Natural Language Processing on Google Cloud

Professional Machine Learning Engineer Certification Learning Path navigate_next Natural Language Processing on Google Cloud navigate_next Text representatation

Text classification using reusable embeddings

2 hours Free

Overview

Duration is 1 min

In this lab, you implement text models to recognize the probable source (Github, Tech-Crunch, or The New-York Times) of titles present in the title dataset, which are created in the respective labs.

Learning objectives

In this lab, you learn how to:

- Use pre-trained TF Hub text modules to generate sentence vectors.
- Incorporate a pre-trained TF-Hub module into a Keras model.
- Deploy and use a text model on CAIP.

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.

Enable the AI Platform Training & Prediction API and Vertex AI API

1. On the Navigation menu, navigate to **APIs & services > Library** and search for `AI Platform Training & Prediction API` in the search box.
2. Click on **AI Platform Training & Prediction API**, then click **Enable**.
3. Then, search for `Vertex AI API` in the search box.
4. Click on **Vertex AI API**, then click **Enable**.

Task 1. Create a Cloud Storage bucket

1. On the Navigation menu, navigate to **Cloud Storage** and Click on **Create bucket**.
2. Set a unique name (use your project ID because it is unique). Then, click **Create**.

Task 2. Launch Vertex AI Notebooks

To create and launch a Vertex AI Workbench notebook:

1. In the Navigation Menu Navigation menu icon, click **Vertex AI > Workbench**.
2. On the **User-Managed Notebook** page, click **Enable Notebooks API** (if it isn't enabled yet), then click **Create New**.
3. In the New instance menu, choose the latest version of **TensorFlow Enterprise 2.6 (with LTS)** in Environment.
4. Name the notebook.
5. Set **Region** to and **Zone** to any zone within the designated region.
6. Leave the remaining fields at their default and click **Create**.

After a few minutes, the Workbench page lists your instance, followed by Open JupyterLab.

7. Click Open JupyterLab to open JupyterLab in a new tab. If you get a message saying beatrix jupyterlab needs to be included in the build, just ignore it.
8. You will see "Build recommended" pop up, click **Build**. If you see the build failed, ignore it.

Task 3. Clone 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
   ```

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. Classify text using reusable embeddings

Duration is 60 min

1. In the notebook interface, navigate to **training-data-analyst > courses > machine_learning > deepdive2 > text_classification > labs > reusable_embeddings.ipynb**.
2. In the notebook interface, click on **Edit > Clear All Outputs** (click on Edit, then in the drop-down menu, select Clear All Outputs).
3. Carefully read through the notebook instructions and fill in lines marked with #TODO where you need to complete the code as needed

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 along. Highlight the text to read the hints (they are in white text).
- If you need more help, you may take a look at the complete solution by navigating to **training-data-analyst > courses > machine_learning > deepdive2 > text_classification > solutions** and opening **reusable_embeddings.ipynb**.

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.