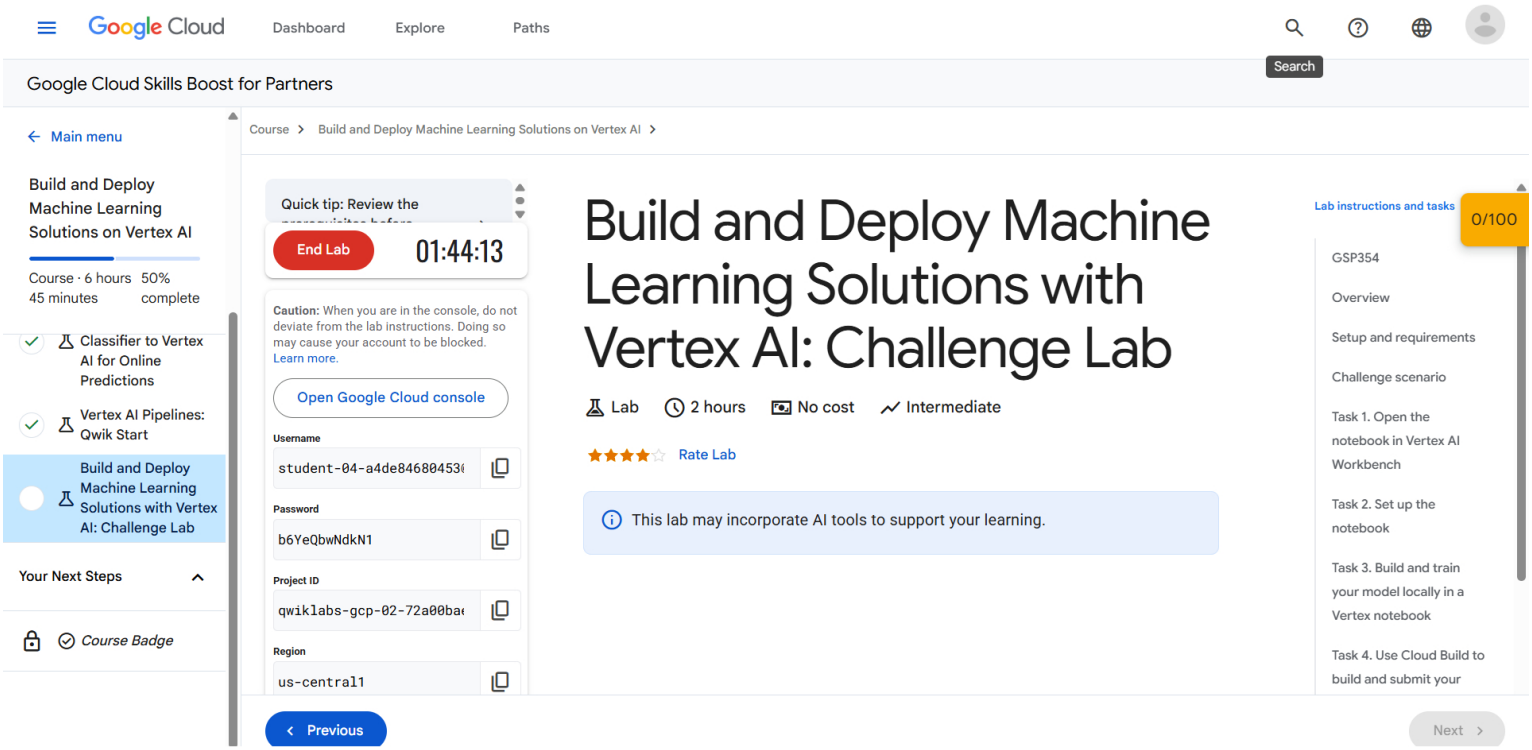


Google Cloud Self-Paced Labs

Overview

In a challenge lab you're given a scenario and a set of tasks. Instead of following step-by-step instructions, you will use the skills learned from the labs in the course to figure out how to complete the tasks on your own! An automated scoring system (shown on this page) will provide feedback on whether you have completed your tasks correctly.

researching error messages to fix your own mistakes.

To score 100% you must successfully complete all tasks within the time period!

This lab is recommended for students who have enrolled in the Build and Deploy Machine Learning Solutions on Vertex AI course. Are you ready for the challenge?

Setup and requirements

Before you click the Start Lab button

starts when you click **Start Lab**, shows how long Google Cloud resources are made available to you.

This hands-on lab lets you do the lab activities in a real cloud environment, not in a simulation or demo environment. It does so by giving you new, temporary credentials 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 (recommended) or private browser window to run this lab. This prevents 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.

account, you may incur charges to that account.

Challenge scenario

You were recently hired as a Machine Learning Engineer at a startup movie review website. Your manager has tasked you with building a machine learning model to classify the sentiment of user movie reviews as positive or negative. These predictions will be used as an input in downstream movie rating systems and to surface top supportive and critical reviews on the movie website application.

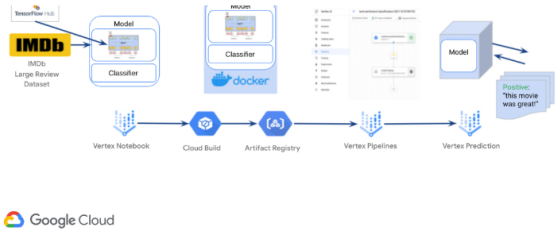
The challenge: your business requirements are that you have just 6 weeks to

your startup's data warehouse, you found that you only have a small dataset of 50k text reviews to build a higher performing solution.

Your challenge

To build and deploy a high performance machine learning model with limited data quickly, you will walk through training and deploying a custom TensorFlow BERT sentiment classifier for online predictions on Google Cloud's Vertex AI platform. Vertex AI is Google Cloud's next generation machine learning development platform where you can leverage the latest ML pre-built components and AutoML to significantly enhance your development productivity, the ability to scale your workflow and decision making with your data, and accelerate time to value.

Lab Architecture Diagram



First, you will progress through a typical experimentation workflow where you will build your model from pre-trained BERT components from TF-Hub and `tf.keras` classification layers to train and evaluate your model in a Vertex Notebook. You will then package your model code into a Docker container to train on Google Cloud's Vertex AI. Lastly, you will define and run a Kubeflow Pipeline on Vertex Pipelines that trains and deploys your model to a Vertex Endpoint that you will query for online predictions.

Task 1. Open the notebook in Vertex AI Workbench

1. In the Google Cloud console, on the **Navigation menu** (☰), click **Vertex AI > Workbench**.

2. Find the `vertex-ai-jupyterlab` instance and click on the **Open JupyterLab** button.

The JupyterLab interface for your Workbench instance opens in a new browser tab.

Task 2. Set up the notebook

1. In your notebook, click the **Terminal**.
2. Install the required packages for the lab:

```
pip3 install -U -r requirements.txt --user
```

3. In the **File Browser** on the left, click on the `vertex-challenge-lab-v1.0.0.ipynb` file.
4. When asked which kernel to use, select the **Python 3 (ipykernel)** kernel.
5. Run through the **Setup** section of the notebook to install the required libraries

- For **Project ID**, use `qwiklabs-gcp-02-72a00bae2210`, and for the **Region**, use `us-central1`.

All the rest of the code to import and pre-process the dataset has been provided for you. The rest of the steps will be inside the notebook file. You should refer back to this lab guide to check your progress and get some hints.

Task 3. Build and train your model locally in a Vertex notebook

In this section, you will train your model locally using TensorFlow.

Note: This lab adapts and extends the official TensorFlow BERT text classification tutorial to utilize Vertex AI services. See the tutorial for additional coverage on fine-tuning BERT models using TensorFlow.

Build and compile a TensorFlow BERT sentiment classifier

1. Fill out the `#TODO` section to add a `hub.KerasLayer` for BERT text preprocessing.
2. Fill out the `#TODO` section to add a `hub.KerasLayer` for BERT text encoding.
3. Fill out the `#TODO` section to save your BERT sentiment classifier locally. You should save it to the `./bert-sentiment-classifier-local` directory.

Click **Check my progress** to verify the objective.

Check my progress

Task 4. Use Cloud Build to build and submit your model container to Artifact Registry

Create Artifact Registry for custom container images

1. Fill out the `#TODO` section to create a Docker Artifact Registry using the gcloud documentation.

Note: Make sure you specify the `location`, `repository-format`, and `description` flags.

Build and submit your container image to Artifact Registry using Cloud Build

1. Fill out the `#TODO` section to use Cloud Build to build and submit your custom model container to Artifact Registry. Learn more about it from the gcloud builds submit documentation.

Note: Make sure the config flag is pointed at `{MODEL_DIR}/cloudbuild.yaml`, defined above, and you include your model directory.

Build and submit your container image to Artifact Registry

Check my progress

Task 5. Define a pipeline using the KFP SDK

1. Fill out the `#TODO` section to add and configure `CustomContainerTrainingJobOp` components for your pipeline.

earlier.

Note: This training can take around 30-40 minutes to train and deploy the model.

Click **Check my progress** to verify the objective.

Define a pipeline using the KFP SDK

Check my progress

Task 6. Query the deployed model using your Vertex endpoint

1. Fill out the `#TODO` section to generate online predictions using your Vertex Endpoint.

Congratulations!

Congratulations! In this lab, you have learned how to build and deploy a custom BERT sentiment classifier using Vertex AI. You have also learned how to use Cloud Build to

build and submit your custom model container to Artifact Registry, and how to define a

Earn your next skill badge

This self-paced lab is part of the Build and Deploy Machine Learning Solutions on Vertex AI course. Completing this skill badge course earns you the badge above, to recognize your achievement. Share your badge on your resume and social platforms, and announce your accomplishment using #GoogleCloudBadge.

This skill badge course is part of Google's Data Scientist/Machine Learning Engineer learning path. If you have already completed the other skill badge courses in this learning path, search the catalog for 20+ other skill badge courses in which you can enroll.

Google Cloud training and 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 August 08, 2024

Lab Last Tested August 08, 2024