# Recommendation Systems on Google Cloud

# Collaborative Filtering on Google Analytics data

1 hour 30 minutes Free

## Overview

This lab shows you how to do collaborative filtering with the Weighted Alternating Least Squares (WALS) matrix refactorization approach.

In this lab, you provide article recommendations for users based on the Kurier.at data. Remember that collaborative filtering doesn't need to know anything about the content. You are only interested in the user-item matrix that defines their relationships.

### Lab objectives

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

- Prepare the user-item matrix and use it with WALS.
- Train a `WALSMatrixFactorization` within TensorFlow locally and on an AI Platform.
- Visualize the embedding vectors with principal components analysis.

## 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.

### Activate Cloud Shell

Cloud Shell is a virtual machine that contains development tools. It offers a persistent 5-GB home directory and runs on Google Cloud. Cloud Shell provides command-line access to your Google Cloud resources. `gcloud` is the command-line tool for Google Cloud. It comes pre-installed on Cloud Shell and supports tab completion.

1. Click the **Activate Cloud Shell** button ( ) at the top right of the console.
2. Click **Continue**.
   It takes a few moments to provision and connect to the environment. When you are connected, you are also authenticated, and the project is set to your *PROJECT_ID*.

**Sample commands**

- List the active account name:

gcloud auth list

(Output)

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

(Example output)

Credentialed accounts: - google1623327_student@qwiklabs.net

- List the project ID:

gcloud config list project

(Output)

[core] project = <project_ID>

(Example output)

[core] project = qwiklabs-gcp-44776a13dea667a6 **Note:** Full documentation of **gcloud** is available in the gcloud CLI overview guide.

### Enable the Vertex AI API

1. In the Google Cloud Console, on the **Navigation menu**, click **Vertex AI > Dashboard**.
2. Click **ENABLE ALL RECOMMENDED API**.

### Create a Vertex AI Notebooks instance

- In Cloud Shell, run the following command:

gcloud notebooks instances create tensorflow-notebook --vm-image-project=deeplearning-platform-release --vm-image-family=tf-1-15-cpu --machine-type=n1-standard-1 --location=us-central1-a **Note:** The notebook creation will take 2 - 3 minutes. You may proceed to the next step while the notebook is being created.

# Task 1. Create a storage bucket

*Duration is 2 min*

1. In the Google Cloud Console, on the **Navigation menu** (≡), click **Cloud Storage**.
2. Click + **Create**.
3. Type a unique name for your bucket, such as your project ID.
4. Click **Create**.
5. Confirm **Enforce public access prevention on this bucket** on "Public access will be prevented" pop-up.

# Task 2. Launch a Vertex AI Notebooks instance

1. In the Google Cloud Console, on the Navigation Menu, click **Vertex AI > Workbench**.
2. On the Notebook instances page, you should see the notebook instance you created earlier. You may need to wait if creation is not complete.
3. Once the notebook is ready, click **Open JupyterLab** next to your notebook name. A JupyterLab window will open in a new tab.
4. You may 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. Train a collaborative filtering recommendation system

1. In the notebook interface, navigate to **training-data-analyst > courses > machine_learning > deepdive2 > recommendation_systems > labs**, and open **wals.ipynb**.
2. In the notebook interface, click **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.

**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. 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 > recommendation_systems > solutions**, and open **wals.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.