

Deploy a BigQuery ML Customer Churn Classifier to Vertex AI for Online Predictions

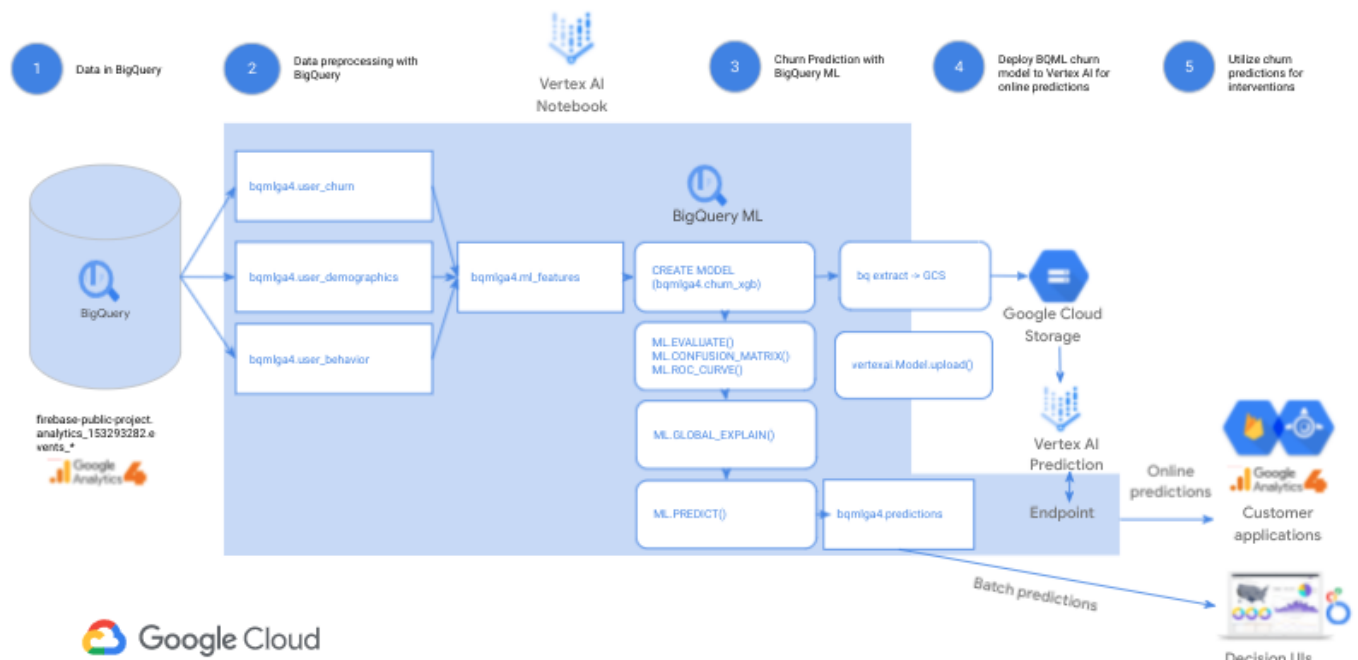
Overview

In this lab, you will train, tune, evaluate, explain, and generate batch and online predictions with a BigQuery ML XGBoost model. You will use a Google Analytics 4 dataset from a real mobile application, Flood it! ([Android app](#), [iOS app](#)), to determine the likelihood of users returning to the application. You will generate batch predictions with your BigQuery ML model as well as export and deploy it to **Vertex AI** for online predictions using the Vertex Python SDK.

[BigQuery ML](#) lets you train and do batch inference with machine learning models in BigQuery using standard SQL queries faster by eliminating the need to move data with fewer lines of code.

[Vertex AI](#) is Google Cloud's complimentary next generation, unified platform for machine learning development. By developing and deploying BigQuery ML machine learning solutions on Vertex AI, you can leverage a scalable online prediction service and MLOps tools for model retraining and monitoring to significantly enhance your development productivity, the ability to scale your workflow and decision making with your data, and accelerate time to value.

Lab Architecture Diagram



Note: BQML is now BigQuery ML.

This lab is inspired by and extends [Churn prediction for game developers using Google Analytics 4 \(GA4\) and BigQuery ML](#). Read the blog post and accompanying tutorial for additional depth on this use case and BigQuery ML.

In this lab, you will go one step further and focus on how Vertex AI extends BigQuery ML's capabilities through online prediction so you can incorporate both customer churn predictions into decision making UIs such as [Looker dashboards](#) but also online predictions directly into customer applications to power targeted interventions such as targeted incentives.

Objectives

- Explore and preprocess a [Google Analytics 4](#) data sample in [BigQuery](#) for machine learning.
- Train a [BigQuery ML XGBoost](#) classifier to predict user churn on a mobile gaming application.
- Tune a BigQuery ML XGBoost classifier using [BigQuery ML hyperparameter tuning features](#).
- Evaluate the performance of a BigQuery ML XGBoost classifier.
- Explain your XGBoost model with [BigQuery ML Explainable AI](#) global feature attributions.
- Generate batch predictions with your BigQuery ML XGBoost model.
- Export a BigQuery ML XGBoost model to a [Google Cloud Storage](#) bucket.
- Upload and deploy a BigQuery ML XGBoost model to a [Vertex AI Prediction](#) Endpoint for online predictions.

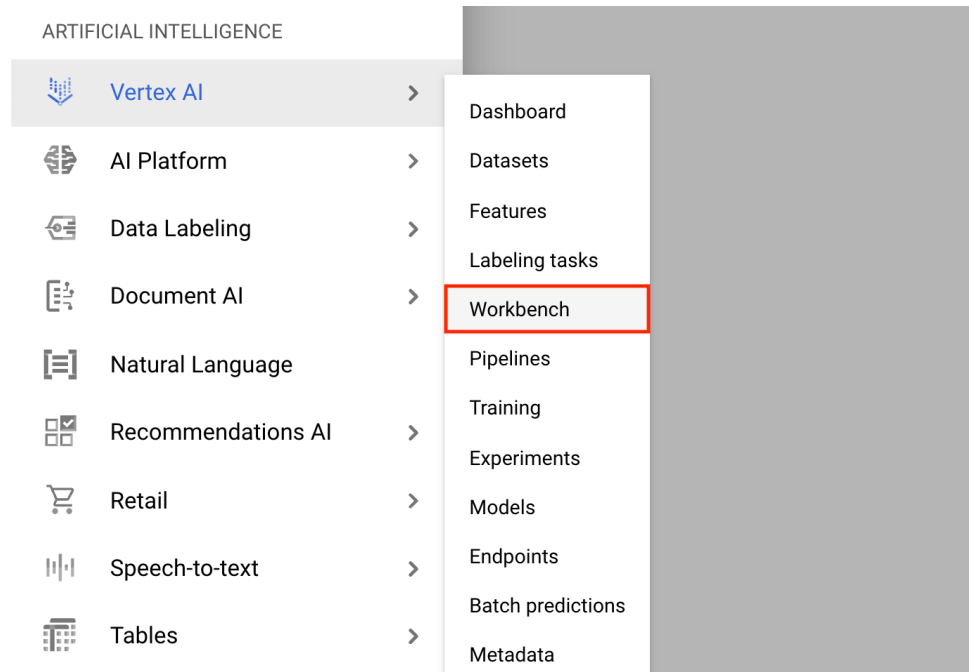
Task 1. Enable Google Cloud services

1. In Cloud Shell, use `gcloud` to enable the services used in the lab:

```
gcloud services enable \  
compute.googleapis.com \  
iam.googleapis.com \  
iamcredentials.googleapis.com \  
monitoring.googleapis.com \  
logging.googleapis.com \  
notebooks.googleapis.com \  
aiplatform.googleapis.com \  
bigquery.googleapis.com
```

Task 2. Deploy Vertex Notebook instance

1. Click on the **Navigation Menu**.
2. Navigate to **Vertex AI**, then to **Workbench**.



3. On the Notebook instances page, navigate to the **User-Managed Notebooks** tab and click **New Notebook**.
4. In the Customize instance menu, select **TensorFlow Enterprise** and choose the version of **TensorFlow Enterprise 2.3 (with LTS) > Without GPUs**.
5. In the **New notebook instance** dialog, for **Region**, select us-central1, for **Zone**, select a zone within the selected region, leave all other fields with their default options, and click **Create**.

After a few minutes, the Vertex AI console will display your instance name, followed by Open Jupyterlab.

6. Click **Open JupyterLab**.

Your notebook is now set up.

Task 3. Clone the lab repository

Next you'll clone the `training-data-analyst` notebook in your JupyterLab instance.

1. In JupyterLab, click the **Terminal** icon to open a new terminal.

Note: If prompted, click **Cancel** for **Build Recommended**.

2. To clone the `training-data-analyst` Github repository, type in the following command, and press **Enter**.

```
cd
git clone https://github.com/GoogleCloudPlatform/training-data-analyst
```

3. To confirm that you have cloned the repository, double-click the `training-data-analyst` directory and confirm that you can see its contents.

Navigate to lab notebook

1. In your notebook, navigate to **training-data-analyst > quests > vertex-ai > vertex-bqml**, and open **lab_exercise.ipynb**.
2. Continue the lab in the notebook, and run each cell by clicking the **Run** (▶) icon at the top of the screen. Alternatively, you can execute the code in a cell with **SHIFT + ENTER**.

Read the narrative and make sure you understand what's happening in each cell.

Task 4. Create a BigQuery dataset

Task 5. Create a BigQuery ML XGBoost churn propensity model

Task 6. Evaluate your BigQuery ML model

Task 7. Batch predict user churn with your BigQuery ML model

Congratulations!

In this lab you trained, tuned, explained, and deployed a BigQuery ML user churn model to Vertex AI to generate high business impact batch and online churn predictions to target customers likely to churn with interventions such as in-game rewards and reminder notifications.