Identifying Bias in Mortgage Data using Cloud AI Platform and the What-if Tool

cloudskillsboost.google/games/2854/labs/17216

GSP709

Overview

In this lab, you use the What-If Tool to identify potential biases in a model that was trained on a mortgage loan applications dataset.

Objectives

In this lab, you will perform the following tasks:

- Build a binary classification model using XGBoost.
- Deploy the model to Cloud AI Platform.
- Use the What-If Tool on the deployed model to search for biases.

Setup and requirements

Before you click the Start Lab button

Read these instructions. Labs are timed and you cannot pause them. The timer, which starts when you click **Start Lab**, shows how long Google Cloud resources will be made available to you.

This hands-on lab lets you do the lab activities yourself in a real cloud environment, not in a simulation or demo environment. It does so by giving you new, temporary credentials that you use to sign in and access Google Cloud for the duration of the lab.

To complete this lab, you need:

Access to a standard internet browser (Chrome browser recommended).

Note: Use an Incognito or private browser window to run this lab. This prevents any conflicts between your personal account and the Student account, which may cause extra charges incurred to your personal account.

Time to complete the lab---remember, once you start, you cannot pause a lab.

Note: If you already have your own personal Google Cloud account or project, do not use it for this lab to avoid extra charges to your account.

How to start your lab and sign in to the Google Cloud Console

1. Click the **Start Lab** button. If you need to pay for the lab, a pop-up opens for you to select your payment method. On the left is the **Lab Details** panel with the following:

    - The **Open Google Console** button
    - Time remaining
    - The temporary credentials that you must use for this lab
    - Other information, if needed, to step through this lab
2. Click **Open Google Console**. The lab spins up resources, and then opens another tab that shows the **Sign in** page.

    Tip: Arrange the tabs in separate windows, side-by-side.

    Note: If you see the **Choose an account** dialog, click **Use Another Account**.
3. If necessary, copy the **Username** from the **Lab Details** panel and paste it into the **Sign in** dialog. Click **Next**.

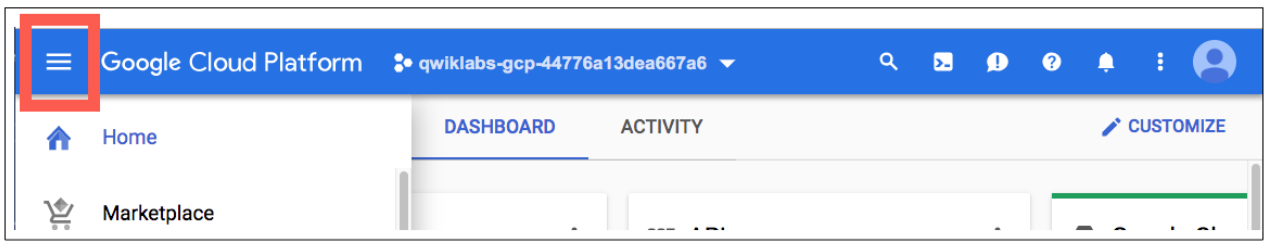
4. Copy the **Password** from the **Lab Details** panel and paste it into the **Welcome** dialog. Click **Next**.

    Important: You must use the credentials from the left panel. Do not use your Google Cloud Skills Boost credentials. **Note:** Using your own Google Cloud account for this lab may incur extra charges.
5. Click through the subsequent pages:

    - Accept the terms and conditions.
    - Do not add recovery options or two-factor authentication (because this is a temporary account).
    - Do not sign up for free trials.

After a few moments, the Cloud Console opens in this tab.

Note: You can view the menu with a list of Google Cloud Products and Services by clicking the **Navigation menu** at the top-left.



Create a storage bucket

Create a bucket using the Cloud Console:

1. In the Cloud Console, on the Navigation menu, click **Cloud Storage**.
2. Click **CREATE BUCKET**.
3. Choose a Regional bucket and set a unique name (use your project ID because it is unique). Then, click **CREATE**.

Click *Check my progress* to verify the objective. Create cloud storage bucket

Start a JupyterLab Notebook instance

1. In the Google Cloud Console, on the **Navigation Menu**, click **Vertex AI > Workbench**.
2. On the Notebook instances page, click **NEW NOTEBOOK**.
3. In the Customize instance menu, select **TensorFlow Enterprise** and choose the version of **TensorFlow Enterprise 2.6 (with LTS)** > **Without GPUs**.
4. In the **New notebook instance** dialog, for **Region**, select `us-central1`, for **Zone**, select a zone within the selected region, leave all other fields with their default options, and click **CREATE**.

After a few minutes, the Vertex AI console will display your instance name, followed by `Open Jupyterlab`.

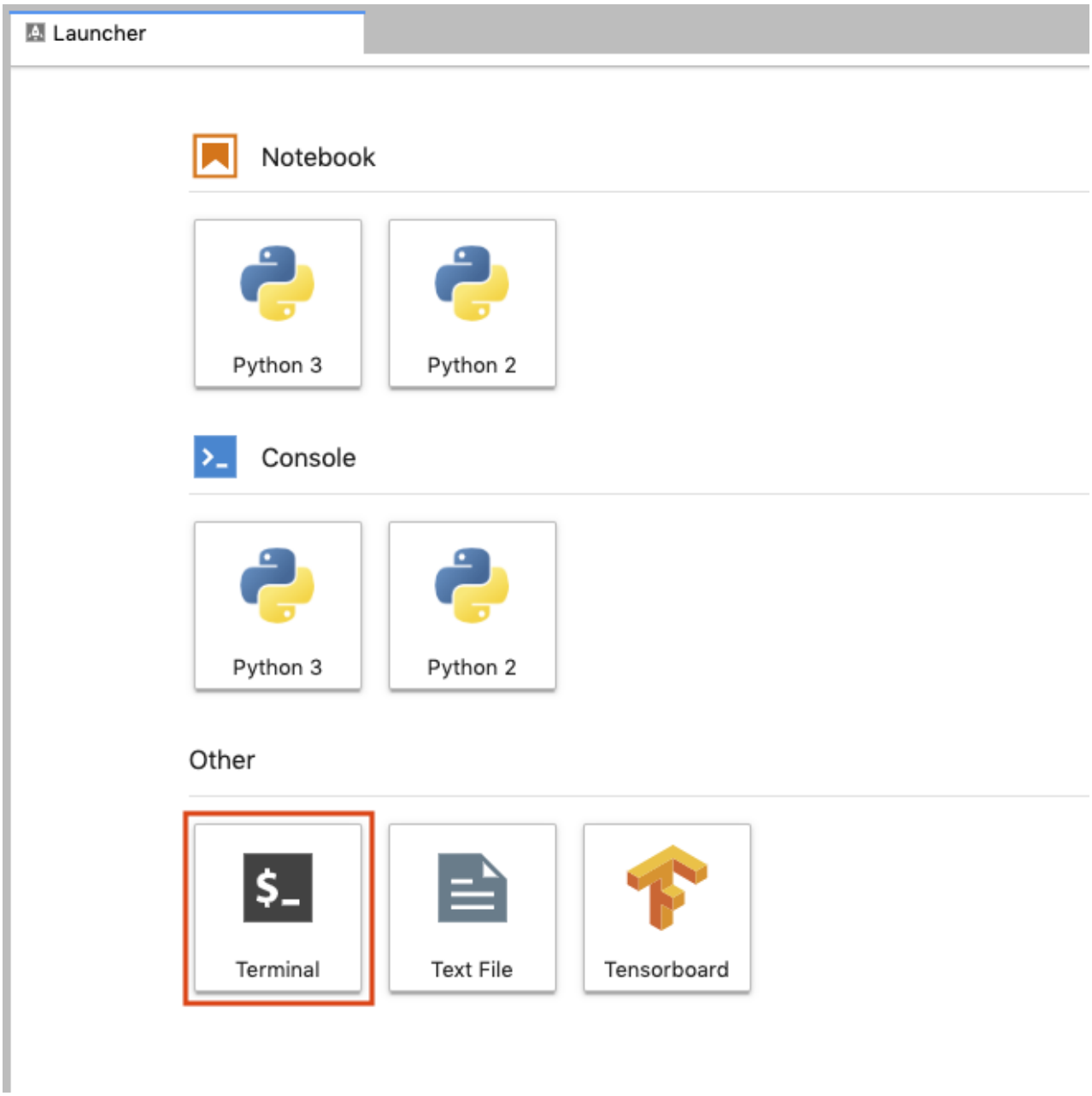
5. Click **Open JupyterLab**. Your notebook is now set up.

Click *Check my progress* to verify the objective. Start a JupyterLab Notebook Instance

Clone the sample code

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

1. In JupyterLab, click the **Terminal** icon to open a new terminal.

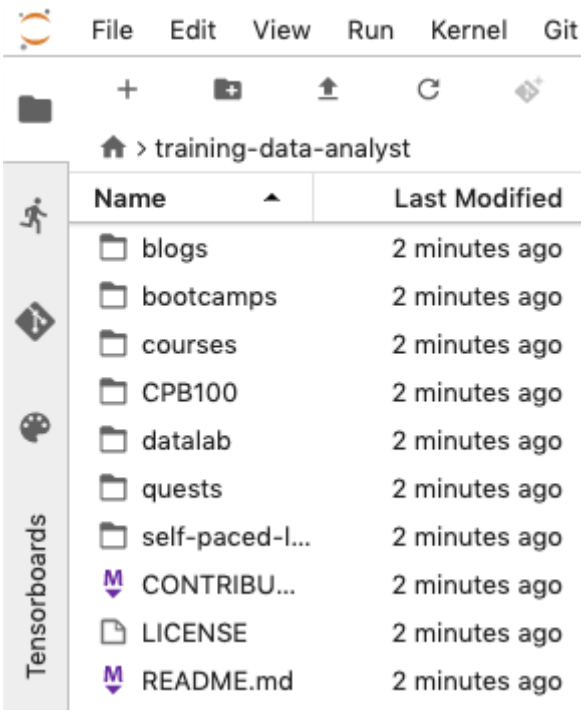


2. To clone the `training-data-analyst` repo, type in the following command, and press **Enter**.

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

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

Click *Check my progress* to verify the objective. Clone the sample code



Explore the What-If Tool

1. Navigate to **training-data-analyst** > **quests** > **dei** and open **xgboost_caip_e2e.ipynb**.
2. Continue the lab in the notebook, and run each cell by clicking the **Run (▶)** icon at the top of the screen. Alternatively, you can execute the code in a cell with **SHIFT + ENTER**. Read the narrative and make sure you understand what's happening in each cell.

Click *Check my progress* to verify the objective. Execute code in JupyterLab notebook

Congratulations!

In this lab you used the What-If Tool to identify potential biases in a model that was trained on a mortgage loan applications dataset.

Finish Your Quest

This self-paced lab is part of the Qwiklabs Explore Machine Learning Models with Explainable AI Quest. A Quest is a series of related labs that form a learning path. Completing a Quest earns you a badge to recognize your achievement. You can make your badge (or badges) public and link to them in your online resume or social media account. Enroll in a Quest and get immediate completion credit if you've taken this lab. See other available Qwiklabs Quests.

Take your next lab

Continue your quest with Explore Machine Learning Models with Explainable AI: Challenge Lab, or check out these suggestions:

Compare Cloud AI Platform Models using the What-If Tool to identify Potential Bias

Google Cloud Training & Certification

...helps you make the most of Google Cloud technologies. Our classes include technical skills and best practices to help you get up to speed quickly and continue your learning journey. We offer fundamental to advanced level training, with on-demand, live, and virtual options to suit your busy schedule. Certifications help you validate and prove your skill and expertise in Google Cloud technologies.

Manual Last Updated December 21, 2021

Lab Last Tested December 21, 2021