

AutoML Tables features and capabilities

This product is covered by the [Pre-GA Offerings Terms \(/terms/service-terms#1\)](/terms/service-terms#1) of the Google Cloud Platform Terms of Service. Pre-GA products may have limited support, and changes to pre-GA products may not be compatible with other versions. For more information, see the [launch stage descriptions \(/products#product-launch-stages\)](/products#product-launch-stages).

This page describes how AutoML Tables enables you and your team to build high-performing models from your tabular data.

See our [Known issues \(/automl-tables/docs/known-issues\)](/automl-tables/docs/known-issues) page for current known issues and how to avoid or recover from them.

AutoML Tables is a Service covered by Google's obligations set forth in the [Data Processing and Security Terms \(/terms/data-processing-terms\)](/terms/data-processing-terms).

Data support

AutoML Tables helps you create clean, effective training data by providing information about missing data, correlation, cardinality, and distribution for each of your features. And because there's no charge for importing your data and viewing information about it, you don't incur charges from AutoML Tables until you start training your model.

Feature engineering

When you kick off training, AutoML Tables automatically performs common feature engineering tasks for you, including:

- Normalize and bucketize numeric features.
- Create one-hot encoding and embeddings for categorical features.
- Perform basic processing for text features.
- Extract date- and time-related features from Timestamp columns.

For more information, see [Data preparation that AutoML Tables does for you](/automl-tables/docs/data-best-practices#tables-does) (/automl-tables/docs/data-best-practices#tables-does).

Model training

Parallel model testing

When you kick off training for your model, AutoML Tables takes your dataset and starts training for multiple model architectures at the same time. This approach enables AutoML Tables to determine the best model architecture for your data quickly, without having to serially iterate over the many possible model architectures. The model architectures AutoML Tables tests include:

- Linear
- Feedforward deep neural network
- Gradient Boosted Decision Tree
- AdaNet
- Ensembles of various model architectures

As new model architectures come out of the research community, we will add those as well.

Model evaluation and final model creation

Using your training and validation sets, we determine the best model architecture for your data. We then train two more models, using the parameters and architecture we determined in the parallel testing phase:

1. A model trained with your training and validation sets.

We use your test set to provide the model evaluation on this model.

2. A model trained with your training, validation, and test sets.

This is the model that we provide to you to use to make predictions.

Choosing between AutoML Tables and BigQuery ML

You might want to use BigQuery ML if you are more focused on rapid experimentation or iteration with what data to include in the model and want to use simpler model types for this purpose (such as logistic regression).

You might want to work directly in the AutoML Tables interface if you have already finalized the data, and you:

- Are optimizing for maximizing model quality (accuracy, low RMSE, and so on) without needing to manually do feature engineering, model selection, ensembling, and so on.
- Are willing to wait longer to attain that model quality. AutoML Tables takes at least an hour to train a model, because it experiments with many modeling options. BigQuery ML potentially returns models in minutes because it sticks with the model architectures and parameter values and ranges you set.
- Have a wide variety of feature inputs (beyond numbers and classes) that would benefit from the additional automated feature engineering that AutoML Tables provides.

Model transparency and Cloud Logging

You can view the structure of your AutoML Tables model using Cloud Logging. In Logging, you can see the final model hyperparameters as well as the hyperparameters and objective values used during model validation.

For more information, see [Logging](/automl-tables/docs/logging) (/automl-tables/docs/logging).

Explainability

We know that you need to be able to explain how your data relates to the final model, and to the predictions it makes. We provide you with two primary ways to gain insight into your model and how it operates:

- [Feature importance](/automl-tables/docs/explain) (/automl-tables/docs/explain)
- [Test data export](#) (#export)

Test data export

You can export your test set, along with the predictions your model made. This capability gives you insight into how your model is performing on individual rows of training data. Examining your test set and its results can help you understand what types of predictions your model performs poorly on, and might provide clues into how you can improve your data for a higher-quality model.

Except as otherwise noted, the content of this page is licensed under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) (<https://creativecommons.org/licenses/by/4.0/>), and code samples are licensed under the [Apache 2.0 License](https://www.apache.org/licenses/LICENSE-2.0) (<https://www.apache.org/licenses/LICENSE-2.0>). For details, see the [Google Developers Site Policies](https://developers.google.com/site-policies) (<https://developers.google.com/site-policies>). Java is a registered trademark of Oracle and/or its affiliates.

Last updated 2020-06-30 UTC.