TFX on Google Cloud Vertex Pipelines

1 hour 30 minutes Free

GSP1023

Overview

Tensorflow Extended (TFX) is Google's end-to-end platform for training and deploying TensorFlow models into production. TFX pipelines orchestrate ordered runs of a sequence of components for scalable, high-performance machine learning tasks in a directed graph. It includes pre-built and customizable components for data ingestion and validation, model training and evaluation, as well as model validation and deployment. TFX is the best solution for taking TensorFlow models from prototyping to production with support on-prem environments and in the cloud such as on Google Cloud's Vertex Pipelines.

Vertex AI Pipelines helps you to automate, monitor, and govern your ML systems by orchestrating your

ML workflow in a serverless manner, and storing your workflow's artifacts using Vertex ML Metadata.

In this lab you will learn how to deploy and run a TFX pipeline on Google Cloud that automates the development and deployment of a TensorFlow 2.7 classification model which predicts the species of penguins.

Objectives

- Create a TFX Pipeline using TFX APIs.
- Define a pipeline runner that uses Vertex Pipelines together with the Kubeflow V2 dag runner.
- Deploy and monitor a TFX pipeline on Vertex Pipelines.

Setup

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.

Access Vertex Notebook

An instance of Vertex Notebooks is used as a primary experimentation/development workbench for this lab.

To launch Vertex Notebooks:

1. Click on the **Navigation Menu** and navigate to **Vertex AI**, then to **Workbench**.
2. You should see `tfx-on-googlecloud` notebook preprovisioned for you. If not, wait a few minutes and refresh the page.
3. Click **Open JupyterLab**. A JupyterLab window will open in a new tab.

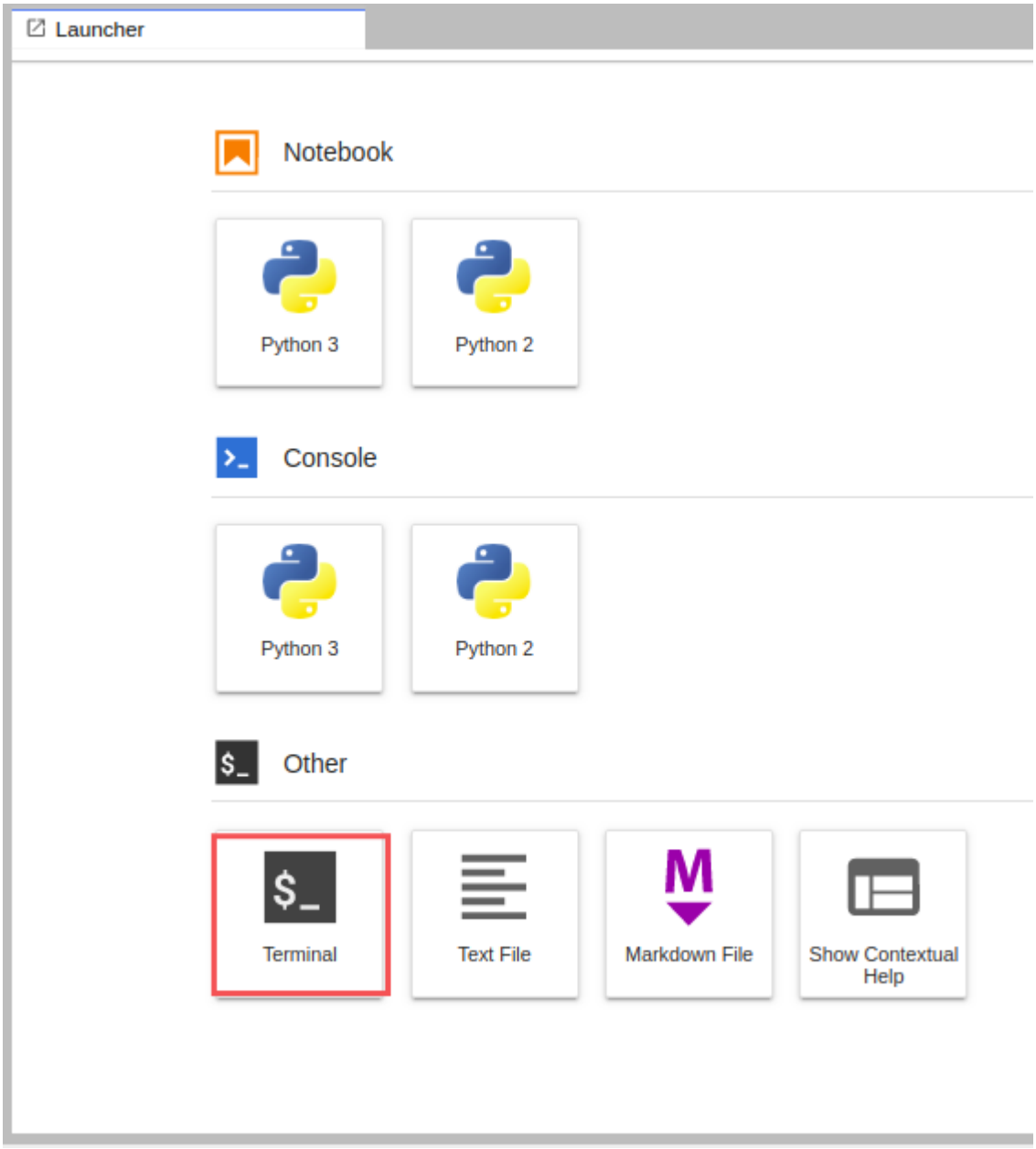
If a **Build Recommended** popup appears when you open your notebook, press **cancel**.

Clone the example repo within your Vertex

Notebooks instance

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

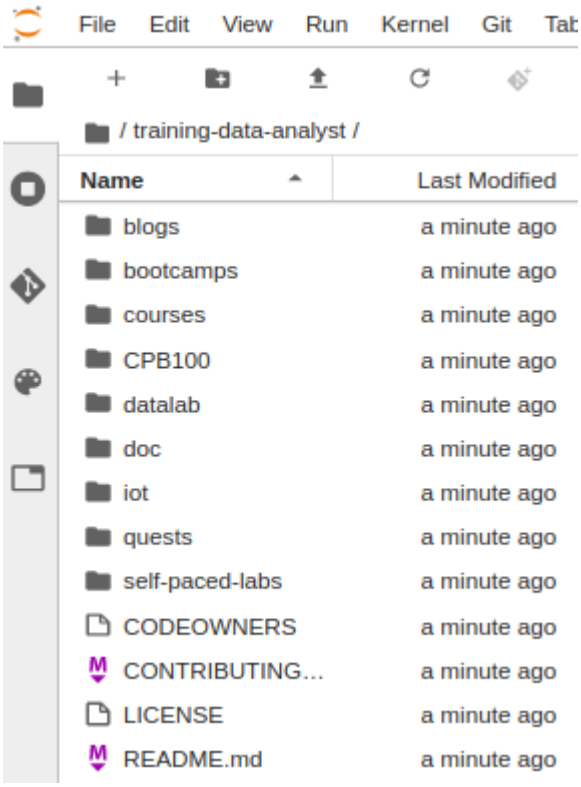
1. In JupyterLab, click the **Terminal** icon to open a new terminal.



2. At the command-line prompt, type the following command and press **Enter**:

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

3. To confirm that you have cloned the repository, in the left panel, double click the `training-data-analyst` folder to see its contents.



Click *Check my progress* to verify the objective.

Clone the example repo within your Vertex Notebooks instance

Navigate to the lab notebook

In your Vertex Notebook, navigate to the following directory:

```
training-data-analyst/self-paced-labs/tfx/tfx-vertex
```

Open `vertex_pipelines_simple.ipynb`.

When prompted, come back to these instructions to *check my progress*. You will need to do this to

receive credit for completing the lab.

Run your training job in the cloud

Click *Check my progress* to verify the objective.

Build and deploy a TFX pipeline to Vertex Pipelines

You may need to wait a few minutes after job completion for progress to be tracked accordingly.

Congratulations!

You have learned how to build and deploy a TFX pipeline to Vertex Pipelines and triggered a pipeline run.

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 May 13, 2022

Lab Last Tested May 13, 2022