

A blue square containing the white text '0.5'. To the right of the square is a yellow triangle pointing downwards.

0.5

Fine-tunable Models

Fine-tunable Models

01	Fine-tunable Models Review
02	Tuning PaLM
03	Vertex AI Pipelines
04	Lab: Running Custom Model Training on Vertex AI Pipelines



Fine-tunable models

Foundation

- PaLM (via Gen AI Studio)
- Stable Diffusion Inpainting
- ControlNet

Image Classification

- EfficientNetV2
- ResNet
- BEiT
- DeiT
- MobileNet
- ... many more!

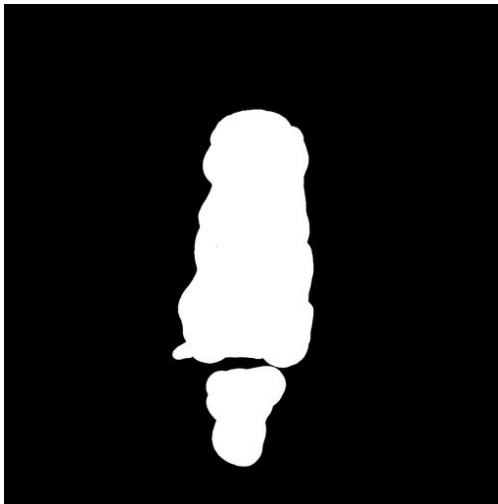
Object Detection

- SpineNet
- YOLO
- Faster R-CNN
- Mask R-CNN
- RetinaNet

Stable Diffusion Inpainting



-



+

"A unicorn
sitting on a
bench"

Stable Diffusion Inpainting



EfficientNetV2



Classification

Mountain: 0.16

Pinyon: 0.074

Geological_formation: 0.073

National Monument: 0.068

Stone_pine: 0.056

DeepLabv3+



■ sidewalk

■ building

■ wall

■ pole

■ polegroup

■ traffic light

■ traffic sign

■ vegetation

■ person

**The models are generally good,
but not good at *your task*.**

Fine-tunable Models

01	Fine-tunable Models Review
02	Tuning PaLM
03	Vertex AI Pipelines
04	Lab: Running Custom Model Training on Vertex AI Pipelines



Let's build a recipe application that is aware of a **food preference** and a **food allergy**.




Create a vegan recipe for the requested dish that does not contain nightshade plants. Format the recipe as JSON.

Dish: <User Input Here>

Pasta

```
{
  "title": "Vegan Pasta with Roasted Vegetables",
  "ingredients": [
    "1 pound gluten-free pasta",
    "1 tablespoon olive oil",
    "1 yellow onion, diced",
    "2 cloves garlic, minced",
    "1 red bell pepper, diced",
    "1 green bell pepper, diced",
    "1 zucchini, diced",
    "1 carrot, diced",
    "1 (15 ounce) can chickpeas, rinsed and
drained",
    "1 (15 ounce) can corn, drained",
    "1/2 cup vegetable broth",
    "1/4 cup chopped fresh basil",
    "1/4 cup chopped fresh parsley",
    "Salt and pepper to taste"
  ],
```



```
    "steps": [
      "Preheat oven to 400 degrees F (200 degrees C). Toss
the vegetables with olive oil and salt and pepper to
taste. Spread them out on a baking sheet and roast for
20-25 minutes, or until tender and slightly browned.",
      "While the vegetables are roasting, cook the pasta
according to the package directions. Drain and set
aside.",
      "In a large pot or skillet, heat the olive oil over
medium heat. Add the onion and cook until softened, about
5 minutes. Add the garlic and cook for 1 minute more.",
      "Add the roasted vegetables, chickpeas, corn,
vegetable broth, basil, and parsley to the pot. Bring to
a boil, then reduce heat and simmer for 10 minutes, or
until the vegetables are heated through.",
      "Add the cooked pasta to the pot and stir to combine.
Serve immediately."
    ],
    "servings": 4
  }
```



Create a vegan recipe for the requested dish that does not contain nightshade plants like tomatoes, potatoes, eggplants, and peppers. Format the recipe as JSON.

Dish: <User Input Here>

Pasta

```
{  
  "name": "Vegan Pasta with Sun-Dried Tomatoes  
and Walnuts",  
  "ingredients": [  
    "1 pound whole wheat pasta",  
    "1/2 cup sun-dried tomatoes, chopped",  
    "1/2 cup walnuts, chopped",  
    "1/4 cup olive oil",  
    "1/4 cup chopped fresh basil",  
    "1/4 cup chopped fresh parsley",  
    "1/4 cup grated Parmesan cheese  
(optional)",  
    "Salt and pepper to taste",  
  ],  
}
```



```
"instructions": [  
  "Bring a large pot of salted water to a boil.  
Add the pasta and cook according to the package  
directions.",  
  "While the pasta is cooking, heat the olive oil  
in a large skillet over medium heat. Add the  
sun-dried tomatoes and walnuts and cook for 2-3  
minutes, or until the walnuts are toasted and the  
tomatoes are softened.",  
  "Drain the pasta and add it to the skillet with  
the sun-dried tomatoes and walnuts. Stir to  
combine.",  
  "Add the basil, parsley, Parmesan cheese (if  
using), salt, and pepper to taste. Stir to  
combine.",  
  "Serve immediately.",  
],  
  "yield": "4 servings"  
}
```



Create a vegan recipe for the requested dish that does not contain nightshade plants like tomatoes, potatoes, eggplants, peppers, chilis, cayenne, and paprika. Replace animal products like milk and eggs with plant-based equivalents like almond milk and flax eggs. Use the following JSON format for the recipe:

```
{ "title": "", "ingredients": [ { "measurement": "", "ingredient": "" } ], "steps": [],  
  "yield": "" }
```

Dish: <User Input Here>

Where do we get lots of examples?

PaLM!

Fine-tuning dataset

```
{"input_text": "PromptExample1", "output_text": "OutputExample1"}  
{"input_text": "PromptExample2", "output_text": "OutputExample2"}  
{"input_text": "PromptExample3", "output_text": "OutputExample3"}
```

input_text

- Should contain instructions
- Should match production user inputs

output_text

- Should contain content you expect
- Should be in the format you want

Dataset JSON

```
{"input_text": 'Create a vegan recipe for the requested dish that does not contain nightshade plants like tomatoes, potatoes, eggplants, peppers, chilis, cayenne, and paprika. Replace animal products like milk and eggs with plant-based equivalents like almond milk and flax eggs. Use the following JSON format for the recipe:\\n{\\n"title": "",\\n"ingredients": [ { "measurement": "", "ingredient": "" } ],\\n"steps": [""],\\n"yield": ""\\n}\\n\\ndish: pasta', "output_text": '{"title": "Hearty Mediterranean Pasta", "ingredients": [{"measurement": "1/2 cup", "ingredient": "cooked chickpeas"}, {"measurement": "1/2 cup", "ingredient": "cooked lentils"}, {"measurement": "1/4 cup", "ingredient": "olive oil"}, {"measurement": "1/4 cup", "ingredient": "balsamic vinegar"}, {"measurement": "1/4 cup", "ingredient": "Italian seasoning"}, {"measurement": "1/4 cup", "ingredient": "fresh parsley, chopped"}, {"measurement": "1/4 cup", "ingredient": "vegan parmesan cheese"}, {"measurement": "1/2 pound", "ingredient": "whole wheat pasta"}], "steps": ["In a large bowl, combine the chickpeas, lentils, olive oil, balsamic vinegar, Italian seasoning, parsley, and vegan parmesan cheese.", "Cook the pasta according to the package directions.", "Drain the pasta and add it to the bowl with the chickpea mixture.", "Toss to combine.", "Serve immediately."], "yield": "4 servings"}'}
```

← Create a tuned model

1 Tuning dataset

2 Model details

START TUNING

Model tuning creates and deploys a new model from an existing one that's better adapted to your use case. Currently, model tuning occurs in limited regions. If you have an organization policy restricting certain regions, model tuning may fail.

Tuning dataset

The tuning dataset is a JSONL file that contains model prompt and responses examples(one per line). It's recommended that you use at least [100-500 samples](#) [↗](#).
[Learn more about dataset formatting](#) [↗](#)

You can upload the file or select one that's already on Cloud Storage.

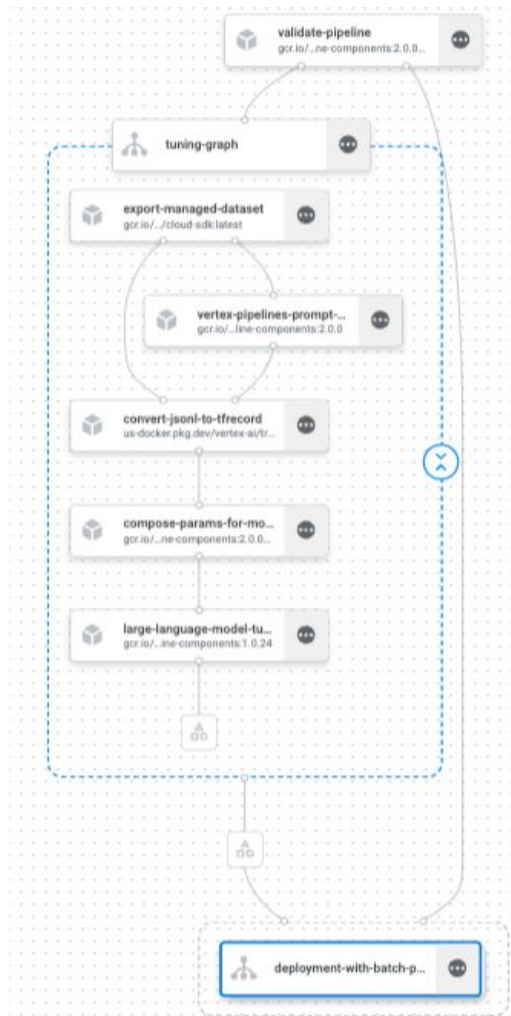
- ☐ Upload JSONL file to Cloud Storage
- ☒ Existing JSONL file on Cloud Storage

GCS file path *

✓ gs:// custom-recipe-app/recipes.jsonl

BROWSE

CONTINUE



Serverless process management!

Use Fine-tuned Model

```
import vertexai
from vertexai.language_models import TextGenerationModel

vertexai.init(project="YOUR-PROJECT-ID", location="us-central1")

model = TextGenerationModel.from_pretrained("YOUR_MODEL")

prompt = """Create a  vegan recipe ...
Dish: """

user_input = "pasta"

response = model.predict(prompt + user_input)

print(f"Response from Model: {response.text}")
```

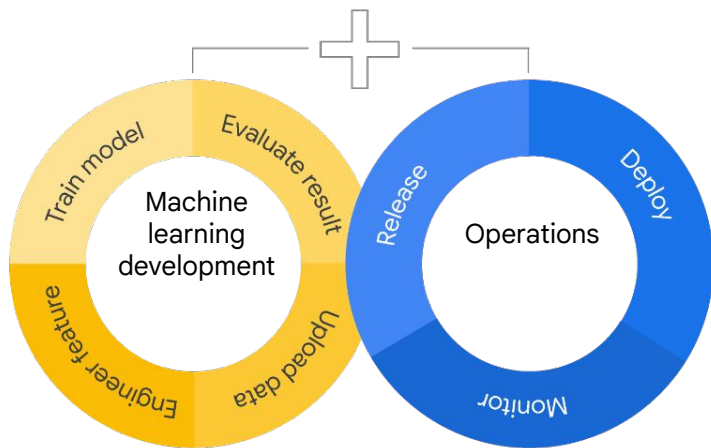
Fine-tunable Models

01	Fine-tunable Models Review
02	Tuning PaLM
03	Vertex AI Pipelines
04	Lab: Running Custom Model Training on Vertex AI Pipelines



Improve models with MLOps

Models need to evolve over time, which requires a process for managing change



MLOps phases:

Continuous integration (CI)

Continuous training (CT)

Continuous delivery (CD)

MLOps

Continuous Integration

Allow the user to tell us when the model is incorrect.

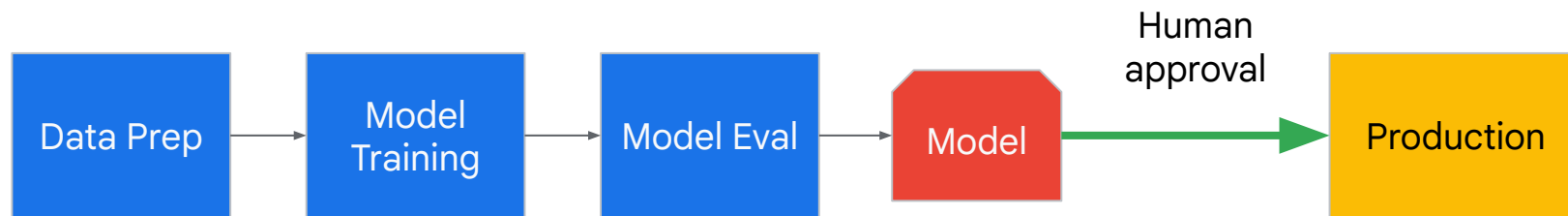
Continuous Training

Train a new version when you've accumulated enough new training examples.

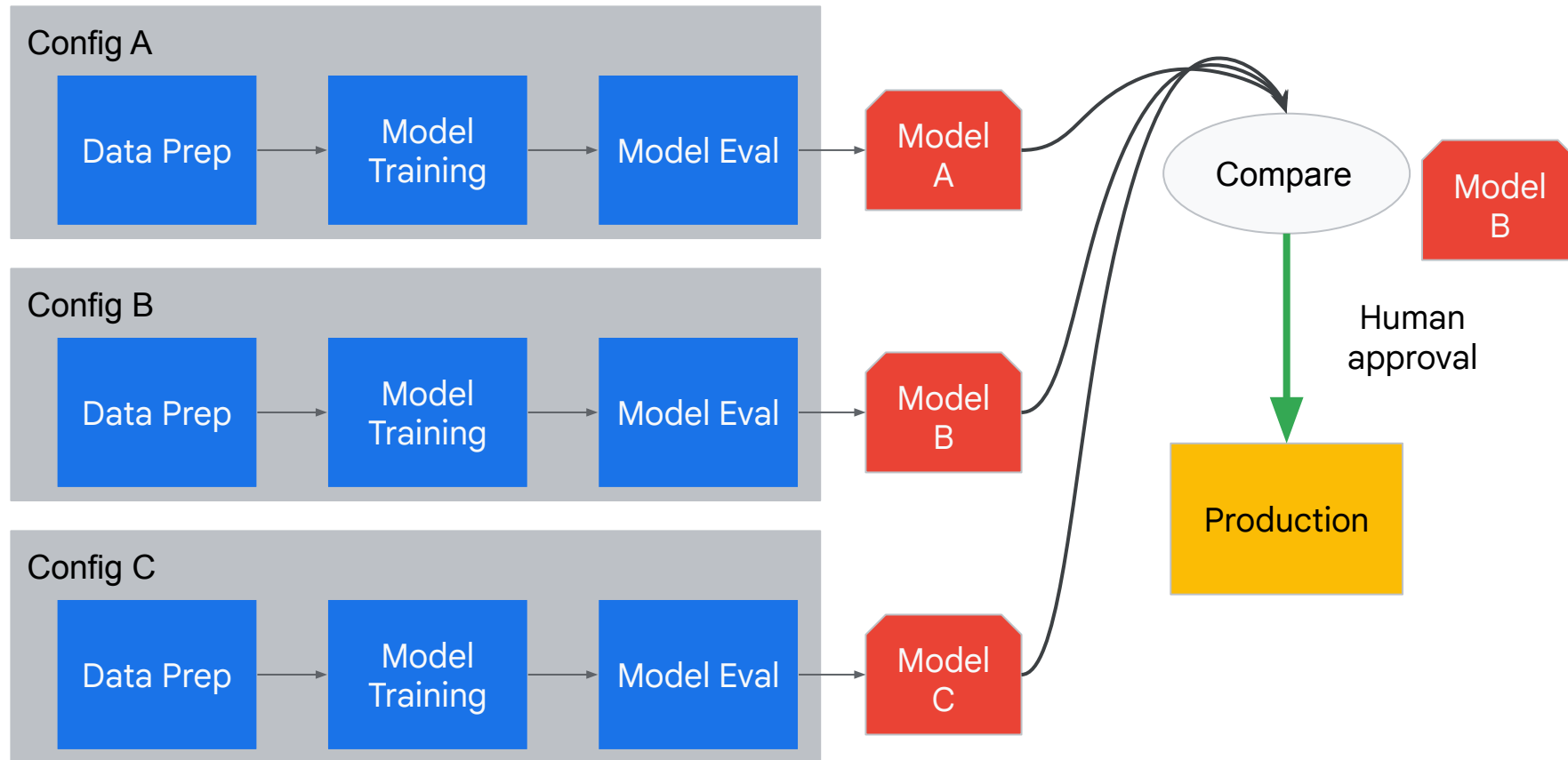
Continuous Delivery

This process usually results in new versions, which are easy to deploy

MLOps



Experiment!





Kubeflow

Open-source platform for machine learning pipelines on Kubernetes



TensorFlow Extended

TFX

Flexible, open-source framework that helps you build, deploy, and manage pipelines.

Kubeflow

Portable

Built on top of
Kubernetes, Kubeflow
runs anywhere!

Reusable

Components are
functions you can use on
your data. Pass in
parameters, get outputs.

Scalable

All pipeline steps are
containers. Handle 1 data
point or 1 million with the
same process.

Vertex AI Pipelines

Serverless

No need to build or manage a K8s cluster!

Experiments

No pipeline run is entirely independent. Keep track of pipeline runs for the same purpose with Experiments.

Ecosystem

Core cloud services like logging, security, and data governance work together with the other Vertex AI tools to provide a complete platform!

Fine-tunable Models

01	Fine-tunable Models Review
02	Tuning PaLM
03	Vertex AI Pipelines
04	Lab: Running Custom Model Training on Vertex AI Pipelines



Take a structured approach to GenAI transformation



Ramp your skills

Continue your learning journey and complete our [Generative AI Learning Path](#), free of charge, on Google Cloud Skills Boost



Organizational readiness

Assess your organization's current status and business needs for generative & traditional AI capabilities



Identify use cases

Select from one of our [Jumpstart GenAI offers](#), and work with Google Cloud to develop a technical design doc and sample code to solve the use case

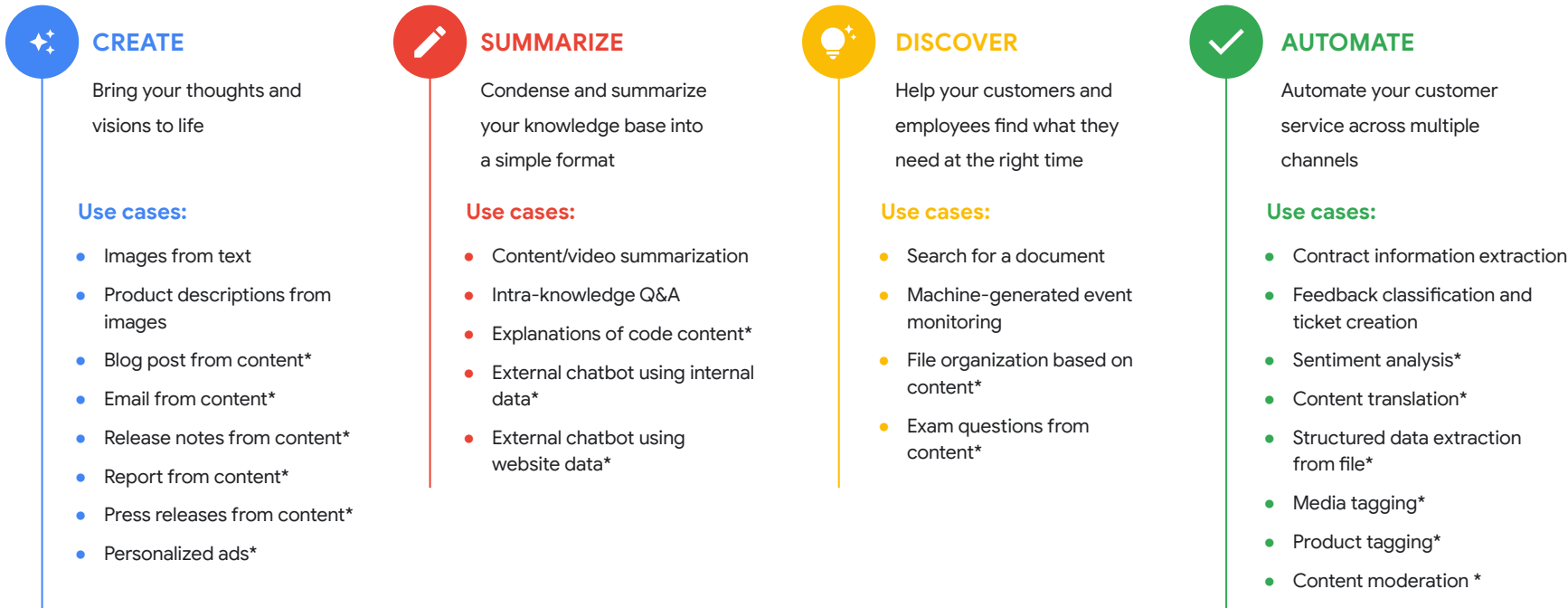


Test and scale

Purchase and implement generative AI solutions. Not all AI is built equal. POC often and fail fast to identify what works for your business.

Contact genai-triage@google.com to learn more

4 GenAI Jumpstart offers - \$25k & 2 weeks per use case



*may require Responsible AI Review

Please mark your attendance at
goo.gle/genai-checkin

Thank you for attending this training!

We would love your feedback! Please take 3-5 minutes to complete our survey and help inform content and program related improvements.

- 1 Scan our QR Code
- 2 Enter the attendance code provided by your instructor
- 3 Complete the survey



Lab: Running Custom Model Training on Vertex AI Pipelines

