Feature Engineering

Feature Engineering navigate_next Feature Engineering

Performing Basic Feature Engineering in BQML

1 hour Free

Overview

Duration is 1 min

In this lab, you utilize feature engineering to improve a model which predicts the fare amount for a taxi ride in New York City.

What you learn

In this lab, you:

- Create SQL statements to evaluate the model
- Extract temporal features
- Perform a feature cross on temporal features

Setup

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.

Task 1. Enable All Recommended API

1. On the **Navigation menu** (☰), click **APIs & services**.
2. Scroll down and confirm that your APIs are enabled.
3. If an API is missing, click **ENABLE APIS AND SERVICES** at the top, search for the API by name, and enable it for your project.

- BigQuery

4. In the Google Cloud Console, on the **Navigation menu**, click **Vertex AI**.
5. Click **Enable All Recommended API**.

Task 2. Launch a Vertex AI Notebooks instance

1. In the Google Cloud Console, on the **Navigation Menu**, click **Vertex AI > Workbench**. Select **User-Managed Notebooks**.
2. On the Notebook instances page, click **New Notebook > TensorFlow Enterprise > TensorFlow Enterprise 2.6 (with LTS) > Without GPUs**.
3. In the **New notebook** instance dialog, confirm the name of the deep learning VM, if you don't want to change the region and zone, leave all settings as they are and then click **Create**. The new VM will take 2-3 minutes to start.
4. Click **Open JupyterLab**.
   A JupyterLab window will open in a new tab.
5. You will see "Build recommended" pop up, click **Build**. If you see the build failed, ignore it.

Task 3. Clone a 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. Performing basic feature engineering in BQML

Duration is 30 min

1. In the notebook interface, navigate to **training-data-analyst > courses > machine_learning > deepdive2 > feature_engineering > labs > 1_bqml_basic_feat_eng_bqml-lab.ipynb**.
2. In the notebook interface, click on **Edit > Clear All Outputs**.
3. Carefully read through the notebook instructions and fill in lines marked with #TODO where you need to complete the code.

Note: To run the current cell, click the cell and hit **SHIFT+ENTER**. Other cell commands are found 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, which are in white text.
- To view the complete solution, navigate to **training-data-analyst > courses > machine_learning > deepdive2 > feature_engineering > solutions** and open **1_bqml_basic_feat_eng.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.