[ML on GCP C10] Content-Based Filtering using Neural Networks

1 hour 30 minutes   1 Credit   Rate Lab

Overview

This lab shows you how to do content-based filtering using DNNs in TensorFlow.

Objectives

In this lab, you learn to perform the following tasks:

- Build the feature columns for a DNN model using `tf.feature_column`

- Create custom evaluation metrics and add them to TensorBoard

- Train a DNN model for content-based recommendations and perform predictions with that model

Introduction

In this lab, you'll be providing building a DNN to make content-based recommendations. Content based filtering uses features of the items and users to generate recommendations. In this example, we'll be making recommendations for users on the Kurier.at news site. To do this, we'll be using features based on the news text, title, author, and recency of the article.

Setup

For each lab, you get a new GCP project and set of resources for a fixed time at no cost.

1. Make sure you signed into Qwiklabs using an **incognito window**.

2. Note the lab's access time (for example, `02:00:00` and make sure you can

   finish in that time block.

   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. You will use them to sign in to Cloud Platform Console.

Open Google Console

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked. **Learn more.**

Username
google2876526_student@qwiklabs.n

Password
TG959yrKDX

GCP Project ID
qwiklabs-gcp-0855e773352d3560

New to labs? View our introductory video!

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 get errors or **incur charges**.

7. Accept the terms and skip the recovery resource page.

   Do not click **End Lab** unless you are finished with the lab or want to restart it.
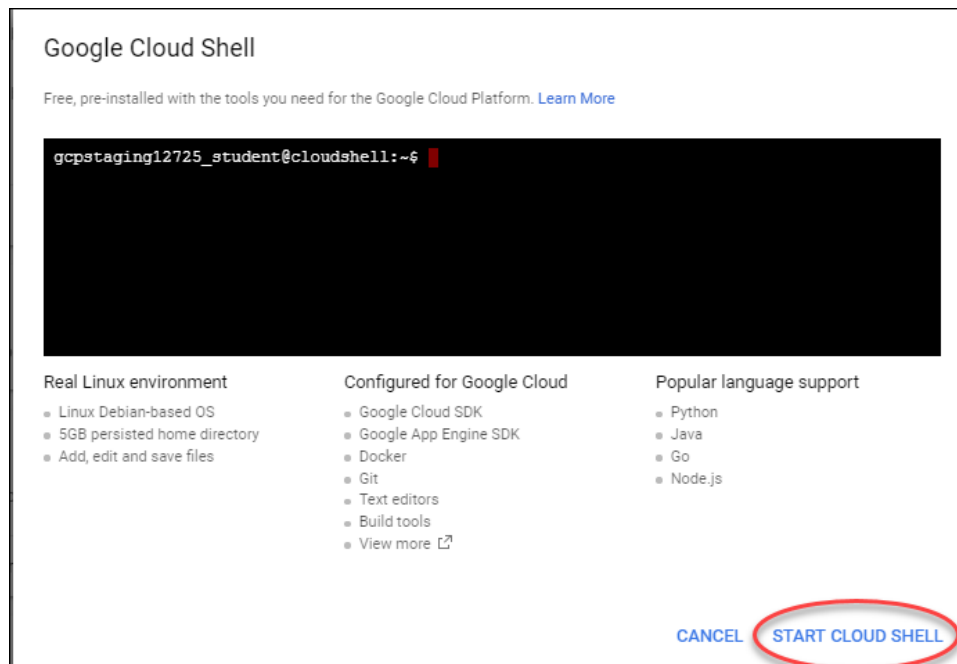   This clears your work and removes the project.

Activate Google Cloud Shell

Google Cloud Shell provides command-line access to your GCP resources.

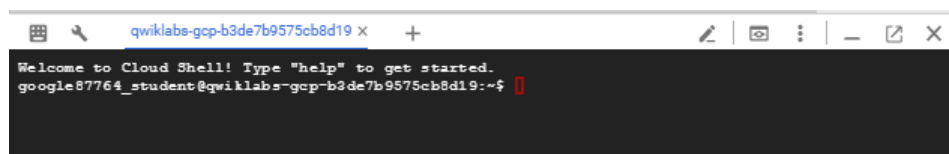From the GCP Console click the **Cloud Shell** icon on the top right toolbar:

Then click **START CLOUD SHELL**:

You can click **START CLOUD SHELL**immediately when the dialog comes up instead of waiting in the dialog until the Cloud Shell provisions.

It takes a few moments to provision and connects to the environment:

The Cloud Shell is a virtual machine loaded with all the development tools you'll need. It offers a persistent 5GB home directory, and runs on the Google Cloud, greatly enhancing network performance and authentication.

Once connected to the cloud shell, you'll see that you are already authenticated and the project is set to your *PROJECT_ID*:

```
gcloud auth list
```

Output:

```
Credentialed accounts:
 - <myaccount>@<mydomain>.com
(active)
```

Note: `gcloud` is the powerful and unified command-line tool for Google Cloud Platform. Full documentation is available on [Google Cloud gcloud Overview](). It comes pre-installed on Cloud Shell and supports tab-completion.

```
gcloud config list project
```

Output:

```
[core]
project = <PROJECT_ID>
```

Launch Cloud Datalab

To launch Cloud Datalab:

1. In **Cloud Shell**, type:

```
gcloud compute zones list
```

2. Pick a zone in a geographically closeby region.

3. In **Cloud Shell**, type:

```
datalab create bdmlvm --zone
<ZONE>
```
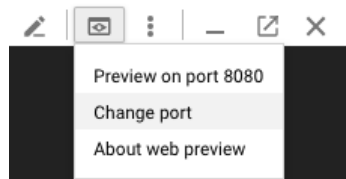
Datalab will take about 5 minutes to start.

```
Note: Follow the prompts during this process.
```

Checkout notebook into Cloud Datalab

If necessary, wait for Datalab to finish launching. Datalab is ready when you see a message prompting you to do a "Web Preview".

1. Click on the **Web Preview** icon on the top-right corner of the Cloud Shell ribbon. Click on the **Change port**. Switch to port **8081** using the **Change Preview Port** dialog box, and then click on **Change and Preview**.

Note: The connection to your Datalab instance remains open for as long as the datalab command is active. If the cloud shell used for running the datalab command is closed or interrupted, the connection to your Cloud Datalab VM will terminate. If that happens, you may be able to reconnect using the command `**datalab connect bdmlvm**` in your new Cloud Shell.

To clone the course repo in your datalab instance:

Step 1

In Cloud Datalab home page (browser), navigate into **notebooks** and add a new notebook using the icon on the top left.
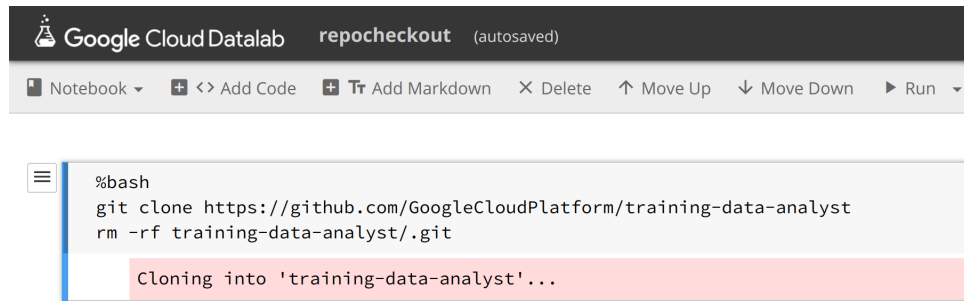
Step 2

Rename this notebook as **repocheckout**.

Step 3

In the new notebook, enter the following commands in the cell, and click on **Run** (on the top navigation bar) to run the commands:

```
%bash
git clone
https://github.com/GoogleCloudPlat
data-analyst
rm -rf training-data-
analyst/.git
```
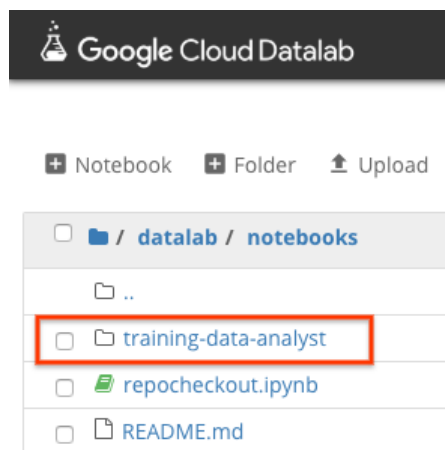
```
%bash
git clone https://github.com/GoogleCloudPlatform/training-data-analyst
rm -rf training-data-analyst/.git
```
```
Cloning into 'training-data-analyst'...
```

Step 4

Confirm that you have cloned the repo by going back to Datalab browser, and ensure you see the **training-data-analyst** directory. All the files for all labs throughout this course are available in this directory.

Open a Datalab notebook

1. In the Datalab browser, navigate to **datalab > notebooks > training-data-analyst > courses > machine_learning > deepdive > 10_recommend > labs > content_based_preproc.ipynb**.

2. Read the commentary, click **Clear | Clear all Cells**, then run the Python snippets (Use **Shift+Enter** to run each piece of code) in the cell, step by step.

3. In the Datalab browser, navigate to **datalab > notebooks > training-data-analyst > courses > machine_learning > deepdive > 10_recommend > labs > content_based_using_neural_networks.ipynb**.

4. Read the commentary, click **Clear | Clear all Cells**, then run the Python snippets (Use **Shift+Enter** to run each piece of code) in the cell, step by step.

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.

Last Tested Date: 12-14-2018

Last Updated Date: 12-14-2018