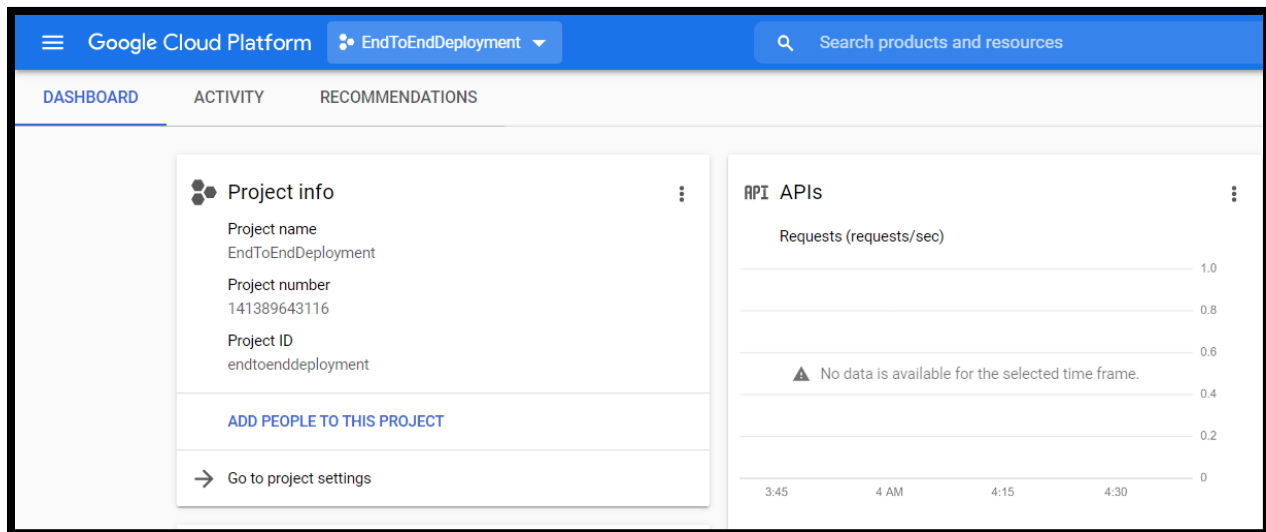
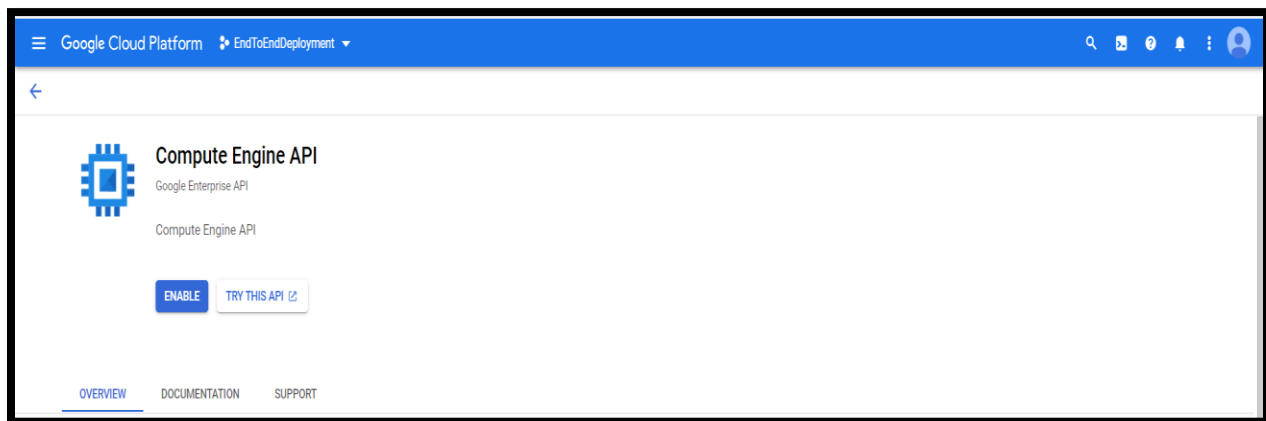


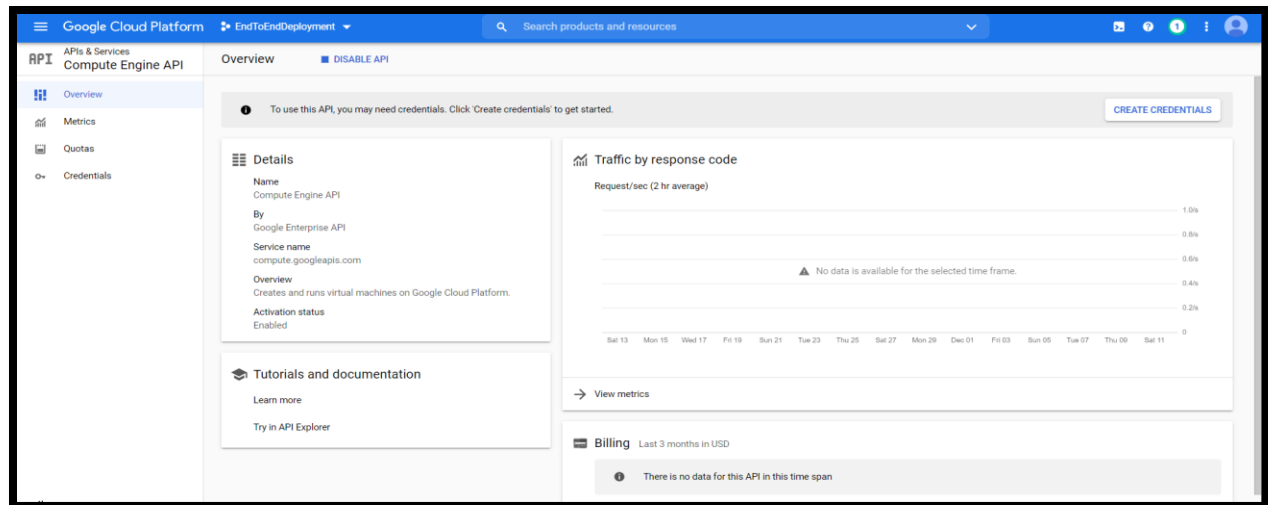
### End-To-End deployment: Fraud detection using Vertex-AI

Step 1: Create new project in GCP.

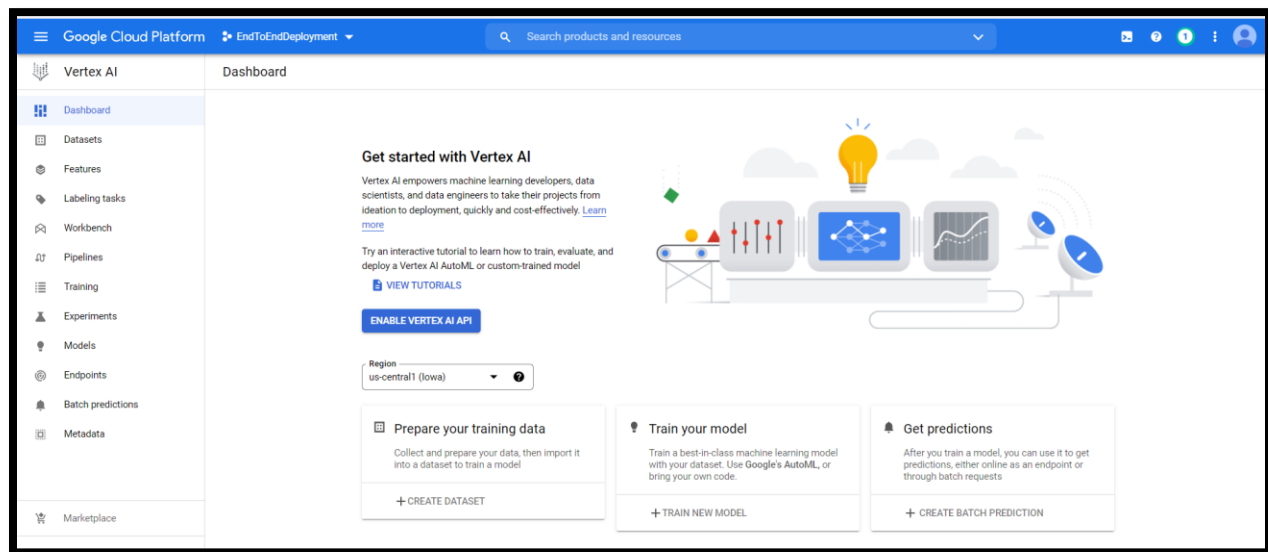


Step 2: Enable compute Engine API





### Step 3: Enable the Vertex AI API



### Step 4: Create a dataset

Google Cloud Platform EndToEndDeployment Search products and resources

Vertex AI

← Create dataset

Dataset name \*  
fraud\_detection  
Can use up to 128 characters.

Select a data type and objective  
First select the type of data your dataset will contain. Then select an objective, which is the outcome that you want to achieve with the trained model. [Learn more about model types](#)

IMAGE TABULAR TEXT VIDEO

☒ Regression/classification  
Predict a target column's value. Supports tables with hundreds of columns and millions of rows.

☐ Forecasting **PREVIEW**  
Predict the likelihood of certain events or demand.

Region  
us-central1 (Iowa)

ADVANCED OPTIONS

CREATE CANCEL

Google Cloud Platform EndToEndDeployment Search products and resources

Vertex AI

← fraud\_detection

SOURCE ANALYZE

Add data to your dataset  
Before you begin, read the [data guide](#) to learn how to prepare your data. Then choose a data source.

Select a data source

- CSV file: Can be uploaded from your computer or on Cloud Storage. [Learn more](#)
- BigQuery: Select a table or view from BigQuery. [Learn more](#)

☐ Upload CSV files from your computer  
☐ Select CSV files from Cloud Storage  
☒ Select a table or view from BigQuery

Select a table or view from BigQuery  
Use a BigQuery table or view as your data source. You'll need [permission to access the dataset](#) and get the [dataset ID](#) and [table ID](#). [Learn more](#)

BigQuery path \*  
☒ bigquery-public-data.ml\_datasets.ulb\_fraud\_detection BROWSE ?  
Enter the qualified id: projectId.datasetId.tableId

What happens next?  
The selected BigQuery table will be associated with your dataset. Making changes to the referenced BigQuery table will affect the dataset before training.

CONTINUE

You can build two model types with tabular data. The model type is automatically chosen based on the data type of your target column.

- Regression models predict a numeric value. For example, predicting home prices or consumer spending.
- Classification models predict a category from a fixed number of categories. Examples include predicting whether an email is spam or not, or classes a student might be interested in attending.

The screenshot displays the Google Cloud Platform Vertex AI interface for the 'fraud\_detection' dataset. The left sidebar contains navigation options: Dashboard, Datasets, Features, Labeling tasks, Workbench, Pipelines, Training, Experiments, Models, Endpoints, Batch predictions, and Metadata. The main content area is divided into 'SOURCE' and 'ANALYZE' tabs, with 'ANALYZE' selected. The 'ANALYZE' tab shows 'Dataset Info' (Created: Dec 12, 2021 4:51 AM, Dataset format: BigQuery, Dataset location(s): bq://bigquery-public\_b.fraud\_detection) and a 'Summary' section with a bar chart showing the distribution of data types: FLOAT (30, 96.77%) and INTEGER (1, 3.23%). Below this is a 'Filter' section with a search bar and a table of dataset columns. The table has columns for Column name, BigQuery type, BigQuery mode, Missing % (count), and Distinct values. The columns listed are Amount, Class, Time, V1, V10, V11, V12, V13, V14, V15, V16, V17, and V18. All columns are of type FLOAT and mode NULLABLE, with missing percentages of 0% and distinct values of 1. A 'GENERATE STATISTICS' button is located to the right of the table. On the far right, the 'Training jobs and models' section provides instructions on how to use the dataset and annotation set to train a new machine learning model with AutoML or custom code, accompanied by a 'TRAIN NEW MODEL' button.

Column name	BigQuery type	BigQuery mode	Missing % (count)	Distinct values
Amount	FLOAT	NULLABLE	-	-
Class	INTEGER	NULLABLE	-	-
Time	FLOAT	NULLABLE	-	-
V1	FLOAT	NULLABLE	-	-
V10	FLOAT	NULLABLE	-	-
V11	FLOAT	NULLABLE	-	-
V12	FLOAT	NULLABLE	-	-
V13	FLOAT	NULLABLE	-	-
V14	FLOAT	NULLABLE	-	-
V15	FLOAT	NULLABLE	-	-
V16	FLOAT	NULLABLE	-	-
V17	FLOAT	NULLABLE	-	-
V18	FLOAT	NULLABLE	-	-

## Step 5: Model Training

### Train new model

1 Training method

2 Model details

3 Training options

4 Compute and pricing

START TRAINING CANCEL

Dataset  
fraud\_detection

Objective \*  
Classification

Please refer to the pricing guide for more details (and available deployment options) for each 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 Training options
- 4 Compute and pricing

START TRAINING CANCEL


Model name \*  
fraud\_detection\_20211212125512 ?

Target column \*  
Class (INTEGER) ▼ ?

☐ Export test dataset to BigQuery

**Data split**

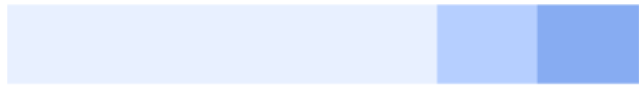
☒ Random assignment  
80% of your data is randomly assigned for training, 10% for validation and 10% for testing.



☐ Manual  
You assign each data row for training, validation, and testing. [Learn more](#)

☐ Chronological assignment  
The earliest 80% of your data is assigned to training, the next 10% for validation and the latest 10% for testing. This option requires a Time column in your dataset. [Learn more](#)

● Training 80% ● Validation 10% ● Testing 10%



Start time End time

**Encryption**

☐ Use a customer-managed encryption key (CMEK)

[^ SHOW LESS](#)

CONTINUE

### Train new model

- ✓ Training method
- ✓ Model details
- 3 Training options**
- 4 Compute and pricing

START TRAINING CANCEL

<input type="checkbox"/>	V3	Automatic ▼	FLOAT	NULLABLE
<input type="checkbox"/>	V4	Automatic ▼	FLOAT	NULLABLE
<input type="checkbox"/>	V5	Automatic ▼	FLOAT	NULLABLE
<input type="checkbox"/>	V6	Automatic ▼	FLOAT	NULLABLE
<input type="checkbox"/>	V7	Automatic ▼	FLOAT	NULLABLE
<input type="checkbox"/>	V8	Automatic ▼	FLOAT	NULLABLE
<input type="checkbox"/>	V9	Automatic ▼	FLOAT	NULLABLE

Total 31 feature columns are included in the training

#### Weight column

Select a column ▼

#### Optimization objective

☐ AUC ROC  
Distinguish between classes

☐ Log loss  
Keeps prediction probabilities as accurate as possible

☒ AUC PRC  
Maximize precision-recall for the less common class


☐ Precision

☐ Recall

[^ SHOW LESS](#)

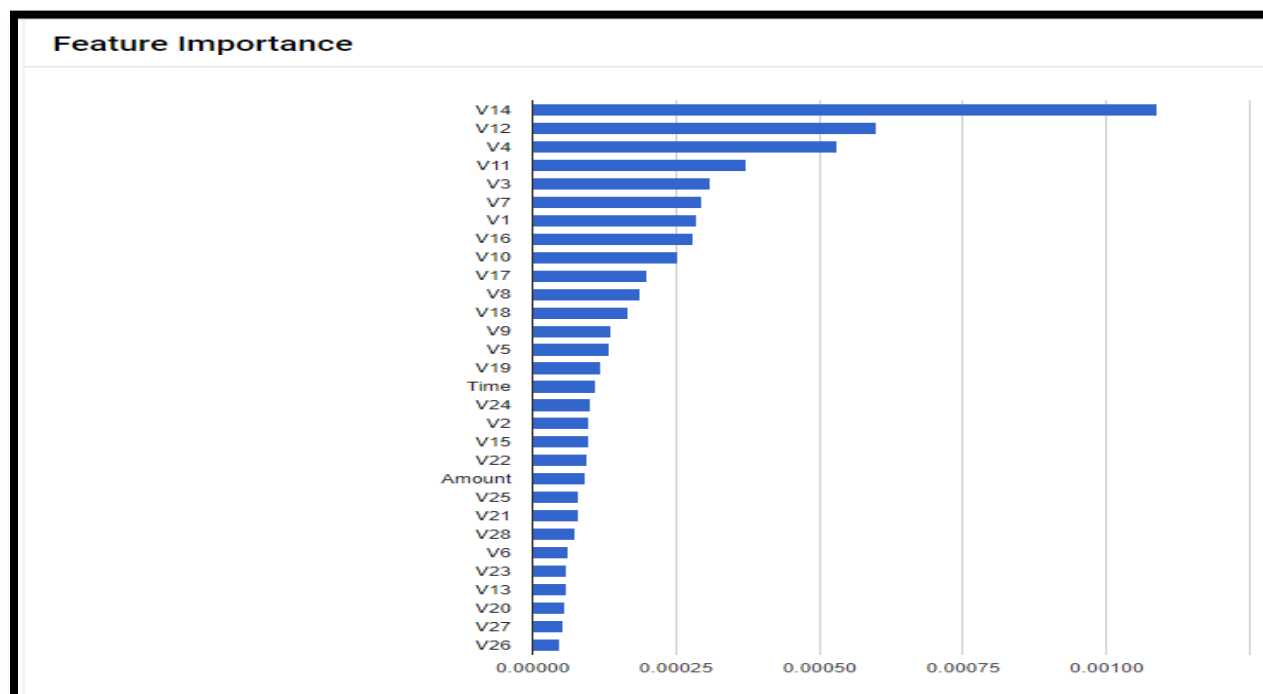
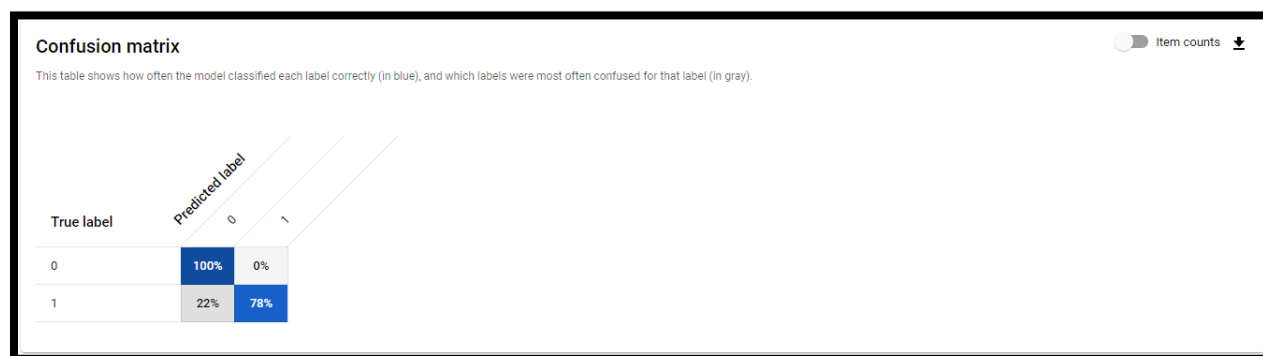
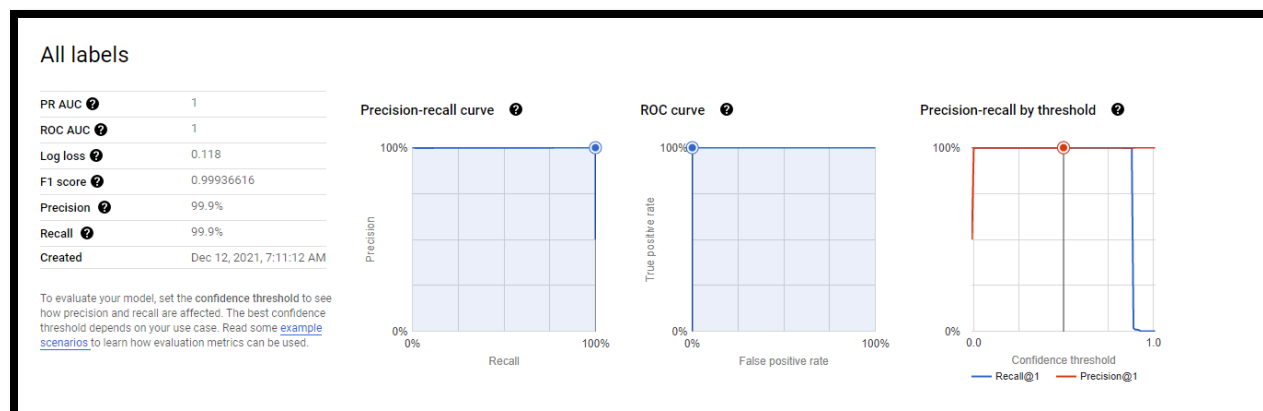
CONTINUE

### Training jobs and models

 **fraud\_detection\_20211212125512**  
Training model...

**TRAIN NEW MODEL**

## Step 6: Understand the confusion matrix after training.





## Step 7: Creating an end point for deployment.

The screenshot shows the Vertex AI console interface. At the top, there's a breadcrumb navigation: [← fraud\\_detection\\_20211212125512](#) [VIEW DATASET](#) [EXPORT](#). Below this is a tab bar with 'EVALUATE', 'DEPLOY & TEST' (selected), 'BATCH PREDICTIONS', and 'MODEL PROPERTIES'. The main content area is titled 'Use your edge-optimized model' and contains a 'Container' section with a description: 'Export your model as a TF Saved Model to run on a Docker container.' Below this is a 'Deploy your model' section with a description: 'Endpoints are machine learning models made available for online prediction requests. Endpoints are useful for timely predictions from many users (for example, in response to an application request). You can also request batch predictions if you don't need immediate results.' A blue 'DEPLOY TO ENDPOINT' button is present. Below the button is a table with columns: Name, ID, Status, Models, Region, Monitoring, Most recent monitoring job, Most recent alerts, Last updated ↓, API, Notification, Labels ⓘ, and Encryption. The table currently shows 'No active endpoints containing this model'. At the bottom, there's a 'Test your model' section with a 'PREVIEW' button and a note: 'In order to test your model, you will need to deploy it first. [Pricing guide](#)'.

The screenshot shows the 'Deploy to endpoint' configuration screen. On the left, there's a sidebar with three steps: 'Define your endpoint' (checked), 'Model settings' (checked), and 'Model monitoring' (selected with a blue bar). Below the sidebar are 'DEPLOY' and 'CANCEL' buttons. On the right, there's a grey box with an information icon and the text: 'Settings in this step apply to all models deployed to the endpoint'. Below this is the 'Model monitoring' section, which includes a description: 'You can monitor the tabular and custom models deployed to this endpoint for changes in feature drift, training-serving skew and other objectives that help you understand how your model is performing to real world data.' At the bottom, there's a toggle switch for 'Enable model monitoring for this endpoint', which is currently turned off.

← fraud\_detection\_20211212125512 [VIEW DATASET](#) [EXPORT](#)

EVALUATE **DEPLOY & TEST** BATCH PREDICTIONS MODEL PROPERTIES

### Deploy your model

Endpoints are machine learning models made available for online prediction requests. Endpoints are useful for timely predictions from many users (for example, in response to an application request). You can also request batch predictions if you don't need immediate results.

[DEPLOY TO ENDPOINT](#)

Name	ID	Status	Models	Region	Monitoring	Most recent monitoring job	Most recent alerts	Last updated ↓	API	Notification	Labels ⓘ	Encryption
<a href="#">fraud_v1</a>	1658529727616385024	Active	1	us-central1	Disabled	—	—	Dec 12, 2021, 4:08:32 PM	<a href="#">Sample request</a>			Google-managed key

### Test your model PREVIEW

Feature column name	Type	Required or optional	Value	Local feature importance
Time	Numerical	Required	<input type="text" value="84730"/>	—
V1	Numerical	Required	<input type="text" value="0.0204233415795775"/>	—
V2	Numerical	Required	<input type="text" value="0.0658476451601173"/>	—
V3	Numerical	Required	<input type="text" value="0.18147056566326897"/>	—
V4	Numerical	Required	<input type="text" value="-0.0181245171894877"/>	—

Predicted column not yet known

Prediction result

—

### Predict label

#### Prediction result

Selected label

Baseline prediction value: 0.8852328062057495

Confidence score: 0.8852328062057495

### Predict label

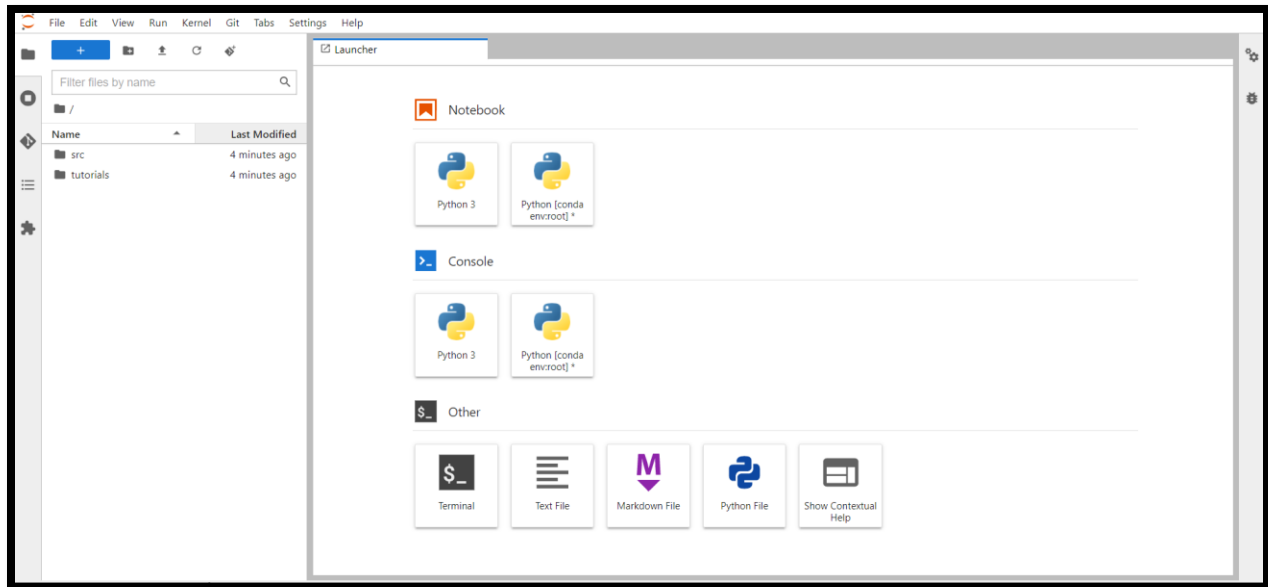
#### Prediction result

Selected label

Baseline prediction value: 0.1147671565413475

Confidence score: 0.1147671565413475

## Step 8: Prediction using vertex AI



```
response = endpoint.predict([test_instance])
print('API response: ', response)

API response: Prediction(predictions=[{'classes': ['0', '1'], 'scores': [0.8859193921089172, 0.1140805631875992]}], deployed_model_id='7420
920635209023488', explanations=None)
```

Complete Code is available in the Notebook PDF attached in same folder