

Natural Language Processing on Google Cloud

[Professional Machine Learning Engineer Certification Learning Path](#) navigate_next [Natural Language Processing on Google Cloud](#) navigate_next NLP with Vertex AI

Text Classification using AutoML

2 hours 30 minutes Free

Overview

In this lab, you use AutoML with [Vertex AI](#) to train a text dataset to predict the source of an article.

Objectives

You learn how to:

- Import a text dataset to AutoML.
- Train the ML model for text classification.
- Evaluate the model performance.
- Deploy the model to an endpoint.
- Get predictions.

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.

What you need

To complete this lab, you need:

- Access to a standard internet browser (Chrome browser recommended).
- Time to complete the lab.

Note: If you have a personal Google Cloud account or project, do not use it for this lab. **Note:** If you are using a Pixelbook, open an Incognito window to run this lab.

Log in to Google Cloud Console

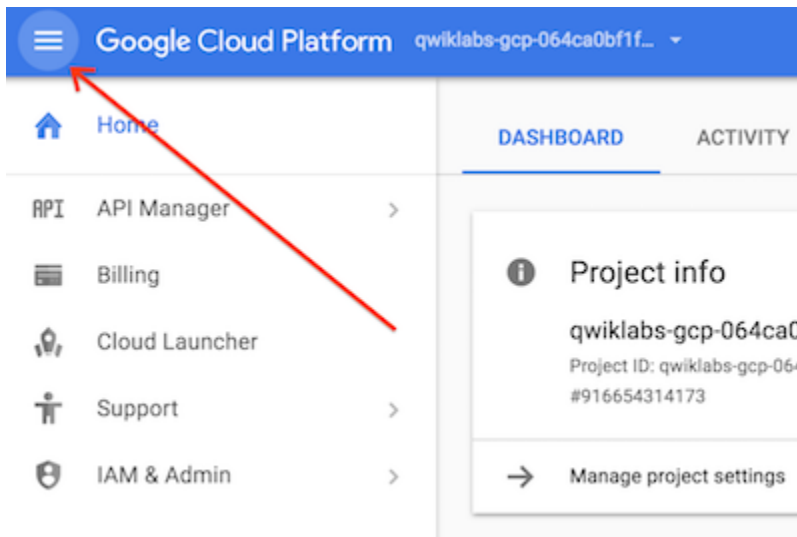
1. Using the browser tab or window you are using for this lab session, copy the **Username** from the **Connection Details** panel and click the **Open Google Console** button.

Note: If you are asked to choose an account, click **Use another account**.

2. Paste in the **Username**, and then the **Password** as prompted.
3. Click **Next**.
4. Accept the terms and conditions.

Since this is a temporary account, which will last only as long as this lab:

- Do not add recovery options
 - Do not sign up for free trials
5. Once the console opens, view the list of services by clicking the **Navigation menu** (≡) at the top-left.



Introduction to Vertex AI

This lab uses [Vertex AI](#), the unified AI platform on Google Cloud, to train and deploy an ML model. Vertex AI offers two options on one platform to build an ML model: a no-code solution with AutoML, and a code-based solution with Custom Training that uses Vertex Workbench. You use AutoML in this lab.

In this lab, you train the model to learn the relationship between article titles and their sources (including GitHub, the New York Times, and TechCrunch). You then use the trained model to predict the source of given article titles.

Task 1. Prepare the training data

The initial Vertex AI dashboard illustrates the major stages to train and deploy an ML model: prepare the data, train the model, and get predictions. Later, the dashboard displays your recent activities, such as the recent datasets, models, predictions, endpoints, and notebook instances.

Create a dataset

1. In the Google Cloud console, in the **Navigation menu**, click **Vertex AI > Datasets**.
2. Click **Create dataset**.
3. On the **Datasets** page, give the dataset a name.
4. For **Data type and objective**, click **Text**, and then select **Text Classification (Multi-label)**.

5. Click **Create**.

Upload data

There are three options for importing text data in Vertex AI:

- Upload text documents from your computer.
- Upload import files from your computer.
- Select import files from Cloud Storage.

For convenience, the dataset is already uploaded to Cloud Storage.

1. For the data source, select **Select import files from Cloud Storage**.
2. For **Import file path**, enter:

cloud-training/OCBL400/title_data_10000.csv

3. Click **Continue**.

Note: The data import may take a few minutes. You can also configure this page by clicking **Datasets** .

(Optional) Analyze import data

1. To see the import data, click **Browse**.
You can find the overall information of the dataset, including the total number of data points and the number for each label/category. You can also use the filter to browse the imported data.
2. To see a brief analysis of the dataset, click **Analyze**.

Note: If you receive a warning that the data can't be imported because of errors, click **Dismiss**. The missing datapoints only count a small percentage of the entire dataset and can be ignored.

Task 2. Train the model

With a dataset uploaded, you're ready to train the model.

1. Click **Train new model**.
2. For **Training method**, select **AutoML**.
3. Click **Continue**.
4. For **Model details**, select **Train new models**.
5. Give the model a name and, optionally, a description.
6. (Optional) Explore **Advanced options** to determine how to assign the training versus testing data and specify the encryption.
7. Click **Start training**.

Depending on the data size and the training objectives, the training can take from a few minutes to a couple of hours. Normally you would receive an email from Google Cloud when the training job is complete. However, in the Qwiklabs environment, you will not receive an email.

Note: To avoid waiting for the model training, later in the lab you download a pre-trained model and get predictions based on the same process you just followed. Task 3 and Task 4 are for demonstration only and can be skipped.

Task 3. Evaluate the model performance (demonstration only)

Vertex AI provides metrics to evaluate the model performance. For text classification, you focus on Precision/Recall curve.

If you had a model trained, you could follow this procedure:

1. Navigate to the **Models** tab.
2. Click the model you just trained.
3. Browse the **Evaluate** tab.

However in this lab, you can skip this step because you use a pre-trained model.

The Precision/Recall curve

The confidence threshold determines how an ML model counts the positive cases. A higher threshold increases precision, but decreases recall. A lower threshold decreases precision, but increases recall. You can manually adjust the threshold to observe its impact on precision and recall and find the best tradeoff point between the two to meet your business needs.

Task 4. Deploy the model (demonstration only)

You will not deploy the model to an endpoint because the model training can take an hour. Here you can review the steps you would perform in a production environment.

Now that you have a trained model, the next step is to create an endpoint in Vertex AI. A model resource in Vertex can have multiple endpoints associated with it, and you can split traffic between endpoints.

Create and define an endpoint

1. On your model page, on the **Deploy and test** tab, click **Deploy to endpoint**.
2. For **Endpoint name**, enter a name.
3. Click **Continue**.

Your endpoint will take a few minutes to deploy. When it's completed, a **green check mark** will appear next to the name.

Now you're ready to get predictions on your deployed model.

Task 5. Configure the environment

1. Click **Activate Cloud Shell** . If prompted click **continue**.
2. To create an Endpoint environment variable, run the following command:

```
ENDPOINT="{{ project_0.startup_script.automl_service_url }}"
```

3. Download the test files from Cloud Storage:

```
gsutil cp gs://cloud-training/OCBL400/CLOUD* .
```

The example files `CLOUD1-JSON` and `CLOUD2-JSON` have content similar to:

```
{ "instances": { "mimeType": "text/plain", "content": "Google's plan for the future of work." } }
```

Task 6. Get predictions

The system now has the test files available, so AutoML can be used to request predictions. The test files include the following two examples:

File	Text Message
CLOUD1-JSON	Google's plan for the future of work.
CLOUD2-JSON	Markdown Cheatsheet

Example One

1. Set `CLOUD1-JSON` as the input file:

```
INPUT_DATA_FILE=CLOUD1-JSON
```

Example Text

Google's plan for the future of work.

2. Request a prediction:

```
curl -X POST -H "Content-Type: application/json" $ENDPOINT/v1 -d "@${INPUT_DATA_FILE}" | jq
```

Example output:

```
{ "predictions": [ { "displayNames": [ "github", "techcrunch", "nytimes" ], "confidences": [ 0.00531658623367548, 0.6435679197311401, 0.4058765470981598 ], "ids": [ "1971075253660549120", "6582761272087937024", "7735682776694784000" ] } ], "deployedModelId": "8749478127236808704", "model": "projects/1030115194620/locations/us-central1/models/4646355819074420736", "modelDisplayName": "new_media", "modelVersionId": "1" }
```

How can you interpret the prediction result? Look at **confidences**. AutoML predicts a 0.5% chance that the article title (Google's plan for the future of work) comes from GitHub, a 64% chance that it comes from Techcrunch, and a 40% chance that it comes from the New York Times. What do you think about the result?

Example Two

1. Set `CLOUD2-JSON` as the input file:

```
INPUT_DATA_FILE=CLOUD2-JSON
```

Example Text

Markdown Cheatsheet

2. Request a prediction:

```
curl -X POST -H "Content-Type: application/json" $ENDPOINT/v1 -d "@${INPUT_DATA_FILE}" | jq
```

Example output:

```
{ "predictions": [ { "ids": [ "1971075253660549120", "6582761272087937024", "7735682776694784000" ],  
"confidences": [ 0.9996405243873596, 0.0005949776386842132, 3.189263225067407e-05 ], "displayNames":  
[ "github", "techcrunch", "nytimes" ] } ], "deployedModelId": "8749478127236808704", "model":  
"projects/1030115194620/locations/us-central1/models/4646355819074420736", "modelDisplayName":  
"new_media", "modelVersionId": "1" } Can you interpret the prediction result? Look at confidences. AutoML  
predicts a 99% chance that the article title (Markdown Cheatsheet) comes from GitHub, a 0.05% chance that it  
comes from Techcrunch, and a 0.003% chance that it comes from the New York Times. Do you agree with the  
result?
```

You can now use Vertex AI to:

- Upload a text dataset.
- Train a text classification model with AutoML.
- Evaluate the model performance.
- Deploy the trained AutoML model to an endpoint.
- Get predictions.

☐ Congratulations! ☐

To learn more about different parts of Vertex AI, see the [Vertex AI documentation](#).

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.

Manual Last Updated August 10, 2023

Lab Last Tested August 10, 2023

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.

- [Overview](#)
- [Setup](#)
- [Introduction to Vertex AI](#)
- [Task 1. Prepare the training data](#)
- [Task 2. Train the model](#)
- [Task 3. Evaluate the model performance \(demonstration only\)](#)
- [Task 4. Deploy the model \(demonstration only\)](#)
- [Task 5. Configure the environment](#)
- [Task 6. Get predictions](#)
- [End your lab](#)