

Google Cloud

Partner Certification Academy



# Professional Machine Learning Engineer

pls-academy-pmle-student-slides-7-2403

The information in this presentation is classified:

---

## Google confidential & proprietary

---

 This presentation is shared with you under NDA.

- Do **not** record or take screenshots of this presentation.
- Do **not** share or otherwise distribute the information in this presentation with anyone **inside** or **outside** of your organization.

Thank you!



# Source Materials

Some of this program's content has been sourced from the following resources:

- [Google Cloud certification site](#)
- [Google Cloud documentation](#)
- [Google Cloud console](#)
- [Google Cloud courses and workshops](#)
- [Google Cloud white papers](#)
- [Google Cloud Blog](#)
- [Google Cloud YouTube channel](#)
- [Google Cloud samples](#)
- [Google codelabs](#)
- [Google Cloud partner-exclusive resources](#)



This material is shared with you under the terms of your Google Cloud Partner **Non-Disclosure Agreement**.



## Google Cloud Skills Boost for Partners

- [Professional Machine Learning Engineer Certification](#)
- [Cloud Skills Boost for Partners Professional Machine Learning Engineer Learning Path](#)
- [Partner Learning Services Instructor-Led PMLE Curriculum](#)

Google Cloud

## Partner Advantage

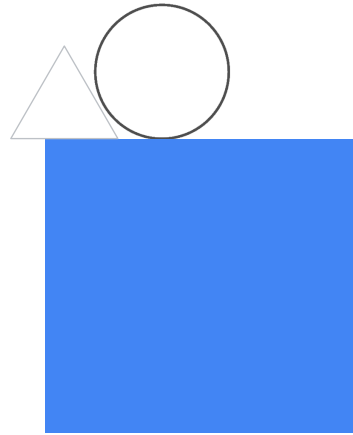
- [Best practices for implementing machine learning on Google Cloud](#)
- [Artificial Intelligence](#)
- [End-to-End MLOps Go-to-Market Kit](#)

# Session Logistics

- When you have a question, please:
  - Click the Raise hand button in Google Meet.
  - Or add your question to the Q&A section of Google Meet.
  - Please note that answers may be deferred until the end of the session.
- These slides are available in the Student Lecture section of your Qwiklabs classroom.
- The session is **not recorded**.
- Google Meet does not have persistent chat.
  - If you get disconnected, you will lose the chat history.
  - Please copy any important URLs to a local text file as they appear in the chat.

# Google Cloud Partner Learning Programs

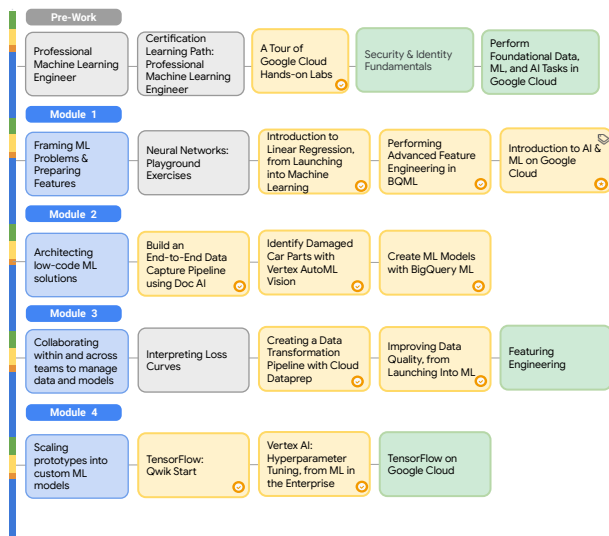
- Partner Certification Academy
- Partner Delivery Readiness Index (DRI)
- Cloud Skills Boost for Partners
- Partner Advantage





## PARTNER CERTIFICATION ACADEMY

# Professional Machine Learning Engineer



A Professional Machine Learning Engineer builds, evaluates, productionizes, and optimizes ML models by using Google Cloud technologies and knowledge of proven models and techniques. The ML Engineer:

- handles large, complex datasets and creates repeatable, reusable code.
- considers responsible AI and fairness throughout the ML model development process, and collaborates closely with other job roles to ensure long-term success of ML-based applications.
- has strong programming skills and experience with data platforms and distributed data processing tools.
- is proficient in the areas of model architecture, data and ML pipeline creation, and metrics interpretation.
- is familiar with foundational concepts of MLOps, application development, infrastructure management, data engineering, and data governance.
- makes ML accessible and enables teams across the organization.

By training, retraining, deploying, scheduling, monitoring, and improving models, the ML Engineer designs and creates scalable, performant solutions.

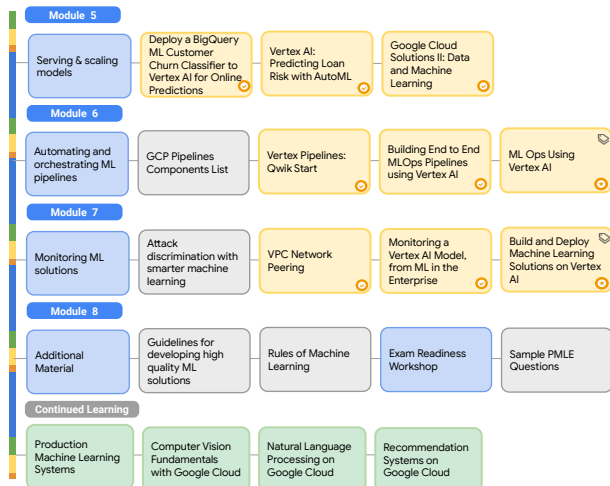
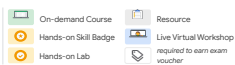
### Recommended candidate:

- Has in-depth experience setting up cloud environments for an organization
- Has experience deploying services and solutions based on business requirements



## PARTNER CERTIFICATION ACADEMY

# Professional Machine Learning Engineer



A Professional Machine Learning Engineer builds, evaluates, productionizes, and optimizes ML models by using Google Cloud technologies and knowledge of proven models and techniques. The ML Engineer:

- handles large, complex datasets and creates repeatable, reusable code.
- considers responsible AI and fairness throughout the ML model development process, and collaborates closely with other job roles to ensure long-term success of ML-based applications.
- has strong programming skills and experience with data platforms and distributed data processing tools.
- is proficient in the areas of model architecture, data and ML pipeline creation, and metrics interpretation.
- is familiar with foundational concepts of MLOps, application development, infrastructure management, data engineering, and data governance.
- makes ML accessible and enables teams across the organization.

By training, retraining, deploying, scheduling, monitoring, and improving models, the ML Engineer designs and creates scalable, performant solutions.

### Recommended candidate:

- Has in-depth experience setting up cloud environments for an organization
- Has experience deploying services and solutions based on business requirements

## Learner Commitment

Each week, learners are to complete the learning path's course content, Cloud Skills Boost for Partner Quests/Challenge Labs and material that the mentor has recommended that will support learning.

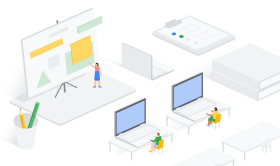
- **Workshop Day:** Meet for the cohort's weekly 'general session'. (≈ 2 hours)
  - **During the week:** Complete the week's course, perform hands-on labs, review any additional material suggested material for the week. (≈ 8 - 16 hours)
- **Important:** Learners must allocate time between each weekly session to study and familiarize themselves with any prerequisite knowledge they may lack. It is also recommended that learners complete the next week's course prior to the scheduled workshop.



## Path to Service Excellence



Certification



Advanced Solutions Training



Delivery Readiness Index

Certification is just one step on your professional journey. Google Cloud also offers our partners access to advanced solutions training, and a new quality-focused program called Delivery Readiness Index (DRI) to help you achieve service excellence with your customers.

## Benchmark your skills with DRI



Google Cloud

DRI helps to benchmark partner proficiency and capability at any point during the customer journey however should be used primarily as a lead measure to predict and prepare for partner delivery success.

DRI assesses and analyzes Partner Consultant GCP proficiency by creating a DRI Profile inclusive of their GCP knowledge, skills, and experience.

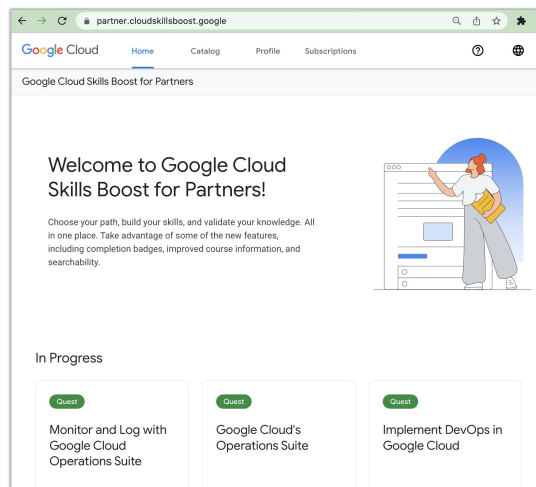
With the DRI insights, we can prescriptively advise the partner project team on the ground and bridge niche capability gaps.

DRI also takes action. For partner consultants, DRI generates a tailored L&D plan that prescribes personalized learning, training, and skill development to build GCP proficiency.

# Google Cloud Skills Boost for Partners

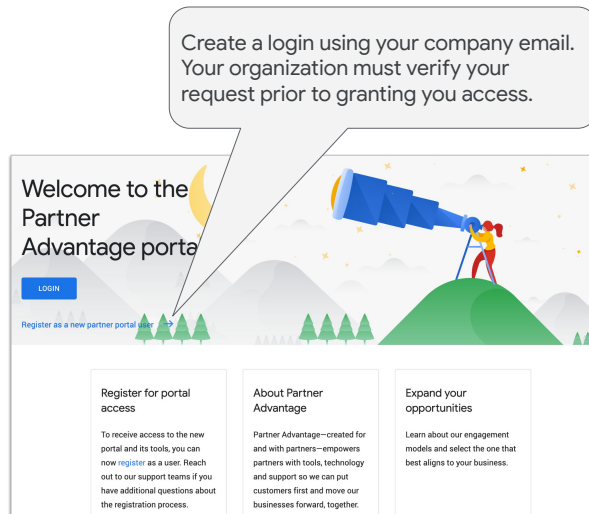
<https://partner.cloudskillsboost.google/>

- On-demand course content
- Hands-on labs
- Skill Badges
- **FREE** to Google Cloud Partners!



# Google Cloud Partner Advantage

- Resources for Google Cloud partner organizations:
  - Recent announcements
  - Solutions/role-based training
  - Live/pre-recorded webinars on various topics
    - [Partner Advantage Live Webinars](#)
- Complements the certification self-study material presented on Google Cloud Skills Boost for Partners
- Helpful Links:
  - [Getting started on Partner Advantage](#)
  - [Join Partner Advantage](#)
  - [Get help accessing Partner Advantage](#)



<https://www.partneradvantage.googlecloud.com/>

Google Cloud

## The getting started link:

<https://support.google.com/googlecloud/topic/9198654#zippy=>

Note the top section, “**Getting Started & User Guides**” and two key documents → Direct Partners to this if they need to enroll into Partner Advantage

1. Logging in to the Partner Advantage Portal - Quick Reference Guide
2. Enrolling in the Partner Advantage Program - Quick Reference Guide

## Focus from this point on:

### Some context on enrolling in PA:

Access to Partner Portal is given in 2 ways

- Partner Admin Led: Partner Administrator at Partner Company can set up users
- User Led: User can go through Self Registration
  - [https://www.partneradvantage.googlecloud.com/GCPPRM/s/partneradvantageportallogin?language=en\\_US](https://www.partneradvantage.googlecloud.com/GCPPRM/s/partneradvantageportallogin?language=en_US)
  - Or directly to the User Registration Form, [https://www.partneradvantage.googlecloud.com/GCPPRM/s/partnerselfregistration?language=en\\_US](https://www.partneradvantage.googlecloud.com/GCPPRM/s/partnerselfregistration?language=en_US)

### Please Note

- After a user self-registers, they receive an email that essentially states:

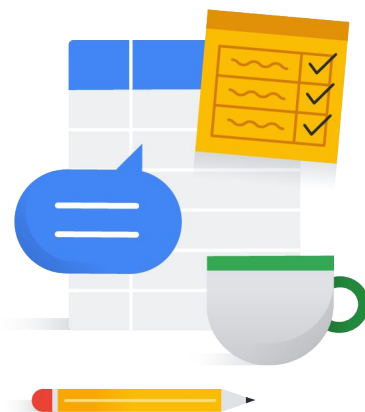
- “Hi {Partner Name}, you are one step away from joining the Google Cloud Partner Advantage Community. Please click to continue with the user registration process. See you in the cloud, The Partner Advantage Team
- Once registered, they can access limited content until their **Partner Administrator approves the user**
- Their Partner Administrator also receive an email notifying them that a member of their organization has registered themselves on their organization's Google Cloud Partner Advantage account.
  - It also states that this user has limited access to the portal
  - They are provided instructions on how to review and provision the appropriate access for the user that has registered
- Once their admin approves the user, they receive an email that states:
  - Hi {User Name}, Your Partner Administrator has updated your access to the Google Cloud Partner Advantage portal. You have been granted edit access to additional account information on the portal on behalf of your organization to help build your business. For additional access needs, please work with your Partner Administrator. See you in the cloud, The Partner Advantage Team

The net takeaway is, on the Support Page (the first link on this slide) [Google Cloud Partner Advantage Support](#), there's a section **“Issue accessing Partner Advantage Portal? Click here for troubleshooting steps”**

- The source of their issue can be related to the different items shown
- Additionally, there's a Partner Administrator / Partner Administrator Team at their partner organization that has to approve their access.. Until that step is completed, they will have access issues/limitation. They will need to identify who this person or team is at their organization

## Program issues or concerns?

- Problems with **accessing** Cloud Skills Boost for Partners
  - [cloud-partner-training@google.com](mailto:cloud-partner-training@google.com)
- Problems with **a lab** (locked out, etc.)
  - [support@qwiklabs.com](mailto:support@qwiklabs.com)
- Problems with accessing Partner Advantage
  - <https://support.google.com/googlecloud/topic/9198654>

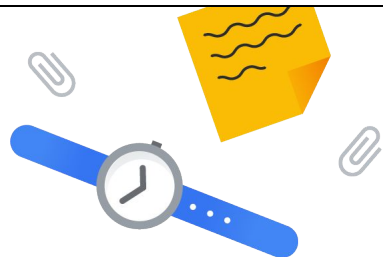


Google Cloud

- Problems with accessing **Cloud Skills Boost for Partners**
  - [cloud-partner-training@google.com](mailto:cloud-partner-training@google.com)
- Problems with **a lab** (locked out, etc.)
  - [support@qwiklab.com](mailto:support@qwiklab.com)
- Problems with accessing **Partner Advantage**
  - <https://support.google.com/googlecloud/topic/9198654>

## Module 7

# Monitoring ML solutions



# Module Agenda

01

Identifying risks to ML solutions

02

Monitoring, testing, and troubleshooting ML solutions







## Identifying risks to ML solutions

**ML systems are susceptible to attack,  
just like other systems.**

## Reference

# Vertex AI Network Security Overview

## Private access options for Vertex AI

Vertex AI supports the following options for accessing Vertex AI endpoints and services privately, without an external IP address:

- [Private Service Connect endpoints for Google APIs](#) let your Google Cloud resources or on-premises systems connect to an endpoint in your VPC network, which forwards requests to Google APIs and services.
- [Private Google Access](#):
  - Lets Google Cloud resources connect to the [standard external IP addresses or Private Google Access domains and VIPs](#) for Google APIs and services through the VPC network's default internet gateway.
  - Lets on-premises hosts connect to Google APIs and services through a Cloud VPN tunnel or VLAN attachment by using one of the [Private Google Access-specific domains and VIPs](#).
- [Private services access](#) lets Google Cloud VM instances connect to a Google or third-party managed VPC network through a VPC Network Peering connection.
- [Private Service Connect endpoints](#) lets Google Cloud VM instances connect to services in another VPC network through an endpoint.

The following table shows the supported access methods for connecting from on-premises and multicloud

Google Cloud

There are a lot of similarly named ways of providing secure private access off the public internet, so let's look at the documentation where you can disambiguate your needs.

Docs: <https://cloud.google.com/vertex-ai/docs/general/netsec-overview>

# VPC Networking Patterns

If you are using a VPC network to secure access to your data:

[Private Services Access](#) will grant your VPC access to most Vertex AI services, such as standard online prediction endpoints.

You will need to [set up VPC network peering](#) to support:

- [Custom training jobs](#)
- [Private prediction endpoints](#)
- [Vector matching online queries](#)

# VPC Service Controls

You can configure [VPC Service Controls for Vertex AI](#) to mitigate the risk of data being exfiltrated.

For example, when you use VPC Service Controls to protect Vertex AI, the following artifacts can't leave your service perimeter:

- Training data for an AutoML model or custom model
- Models that you created
- Models that you searched by using Neural Architecture Search
- Requests for online predictions
- Results from a batch prediction request

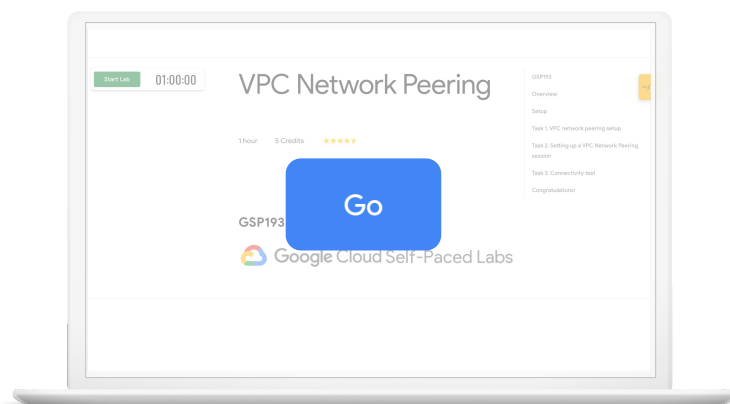
Recommended Lab

## VPC Network Peering

From the Course

Security & Identity Fundamentals

Select this lab from the section:  
"Security & Identity  
Fundamentals"

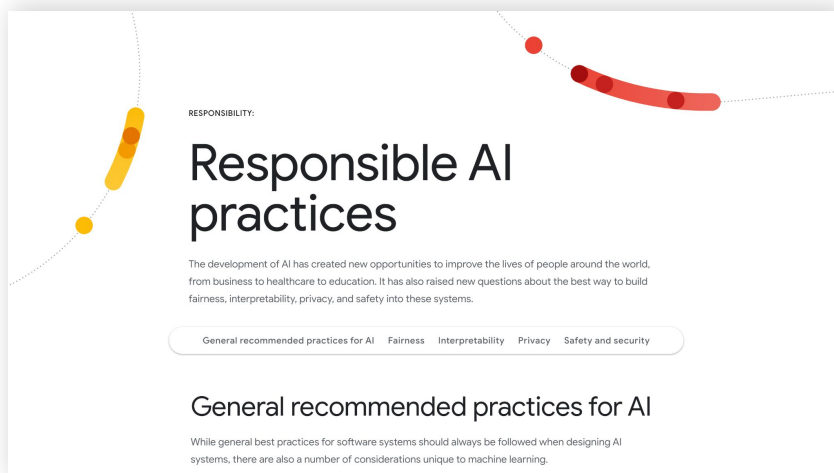


Select the VPC Network Peering lab from the Security & Identity Fundamentals  
Course: [https://partner.cloudskillsboost.google/course\\_templates/770](https://partner.cloudskillsboost.google/course_templates/770)

**ML systems are often used to advise decision-making. This should be done responsibly & fairly.**

## Reference

# Google Responsible AI Practices



[Google Responsible AI Practices](#)



## Demo

# Attack discrimination with smarter machine learning

## Simulating loan decisions for different groups

Drag the black threshold bars left or right to change the cut-offs for loans.  
Click on different preset loan strategies.

### Loan Strategy

Maximize profit with:

**MAX PROFIT**

No constraints

**GROUP UNAWARE**

Blue and orange thresholds  
are the same

**DEMOGRAPHIC  
PARITY**

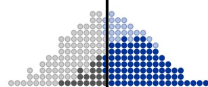
Same fractions blue / orange loans

**EQUAL  
OPPORTUNITY**

### Blue Population

0 10 20 30 40 50 60 70 80 90 100

loan threshold: 55

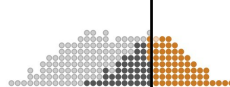


denied loan / would default    granted loan / defaults  
denied loan / would pay back    granted loan / pays back

### Orange Population

0 10 20 30 40 50 60 70 80 90 100

loan threshold: 55



denied loan / would default    granted loan / defaults  
denied loan / would pay back    granted loan / pays back

DEMO: [Attack discrimination with smarter machine learning](#)

**Often, fair decision making requires that decisions are explainable.**

## Vertex Explainable AI overview

- Can show you examples that appear similar to a given example in your data.
- Can tell you how much each feature contributes in the predicted result
- Useful to improve model prediction (removing useless features)
- Helps to recognize bias
- Works for both structured and unstructured datasets and non-linear problems

Vertex Explainable AI helps you understand your model's outputs for classification and regression tasks. Vertex AI tells you how much each feature in the data contributed to the predicted result. You can then use this information to verify that the model is behaving as expected, recognize bias in your models, and get ideas for ways to improve your model and your training data.

Vertex AI supports Vertex Explainable AI for the following [types of models](#):

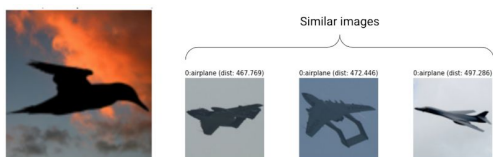
- AutoML image models (classification models only)
- AutoML tabular models (classification and regression models only)
- Custom-trained models based on tabular data
- Custom-trained models based on image data

For AutoML tabular models, feature attributions are displayed in Google Cloud console as *feature importance*. You can see [model feature importance](#) for the model overall, and *local feature importance* for both [online](#) and [batch](#) predictions.

Docs: <https://cloud.google.com/vertex-ai/docs/explainable-ai/overview>

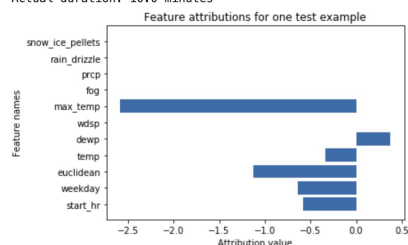
# Two kinds of explainability

## Example-based



## Feature-based

Predicted duration: 11.1651134 minutes  
Actual duration: 10.0 minutes



## Example-based

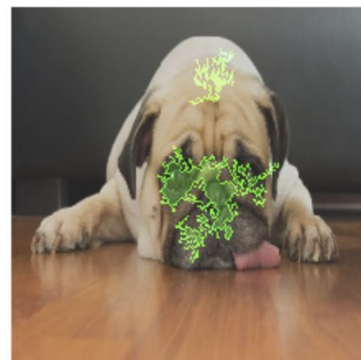
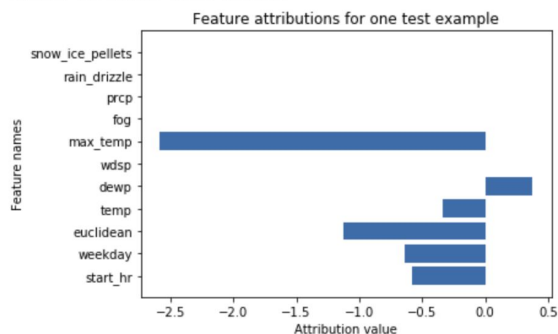
Uses nearest neighbor search to return a list of examples (typically from the training set) that are most similar to the input.

## Feature-based

Feature attributions indicate how much each feature in your model contributed to the predictions for each given instance. When you request predictions, you get predicted values as appropriate for your model. When you request explanations, you get the predictions along with feature attribution information.

## Feature attributions

Predicted duration: 11.1651134 minutes  
Actual duration: 10.0 minutes



Google Cloud

Source <https://cloud.google.com/vertex-ai/docs/explainable-ai/overview>

Feature attributions indicate how much each feature in your model contributed to the predictions for each given instance. When you request predictions, you get predicted values as appropriate for your model. When you request *explanations*, you get the predictions along with feature attribution information.

Feature attributions work on tabular data, and include built-in visualization capabilities for image data. Consider the following examples:

- A deep neural network is trained to predict the duration of a bike ride, based on weather data and previous ride sharing data. If you request only predictions from this model, you get predicted durations of bike rides in number of minutes. If you request *explanations*, you get the predicted bike trip duration, along with an attribution score for each feature in your explanations request. The attribution scores show how much the feature affected the change in prediction value, relative to the baseline value that you specify. Choose a meaningful baseline that makes sense for your model - in this case, the median bike ride duration. You can plot the feature attribution scores to see which features contributed most strongly to the resulting prediction
- An image classification model is trained to predict whether a given

- image contains a dog or a cat. If you request predictions from this model on a new set of images, then you receive a prediction for each image ("dog" or "cat"). If you request *explanations*, you get the predicted class along with an overlay for the image, showing which pixels in the image contributed most strongly to the resulting prediction:



## Monitoring, testing, and troubleshooting ML solutions

# Model Monitoring: Compare Prediction Distributions to Training Distributions

## Visualize feature skew and drift

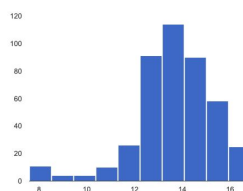
Visualizing data distribution as histograms lets you quickly focus on the changes that occurred in the data. Afterward, you might decide to adjust your feature generation pipeline or retrain the model.

### Feature distribution

Synthetic of distributions when job ran at Apr 29, 2021, 10:00:00 AM

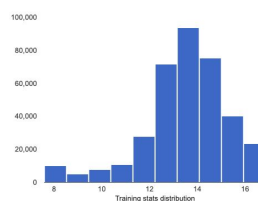
### Latest prediction stats distribution

Hover over the chart to view stats



### Monitoring jobs (up to last 50)

- Apr 29, 2021, 10:00:00 AM
- Apr 29, 2021, 9:00:00 AM
- Apr 29, 2021, 8:00:00 AM
- Apr 19, 2021, 9:00:00 PM
- Apr 19, 2021, 8:00:00 PM
- Apr 19, 2021, 7:00:00 PM
- Apr 19, 2021, 6:00:00 PM
- Apr 19, 2021, 5:00:00 PM
- Apr 19, 2021, 4:00:00 PM
- Apr 16, 2021, 10:00:00 PM
- Apr 16, 2021, 9:00:00 PM
- Apr 16, 2021, 8:00:00 PM
- Apr 16, 2021, 7:00:00 PM
- Apr 16, 2021, 6:00:00 PM
- Apr 16, 2021, 5:00:00 PM



[Model Monitoring Documentation](#)  
and [a nice demo video](#).

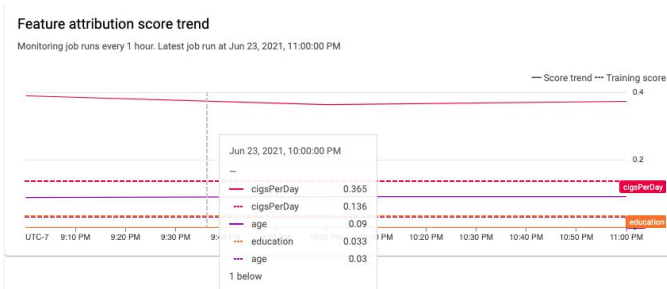
Recommendation: play the demo video.



# Visualize Changes in Feature Attribution

## Visualize feature attribution skew and drift

In a stable machine learning system, features' relative importance generally remains relatively stable over time. If an important feature drops in importance, it might signal that something about that feature has changed.



[Model Monitoring Documentation](#)  
and [a nice demo video](#).

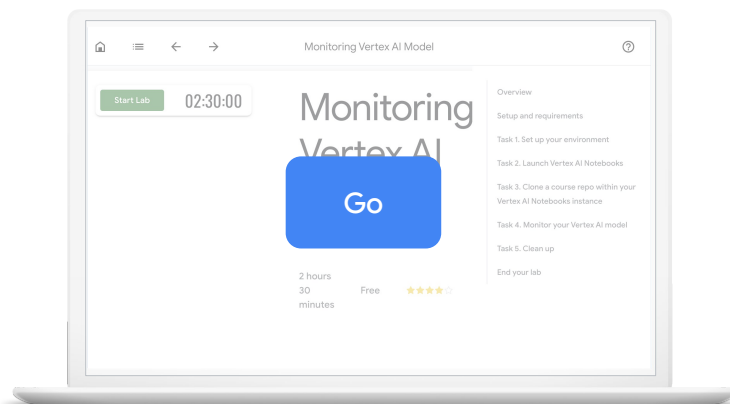
Recommendation: play the demo video.

Recommended Lab

## Monitoring a Vertex AI Model

From the Course  
[Machine Learning in the Enterprise](#)

Select this lab from the section:  
"Prediction and Model Monitoring  
Using Vertex AI"



Google Cloud

### [Monitoring a Vertex AI Model](#)

This is a long lab.

Alternate Colab notebooks follow on next slides are:

[Colab: Debugging in Classification](#)

[Colab: Debugging in Regression](#)

From [Debugging Models](#) in ML Foundational Course

# Questions and answers



## Thank you for attending this training!

We love your feedback! Please take  
a minute to complete the survey and  
help us improve our courses.

