task 5. Clean up

1. In the Google Cloud Console, on the Navigation menu, click Vertex AI > Endpoints.
2. Click on your endpoint (for example, churn), and then click Undeploy model from endpoint.
3. Click UNDEPLOY to confirm the undeployment.
4. Click the Back arrow, select your endpoint, and then click Delete. Click Confirm to delete the endpoint.
5. In the Vertex AI navigation menu, click Models, click the overflow menu (⋮), and then click Delete model.
6. Click Delete to confirm the deletion.

End your lab

When you have completed your lab, click End Lab. Qwiklabs removes the resources you've used and cleans the account for you.

You will be given an opportunity to rate the lab experience. Select the applicable number of stars, type a comment, and then click Submit.

The number of stars indicates the following:

- **1 star = Very dissatisfied**
- **2 stars = Dissatisfied**
- **3 stars = Neutral**
- **4 stars = Satisfied**
- **5 stars = Very satisfied**

You can close the dialog box if you don't want to provide feedback.

For feedback, suggestions, or corrections, please use the Support tab.

Copyright 2022 Google LLC All rights reserved. Google and the Google logo are trademarks of Google LLC. All other company and product names may be trademarks of the respective companies with which they are associated.