

VERTEX AI PIPELINES

SCREENSHOTS:

CREATED A SIMPLE PIPELINE FIRST:

The screenshot shows the Google Cloud Platform Vertex AI Pipelines interface. The left sidebar is titled 'Vertex AI' and includes options like Dashboard, Datasets, Features, Labeling tasks, Workbench, Pipelines (which is selected), Training, Experiments, Models, Endpoints, Batch predictions, and Metadata. The main area is titled 'hello-world-pi...' and shows a 'Runtime Graph' with three steps: 'emoji' (python 3.7), 'build-sentence' (python 3.7), and 'product-name' (python 3.9). The status bar indicates '3/3 steps completed' and '55%'. To the right, there's a 'Pipeline run analy' card with a summary of the run, basic info (Duration: 2 sec, Started: Dec 10, 2021, 11:42:21 PM, Completed: Dec 10, 2021, 11:42:23 PM, Run name: hello-world-pipeline-20211211074218, Pipeline name: hello-world, Runtime environment: Serverless, Region: us-central1, Service account: 376556357571-compute@developer.gserviceaccount.com), and run parameters (emoji_str: string, value: sparkles; text: string, value: Vertex Pipelines). The bottom navigation bar has tabs for Logs and Marketplace.

Avinash Ramesh

VERTEX AI PIPELINES

The screenshot shows the Google Cloud Platform Vertex AI Pipeline interface. On the left, the sidebar navigation includes 'Dashboard', 'Datasets', 'Features', 'Labeling tasks', 'Workbench', 'Pipelines' (selected), 'Training', 'Experiments', 'Models', 'Endpoints', 'Batch predictions', and 'Metadata'. The main area displays a 'Runtime Graph' for the 'hello-world-pipeline'. The graph shows three steps: 'emoji' (python 3.7), 'product-name' (python 3.9), and 'build-sentence' (python 3.7). The 'build-sentence' step has two inputs from the previous steps and one output. The pipeline run status is '3/3 steps completed'. The 'Pipeline run analysis' section on the right shows the 'SUMMARY' tab selected, displaying execution info for the 'build-sentence' step. The step details include:

Display name	build-sentence
Name	build-sentence
Type	system.ContainerExecution
Duration	0 sec
Started	Dec 10, 2021, 11:42:23 PM
Completed	Dec 10, 2021, 11:42:23 PM

Below the summary are sections for 'Input Parameters' and 'Output Parameters'.

Input Parameters

Parameter	Type	Value
emojitext	string	sparkles
emoji	string	:+)
product	string	Vertex Pipelines

Output Parameters

Parameter	Type	Value
Output	string	Vertex Pipelines is :+)

CREATED AN END-TO-END PIPELINE:

The screenshot shows the Google Cloud Platform Vertex AI Pipeline interface. The sidebar navigation is identical to the previous screenshot. The main area displays a 'Runtime Graph' for the 'automl-tab-beans-training-v2' pipeline. The graph shows several steps: 'tabledataset-create', 'automltabbeans-trainingjob', 'classification-model-eval...', and 'condition-deploy-decision...'. The pipeline run status is '1/6 steps completed'. The 'Pipeline run analysis' section on the right shows the 'SUMMARY' tab selected, displaying basic info for the run. The run details include:

Duration	11 min 56 sec
Started	Dec 10, 2021, 11:42:26 PM
Completed	automl-tab-beans-training-v2-20211211074224
Run name	automl-tab-beans-training-v2-20211211074224
Pipeline name	automl-tab-beans-training-v2
Runtime environment	Serverless
Region	us-central1
Service account	376556357571-compute@developer.gserviceaccount.com
Debugging info	View pipeline proto

Below the basic info are sections for 'Run Parameters' and 'Pipeline parameter values used for this run'.

Run Parameters

Parameter	Type	Value
api_endpoint	string	us-central1-aiplatform.googleapis.com
gcp_region	string	us-central1
thresholds_dict_str	string	{"wRoC": 0.95}
display_name	string	automl-beans1639208542
project	string	cmpe256elp
bq_source	string	bq://gs-dev-demos-beans.beans1

EXPANDED ARTIFACTS:

Avinash Ramesh

VERTEX AI PIPELINES

The screenshot shows the Google Cloud Platform Vertex AI Pipelines interface. On the left, the sidebar navigation includes: Dashboard, Datasets, Features, Labeling tasks, Workbench, Pipelines (selected), Training, Experiments, Models, Endpoints, Batch predictions, Metadata, and Marketplace. The main area displays a runtime graph for a pipeline named "automl-tab-beans-training-v2". The graph consists of several components: "tabulardataset-create", "automltabulartrainingjob...", "model", "classification-model-eval...", "condition-deploy-decision...", and "metricsc" and "metrics". The status bar indicates "1/6 steps completed" and "65%". To the right, the "Pipeline run analysis" panel is open, showing the "SUMMARY" tab. The summary details the run's duration (15 min 9 sec), start time (Dec 10, 2021, 11:42:26 PM), completion, run name (automl-tab-beans-training-v2-20211211074224), pipeline name (automl-tab-beans-training-v2), runtime environment (Serverless), region (us-central1), service account (376556357571-compute@developer.gserviceaccount.com), and debugging info (View pipeline proto). Below this, the "Run Parameters" section lists pipeline parameter values used for this run, including api_endpoint, gcp_region, thresholds_dict_str, display_name, project, and bq_source.

COMPLETE ALL STEPS

This screenshot shows the same Vertex AI Pipelines interface as the previous one, but with a different pipeline run. The runtime graph now shows "6/6 steps completed" and "75%". The "Pipeline run analysis" panel shows the "SUMMARY" tab with the following details: Duration (2 hr 46 min), Started (Dec 10, 2021, 11:42:26 PM), Completed (Dec 11, 2021, 2:28:55 AM), Run name (automl-tab-beans-training-v2-20211211074224), Pipeline name (automl-tab-beans-training-v2), Runtime environment (Serverless), Region (us-central1), Service account (376556357571-compute@developer.gserviceaccount.com), and Debugging info (View pipeline proto). The "Run Parameters" section is identical to the previous run. The "Run metrics" section shows a confusion matrix with the following data:

	0	1
0	331.0	16.0
1	6.0	0.0

METRICS:

Avinash Ramesh

VERTEX AI PIPELINES

Google Cloud Platform CMPE256NLP

Vertex AI Pipeline: automl-tab-beans-training-v2-20211211074224?project=cmpe256nlp

Pipeline Graph:

```

graph TD
    A[tabulardataset-create] --> B[automltablabeelingjob...]
    B --> C[classification-model-eval...]
    C --> D[model-deploy]
    C --> E[condition-deploy-decision...]
    E --> F[endpoint-create]
    F --> G[model-deploy]
    
```

Metrics:

Name	metricsc
Type	system.ClassificationMetrics
URI	gs://cmpe256nlp-bucket/pipeline_root/376556357571/automl-tab-beans-training-v2-20211211074224/classification-model-eval-metrics_-72264362041493504/metricsc

Confusion matrix:

True label	Predicted label						
	DERMASON	SIRA	SEKER	HOROZ	CALI	BARBUNYA	BOMBAY
DERMASON	94%	5%	2%	—	—	—	—
SIRA	7%	90%	1%	2%	—	1%	—
SEKER	1%	3%	96%	—	—	—	—
HOROZ	1%	3%	—	96%	1%	—	—
CALI	—	1%	—	4%	94%	2%	—
BARBUNYA	—	3%	1%	—	5%	91%	—
BOMBAY	—	—	—	—	—	—	100%

ROC Curve:

Receiver operating characteristic (ROC) curve is a graph showing the performance of a classification model. The curve plots TPR vs. FPR at different classification thresholds.

True positive rate

False positive rate

Avinash Ramesh