



Light House Projects Season-2

GCP- HANDS ON

AI MODEL BUILDING AND DEPLOYMENT

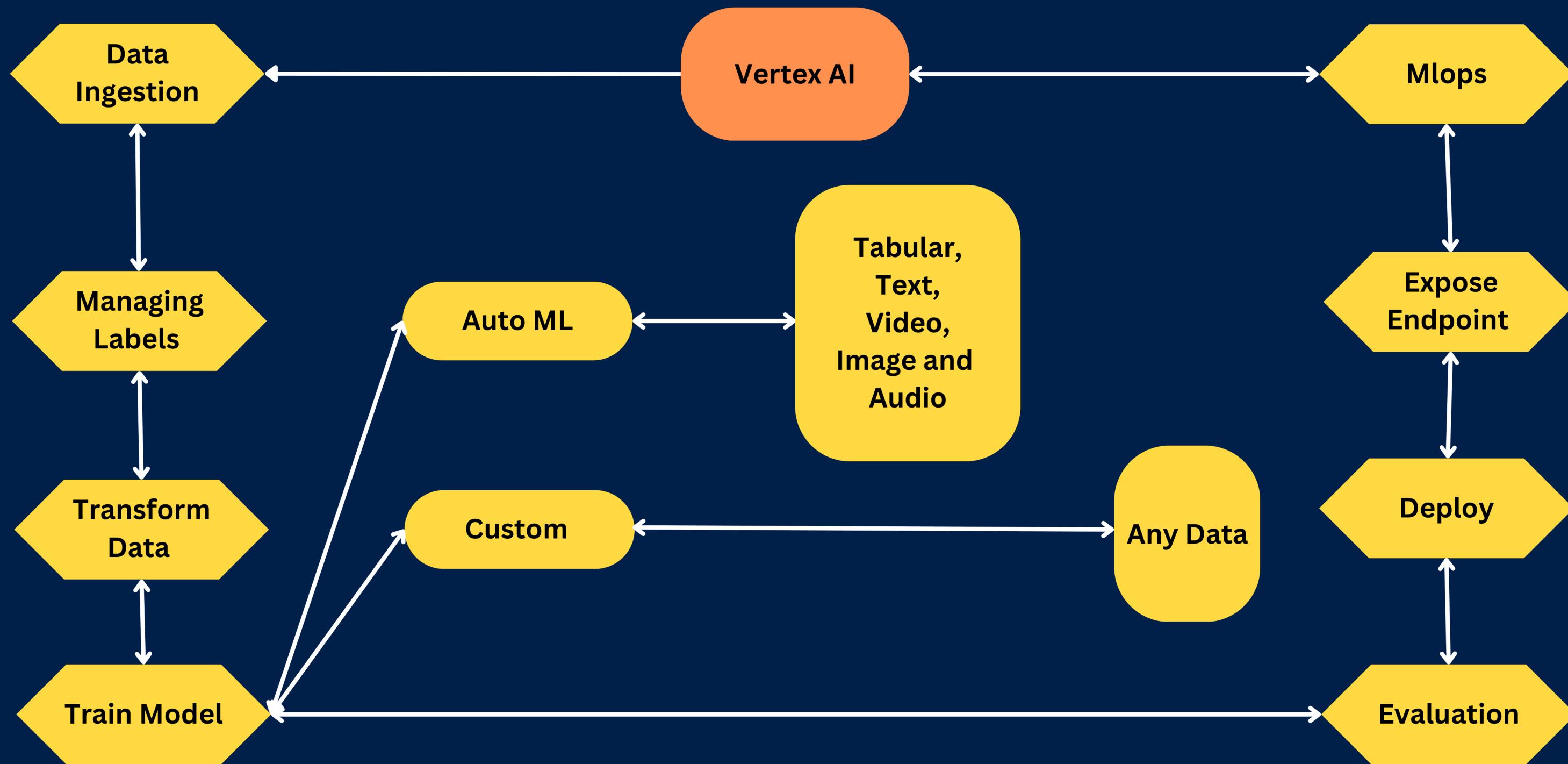


ROOPESH BHARATWAJ K R.
DATA SCIENTIST-II

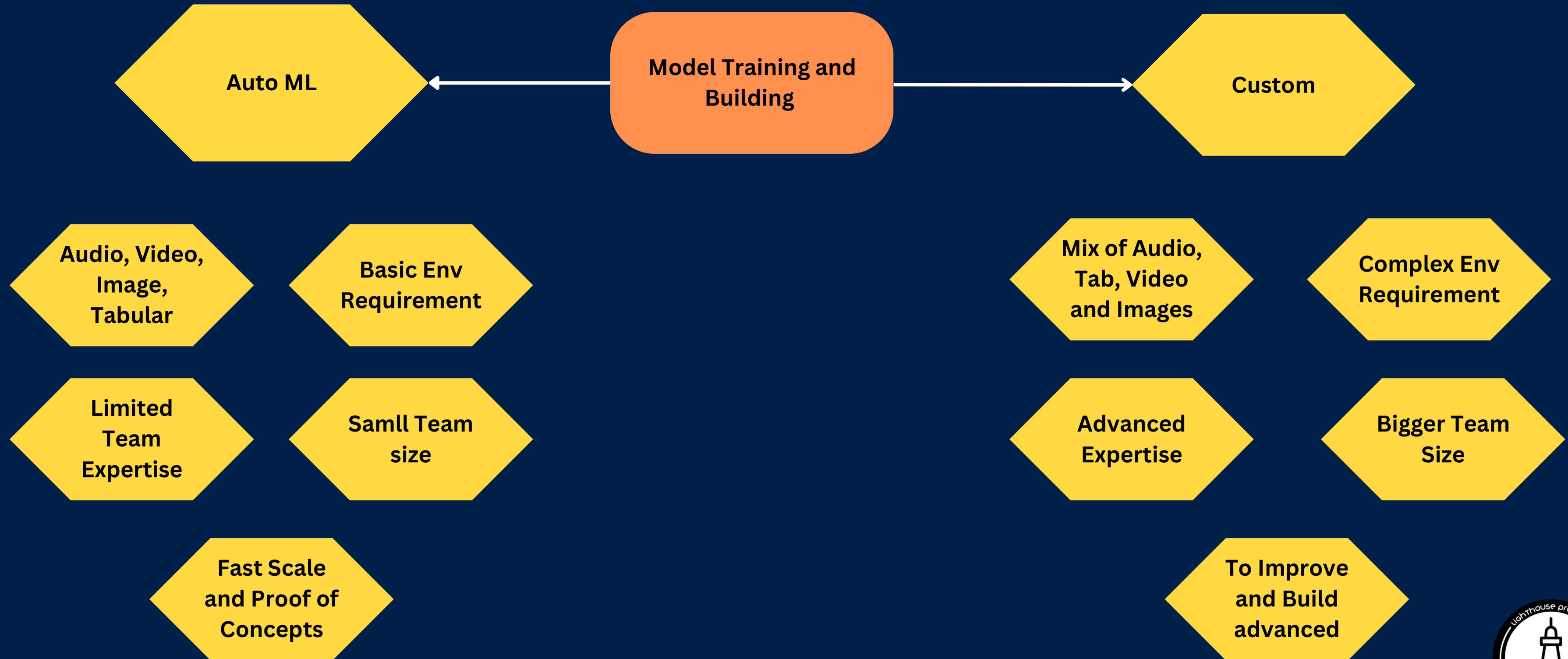




Vertex AI - Life Cycle



Building AI Modells



Auto ML



Step-1. Data Ingestion

The screenshot shows the Google Cloud Vertex AI interface for creating a new dataset. The top navigation bar includes 'Google Cloud' and 'My First Project'. A search bar at the top right says 'Search for resources, documents, products and more (/)' with a 'Seek' button. On the left, a sidebar lists 'Vertex AI' under 'TOOLS', and sections for 'Generative AI Studio', 'Data', and 'Model Development'. Under 'Data', 'Datasets' is selected. The main area is titled 'Create data set' and shows a 'dataset name' input field containing 'untitled_1688117511499'. Below it, a note says 'Can contain up to 128 characters.' A section titled 'Select data type and destination' provides options: 'PICTURE', 'TABULAR' (which is selected), 'TEXT', and 'VIDEO'. Under 'TABULAR', two options are shown: 'regression/classification' (selected) and 'forecast'. A note for 'regression/classification' says: 'Predict the value of a target column. Tables with hundreds of columns and millions of rows are supported.' A note for 'forecast' says: 'Predict the likelihood of certain events or the need.' At the bottom, a 'region' dropdown is set to 'us-west1 (Oregon)'.

Step-2. Select Data Source

The screenshot shows a machine learning model configuration interface. At the top, there are tabs for 'sample' (with a back arrow), 'SOURCE' (selected), and 'ANALYZE'. A 'TRAIN NEW MODEL' button is in the top right. The 'SOURCE' tab contains a 'Change data source' section with a note: 'Read the [guide](#) first so you know how to prepare the data. Then select a data source.' Below it is a 'Select data source' section with two bullet points: 'CSV file : Can be uploaded from your computer or on Cloud Storage. [learn more](#)' and 'BigQuery : Select a table or view from BigQuery. [learn more](#)'. Underneath are three radio button options: 'Upload CSV files from your computer' (selected), 'Select CSV files from Cloud Storage', and 'Select a table or view from BigQuery'. The 'Conclusion' section on the right shows two house icons with price tags: '\$625,000' and '\$975,000'. A note below says: 'You can create two types of models with tabular data. The model type is automatically selected based on the data type of the target column.' At the bottom, a 'CHOOSE FILES' button is visible.



Step-3. Selected Data

Click Train New Model



Step-4. Auto ML Training

Train new model

- Training method
- Model details
- Training options
- Compute and pricing

Dataset: fraud-bigquery
Objective *: Classification

Please refer to the pricing guide for more details (and available deployment options) for each method.

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

START TRAINING CANCEL

Step-5. Select Model Name and Target Column

Train new model

Training method
 Model details
 Training options
 Compute and pricing

Train new model

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 *: fraud-bigquery

Description

Target column *: Class (INTEGER)

Export test dataset to BigQuery

ADVANCED OPTIONS

CONTINUE



Step-6. Select Advanced Options

Train new model

Training method

Model details

3 Training options

4 Compute and pricing

START TRAINING CANCEL

	V18	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V19	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V2	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V20	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V21	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V22	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V23	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V24	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V25	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V26	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V27	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V28	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V3	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V4	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V5	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V6	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V7	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V8	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V9	Numeric ▾	FLOAT	NULLABLE	-

Rows per page: 50 ▾

Total of 31 feature columns are included in the training

ADVANCED OPTIONS

CONTINUE

Step-7. Select AUC PRC

Train new model

Training method

Model details

3 Training options

4 Compute and pricing

START TRAINING CANCEL

	V4	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V5	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V6	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V7	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V8	Numeric ▾	FLOAT	NULLABLE	-
<input type="checkbox"/>	V9	Numeric ▾	FLOAT	NULLABLE	-

Rows per page: 50 ▾

Total of 31 feature columns are included in the training

Weight column

Select a column to specify how to weight each row of the training data. By default, each row of your training data is weighted equally. [?](#)

Optimisation objective *

AUC ROC
Distinguish between classes

Log loss
Keeps prediction probabilities as accurate as possible

AUC PRC
Maximise precision-recall for the less common class

Precision at recall
Maximise precision for the less common class

Recall at precision
Maximise recall for the less-common class

[^ SHOW LESS](#)

CONTINUE



Step-8. Enter Budget as 1 and select Early stopping

Train new model

- training method
- model details
- training options
- 4** Computing and Prices

START TRAINING INTERRUPT

Enter the maximum number of node hours you want to use to train the model.

You can already train with only 1 knot hour . You may also be entitled to use free node hours for training. [price overview ↗](#)

budget * Maximum node hours [?](#)

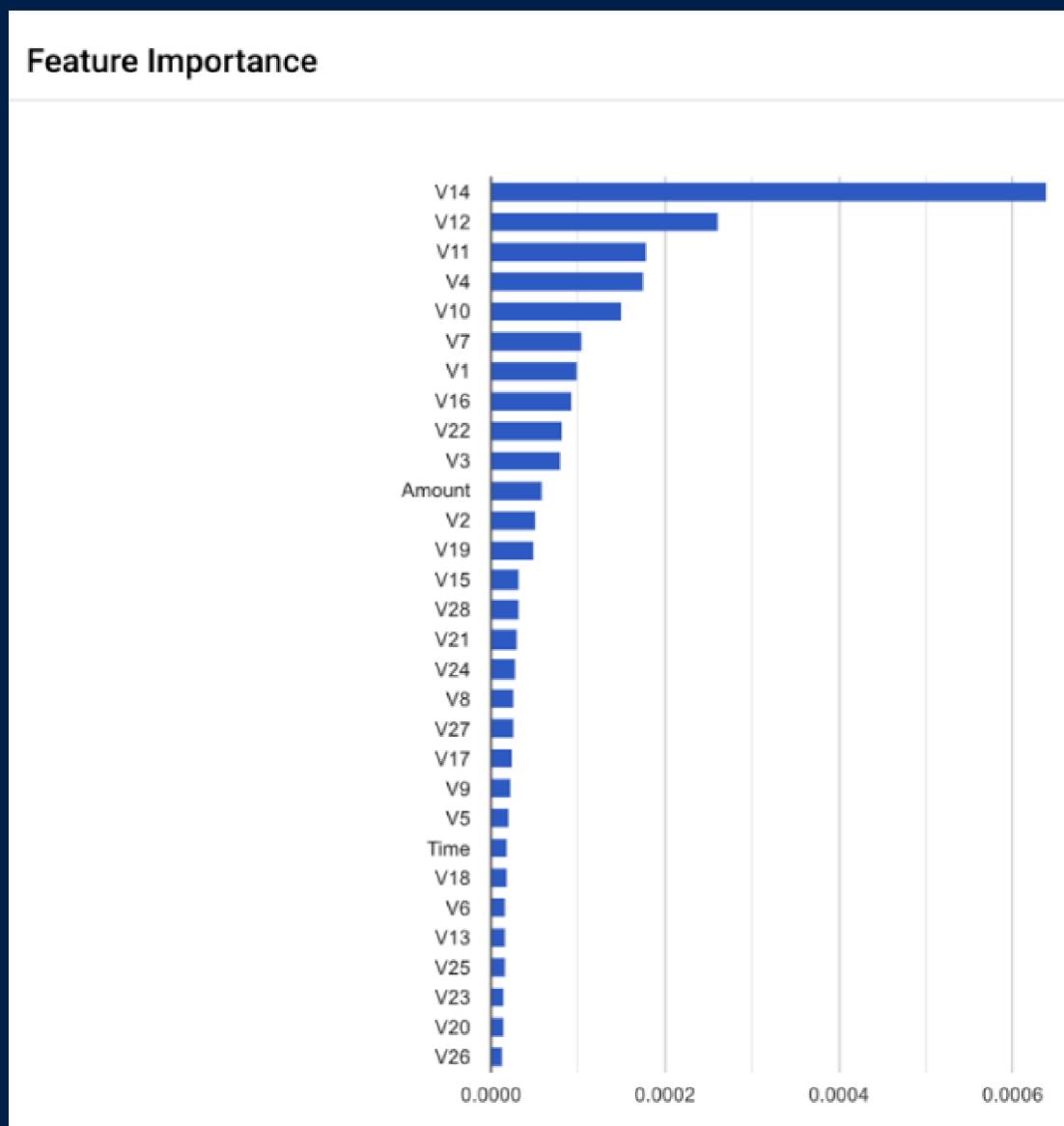
Estimated Completion: June 30, 2023 2PM GMT+1

Enable early termination
Stops model training when no more improvements are possible and refunds the remaining training budget. If early termination is disabled, the training will continue until the budget is used up.



Step-9. Model Evaluations and Metrics

Feature Importance



Confusion Matrix



Step-10 Deploy the Model to Endpoint

The screenshot shows the Google Cloud Vertex AI interface for a project named "My First Project". The main navigation bar includes "Google Cloud", "My First Project", a search bar, and various icons. On the left, there's a sidebar with sections like "Vertex AI", "TOOLS" (Dashboard, Model Garden, workbench, pipelines), "GENERATIVE AI STUDIO" (Overview, Language, speech), "DATA" (feature store, data sets, Label creation tasks), "MODEL DEVELOPMENT" (Training, Testing, metadata), and "DEPLOY AND USE" (Marketplace). The "DEPLOY AND TEST" tab is currently selected. Below it, there's a section titled "Use your edge-optimized model" with a "Container" option. The main area is titled "Deploy your model" and contains a description of what endpoints are and a prominent blue "DEPLOY TO ENDPOINT" button. A red circle highlights this button. In the bottom right, there's a table showing two models: "house" and "housing", with columns for status, models, Resource pool for deployment, region, monitoring, Last monitoring job, Recent Notifications, Last updated, API, notification, labels, and encryption. The "house" model is listed as "Aktiv". At the bottom, there's a "test model" section with a "PREVIEW" button and a table showing function column name, Type, Value, and Importance of local feature.

Name	ID	Status	Models	Resource pool for deployment	Region	Monitoring	Last monitoring job	Recent Notifications	Last updated	API	Notification	Labels	Encryption
house	8458951970806300672	Aktiv	1	--	us-central1	Aktiviert	—	0 Benachrichtigungen	30.06.2023, 13:13:36	Beispielanfrage	Von Google verwaltet		
housing	2199968835551887360	Aktiv	1	--	us-central1	Aktiviert	—	0 Benachrichtigungen	30.06.2023, 12:10:11	Beispielanfrage	Von Google verwaltet		



Step-11. Deploy to Endpoint- Endpoint Name

Deploy to endpoint

- 1 Define end point
- 2 model settings
- 3 model monitoring
- 4 monitoring goals

Create new endpoint

endpoint name *

Location

region

access

Determines how the endpoint can be accessed. By default, endpoints are available for predictions via a REST API. Access to an endpoint cannot be changed after it is created.

default
Exposes the endpoint for predictions via a REST API. AutoML and custom trained models can be added to standard endpoints.

Private
Create a VPC network and [access private services](#) a private connection to that endpoint. Only custom trained and tabular models can be added to private endpoints.
[Further information](#)

EXPANDED OPTIONS

[FURTHER](#)

Step-12. Model Settings

Deploy to endpoint

- 1 Define end point
- 2 model settings
- 3 model monitoring
- 4 monitoring goals

New model
house details (version 1)

traffic sharing *

calculate resources

Specify how compute resources will provide prediction traffic for the model

- Autoscaling : If you set a minimum and maximum, compute nodes will scale to meet traffic demand within those boundaries
- No scaling : If you only set a minimum, then that number of compute nodes will always run regardless of traffic demand (the maximum will be auto to minimum)

filter Enter to filter

option	compute node	memory
default	n1-standard-2	2 vCPUs, 7.5 GiB memory
Big storage	n1-standard-4	4 vCPUs, 15 GiB memory
Powerful CPU	n1-standard-8	8 vCPUs, 30 GiB memory
Powerful GPU	n1-standard-16	16 vCPUs, 60 GiB memory
Mega GPU		

logging

The logging settings apply permanently to this endpoint. Logging fees apply. If you want to change the logging setting in the future, create a new endpoint. [Further information](#)

Enable access logging for this endpoint

Disable existing logging for this endpoint



Step-13. Model Monitoring

Deploy to endpoint

- Define end point
- model settings
- model monitoring
- monitoring goals

① Model monitoring applies to all models deployed on this endpoint

monitoring goal

Detection of discrepancies between training and deployment
Differences between training and deployment occur when the distribution of feature data in production differs from the distribution of feature data in model training

Detection of prediction drift
Prediction drift occurs when the distribution of feature data in production changes significantly over time

Detection of discrepancies between training and deployment

Source of training data

In order for the monitoring job to detect discrepancies between training and deployment, it must compare the model training data to the dataset used to train the model

Cloud Storage bucket

BigQuery table

Vertex AI dataset

Vertex AI dataset *

The column name from the training data that you want the model to predict. This column is ignored when capturing feature deviations.

target column *

Notification thresholds (optional)

Determines which features to monitor and how much distance there is between the distribution of input features and the reference. At the end of each monitoring run, you will receive a notification email if a threshold has been exceeded. [Learn more](#)

Step-14. Monitoring Goals

Deploy to endpoint

- Define end point
- model settings
- model monitoring
- monitoring goals

② Detection of prediction drift
Prediction drift occurs when the distribution of feature data in production changes significantly over time

Detection of discrepancies between training and deployment

Source of training data

In order for the monitoring job to detect discrepancies between training and deployment, it must compare the model training data to the dataset used to train the model

Cloud Storage bucket

BigQuery table

Vertex AI dataset

Location of training data *

house-details/propertydataN.csv

SEARCH

Supported file types: .CSV, .TFRecord

target column

The column name from the training data that you want the model to predict. This column is ignored when capturing feature deviations.

target column *

Notification thresholds (optional)

Determines which features to monitor and how much distance there is between the distribution of input features and the reference. At the end of each monitoring run, you will receive a notification email if a threshold has been exceeded. [Learn more](#)

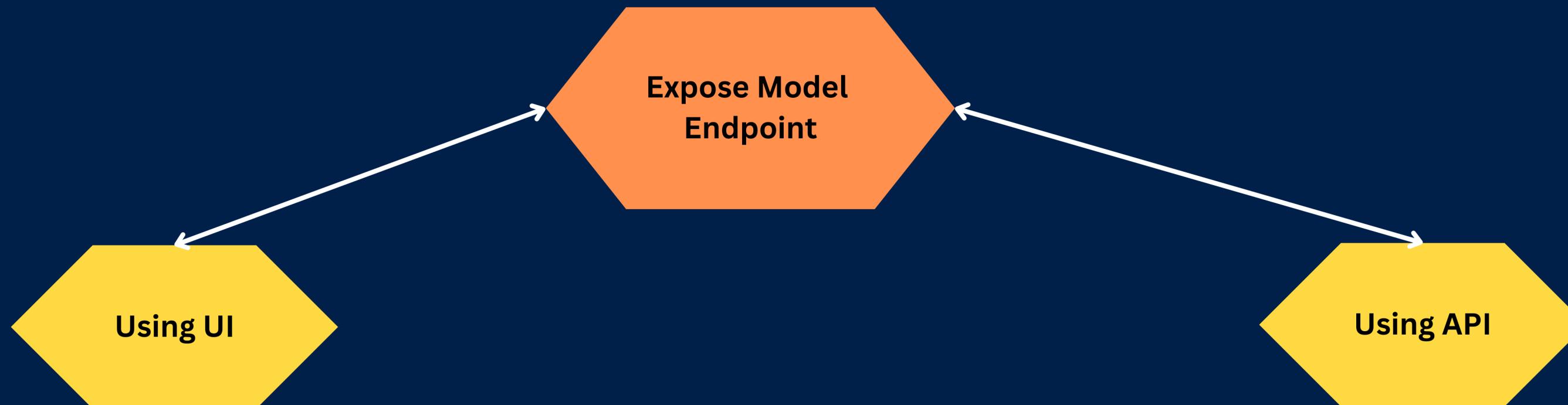
If you leave the field blank, all features are monitored and the notification threshold is 0.3.

JSON for notification limits

Train models configured to receive attribution scores via Explainable AI

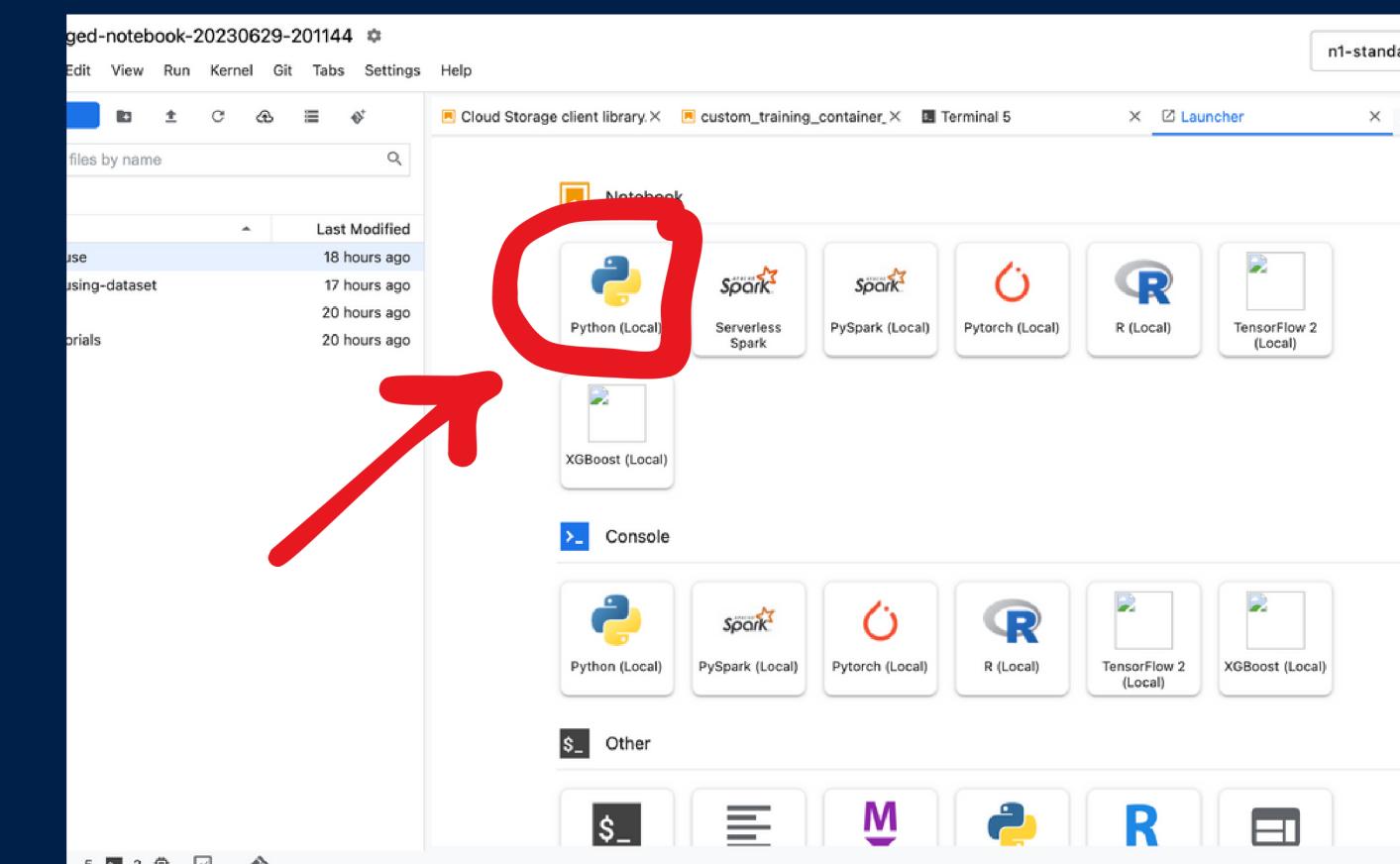


Step-15. Getting Predictions on the Deployed Model



Test your model

Feature column name	Type	Required or optional	Value	Local feature importance
Amount	Numerical	Required	17.99	--
Time	Numerical	Required	80422	--
V1	Numerical	Required	-0.24122594937617	--
V10	Numerical	Required	-0.18657577690678898	--



Step-16. Open Jupyter Notebook to Test Model

Expose Model to End Point using VerteX AI -SDK

This phase is where you have finished successfully trained your Model and Now we are going to expose the Trained model as an API

Step-1: In your notebook, run the following in a cell to install the Vertex SDK:

```
Python  
!pip3 install google-cloud-aiplatform --upgrade --user
```

Step-2: Import SDK and Create Reference to the EndPoints you just deployed.

```
Python  
from google.cloud import aiplatform  
  
endpoint = aiplatform.Endpoint()  
  
endpoint_name="projects/YOUR-PROJECT-NUMBER/locations/us-central1  
/endpoints/YOUR-ENDPOINT-ID"  
)
```

This Document has been provided and all detailed code is available



Finally we have Deployed





Thank You

Stay Tuned for For More Information

Scan The QR Code for Light House Project NewsLetter

