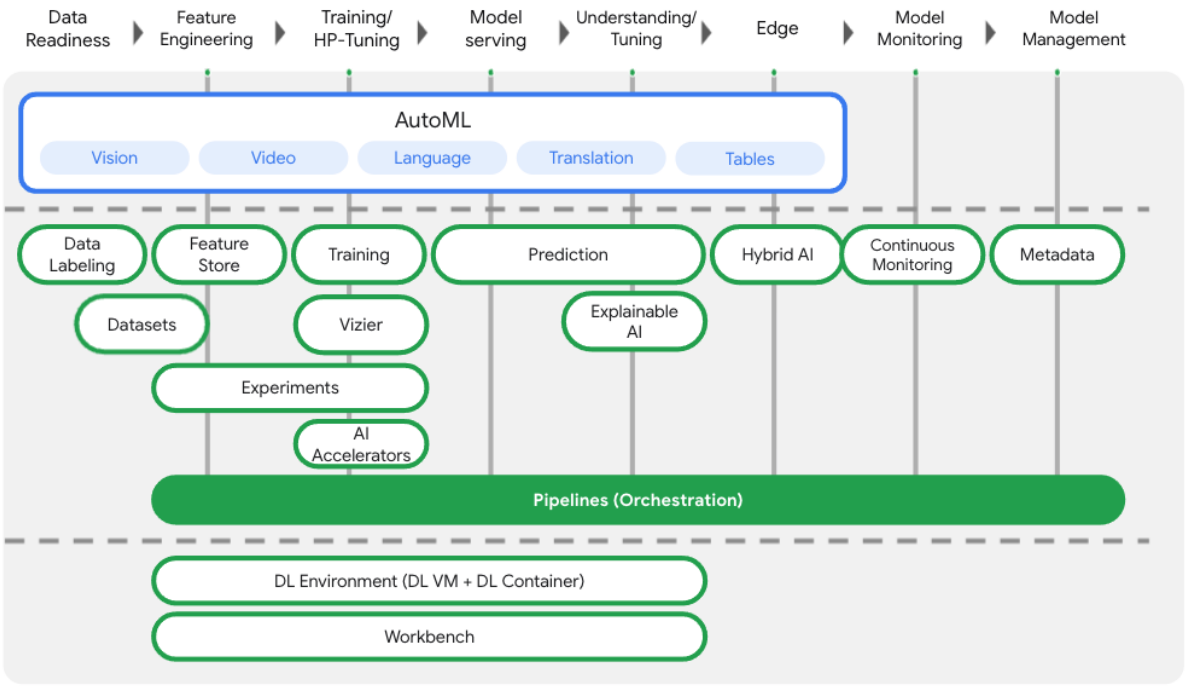
a) Vertex AI pipeline Intro

Reference: https://codelabs.developers.google.com/vertex-pipelines-intro#0

Objectives:

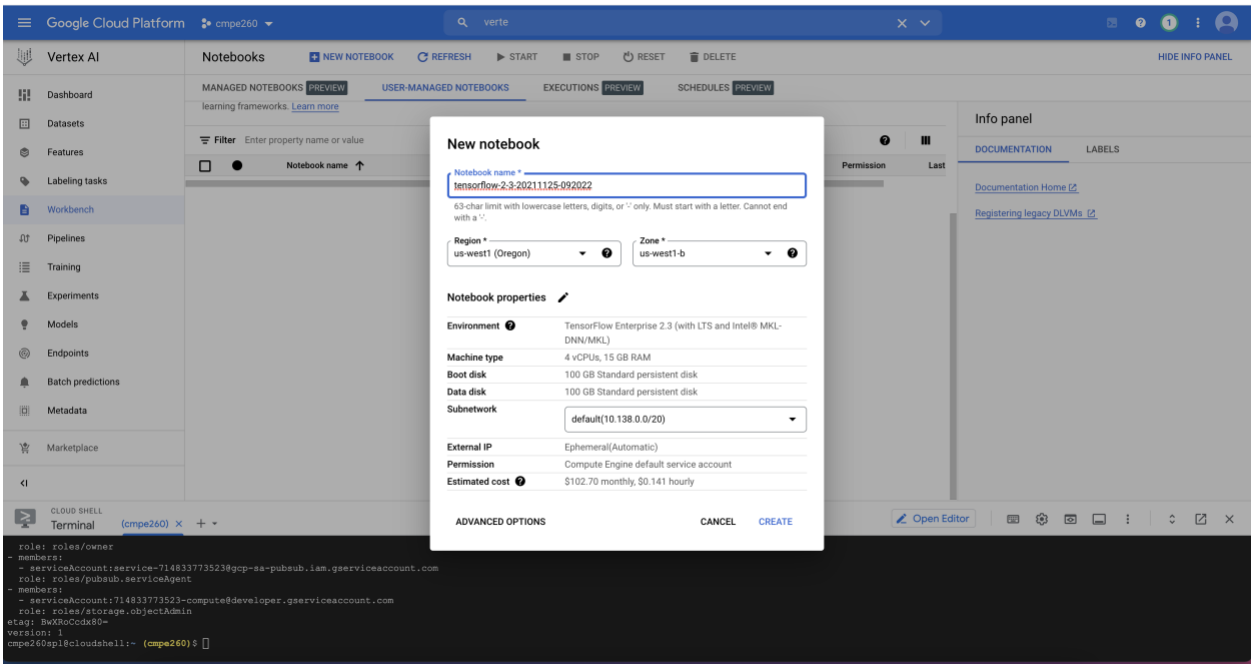
- Use the Kubeflow Pipelines SDK to build scalable ML pipelines
- Create and run a 3-step intro pipeline that takes text input
- Create and run a pipeline that trains, evaluates, and deploys an AutoML classification model
- Use pre-built components for interacting with Vertex AI services, provided through the google_cloud_pipeline_components library
- Schedule a pipeline job with Cloud Scheduler



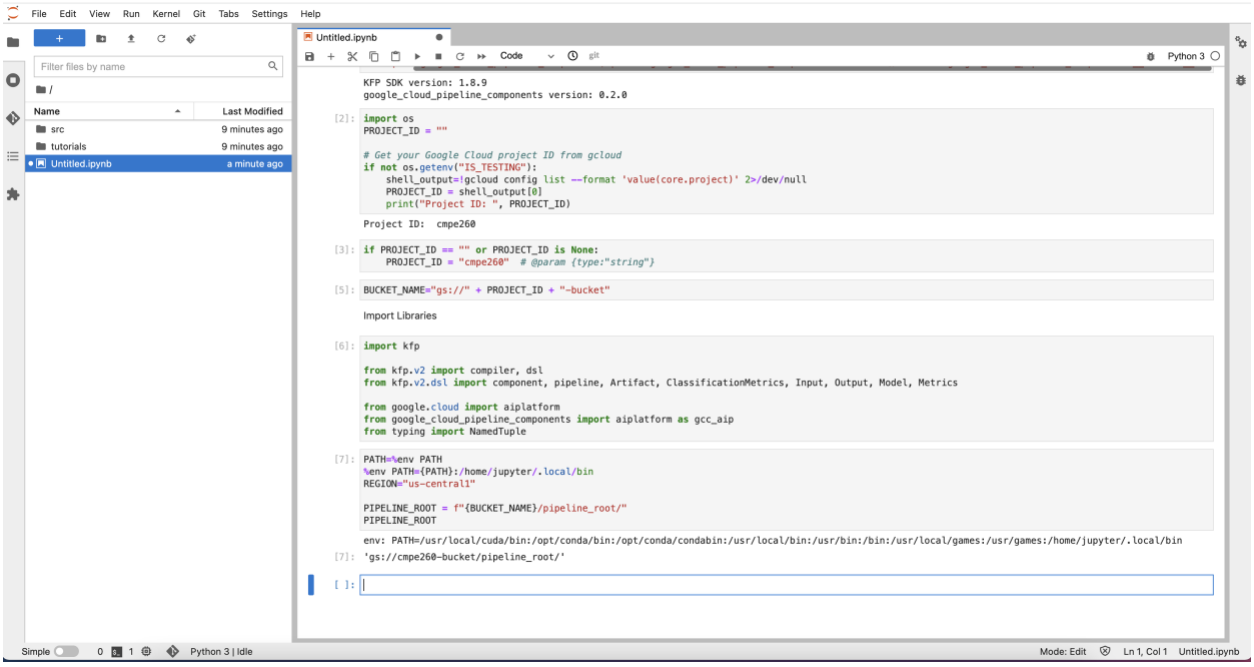
Cloud environment setup:

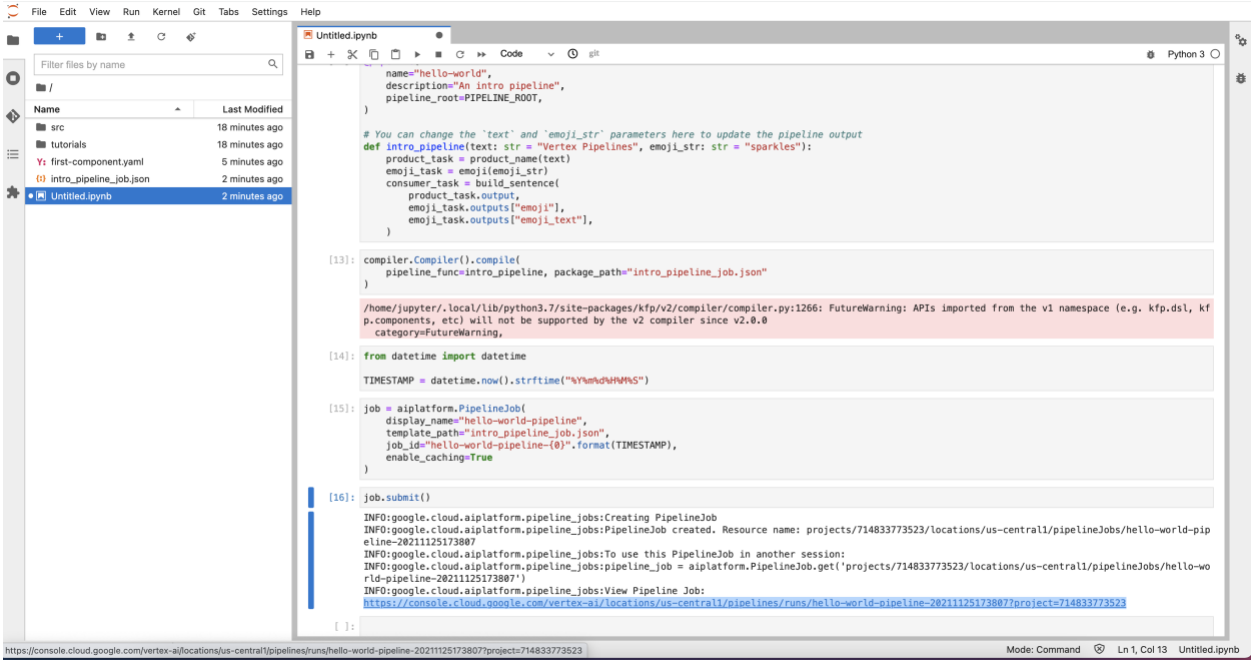
Enable API services and create google bucket:

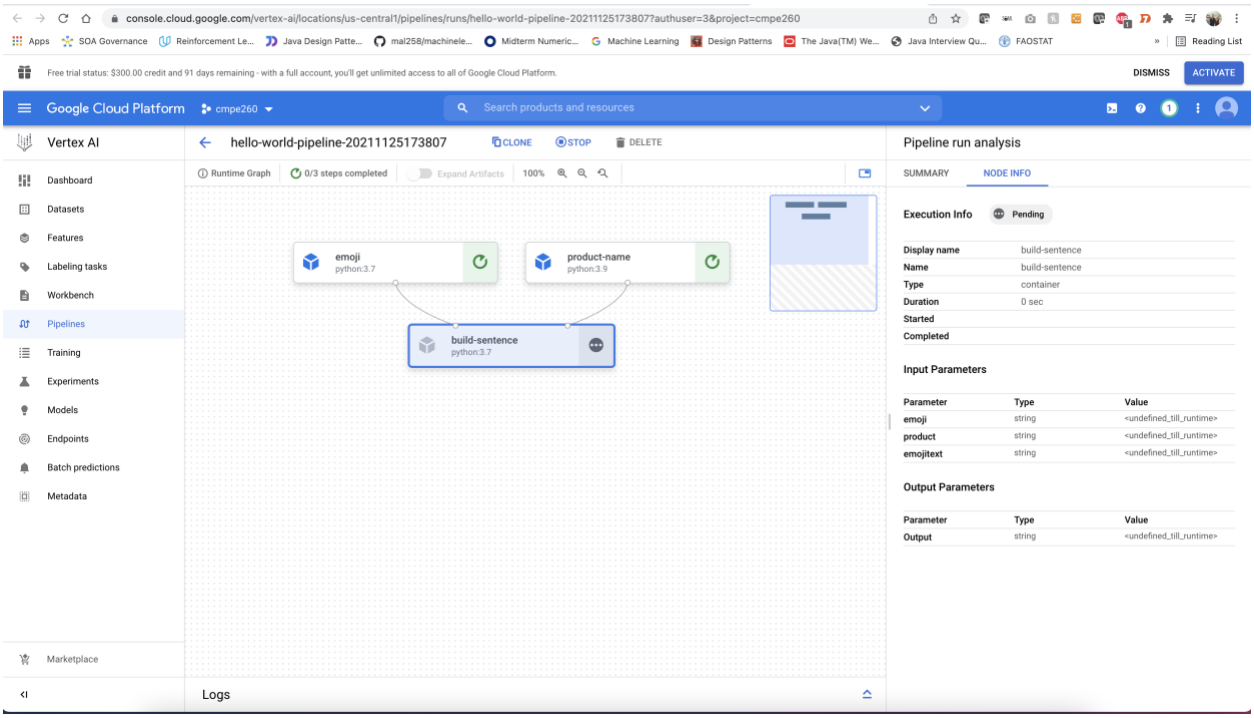
Create a Vertex AI workbench instance:



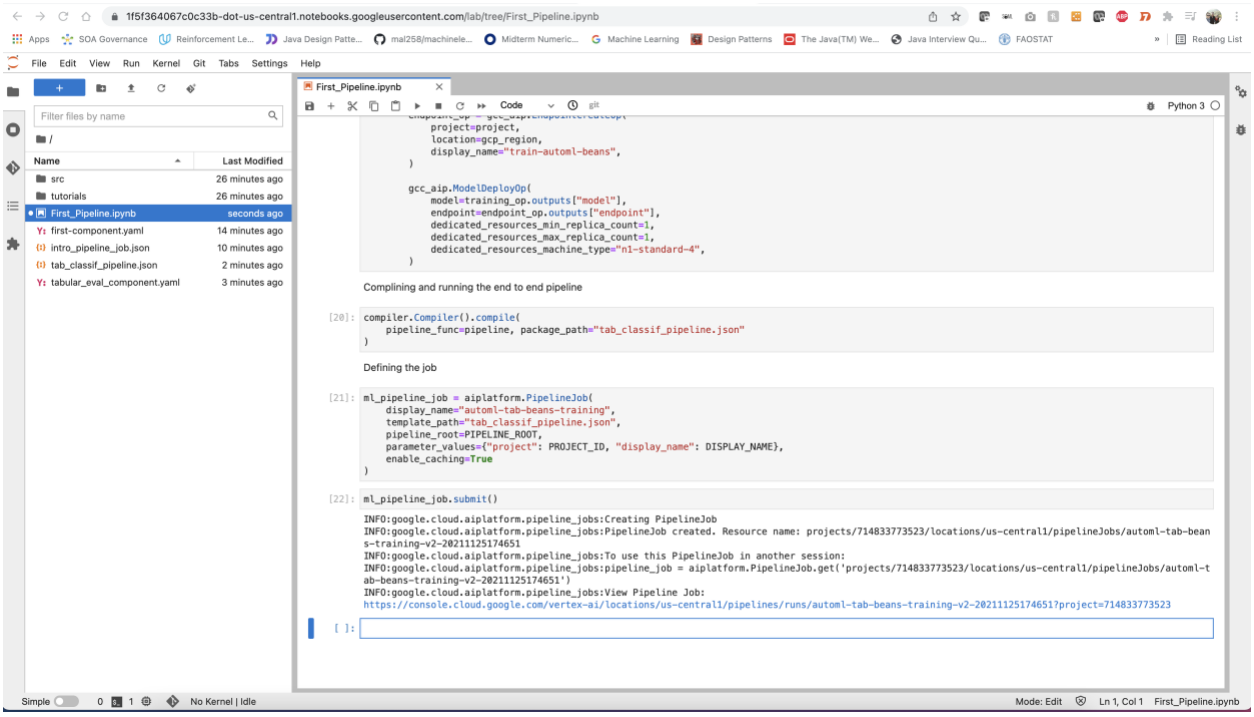
Vertex pipeline setup


File Edit View Run Kernel Git Tabs Settings Help
Filter files by name
/
Name
Last Modified
src 9 minutes ago
tutorials 9 minutes ago
Untitled.ipynb a minute ago
Untitled.ipynb
Code
Python 3
KFP SDK version: 1.8.9
google_cloud_pipeline_components version: 0.2.0
[2]:
import os
PROJECT_ID = ""
# Get your Google Cloud project ID from gcloud
if not os.getenv("IS_TESTING"):
shell_output=!gcloud config list --format 'value(core.project)' 2>/dev/null
PROJECT_ID = shell_output[0]
print("Project ID: ", PROJECT_ID)
Project ID: cmpe260
[3]:
if PROJECT_ID == "" or PROJECT_ID is None:
PROJECT_ID = "cmpe260" # @param {type:"string"}
[5]:
BUCKET_NAME="gs://" + PROJECT_ID + "-bucket"
Import Libraries
[6]:
import kfp
from kfp.v2 import compiler, dsl
from kfp.v2.dsl import component, pipeline, Artifact, ClassificationMetrics, Input, Output, Model, Metrics
from google.cloud import aiplatform
from google_cloud_pipeline_components import aiplatform as gcc_aip
from typing import NamedTuple
[7]:
PATH=%env PATH
%env PATH={PATH}:/home/jupyter/.local/bin
REGION="us-central1"
PIPELINE_ROOT = f"{BUCKET_NAME}/pipeline_root/"
PIPELINE_ROOT
env: PATH=/usr/local/cuda/bin:/opt/conda/bin:/opt/conda/condabin:/usr/local/bin:/usr/bin:/bin:/usr/local/games:/usr/games:/home/jupyter/.local/bin
[7]:
'gs://cmpe260-bucket/pipeline_root/'
[ ]:
Simple 0 1 Python 3 | Idle
Mode: Edit Ln 1, Col 1 Untitled.ipynb

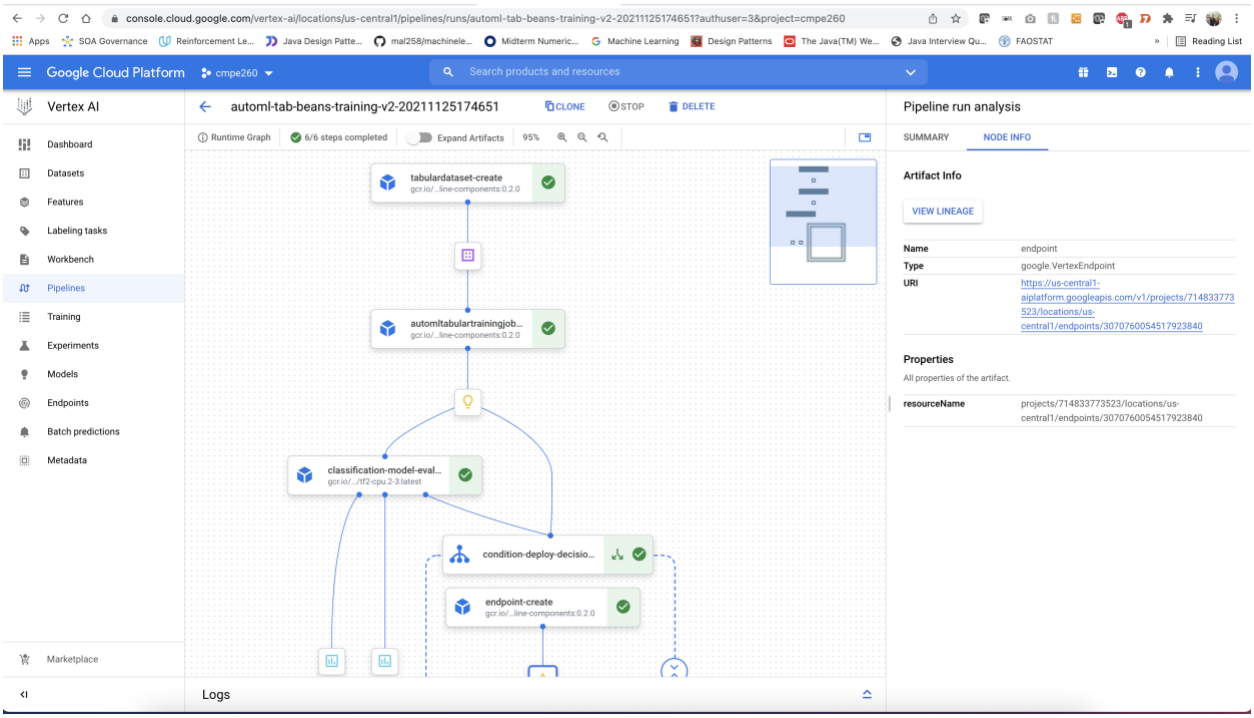

File Edit View Run Kernel Git Tabs Settings Help
Filter files by name
/
Name
Last Modified
src 18 minutes ago
tutorials 18 minutes ago
first-component.yaml 5 minutes ago
intro_pipeline_job.json 2 minutes ago
Untitled.ipynb 2 minutes ago
Untitled.ipynb
Code
Python 3
name="hello-world",
description="An intro pipeline",
pipeline_root=PIPELINE_ROOT,
)
# You can change the `text` and `emoji_str` parameters here to update the pipeline output
def intro_pipeline(text: str = "Vertex Pipelines", emoji_str: str = "sparkles"):
product_task = product_name(text)
emoji_task = emoji(emoji_str)
consumer_task = build_sentence(
product_task.output,
emoji_task.outputs["emoji"],
emoji_task.outputs["emoji_text"],
)
[13]:
compiler.Compiler().compile(
pipeline_func=intro_pipeline, package_path="intro_pipeline_job.json"
)
/home/jupyter/.local/lib/python3.7/site-packages/kfp/v2/compiler/compiler.py:1266: FutureWarning: APIs imported from the v1 namespace (e.g. kfp.dsl, kfp.components, etc) will not be supported by the v2 compiler since v2.0.0
category=FutureWarning,
[14]:
from datetime import datetime
TIMESTAMP = datetime.now().strftime("%Y%m%d%H%M%S")
[15]:
job = aiplatform.PipelineJob(
display_name="hello-world-pipeline",
template_path="intro_pipeline_job.json",
job_id="hello-world-pipeline-{0}".format(TIMESTAMP),
enable_caching=True
)
[16]:
job.submit()
INFO:google.cloud.aiplatform.pipeline_jobs:Creating PipelineJob
INFO:google.cloud.aiplatform.pipeline_jobs:PipelineJob created. Resource name: projects/714833773523/locations/us-central1/pipelineJobs/hello-world-pipeline-20211125173807
INFO:google.cloud.aiplatform.pipeline_jobs:To use this PipelineJob in another session:
INFO:google.cloud.aiplatform.pipeline_jobs:pipeline_job = aiplatform.PipelineJob.get('projects/714833773523/locations/us-central1/pipelineJobs/hello-world-pipeline-20211125173807')
INFO:google.cloud.aiplatform.pipeline_jobs:View Pipeline Job:
https://console.cloud.google.com/vertex-ai/locations/us-central1/pipelines/runs/hello-world-pipeline-20211125173807?project=714833773523
[ ]:
https://console.cloud.google.com/vertex-ai/locations/us-central1/pipelines/runs/hello-world-pipeline-20211125173807?project=714833773523
Mode: Command Ln 1, Col 13 Untitled.ipynb



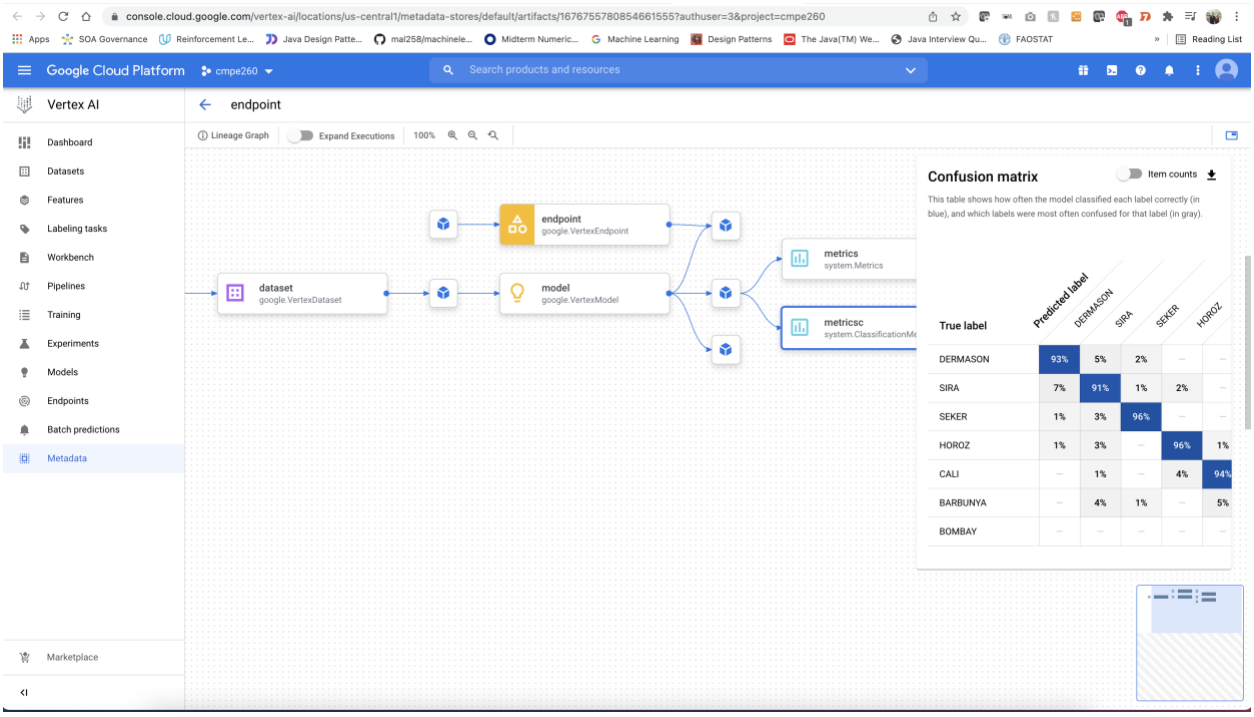
Creating an end to end ML pipeline

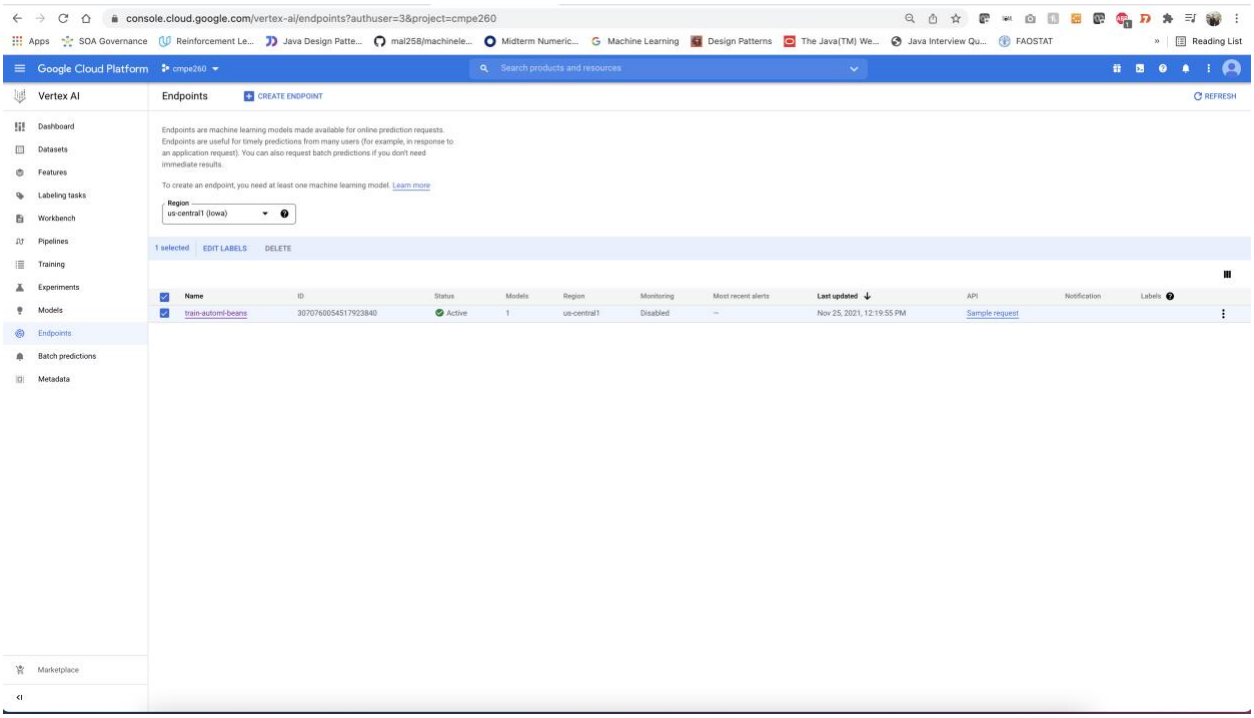


Compiling and running the pipeline

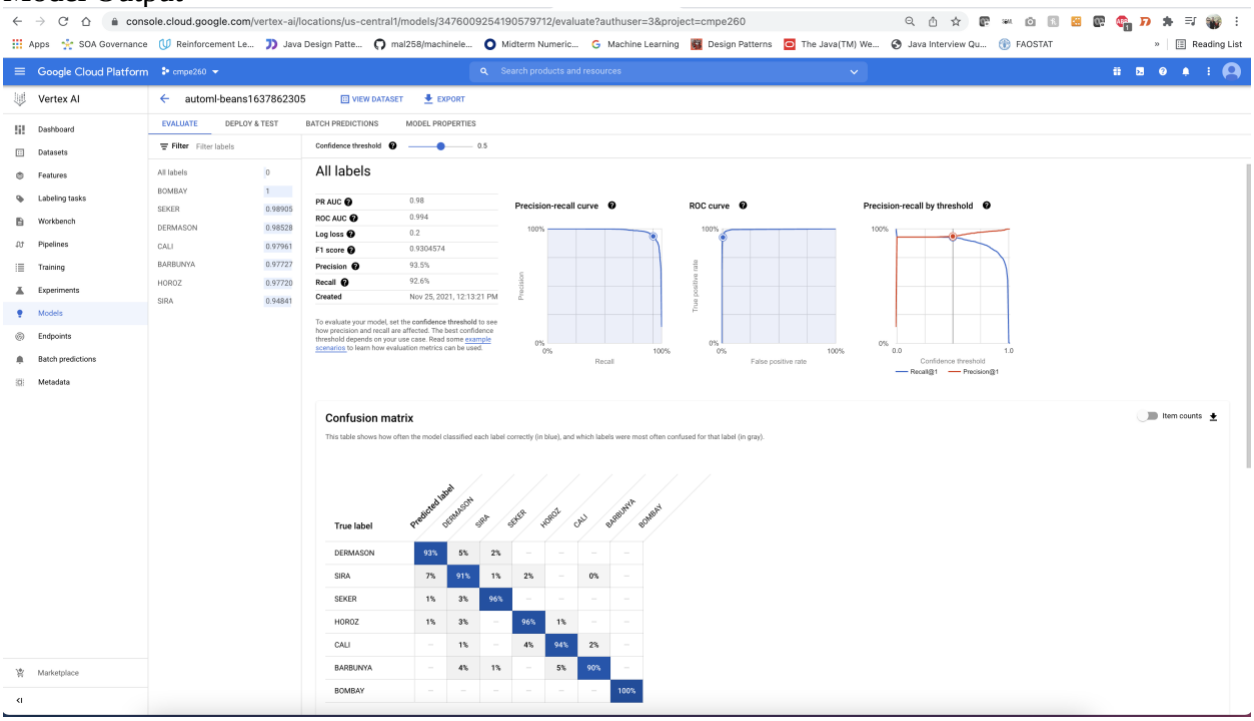


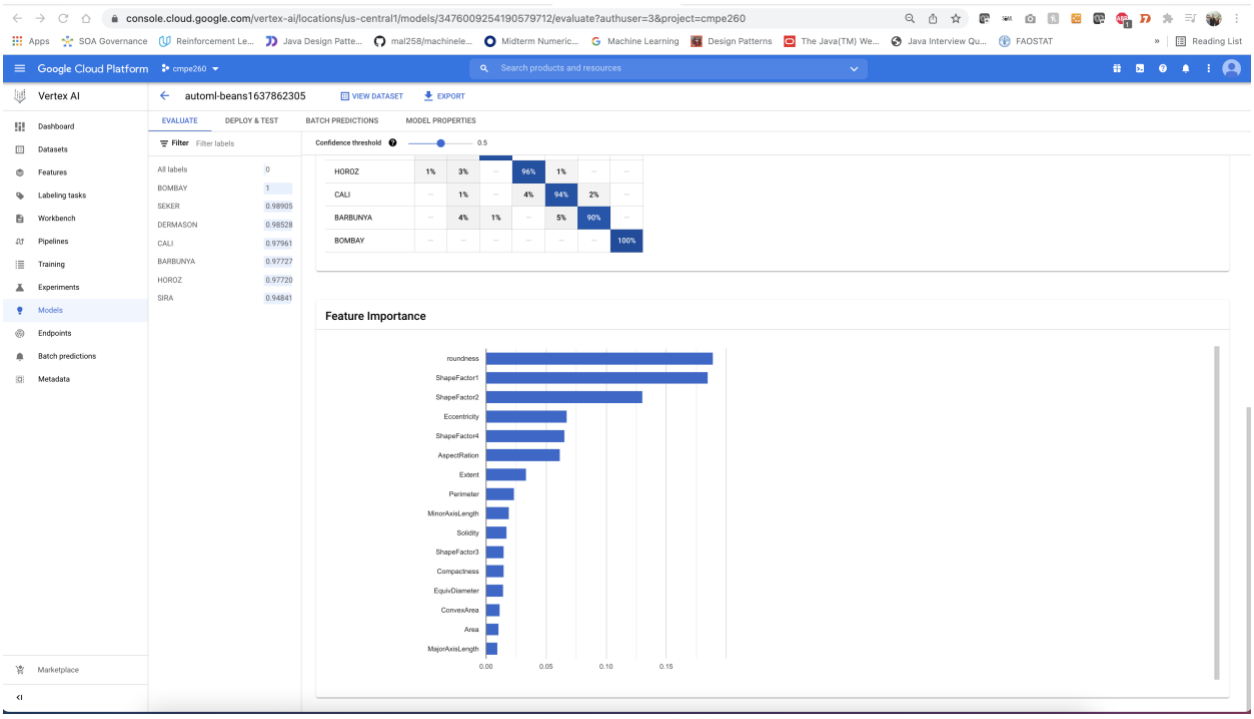
Data Set and End Point View

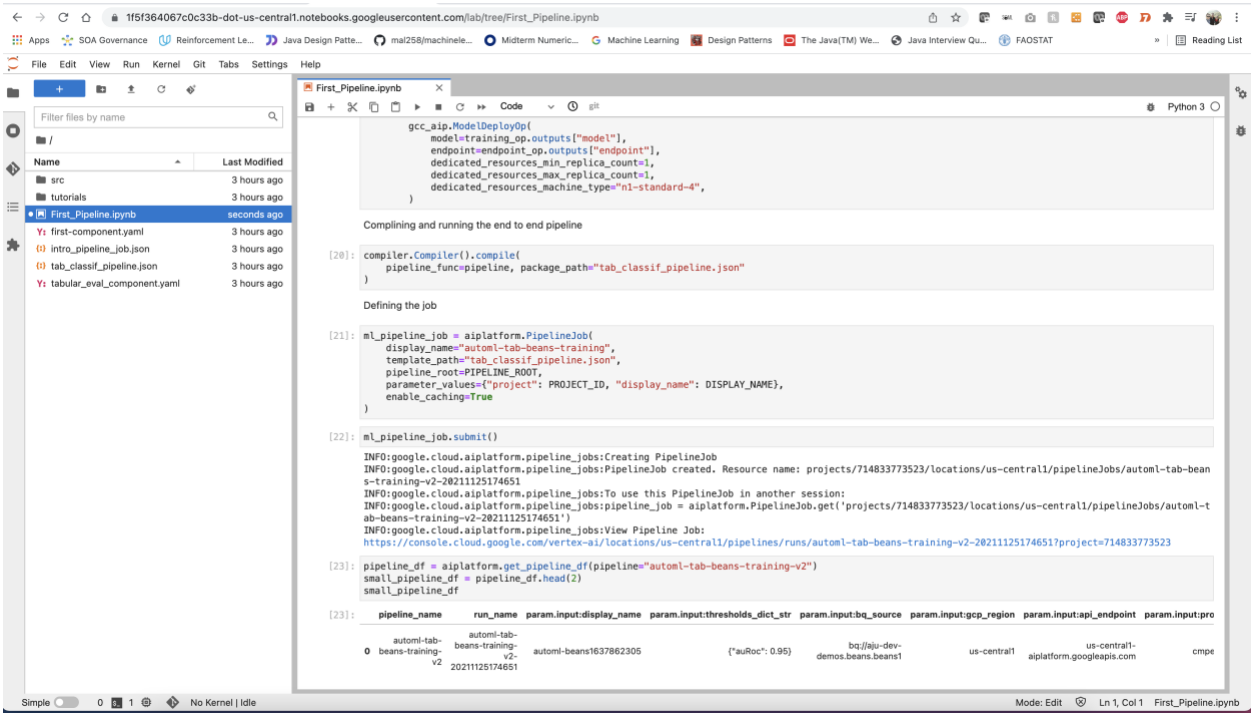




Model Output




Vertex AI
automl-beans1637862305
VIEW DATASET
EXPORT
EVALUATE
DEPLOY & TEST
BATCH PREDICTIONS
MODEL PROPERTIES
Feature Importance
roundness
ShapeFactor1
ShapeFactor2
Eccentricity
ShapeFactor4
AspectRation
Extent
Perimeter
MinorAxisLength
Solidity
ShapeFactor3
Compactness
EquivDiameter
ConvexArea
Area
MajorAxisLength


First_Pipeline.ipynb
Complining and running the end to end pipeline
Defining the job
INFO:google.cloud.aiplatform.pipeline_jobs:Creating PipelineJob
INFO:google.cloud.aiplatform.pipeline_jobs:PipelineJob created. Resource name: projects/714833773523/locations/us-central1/pipelineJobs/automl-tab-beans-training-v2-20211125174651
INFO:google.cloud.aiplatform.pipeline_jobs:To use this PipelineJob in another session:
INFO:google.cloud.aiplatform.pipeline_jobs:pipeline_job = aiplatform.PipelineJob.get('projects/714833773523/locations/us-central1/pipelineJobs/automl-tab-beans-training-v2-20211125174651')
INFO:google.cloud.aiplatform.pipeline_jobs:View Pipeline Job:
https://console.cloud.google.com/vertex-ai/locations/us-central1/pipelines/runs/automl-tab-beans-training-v2-20211125174651?project=714833773523