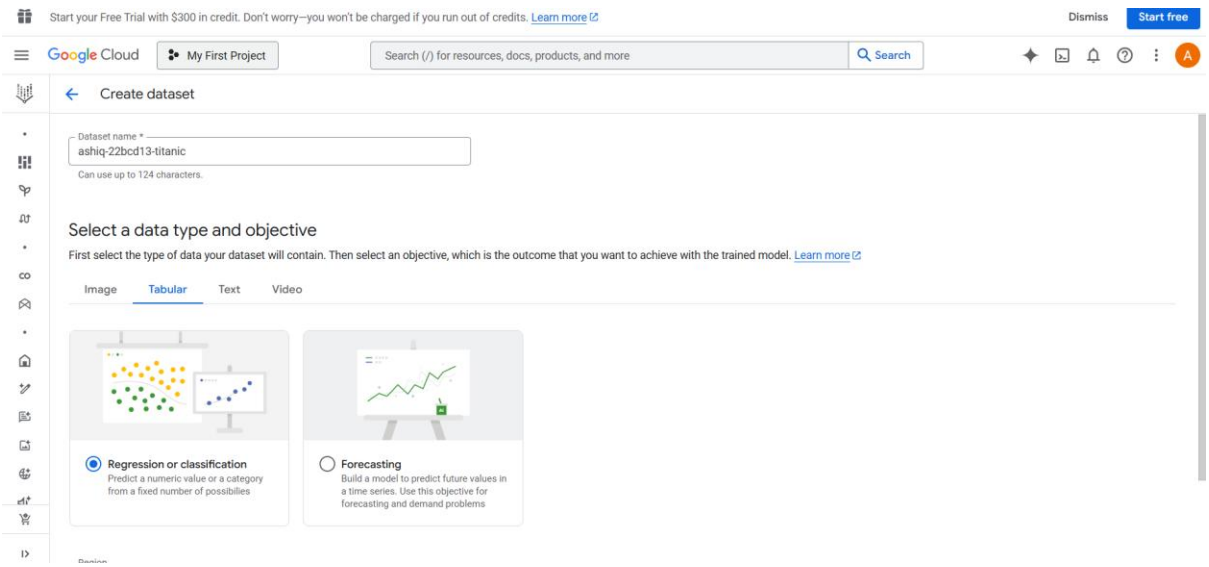
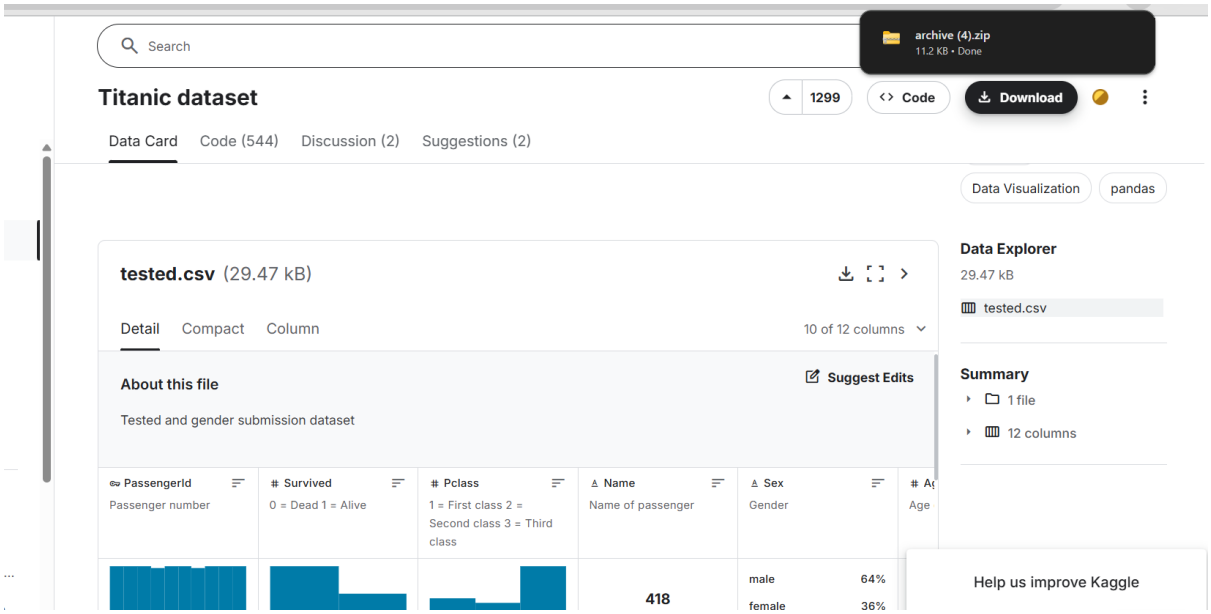
Lab 8
Web Service and Machine Learning in GCP

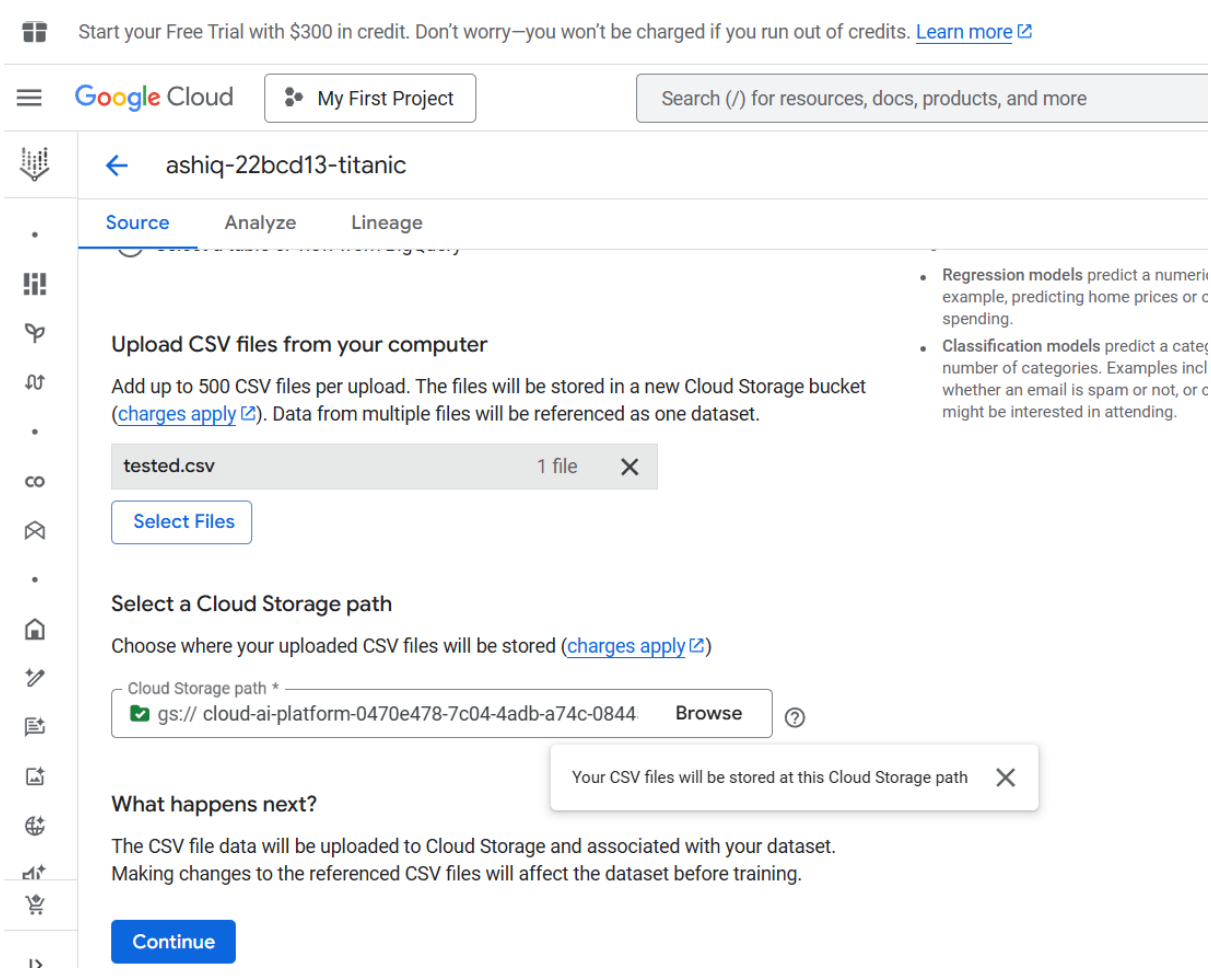
Name: Ashiq Firoz

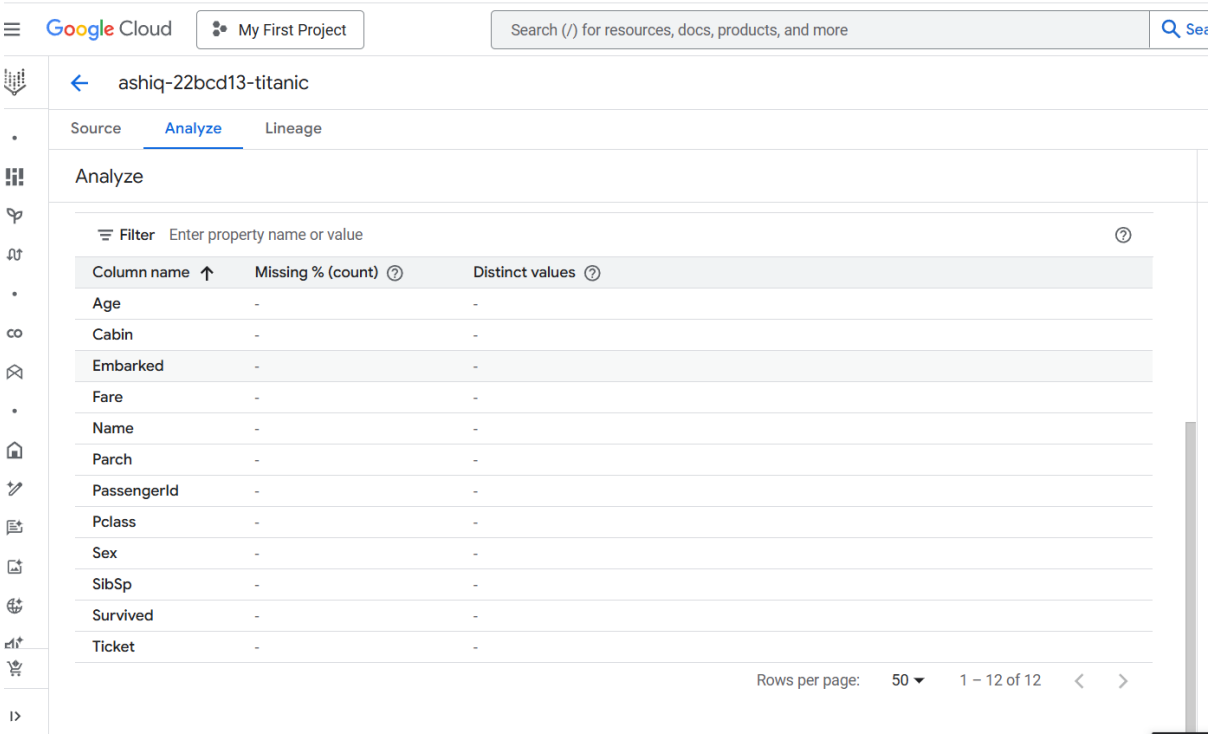
Roll: 2022BCD0013

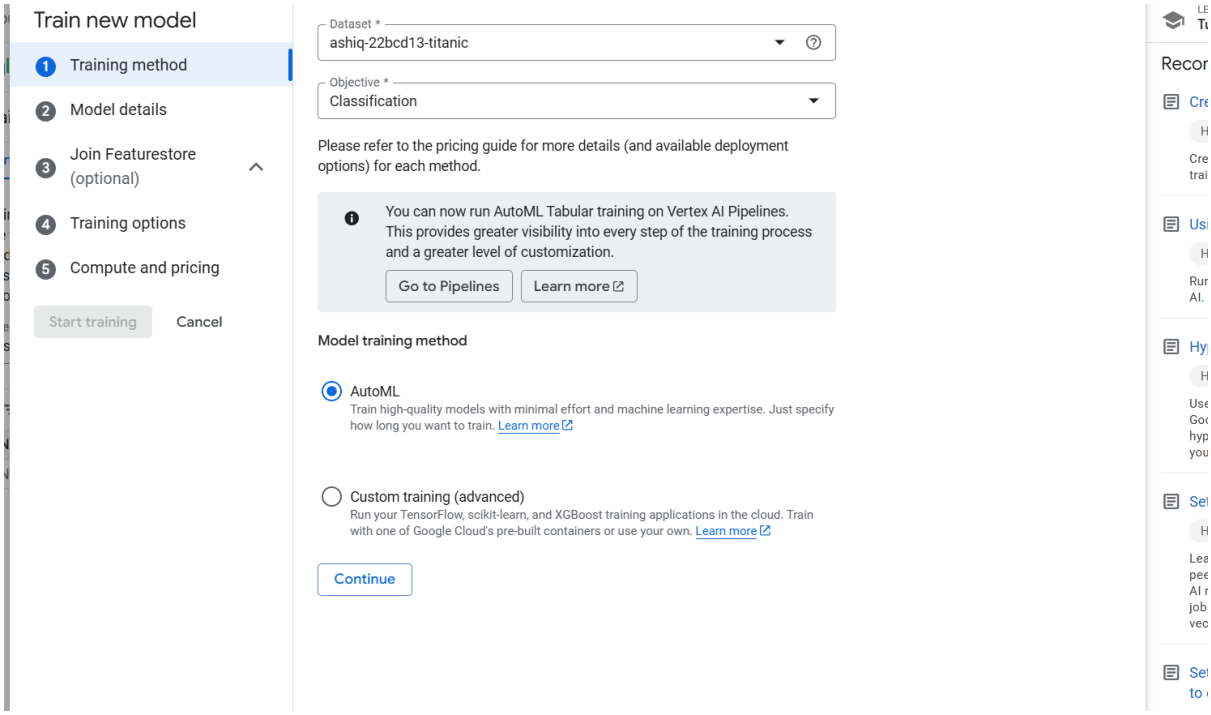
2. Work with a Tabular Dataset to Build and Analyze Machine Learning Models a. Upload and preprocess the dataset, performing Exploratory Data Analysis (EDA) to understand its structure, identify patterns, and handle missing values. b. Utilize Google AutoML to train a machine learning model on the dataset and generate predictions based on the trained models.

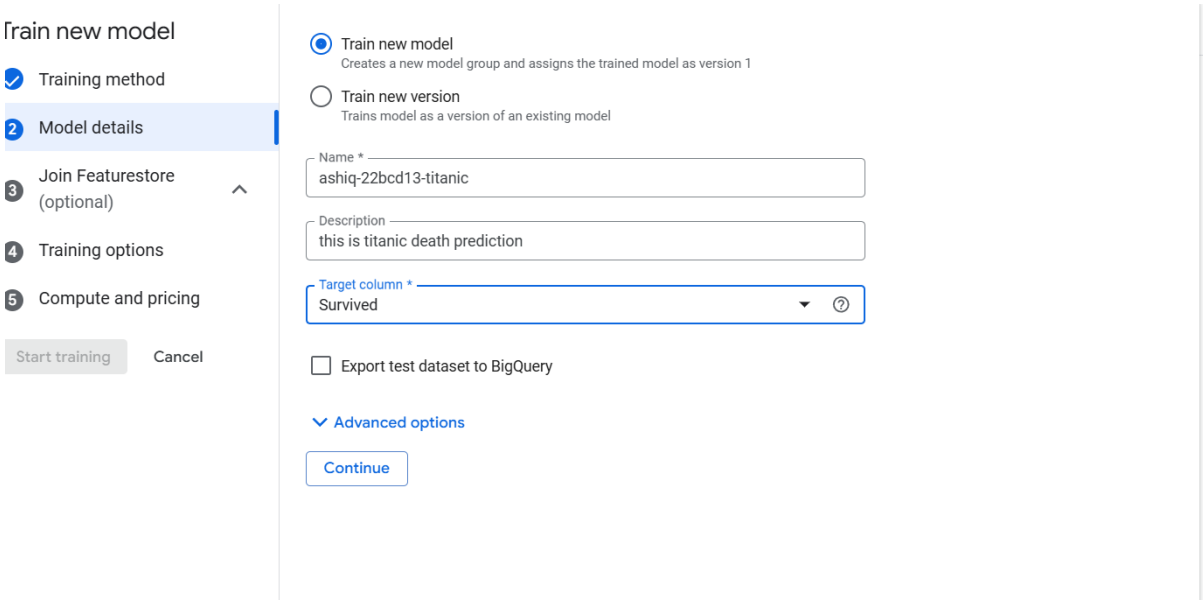


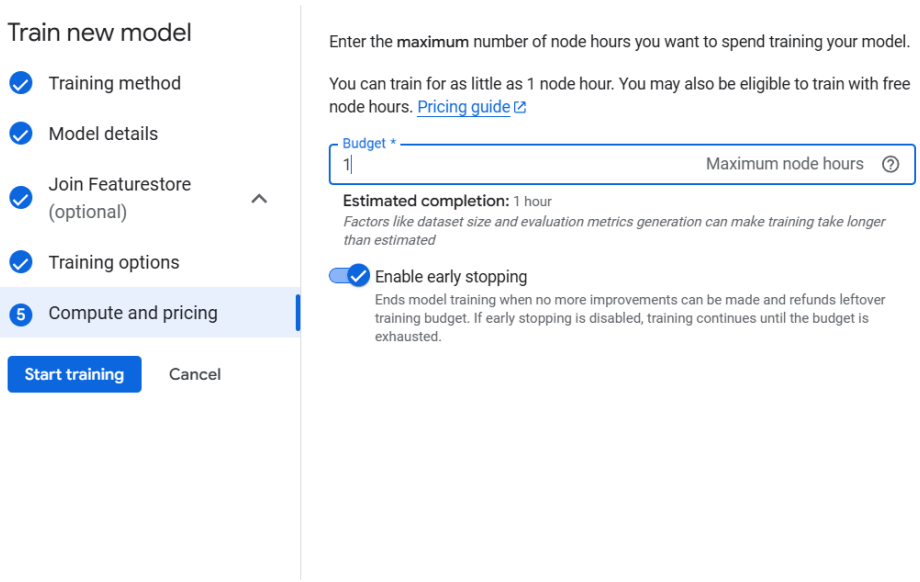

Search
archive (4).zip
11.2 KB • Done
Titanic dataset
1299
Code
Download
Data Card
Code (544)
Discussion (2)
Suggestions (2)
Data Visualization
pandas
tested.csv (29.47 kB)
Detail
Compact
Column
10 of 12 columns
About this file
Suggest Edits
Tested and gender submission dataset
PassengerId
Passenger number
Survived
0 = Dead 1 = Alive
Pclass
1 = First class 2 = Second class 3 = Third class
Name
Name of passenger
Sex
Gender
418
male 64%
female 36%
Data Explorer
29.47 kB
tested.csv
Summary
1 file
12 columns
Help us improve Kaggle

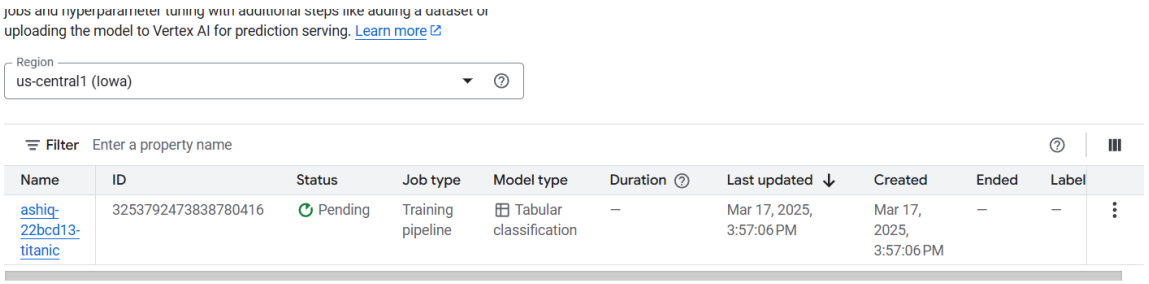

Start your Free Trial with $300 in credit. Don't worry—you won't be charged if you run out of credits. Learn more
Google Cloud
My First Project
Search (/) for resources, docs, products, and more
ashiq-22bcd13-titanic
Source
Analyze
Lineage
Upload CSV files from your computer
Add up to 500 CSV files per upload. The files will be stored in a new Cloud Storage bucket (charges apply). Data from multiple files will be referenced as one dataset.
tested.csv
1 file
Select Files
Select a Cloud Storage path
Choose where your uploaded CSV files will be stored (charges apply)
Cloud Storage path *
gs:// cloud-ai-platform-0470e478-7c04-4adb-a74c-0844
Browse
Your CSV files will be stored at this Cloud Storage path
What happens next?
The CSV file data will be uploaded to Cloud Storage and associated with your dataset.
Making changes to the referenced CSV files will affect the dataset before training.
Continue
Regression models predict a numeri
example, predicting home prices or c
spending.
Classification models predict a categ
number of categories. Examples incl
whether an email is spam or not, or c
might be interested in attending.

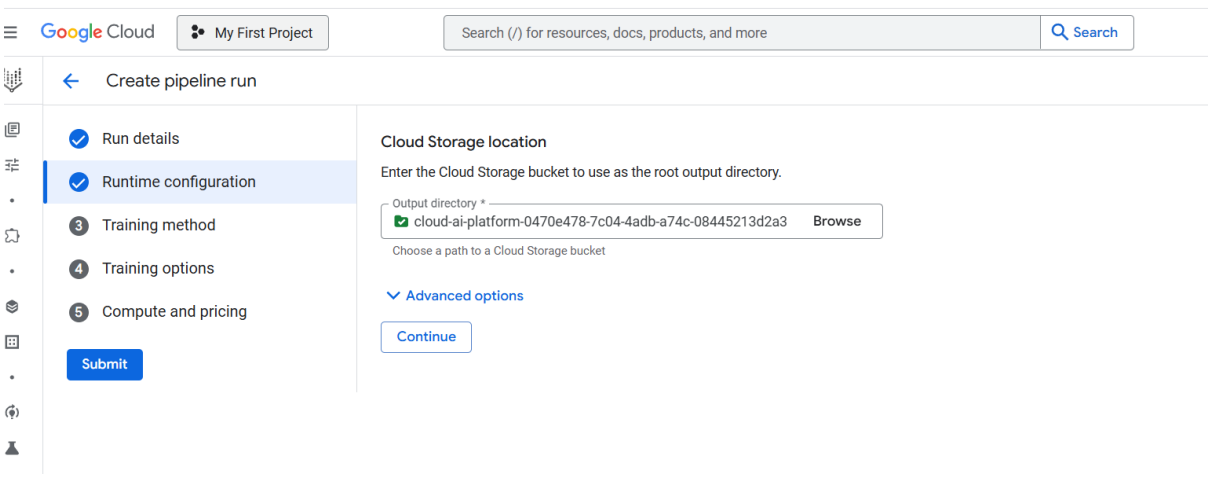

Google Cloud
My First Project
Search (/) for resources, docs, products, and more
ashiq-22bcd13-titanic
Source
Analyze
Lineage
Analyze
Filter
Enter property name or value
Column name
Missing % (count)
Distinct values
Age - -
Cabin - -
Embarked - -
Fare - -
Name - -
Parch - -
PassengerId - -
Pclass - -
Sex - -
SibSp - -
Survived - -
Ticket - -
Rows per page: 50
1 – 12 of 12

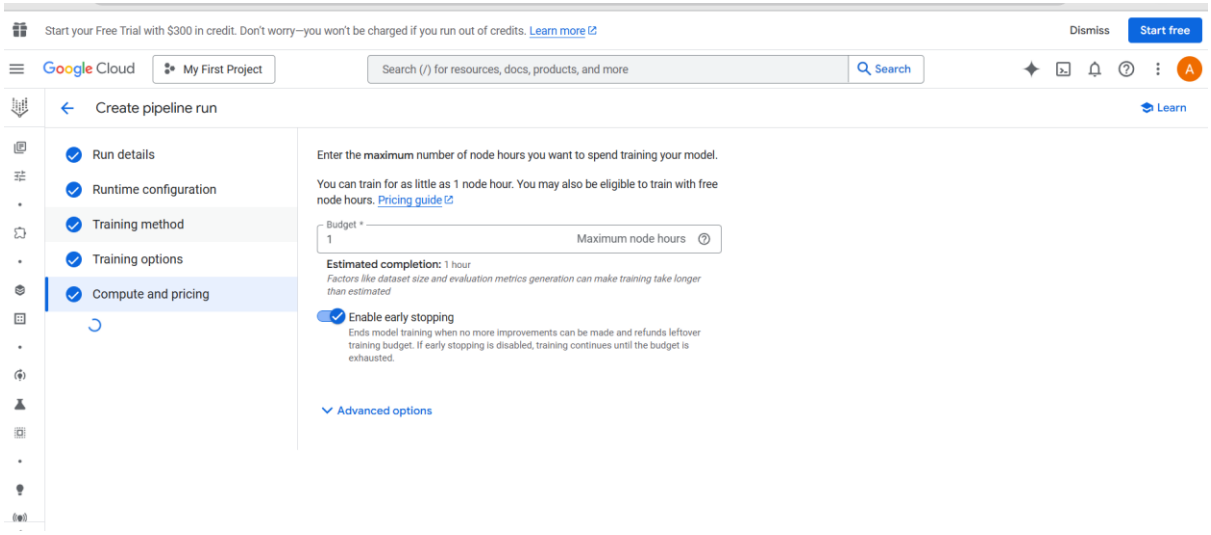

Train new model
1 Training method
2 Model details
3 Join Featurestore (optional)
4 Training options
5 Compute and pricing
Start training
Cancel
Dataset *
ashiq-22bcd13-titanic
Objective *
Classification
Please refer to the pricing guide for more details (and available deployment options) for each method.
You can now run AutoML Tabular training on Vertex AI Pipelines. This provides greater visibility into every step of the training process and a greater level of customization.
Go to Pipelines
Learn more
Model training method
AutoML
Train high-quality models with minimal effort and machine learning expertise. Just specify how long you want to train. Learn more
Custom training (advanced)
Run your TensorFlow, scikit-learn, and XGBoost training applications in the cloud. Train with one of Google Cloud's pre-built containers or use your own. Learn more
Continue

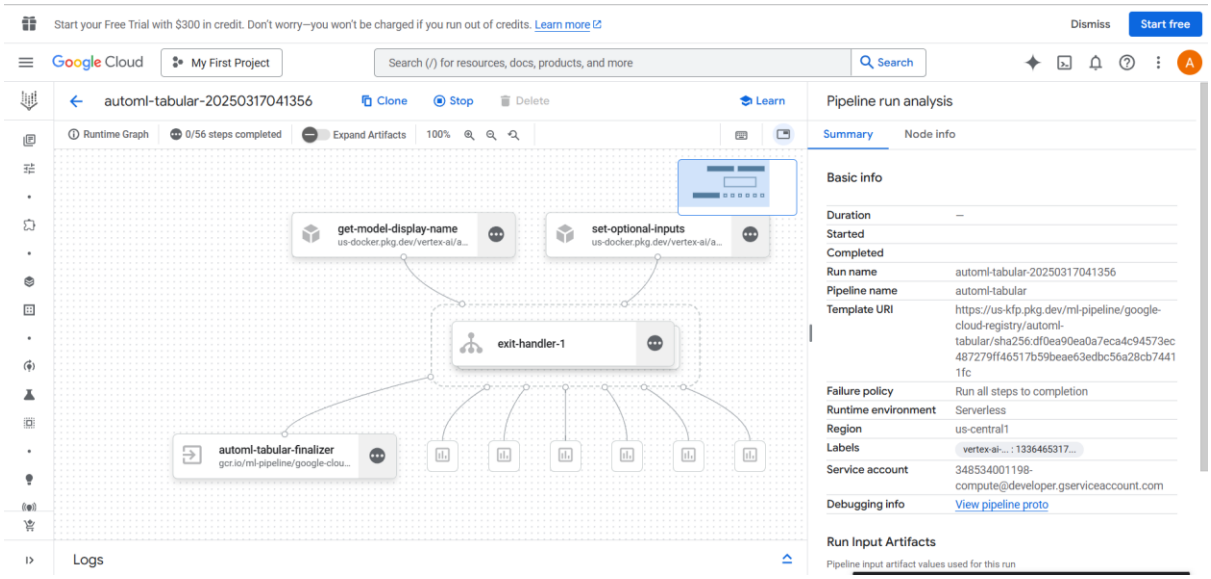

Train new model
Training method
2 Model details
3 Join Featurestore (optional)
4 Training options
5 Compute and pricing
Start training
Cancel
Train new model
Creates a new model group and assigns the trained model as version 1
Train new version
Trains model as a version of an existing model
Name *
ashiq-22bcd13-titanic
Description
this is titanic death prediction
Target column *
Survived
Export test dataset to BigQuery
Advanced options
Continue


Train new model
Training method
Model details
Join Featurestore (optional)
Training options
5 Compute and pricing
Start training
Cancel
Enter the maximum number of node hours you want to spend training your model.
You can train for as little as 1 node hour. You may also be eligible to train with free node hours. Pricing guide
Budget *
1
Maximum node hours
Estimated completion: 1 hour
Factors like dataset size and evaluation metrics generation can make training take longer than estimated
Enable early stopping
Ends model training when no more improvements can be made and refunds leftover training budget. If early stopping is disabled, training continues until the budget is exhausted.


uploading the model to Vertex AI for prediction serving. Learn more
Region
us-central1 (Iowa)
Filter
Enter a property name
Name
ID
Status
Job type
Model type
Duration
Last updated
Created
Ended
Label
ashiq-22bcd13-titanic
3253792473838780416
Pending
Training pipeline
Tabular classification
–
Mar 17, 2025, 3:57:06 PM
Mar 17, 2025, 3:57:06 PM
–
–


Google Cloud
My First Project
Search (/) for resources, docs, products, and more
Search
Create pipeline run
Run details
Runtime configuration
Training method
Training options
Compute and pricing
Submit
Cloud Storage location
Enter the Cloud Storage bucket to use as the root output directory.
Output directory *
cloud-ai-platform-0470e478-7c04-4adb-a74c-08445213d2a3
Browse
Choose a path to a Cloud Storage bucket
Advanced options
Continue


Start your Free Trial with $300 in credit. Don't worry—you won't be charged if you run out of credits. Learn more
Dismiss
Start free
Google Cloud
My First Project
Search (/) for resources, docs, products, and more
Search
Create pipeline run
Learn
Run details
Runtime configuration
Training method
Training options
Compute and pricing
Enter the maximum number of node hours you want to spend training your model.
You can train for as little as 1 node hour. You may also be eligible to train with free node hours. Pricing guide
Budget *
1
Maximum node hours
Estimated completion: 1 hour
Factors like dataset size and evaluation metrics generation can make training take longer than estimated
Enable early stopping
Ends model training when no more improvements can be made and refunds leftover training budget. If early stopping is disabled, training continues until the budget is exhausted.
Advanced options


Start your Free Trial with $300 in credit. Don't worry—you won't be charged if you run out of credits. Learn more
Dismiss
Start free
Google Cloud
My First Project
Search (/) for resources, docs, products, and more
Search
automl-tabular-20250317041356
Clone
Stop
Delete
Learn
Runtime Graph
0/56 steps completed
Expand Artifacts
100%
get-model-display-name
us-docker.pkg.dev/vertex-ai/a...
set-optional-inputs
us-docker.pkg.dev/vertex-ai/a...
exit-handler-1
automl-tabular-finalizer
gcr.io/ml-pipeline/google-clou...
Logs
Pipeline run analysis
Summary
Node info
Basic info
Duration
Started
Completed
Run name
automl-tabular-20250317041356
Pipeline name
automl-tabular
Template URI
https://us-kfp.pkg.dev/ml-pipeline/google-cloud-registry/automl-tabular/sha256:df0ea90ea0a7eca4c94573ec487279ff46517b59beae63edbc56a28cb74411fc
Failure policy
Run all steps to completion
Runtime environment
Serverless
Region
us-central1
Labels
Service account
348534001198-compute@developer.gserviceaccount.com
Debugging info
View pipeline proto
Run Input Artifacts
Pipeline input artifact values used for this run