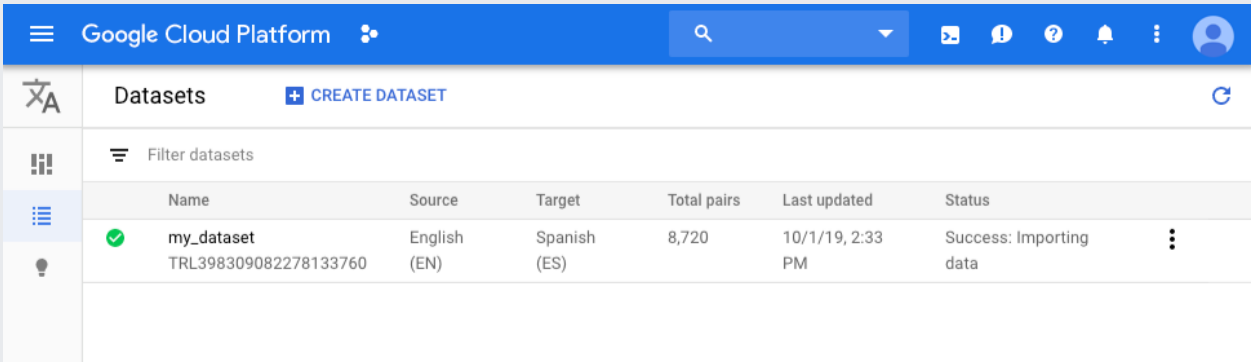


This page shows you how to train and make use of a custom machine translation model by using AutoML Translation UI. It trains a custom English-to-Spanish translation model using technology-oriented sentence pairs from software localization.

Open the [AutoML Translation UI](https://console.cloud.google.com/translation) (<https://console.cloud.google.com/translation>) and select your project from the drop-down list in the upper right of the title bar. (You must have at least **roles/editor** access to the project.) The application walks you through the necessary set-up steps, which are also described in [Before you begin](/translate/automl/docs/before-you-begin) (</translate/automl/docs/before-you-begin>).

1. [Download](/translate/automl/docs/sample/automl-translation-data.zip) (</translate/automl/docs/sample/automl-translation-data.zip>) the archive file containing the sample data for training the model, and extract the file `en-es.tsv`.
2. Visit the [AutoML Translation UI](https://console.cloud.google.com/translation) (<https://console.cloud.google.com/translation>).
3. Select the project for which you enabled AutoML Translation.



The screenshot shows the Google Cloud Platform interface for the 'Datasets' section. The header includes the Google Cloud Platform logo, a search bar, and user profile icons. Below the header, there is a 'Filter datasets' section and a table listing datasets. The table has columns for Name, Source, Target, Total pairs, Last updated, and Status. A single dataset is listed: 'my\_dataset' with 8,720 pairs, last updated on 10/1/19 at 2:33 PM, and a status of 'Success: Importing data'.

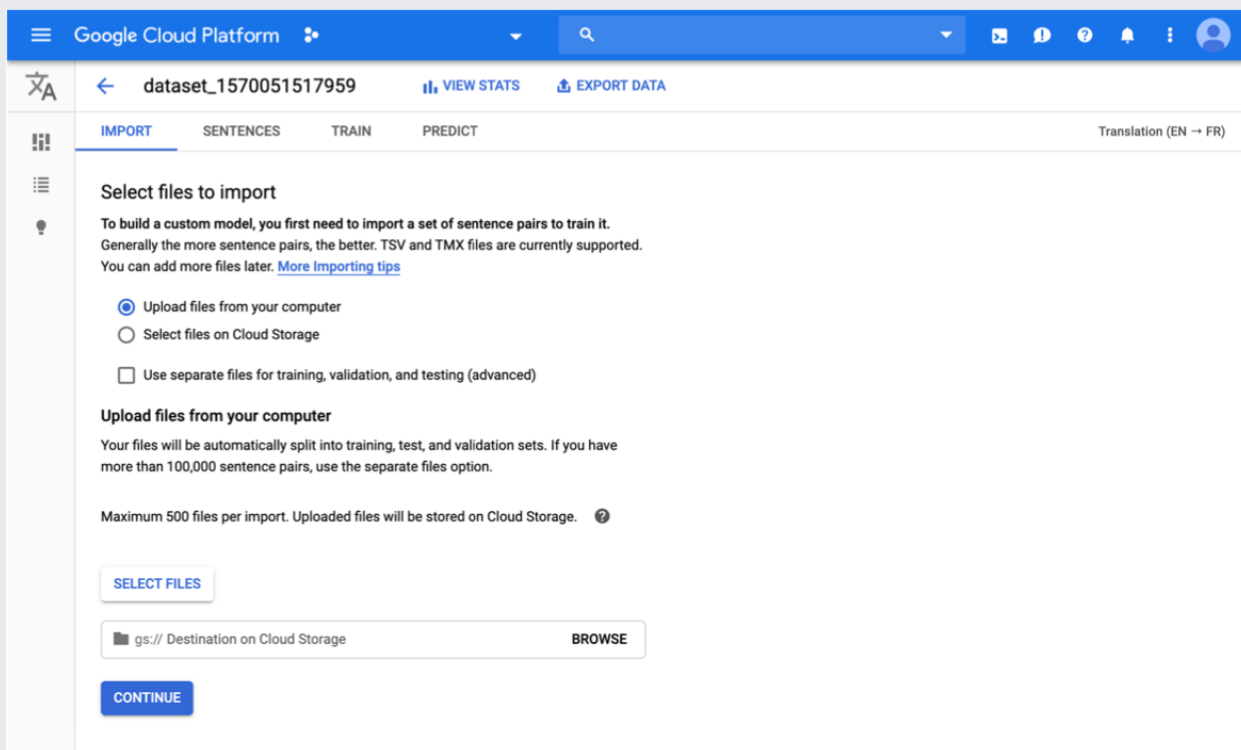
Name	Source	Target	Total pairs	Last updated	Status
my_dataset TRL398309082278133760	English (EN)	Spanish (ES)	8,720	10/1/19, 2:33 PM	Success: Importing data

4. Click the **Create Dataset** button.
5. On the **Create dataset** page, enter a name for the dataset and select the source and target languages.

When you select **English** as the **Translate from** language, the available **Translate to** languages appear. Select **Spanish**.

6. Click **Create**.

7. On the **Import** tab for your dataset, do the following:



- Select **Upload files from your computer**, click **Select Files**, and choose the `en-es.tsv` file you downloaded previously.
- When choosing files from local, you must specify the gcs path where the uploaded files are to be stored. The gcs bucket region must be `us-central1`.

8. Click **Continue**.

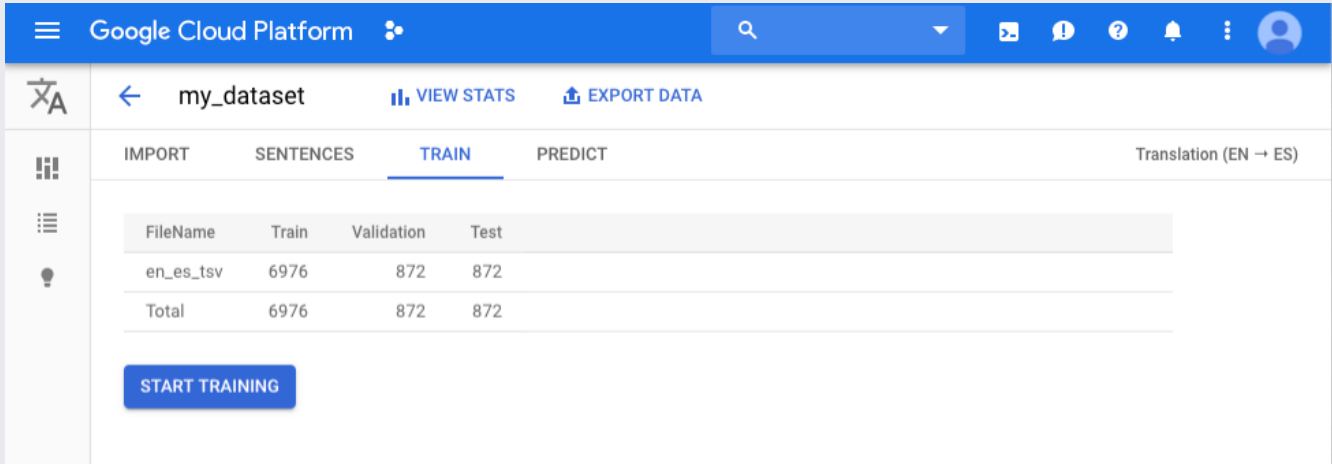
You're returned to the **Datasets** page; your dataset will show an in progress animation while your documents are being imported. When your dataset has been successfully uploaded, you will receive a message at the email address that you used to sign up for the program.

9. Review the dataset.

After your data has been successfully imported, select the dataset from the dataset listing page (or click the link in the email notification) to see the details about the dataset. The

name of the selected dataset appears in the title bar, and the page lists the sentence pairs and which stage of processing they will be used for (**TRAIN, VALIDATION, TEST**).

To begin training your custom model, click the **Train** tab just below the title bar, then the **Start Training** button. Accept the default base model **Google NMT** in the **Train New Model** dialog box.



The screenshot shows the Google Cloud Platform interface for the AutoML Translation service. The top navigation bar is blue and contains the Google Cloud Platform logo, a search bar, and several utility icons. Below the navigation bar, the breadcrumb path is 'my\_dataset', followed by 'VIEW STATS' and 'EXPORT DATA' links. The main content area has four tabs: 'IMPORT', 'SENTENCES', 'TRAIN' (which is selected and underlined), and 'PREDICT'. The 'TRAIN' tab displays a table with the following data:

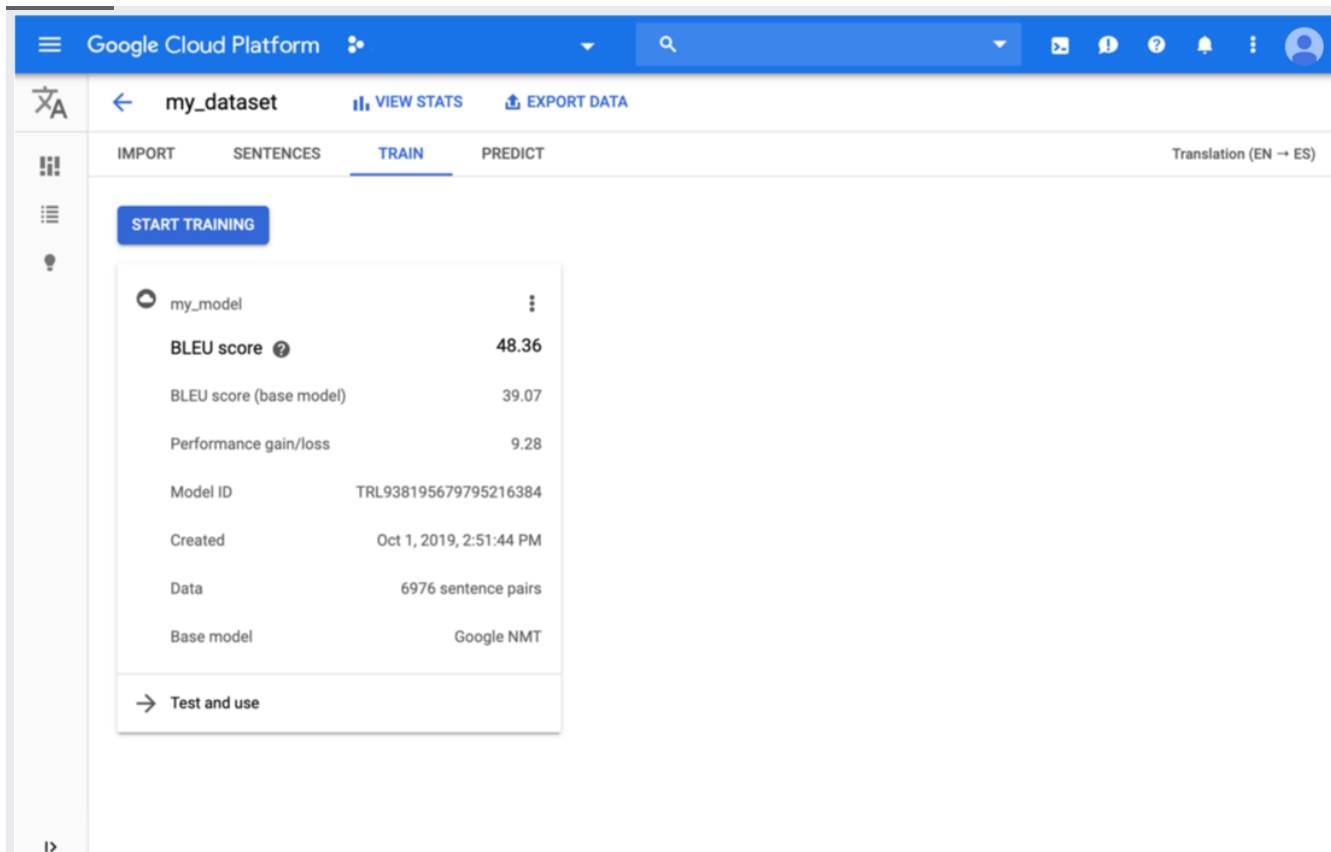
FileName	Train	Validation	Test
en_es_tsv	6976	872	872
Total	6976	872	872

Below the table is a blue button labeled 'START TRAINING'. The top right corner of the main content area indicates the translation direction: 'Translation (EN → ES)'.

Training a model can take several hours to complete. After the model is successfully trained, you will receive a message at the email address you used to sign up for the program.

When you receive notification that training is complete, open the email message and click the link to return to the AutoML Translation UI. The **Train** page shows high-level metrics for the model, most notably its BLEU score. The BLEU (Bilingual Evaluation Understudy).

(/translate/automl/docs/evaluate#bleu) score indicates how similar the candidate text is to the reference texts, with values closer to one representing more similar texts.



The screenshot shows the Google Cloud Platform interface for AutoML Translation. The top navigation bar includes the Google Cloud Platform logo, a search bar, and user profile icons. The main content area is titled 'my\_dataset' and has tabs for 'IMPORT', 'SENTENCES', 'TRAIN', and 'PREDICT'. The 'TRAIN' tab is active, displaying a 'START TRAINING' button and a table of model performance metrics.

Metric	Value
my_model	48.36
BLEU score	48.36
BLEU score (base model)	39.07
Performance gain/loss	9.28
Model ID	TRL938195679795216384
Created	Oct 1, 2019, 2:51:44 PM
Data	6976 sentence pairs
Base model	Google NMT

Below the table, there is a link labeled 'Test and use'.

Click the **Predict** tab just below the title bar or the **Test and use** link below the model information. Enter some text to translate and click the **Translate** button. You can compare the results from your custom model to the Google NMT model.

To avoid unnecessary Google Cloud charges, use the [Cloud Console](https://console.cloud.google.com/) (<https://console.cloud.google.com/>) to delete your project if you do not need it.

- When you're ready to create your own dataset to create an AutoML Translation model, [read the instructions on how to prepare your data](/translate/automl/docs/prepare) (/translate/automl/docs/prepare).

