

# Cloud Computing Machine Learning / Generative AI

---

CIS437

Erik Fredericks // [frederer@gvsu.edu](mailto:frederer@gvsu.edu)

*Adapted from Google Cloud Computing Foundations, Overview of Cloud Computing (Wufka & Canonic)*

WEEEEEE

[https://www.youtube.com/watch?v=tF\\_4baiUiQ](https://www.youtube.com/watch?v=tF_4baiUiQ)

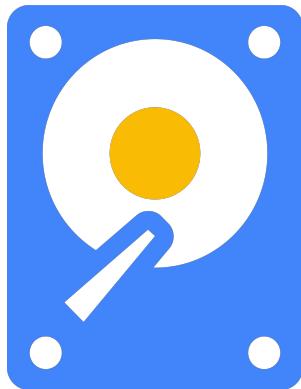


@teenybiscuit

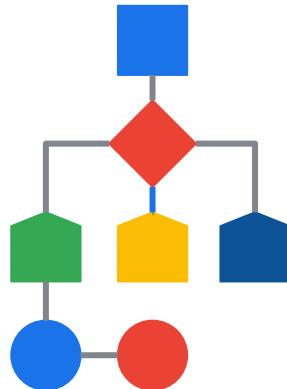


Watch separately - fluff piece on what AI is

# ML uses standard algorithms to derive predictive insights from data and make repeated decisions



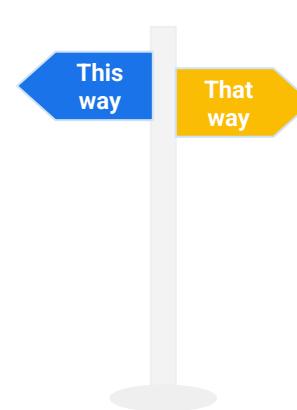
Data



Algorithm



Predictive insight

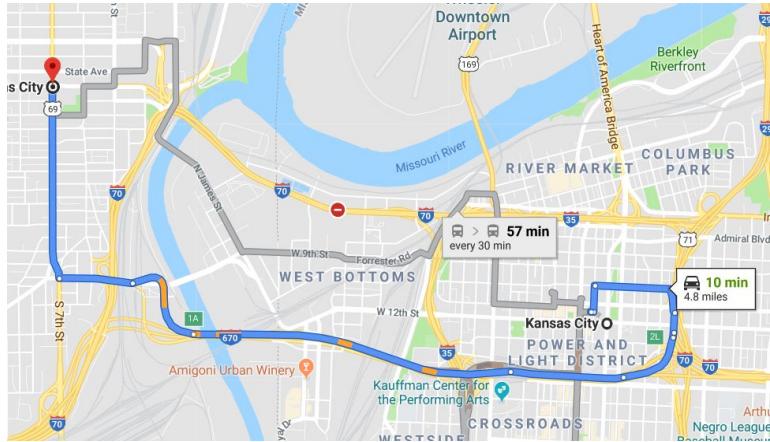


Decision

# ML uses standard algorithms



Estimate the tax I owe



How long will it take me to get home?

# Model training requires examples

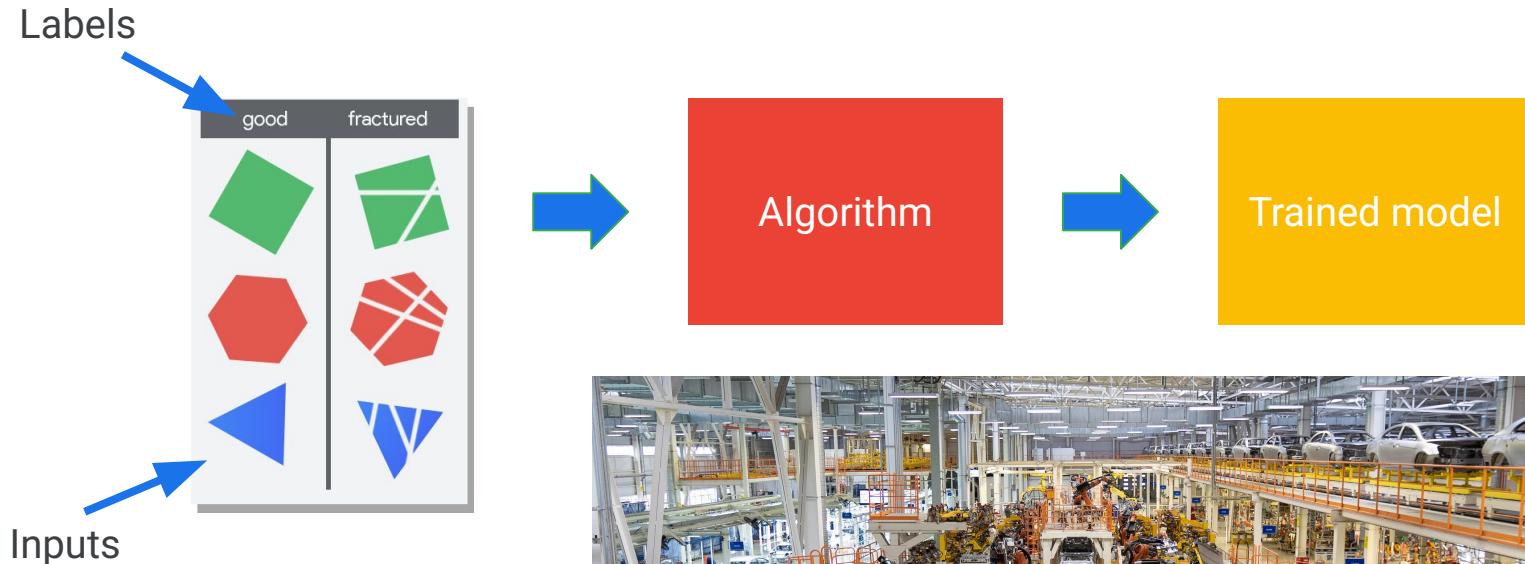


Examples of tax filings

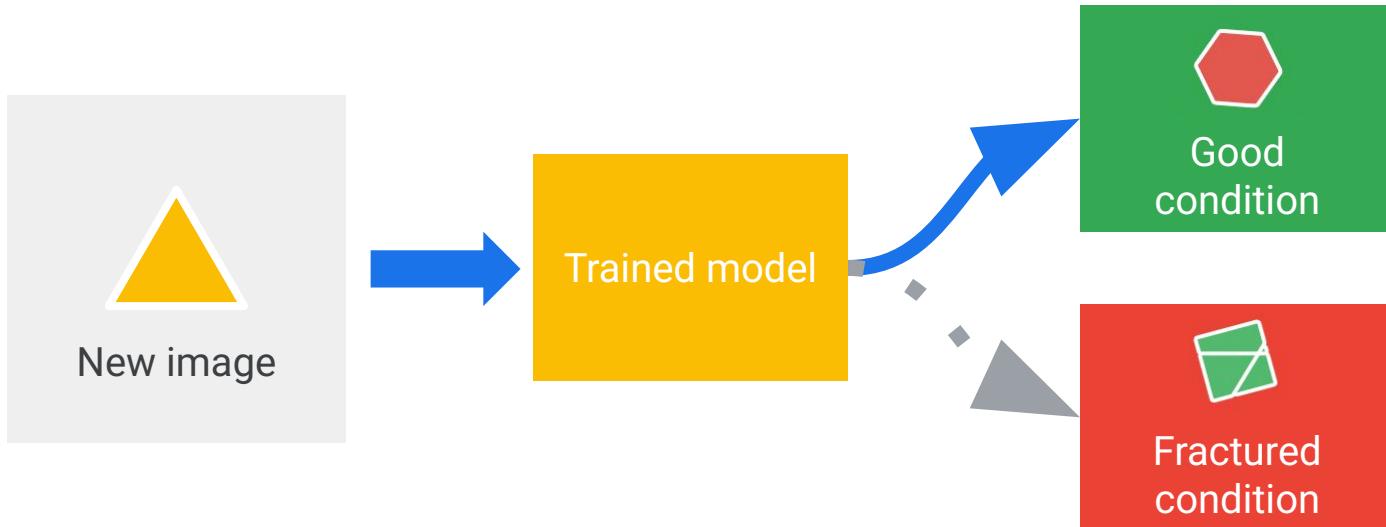


Examples of trips

# Train an ML model with examples



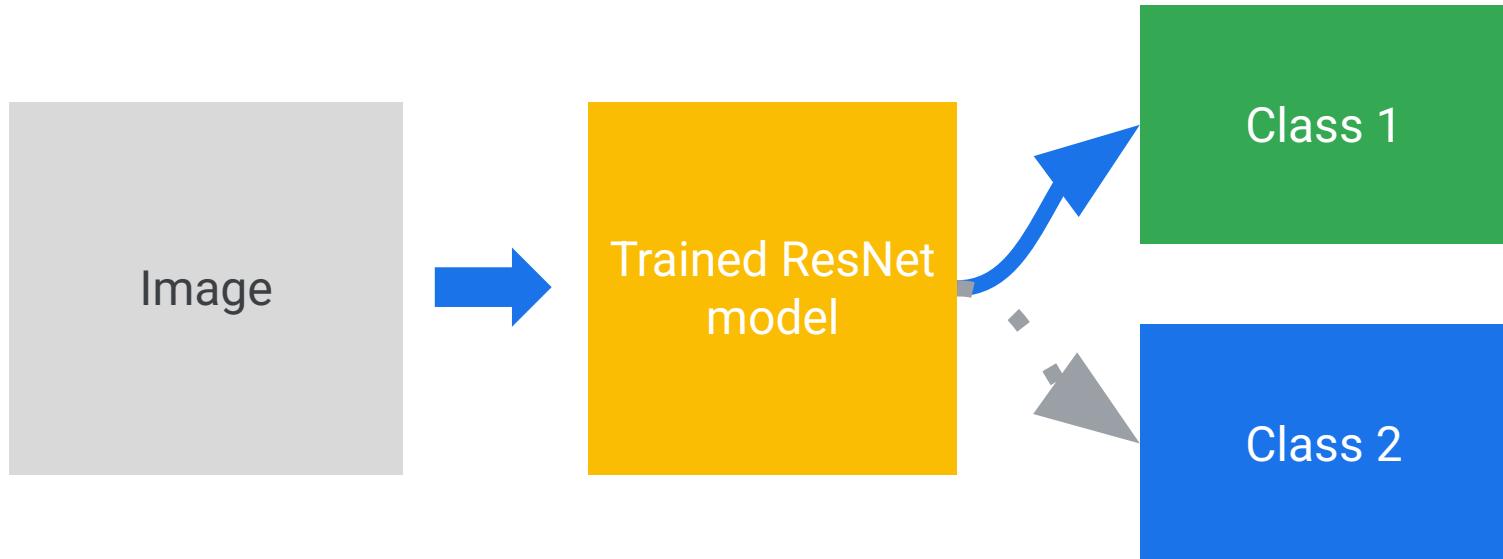
# Predict with a trained model



# Standard algorithm use cases

- 1 Detect a pattern in an image.
- 2 Predict the future of a time series.
- 3 Understand or transcribe human speech or text.

# A standard algorithm for image classification



---

The same algorithm applied to other data yields a different model



---

The algorithm is the same, but the trained model is different



Image model  
(trained)



Water lily

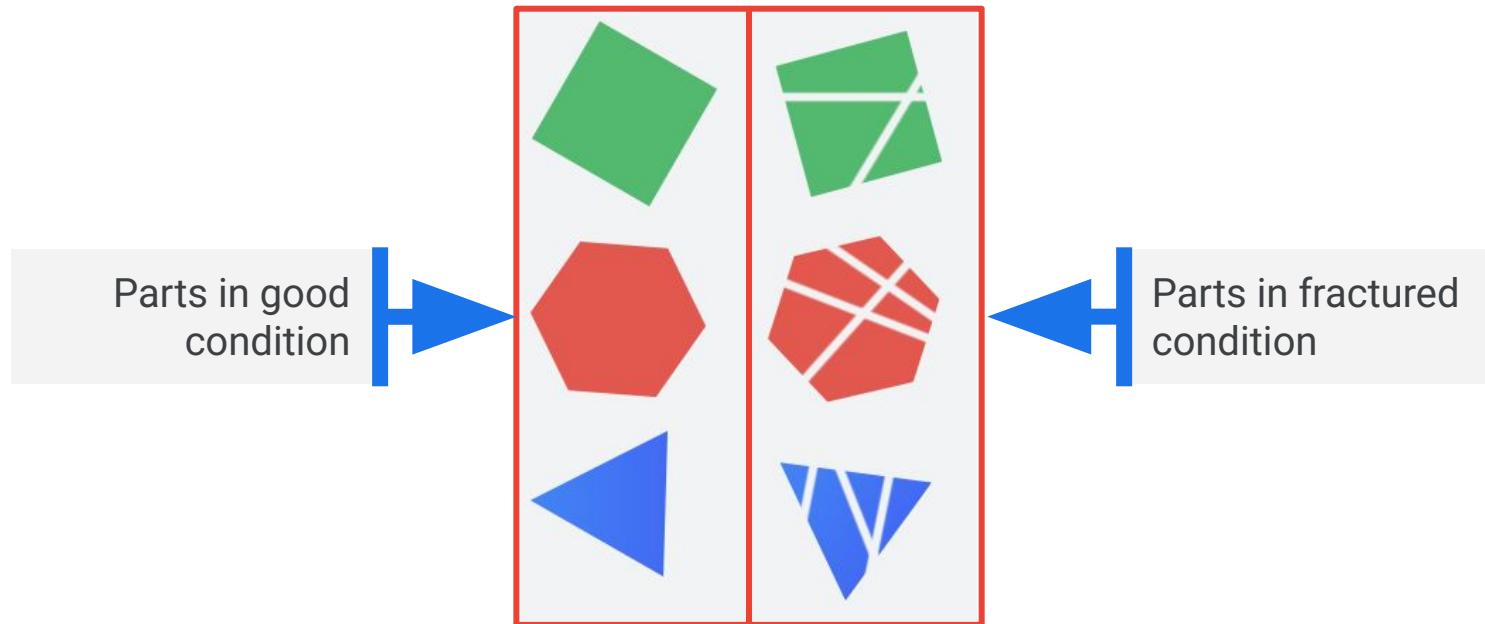


Image model  
(trained)

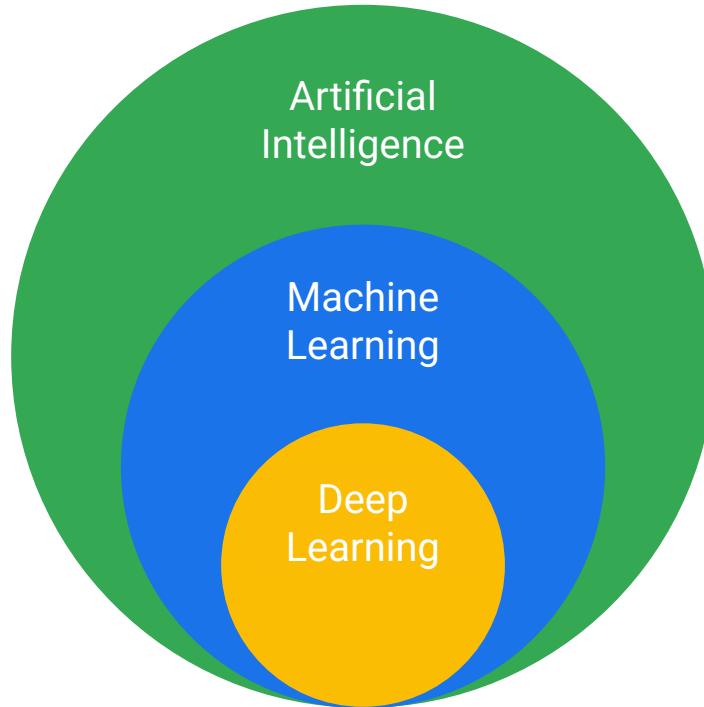


Fractured part

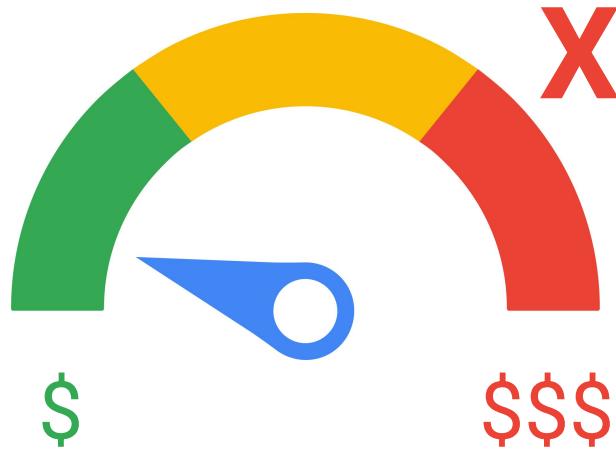
# What data do you need?



# ML is a type of AI



# The impact of ML is scale



Not about saving money

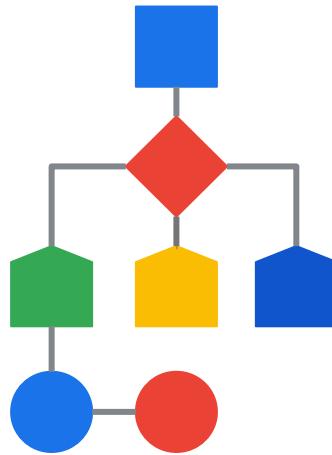


But about doing it at greater scale

# Barriers to entry have now fallen



Data

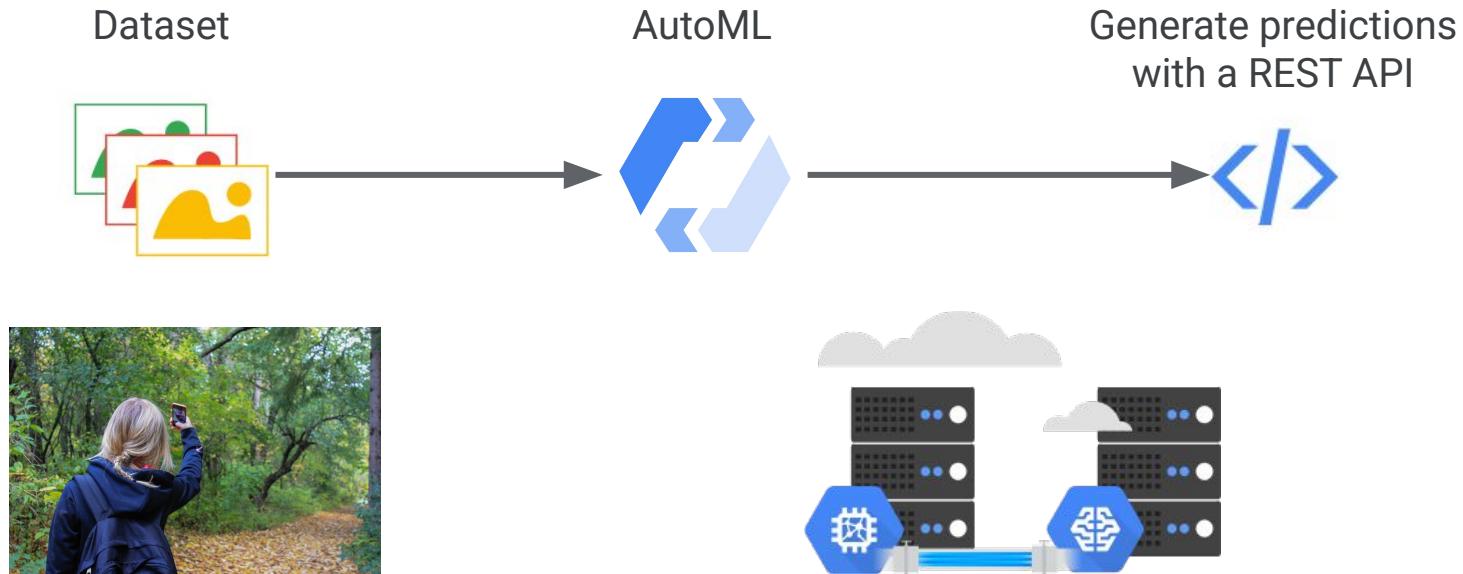


Algorithm



Hardware  
Software

# How does this translate to building an ML model to identify diseased leaves?



# The Google Cloud machine learning spectrum

## Pre-trained ML models

Ready to go



Vision  
API



Speech - to -  
Text API



Cloud Talent  
Solution API



Cloud  
Translation  
API



Cloud Natural  
Language  
API



Video  
Intelligence  
API

## AutoML

Bring your own data



Vision  
Video Intelligence  
Natural language  
Translation  
Data tables

## ML frameworks

More control for advanced users

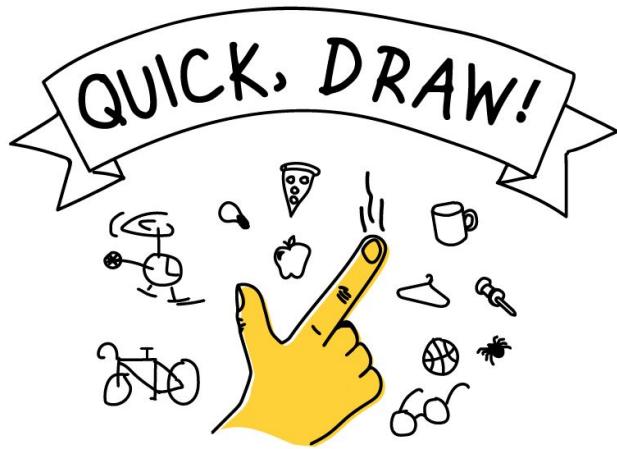


TensorFlow



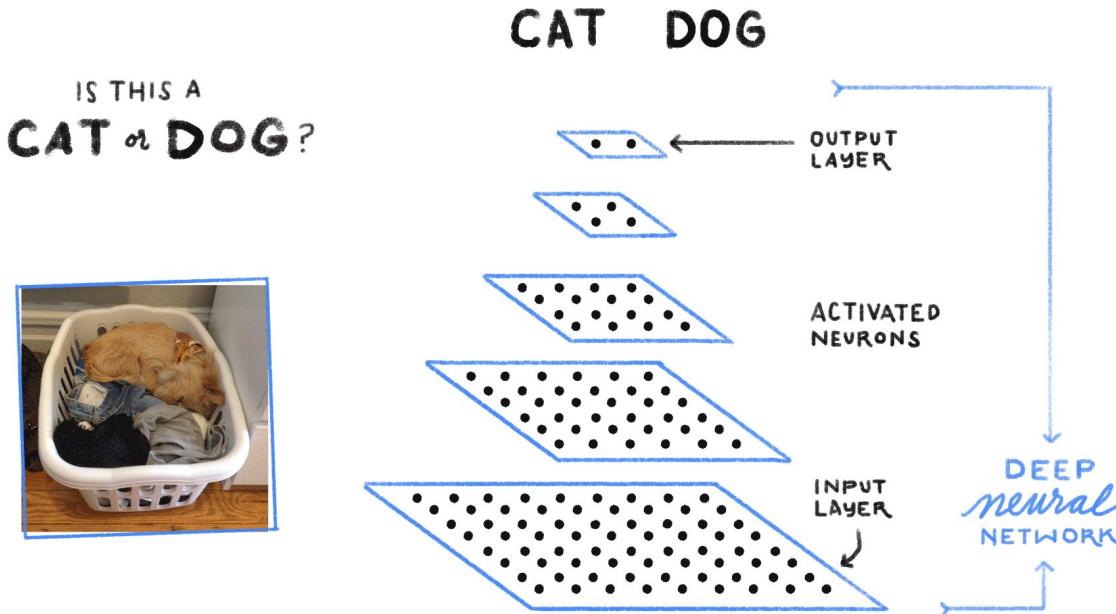
Vertex AI

# Having fun with ML: Quick, Draw



[quickdraw.withgoogle.com](https://quickdraw.withgoogle.com)

# Modern AI applications use deep learning



Epoch  
000,000

Learning rate: 0.03 Activation: Tanh Regularization: None Regularization rate: 0 Problem type: Classification

**DATA**  
Which dataset do you want to use?  
   
   
Ratio of training to test data: 50%  
Noise: 0  
Batch size: 10  
REGENERATE

**FEATURES**  
Which properties do you want to feed in?  
 $X_1$    
 $X_2$    
 $X_1^2$    
 $X_2^2$    
 $X_1 X_2$    
 $\sin(X_1)$    
 $\sin(X_2)$  

**HIDDEN LAYERS**  
+ - 4 HIDDEN LAYERS  
+ - 2 neurons   + - 2 neurons   + - 6 neurons   + - 5 neurons  
*This is the output from one neuron. Hover to see it larger.*

**OUTPUT**  
Test loss 0.518  
Training loss 0.504

The outputs are mixed with varying weights, shown by the thickness of the lines.

Colors shows data, neuron and weight values.  
 Show test data  Discretize output

<http://playground.tensorflow.org>

---

## Interim Demo - ML APIs (handmade, though reflective of the ones at the end)

[https://codelabs.cs.pdx.edu/labs/C07.4g\\_mlapis/index.html?index=..%2F..cs430](https://codelabs.cs.pdx.edu/labs/C07.4g_mlapis/index.html?index=..%2F..cs430)

# Agenda

Introduction to Machine Learning  
in the Cloud

**Building Bespoke Machine  
Learning Models with Vertex AI**

Lab: AI Platform Qwik Start

AutoML

Google's Pre-trained Machine  
Learning APIs



# For the experts!

## Pre-trained ML models

Ready to go



Vision  
API



Speech - to -  
Text API



Cloud Talent  
Solution API



Cloud  
Translation  
API



Cloud Natural  
Language  
API



Video  
Intelligence  
API

## AutoML

Bring your own data



Vision  
Video Intelligence  
Natural language  
Translation  
Data tables

## ML frameworks

More control for advanced users



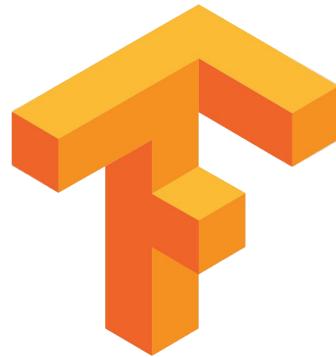
TensorFlow



Vertex AI

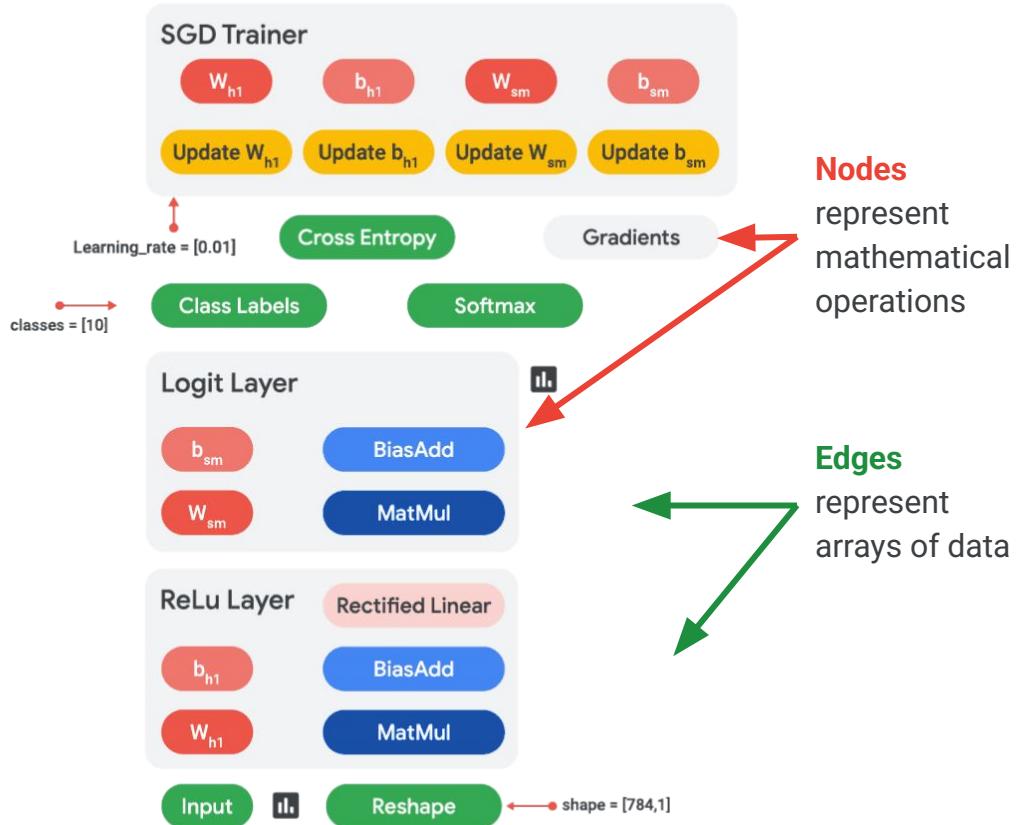
---

# Create custom ML models with TensorFlow



# TensorFlow

# How TensorFlow works



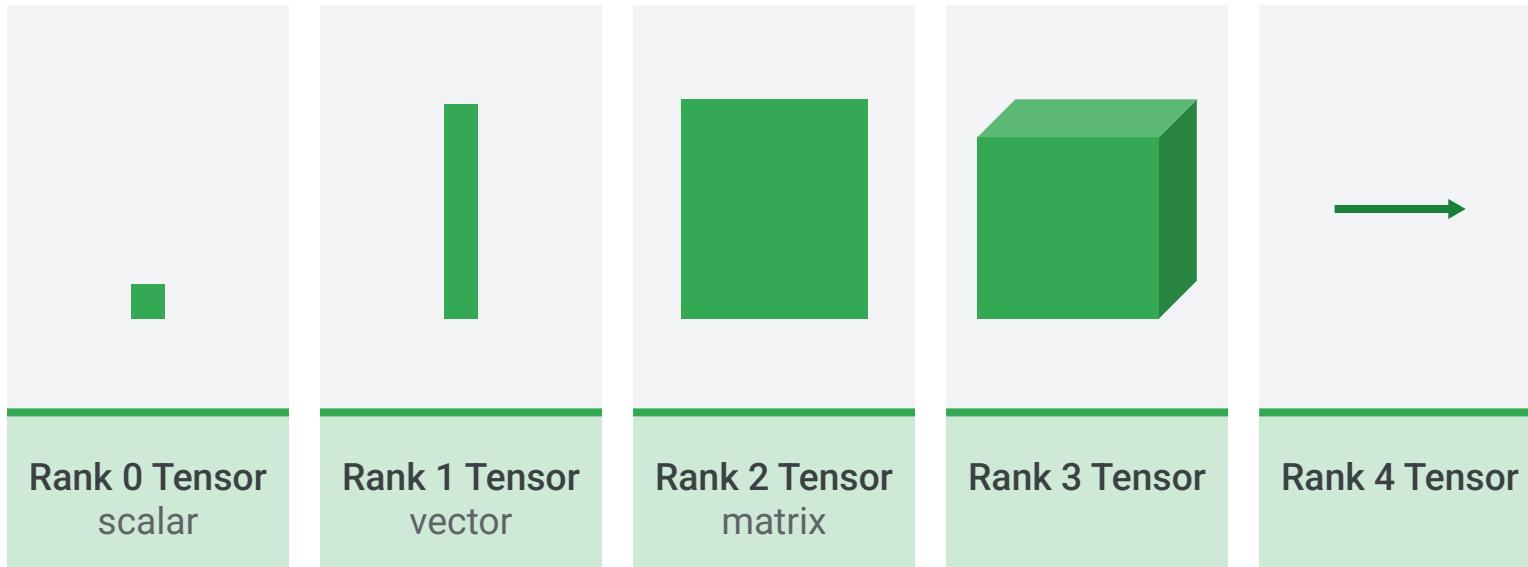
## Nodes

represent mathematical operations

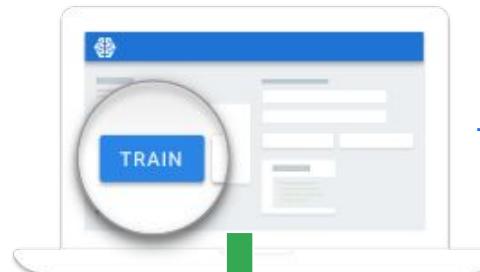
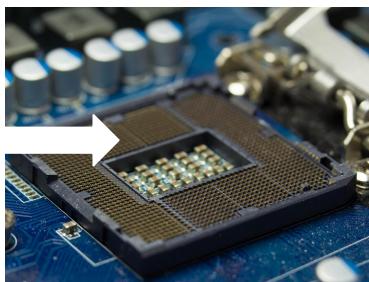
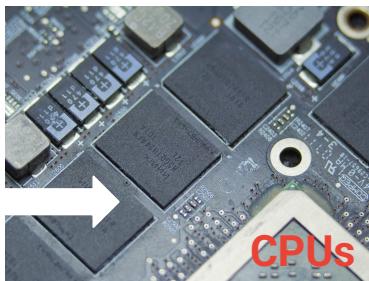
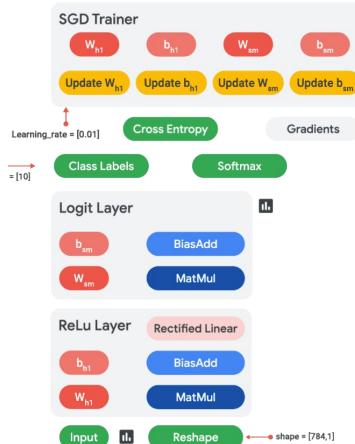
## Edges

represent arrays of data

# A tensor is an N-dimensional array of data



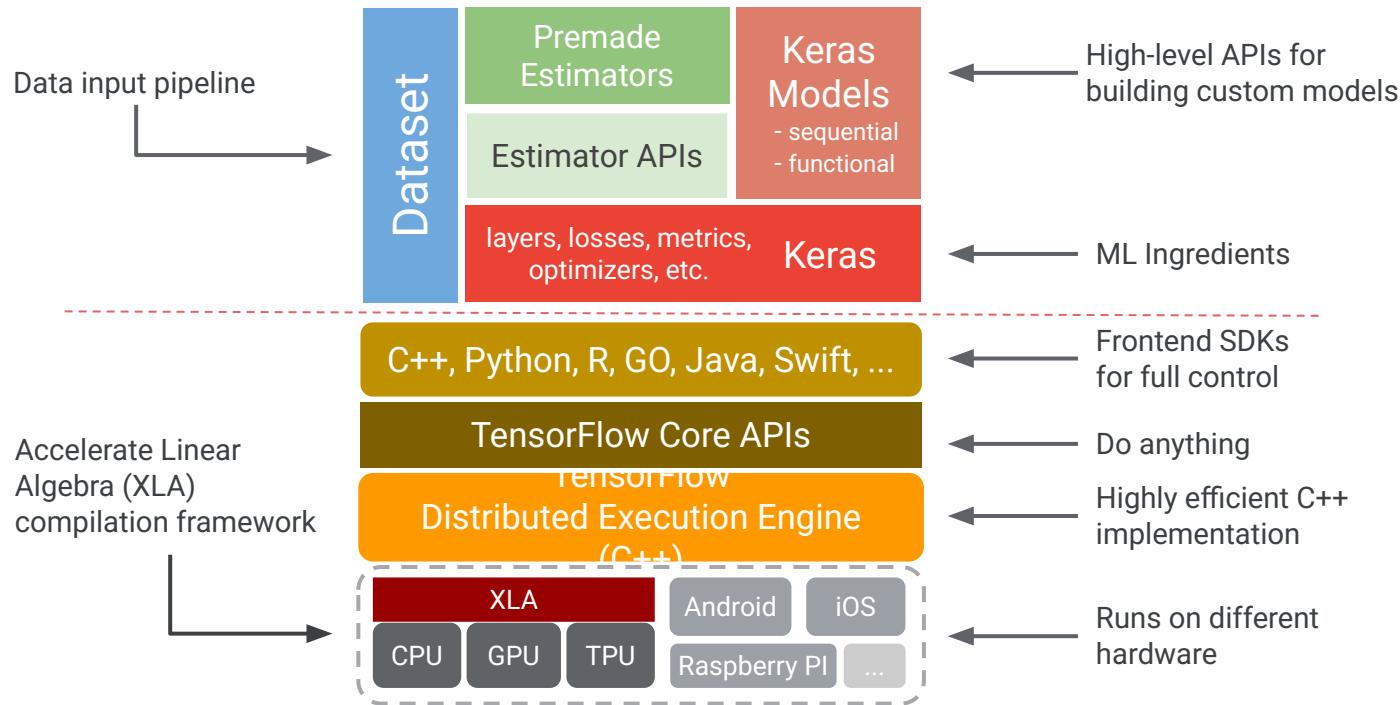
# TensorFlow graphs are portable between different devices



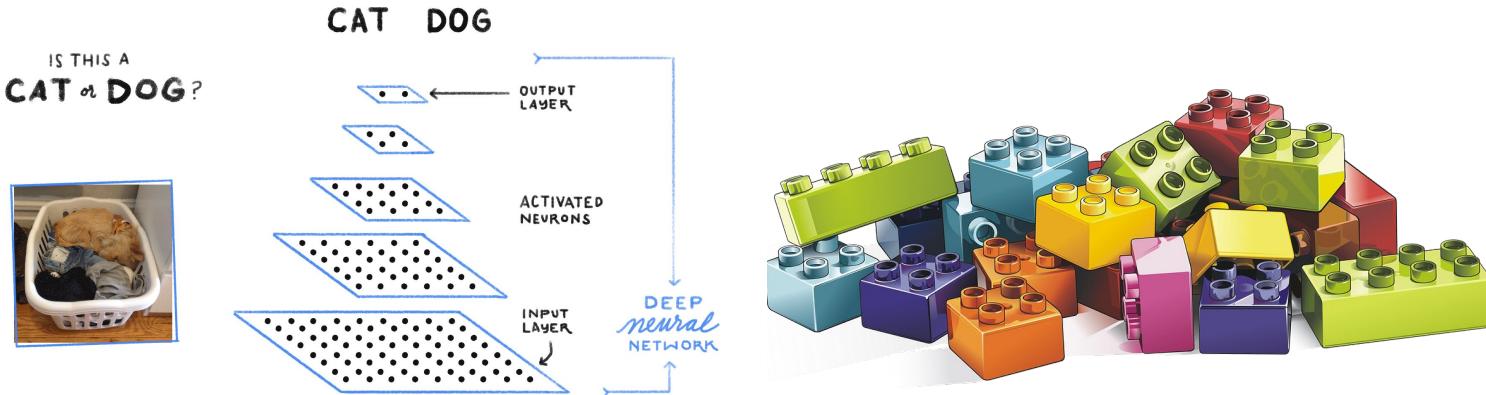
Train on cloud.

Run inference on  
iOS, Android,  
Raspberry Pi, etc.

# TensorFlow contains multiple abstraction layers

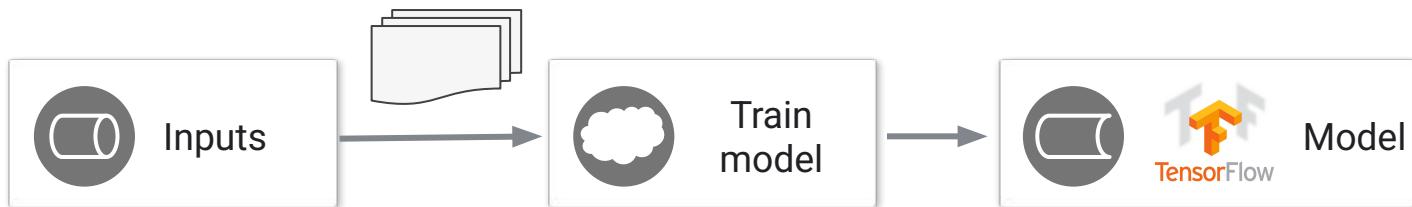
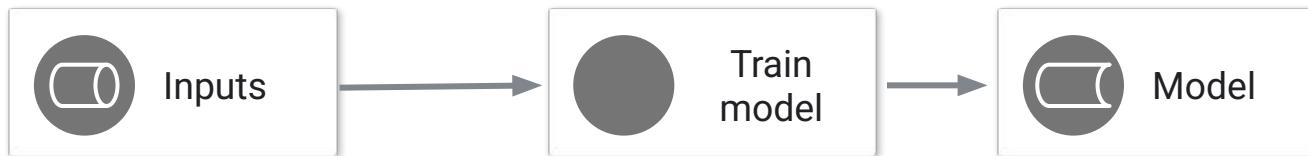


# Keras is a friendly high-level API for DNNs

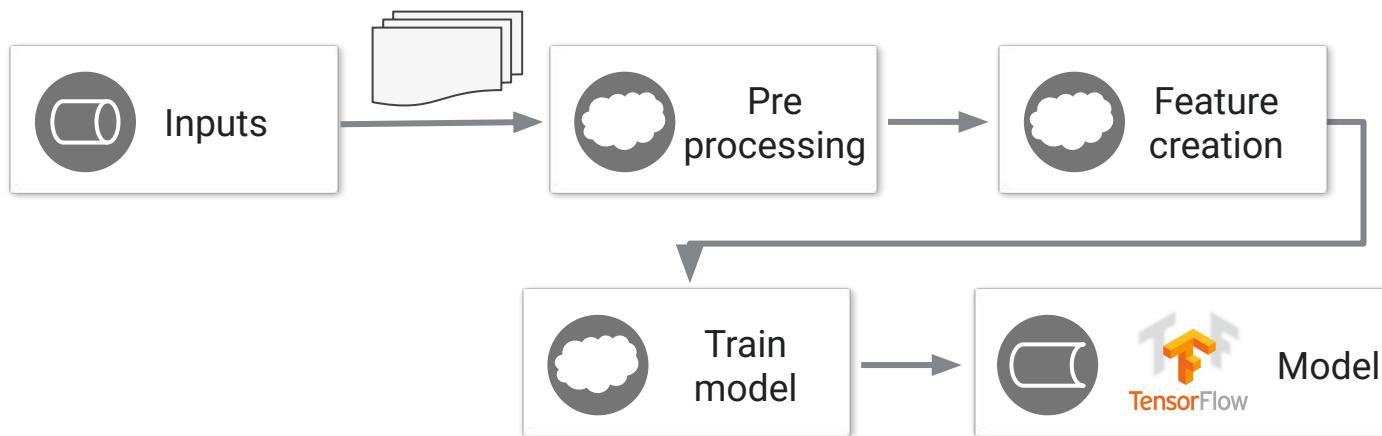


```
from keras.layers import Dense  
  
model.add(Dense(units=64, activation='relu', input_dim=100))  
model.add(Dense(units=10, activation='softmax'))
```

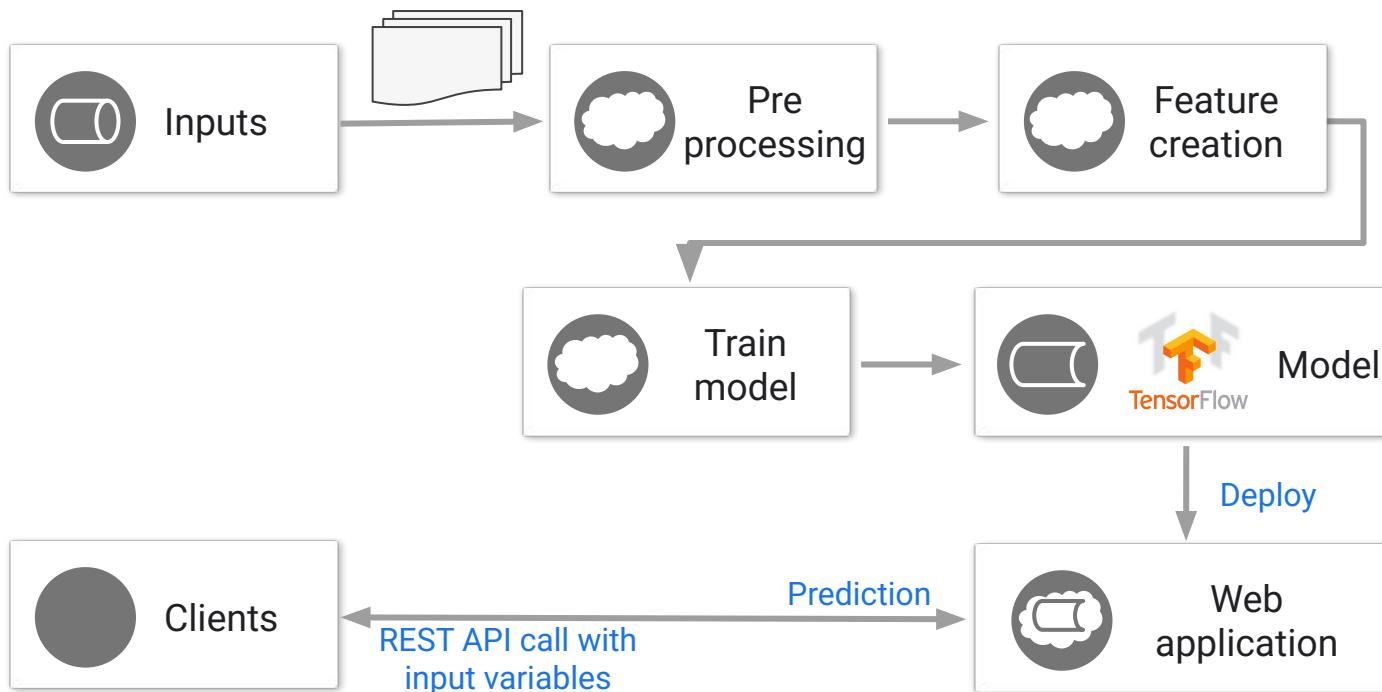
As your data size increases, batching and distribution become important



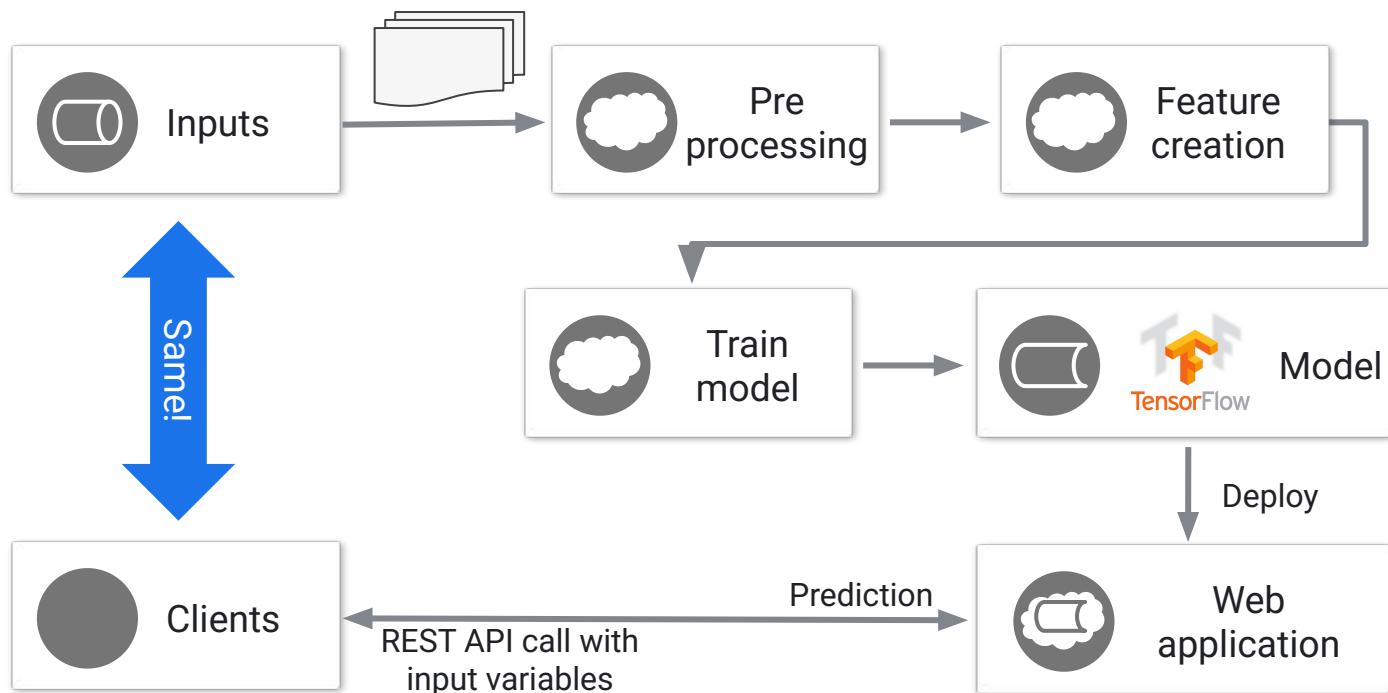
# Input necessary transformations



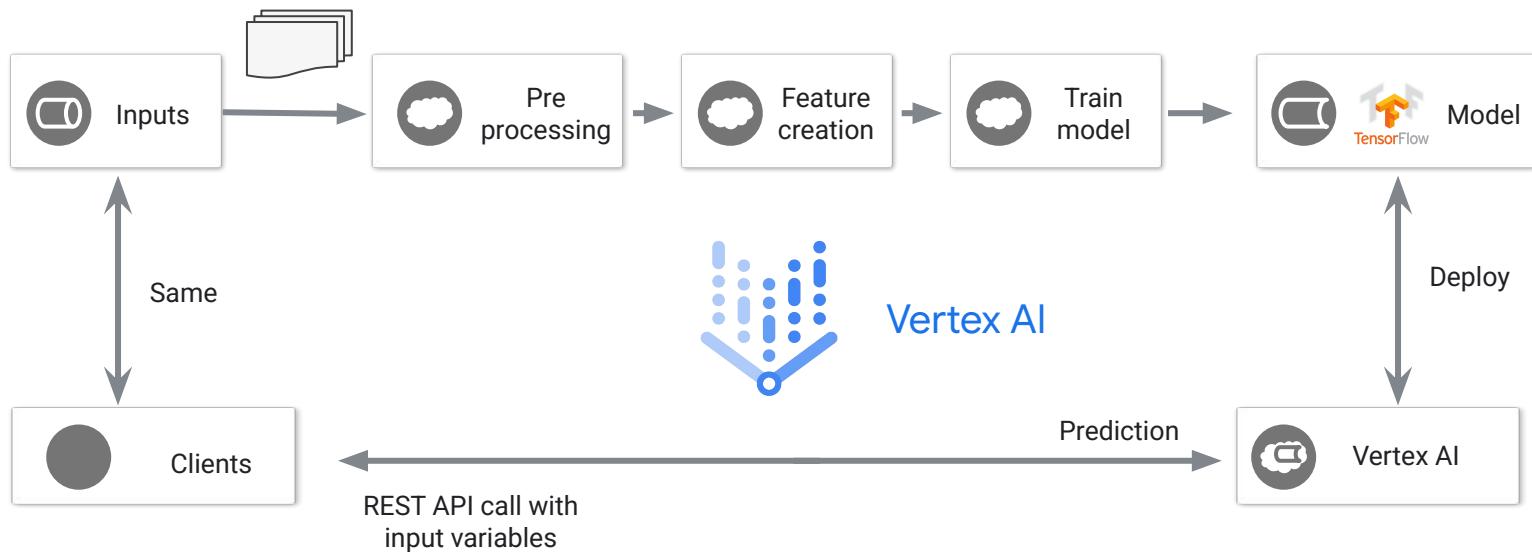
# ML code needs to scale



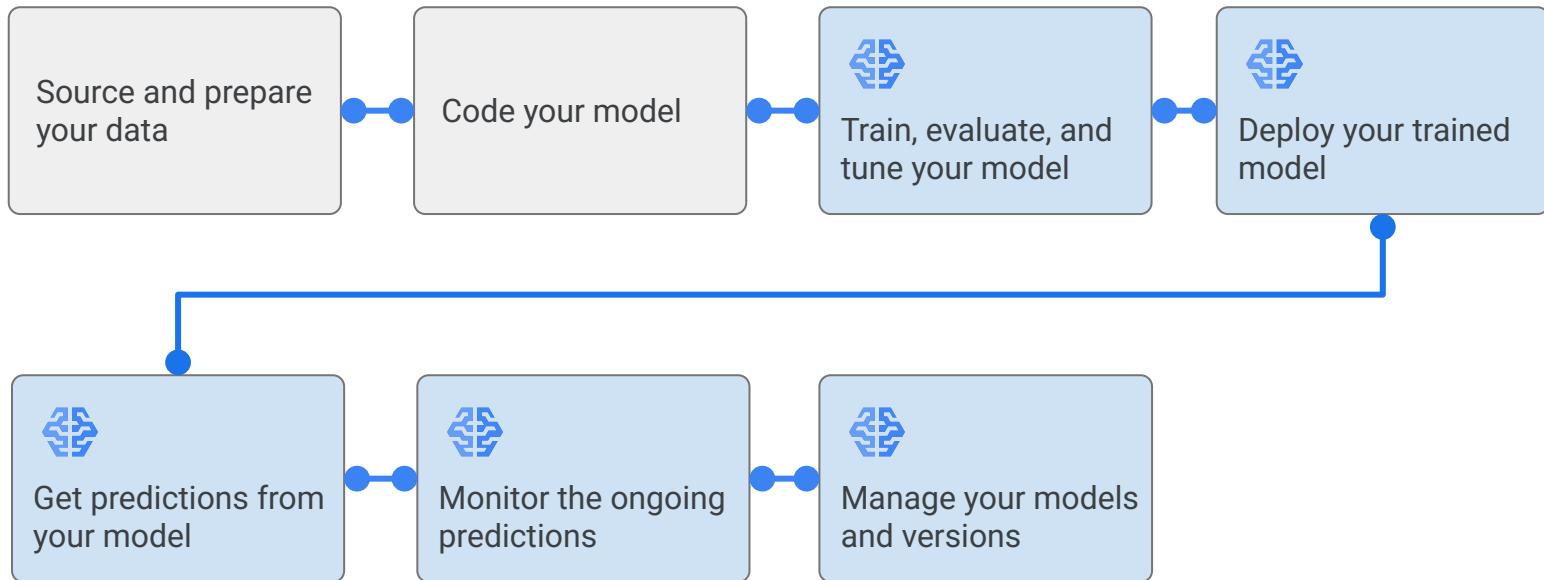
# Preprocessing may be required



# Vertex AI: Repeatable, scalable, tuned



# Vertex AI and the ML workflow



# Agenda

Introduction to Machine Learning  
in the Cloud

Building Bespoke Machine  
Learning Models with Vertex AI

[Lab: AI Platform Qwik Start](#)

AutoML

Google's Pre-trained Machine  
Learning APIs



---

# Lab Intro

## AI Platform: Qwik Start

Train and deploy a TensorFlow model to AI Platform for serving (prediction).

The lab can be found [here](#).



(If you want to play  
with TF)

# Machine Learning with TensorFlow in Vertex AI

[https://www.cloudskillsboost.google/focuses/3391?catalog\\_rank=%7B%22rank%22%3A6%2C%22num\\_filters%22%3A0%2C%22has\\_search%22%3Atrue%7D&parent=catalog&search\\_id=39816913](https://www.cloudskillsboost.google/focuses/3391?catalog_rank=%7B%22rank%22%3A6%2C%22num_filters%22%3A0%2C%22has_search%22%3Atrue%7D&parent=catalog&search_id=39816913)

## Lab objectives (1/2)

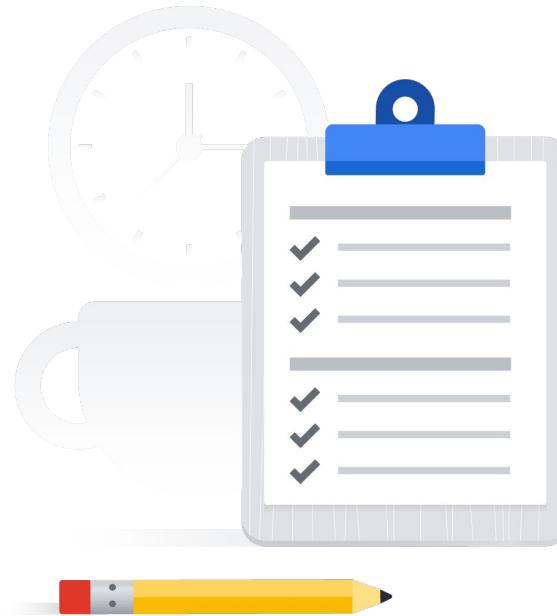
Create a TensorFlow training application and validate it locally.

Run your training job on a single worker instance in the cloud.

Run your training job as a distributed training job in the cloud.

Optimize your hyperparameters by using hyperparameter tuning.

Deploy a model to support prediction.



---

## Lab objectives (2/2)

Request an online prediction and see the response.

Request a batch prediction.

---

# Lab Intro

~~Scikit learn Model Serving with  
Online Prediction Using AI Platform  
(Alternative)~~

~~Train a simple scikit learn model, deploy the  
model to AI Platform Prediction, and make  
online predictions against that model.~~

~~The lab can be found [here](#).~~



# Bracketology with Google Machine Learning

[https://www.skills.google/focuses/  
4337?catalog\\_rank=%7B%22rank%  
22%3A8%2C%22num\\_filters%22%3  
A0%2C%22has\\_search%22%3Atru  
e%7D&parent=catalog&search\\_id=  
60097245](https://www.skills.google/focuses/4337?catalog_rank=%7B%22rank%22%3A8%2C%22num_filters%22%3A0%2C%22has_search%22%3Atrue%7D&parent=catalog&search_id=60097245)



# Lab objectives

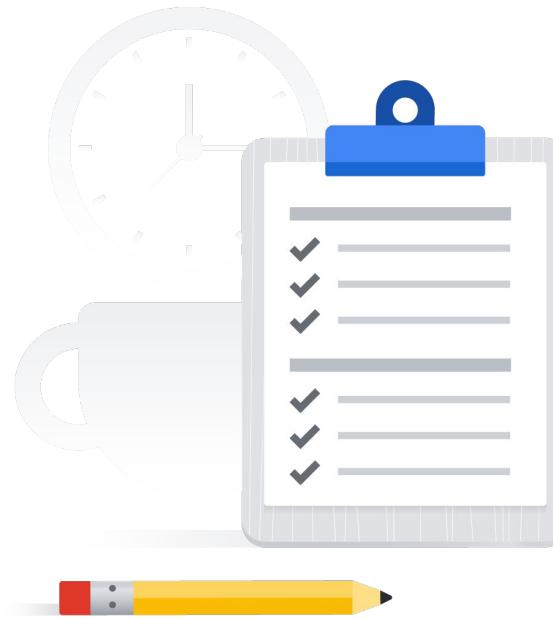
Create and save a model to a file.

Upload a saved model to Google Cloud Storage.

Create a model resource in AI Platform.

Create a model version (linking your scikit-learn model).

Make an online prediction.



# Agenda

Introduction to Machine Learning  
in the Cloud

Building Bespoke Machine  
Learning Models with Vertex AI

Lab: AI Platform Qwik Start

[AutoML](#)

Google's Pre-trained Machine  
Learning APIs



# Machine learning for the masses!

## Pre-trained ML models

Ready to go



Vision  
API



Speech - to -  
Text API



Cloud Talent  
Solution API



Cloud  
Translation  
API



Cloud Natural  
Language  
API



Video  
Intelligence  
API

## AutoML

Bring your own data



Vision  
Video Intelligence  
Natural language  
Translation  
Data tables

## ML frameworks

More control for advanced users



TensorFlow



Vertex AI

# What's required to solve an ML problem?

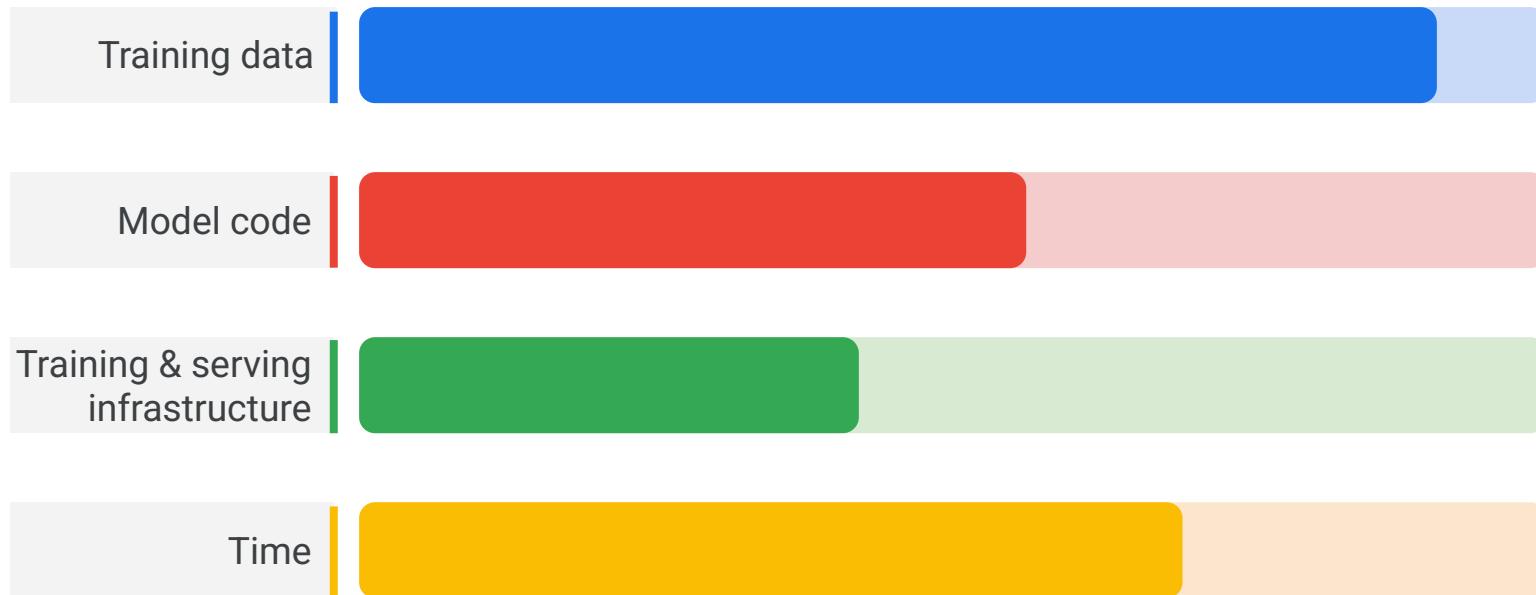
Training data

Model code

Training & serving  
infrastructure

Time

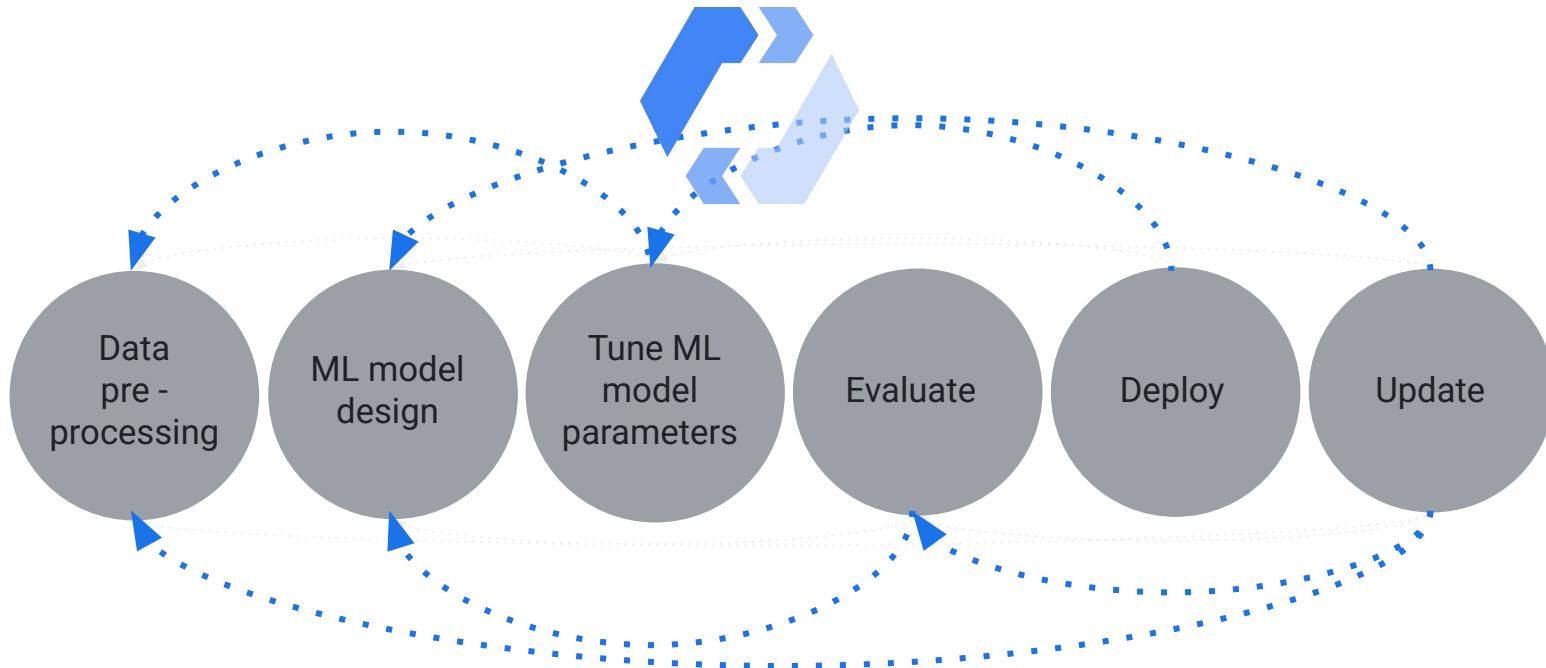
# What's required when using a managed service?



# What's required when using AutoML?



# AutoML simplifies the process



# Use AutoML for what you can see



## AutoML Vision

Derive insights from images in the cloud or at the edge.



## AutoML Video Intelligence

Enable powerful content discovery and engaging video experiences.

# Use AutoML for what you can hear



## AutoML Natural Language

Reveal the structure and meaning of text through machine learning.



## AutoML Translation

Dynamically translate between languages.

---

# Use AutoML to turn structured data into predictive insights



## AutoML Tables

Automatically build and deploy state-of-the-art machine learning models on structured data.

# Agenda (1/3)

Introduction to Machine Learning  
in the Cloud

Building Bespoke Machine  
Learning Models with Vertex AI

Lab: AI Platform Qwik Start

AutoML

Google's Pre-trained Machine  
Learning APIs



# Access pre-trained ML APIs for common applications

## Pre-trained ML models

Ready to go



Vision  
API



Speech - to -  
Text API



Cloud Talent  
Solution API



Cloud  
Translation  
API



Cloud Natural  
Language  
API



Video  
Intelligence  
API

## AutoML

Bring your own data



Vision  
Video Intelligence  
Natural language  
Translation  
Data tables

## ML frameworks

More control for advanced users

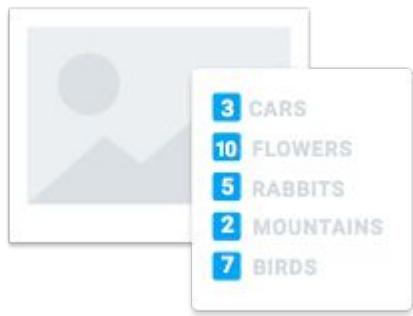


TensorFlow



Vertex AI

# Use the Vision API to understand image content



Detect and label

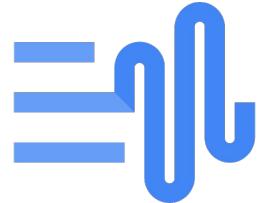
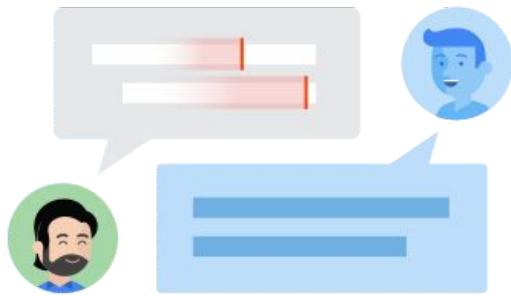


Extract text

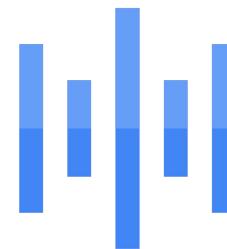


Identify entities

# Convert speech to text and vice versa



Text-to-Speech



Speech-to-Text

# Dynamically translate between languages using the Cloud Translation API

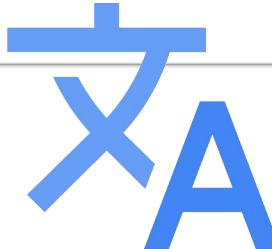
Source Language  
French (fr)

Target Language  
English (en)

Sample text. Enter your own text to translate.

Il ne faut avoir aucun regret pour le passé, aucun remords pour le présent, et une confiance inébranlable pour

There must be no regrets for the past, no remorse for the present, and unshakable confidence for the future.



A photograph showing four professionals in a modern office environment, likely a trading floor or data center. Two men and two women are visible, all dressed in business attire. They are gathered around a large, curved screen displaying various data visualizations and charts, possibly related to finance or technology. The room has a high ceiling with recessed lighting and large windows in the background.

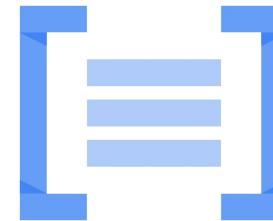
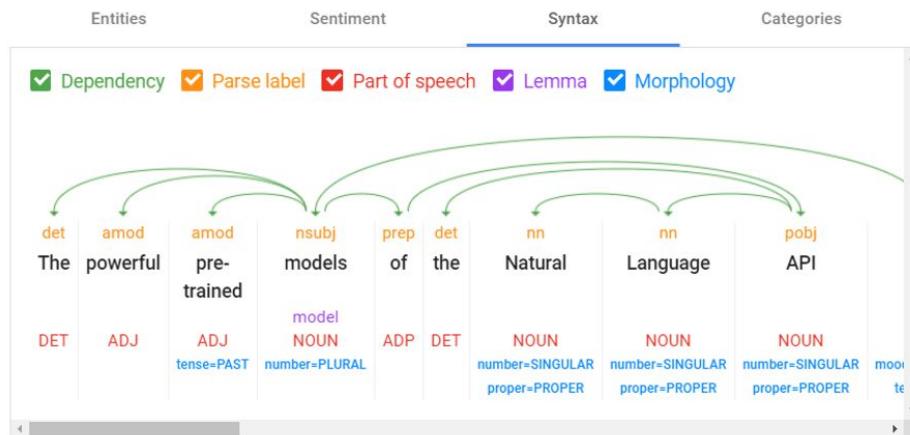
**Google Cloud + Bloomberg**

# Derive insights from unstructured text with the Cloud Natural Language API

The powerful pre-trained models of the Natural Language API let developers work with natural language understanding features including sentiment analysis, entity analysis, entity sentiment analysis, content classification, and syntax analysis.

RESET

[See supported languages](#)



# Make your media more discoverable with the Video Intelligence API

Labels      Shots      Explicit Content      API

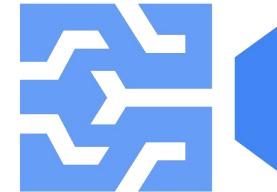
**Video Labels**  
Detect and label entities, such as dogs, flowers, and people, throughout the entire video.



▶ 00:00:30 / 00:00:42      🔍      🔊

dinosaur	80%
vehicle	79%
tree	40%
plant	
bicycle	36%
vehicle	
land vehicle	36%
vehicle	

**Note:** shot-level label annotations are present on the shots tab.



---

# Agenda

Lab: Cloud Natural Language API:  
Qwik Start

Lab: Google Cloud Speech API:  
Qwik Start

Lab: Video Intelligence: Qwik Start

Lab: Reinforcement Learning:  
Qwik Start

Quiz

Summary



---

# Lab Intro

## Cloud Natural Language API: Qwik Start

Apply the analyze-entities method to ask the Cloud Natural Language API to extract "entities" from a snippet of text.

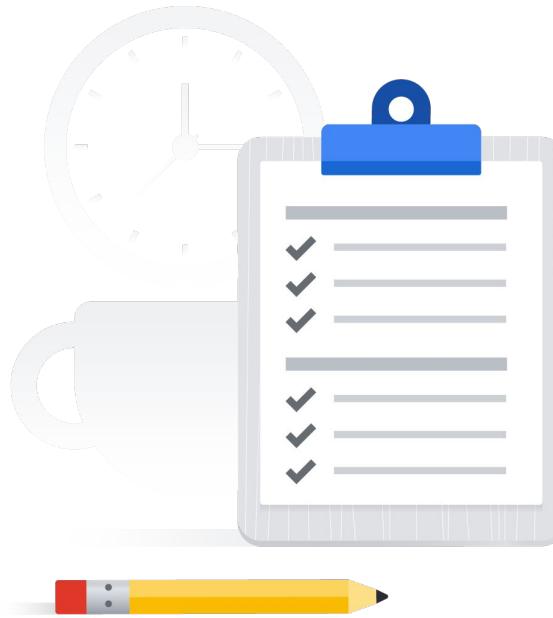
You can find the lab [here](#).



# Lab objectives

Create an API key.

Make an entity analysis request.



---

# Lab Intro

Entity and Sentiment Analysis with  
the Natural Language API  
(Alternative)

Use the Natural Language API to analyze  
entities, sentiment, and syntax.

You can find the lab [here](#).



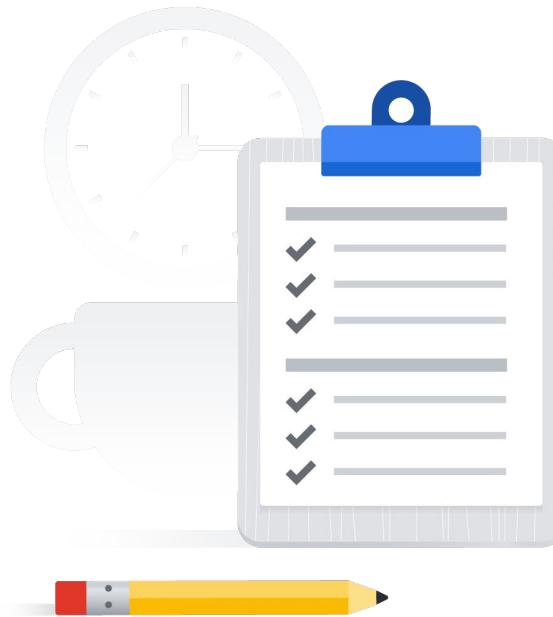
# Lab objectives

Create a Natural Language API request and call the API with curl.

Extract entities and run sentiment analysis on text with the Natural Language API.

Perform linguistic analysis on text with the Natural Language API.

Create a Natural Language API request in a different language.



---

# Agenda

Lab: Cloud Natural Language API:  
Qwik Start

[Lab: Google Cloud Speech API:  
Qwik Start](#)

Lab: Video Intelligence: Qwik Start

Lab: Reinforcement Learning:  
Qwik Start

Quiz

Summary



---

# Lab Intro

Google Cloud Speech API:  
Qwik Start

Create and call a Cloud Speech API request.

The lab can be found [here](#).

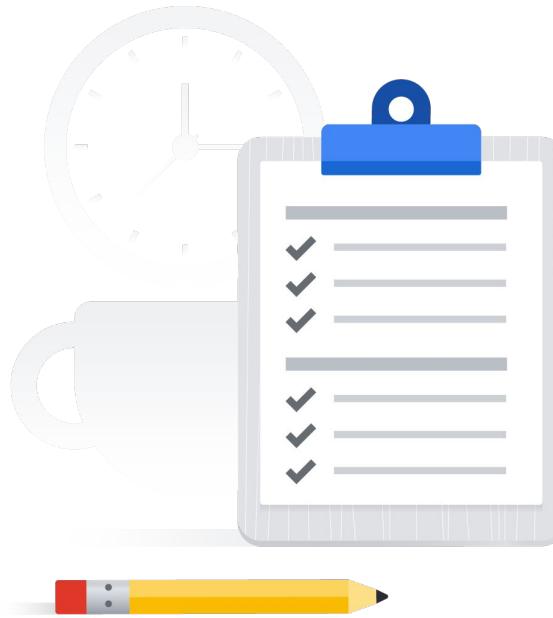


# Lab objectives

Create an API key.

Create a Speech API request.

Call the Speech API request.



---

## Lab Intro

### Speech to Text Transcription with the Cloud Speech API (Alternative)

Send an audio file to the Cloud Speech API for transcription.

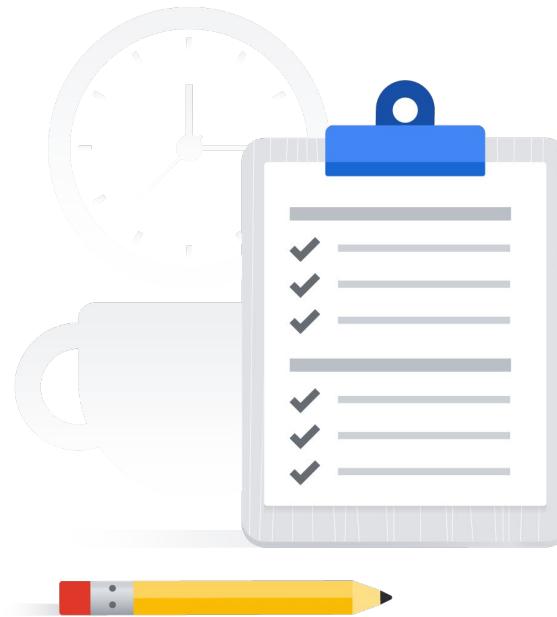
The lab can be found [here](#).



# Lab objectives

Create a Speech API request and calling the API with curl.

Call the Speech API with audio files in different languages.



---

# Agenda

Lab: Cloud Natural Language API:  
Qwik Start

Lab: Google Cloud Speech API:  
Qwik Start

[Lab: Video Intelligence: Qwik Start](#)

Lab: Reinforcement Learning:  
Qwik Start

Quiz

Summary



---

# Lab Intro

## Video Intelligence: Qwik Start

Make an annotate video request using the Video Intelligence API.

The lab can be found [here](#).

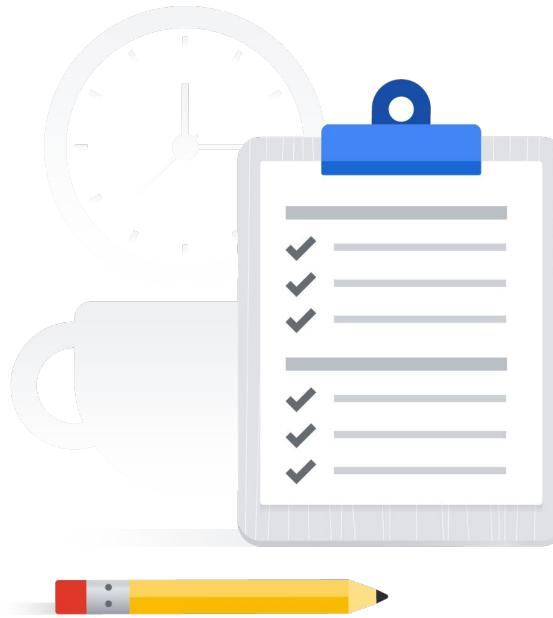


# Lab objectives

Enable the Video Intelligence API.

Set up authorization.

Make an annotate video request.



---

## Lab Intro

Scanning User-generated Content  
Using the Cloud Video Intelligence  
and Cloud Vision APIs (Alternative)

Deploy a set of Cloud Functions in order to  
process images and videos with the Vision  
API and Video Intelligence API.

The lab can be found [here](#).

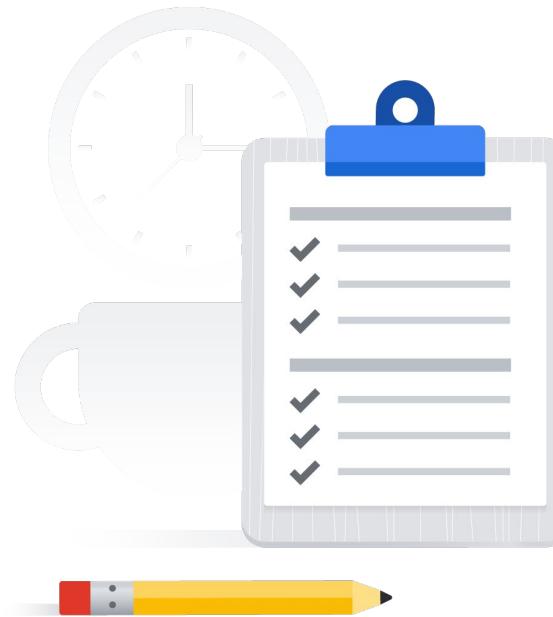


# Lab objectives

Deploy four Cloud Functions.

Create the supporting Cloud Storage buckets, Pub/Sub topics, and Cloud Storage Pub/Sub Notifications.

Create the supporting BigQuery dataset and table.



---

# Agenda

Lab: Cloud Natural Language API:  
Qwik Start

Lab: Google Cloud Speech API:  
Qwik Start

Lab: Video Intelligence: Qwik Start

Lab: Reinforcement Learning:  
Qwik Start

Quiz

Summary



---

# Lab Intro

## Reinforcement Learning: Qwik Start

Learn the basics of reinforcement learning by building a simple game.

The lab can be found [here](#).



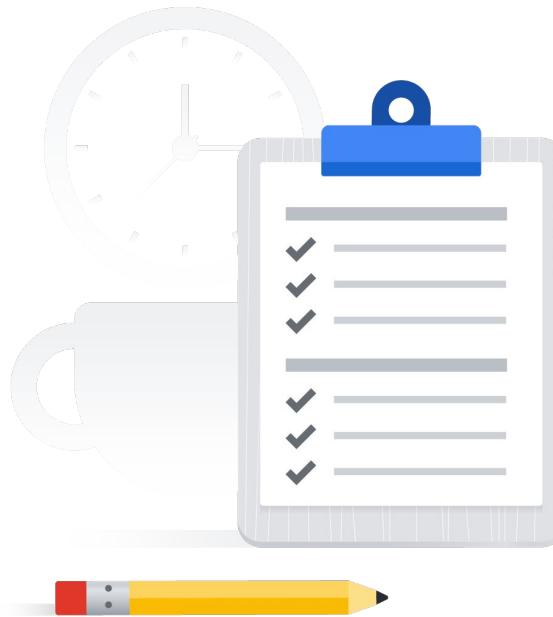
# Lab objectives

Understand the fundamental concepts of reinforcement learning.

Create an AI Platform Tensorflow 2.1 Notebook.

Clone the sample repository from the training data analyst repo found on Github.

Read, understand, and run the steps found in the notebook.



---

# Agenda

Lab: Cloud Natural Language API:  
Qwik Start

Lab: Google Cloud Speech API:  
Qwik Start

Lab: Video Intelligence: Qwik Start

Lab: Reinforcement Learning:  
Qwik Start

[Quiz](#)

Summary



---

# Scenario #1

## Question

Which machine learning tool would be the best option for someone that wants a custom model but has limited application development or data science skills?

- A. Vertex AI
- B. AutoML
- C. Tensorflow
- D. Speech API

---

# Scenario #1

## Answer

Which machine learning tool would be the best option for someone that wants a custom model but has limited application development or data science skills?

- A. Vertex AI
- B. AutoML**
- C. Tensorflow
- D. Speech API

---

## Scenario #2

### Question

What Google machine learning API can be used to gain meaning and sentiment from text?

- A. Natural Language API
- B. Speech-to-Text API
- C. Vision API
- D. Video Intelligence API

---

## Scenario #2

### Answer

What Google machine learning API can be used to gain meaning and sentiment from text?

- A. Natural Language API
- B. Speech-to-Text API
- C. Vision API
- D. Video Intelligence API

---

# Scenario #3

## Question

Which machine learning service can run Tensorflow at scale?

- A. Pre-trained machine learning APIs
- B. AutoML
- C. Vertex AI
- D. Tensorflow

---

## Scenario #3

### Answer

Which machine learning service can run Tensorflow at scale?

- A. Pre-trained machine learning APIs
- B. AutoML
- C. Vertex AI
- D. Tensorflow

---

# Agenda

Lab: Cloud Natural Language API:  
Qwik Start

Lab: Google Cloud Speech API:  
Qwik Start

Lab: Video Intelligence: Qwik Start

Lab: Reinforcement Learning:  
Qwik Start

Quiz

Summary



---

## Summary (1/2)

- ML is a way to get predictive insights from data to make repeated decisions.
- ML uses standard algorithms.
- Model training requires examples.
- Trained ML models can be used to make predictions.
- An algorithm can be applied to other data to yield a different model.
- ML models need good data.

---

## Summary (2/2)

- TensorFlow was developed by Google and has become the leading open-source tool for building ML models.
- Vertex AI allows you to train, evaluate, and tune your model, deploy your trained model and get predictions, and monitor predictions on an ongoing basis.
- AutoML allows you to train high-quality custom machine learning models with minimal effort or machine learning expertise.
- Google has a range of pre-trained ML APIs that can be used for common applications.

# Another use case for ML

Extract, Analyze, and Translate Text from Images with the Cloud ML APIs

[https://www.cloudskillsboost.google/course\\_templates/739/labs/491251](https://www.cloudskillsboost.google/course_templates/739/labs/491251)

ECU.UM OPENAI

TAU

# What is Generative AI but a wild application of machine learning

<https://cloud.google.com/ai/generative-ai>

and

<https://cloud.google.com/vertex-ai/generative-ai/docs/learn/overview>

# Seems like fun

Build an AI Image Generator app using Imagen on Vertex AI

[https://www.cloudskillsboost.google/focuses/96792?catalog\\_rank=%7B%22rank%22%3A4%2C%22num\\_filters%22%3A0%2C%22has\\_search%22%3Atrue%7D&parent=catalog&search\\_id=39778861](https://www.cloudskillsboost.google/focuses/96792?catalog_rank=%7B%22rank%22%3A4%2C%22num_filters%22%3A0%2C%22has_search%22%3Atrue%7D&parent=catalog&search_id=39778861)

# If you want to play with a RAG

Multimodal Retrieval Augmented Generation (RAG) using the Vertex AI Gemini API

[https://www.cloudskillsboost.google/focuses/85643?catalog\\_rank=%7B%22rank%22%3A2%2C%22num\\_filters%22%3A0%2C%22has\\_search%22%3Atrue%7D&parent=catalog&search\\_id=39817455](https://www.cloudskillsboost.google/focuses/85643?catalog_rank=%7B%22rank%22%3A2%2C%22num_filters%22%3A0%2C%22has_search%22%3Atrue%7D&parent=catalog&search_id=39817455)

# Or the Gemini CLI

<https://codelabs.developers.google.com/gemini-cli-hands-on#1>

Or ollama to GKE (use local terminal, not cloud shell)

<https://codelabs.developers.google.com/codelabs/production-ready-ai-with-qc/5-deploying-agents/deploying-open-models-gke>