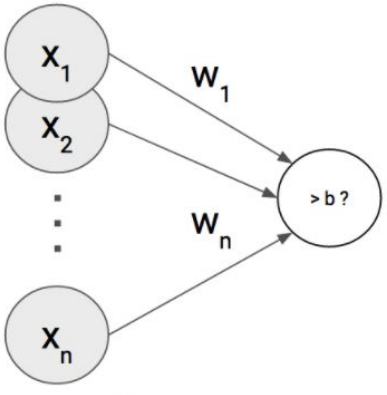
Google Cloud
Summit

Real-world Machine Learning with TensorFlow and Cloud ML

Cloud-based ML solutions in Action

The Buzz Words

AI, ML and Neural Network



Artificial Intelligence

The science of making things smart

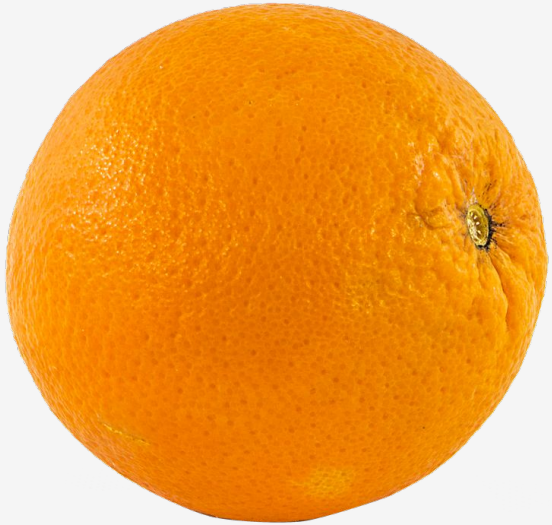
Machine Learning

Building machines that can learn

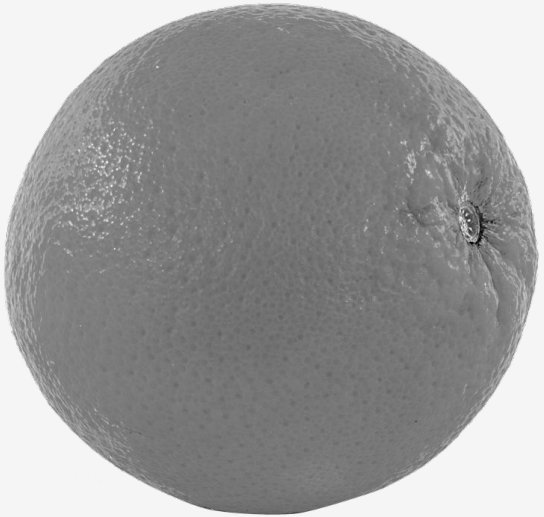
Neural Network

A type of algorithm in machine learning

How do you classify this with your code?



How do you classify this with your code?



Activity Detection

```
if(speed<4){
    status=WALKING;
}
```

```
if(speed<4){
    status=WALKING;
} else {
    status=RUNNING;
}
```

```
if(speed<4){
    status=WALKING;
} else if(speed<12){
    status=RUNNING;
} else {
    status=BIKING;
}
```

```
// Oh crap
```

Activity Detection

```
0101001010100101010
1001010101001011101
0100101010010101001
0101001010100101010
```

Label = WALKING

```
1010100101001010101
0101010010010010001
0010011111010101111
1010100100111101011
```

Label = RUNNING

```
1001010011111010101
1101010111010101110
1010101111010101011
1111110001111010101
```

Label = BIKING

```
1111111111010011101
0011111010111110101
0101110101010101110
1010101010100111110
```

Label = GOLFING

How about these?



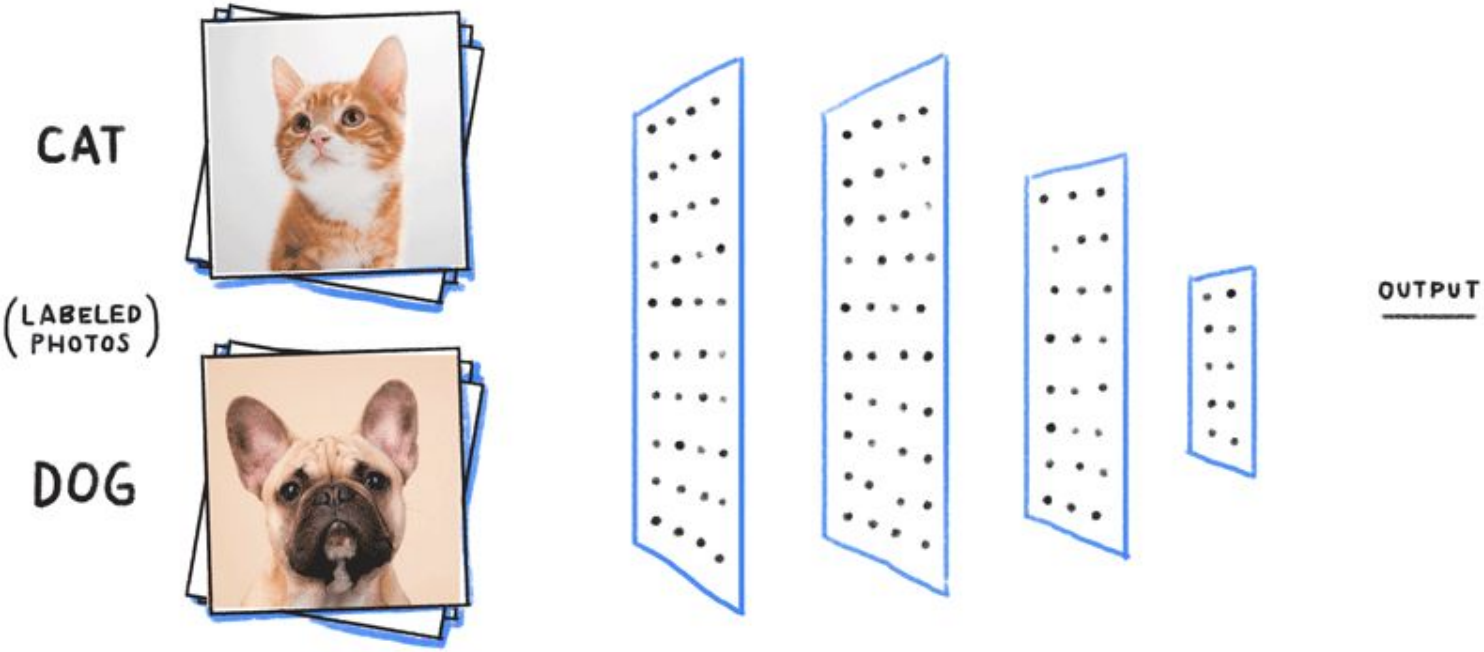




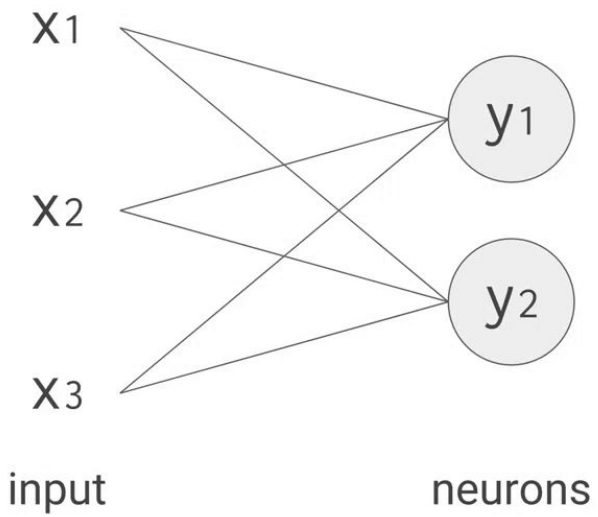

With Cloud Vision API...

The mechanics of Neural Network

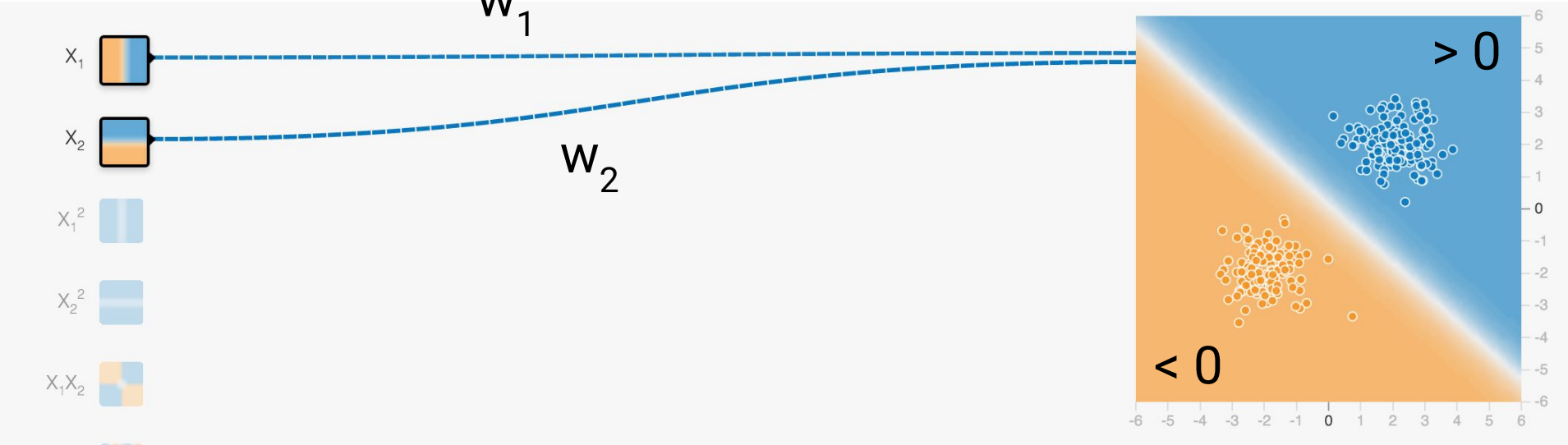
Neural Network is a **function** that can **learn**



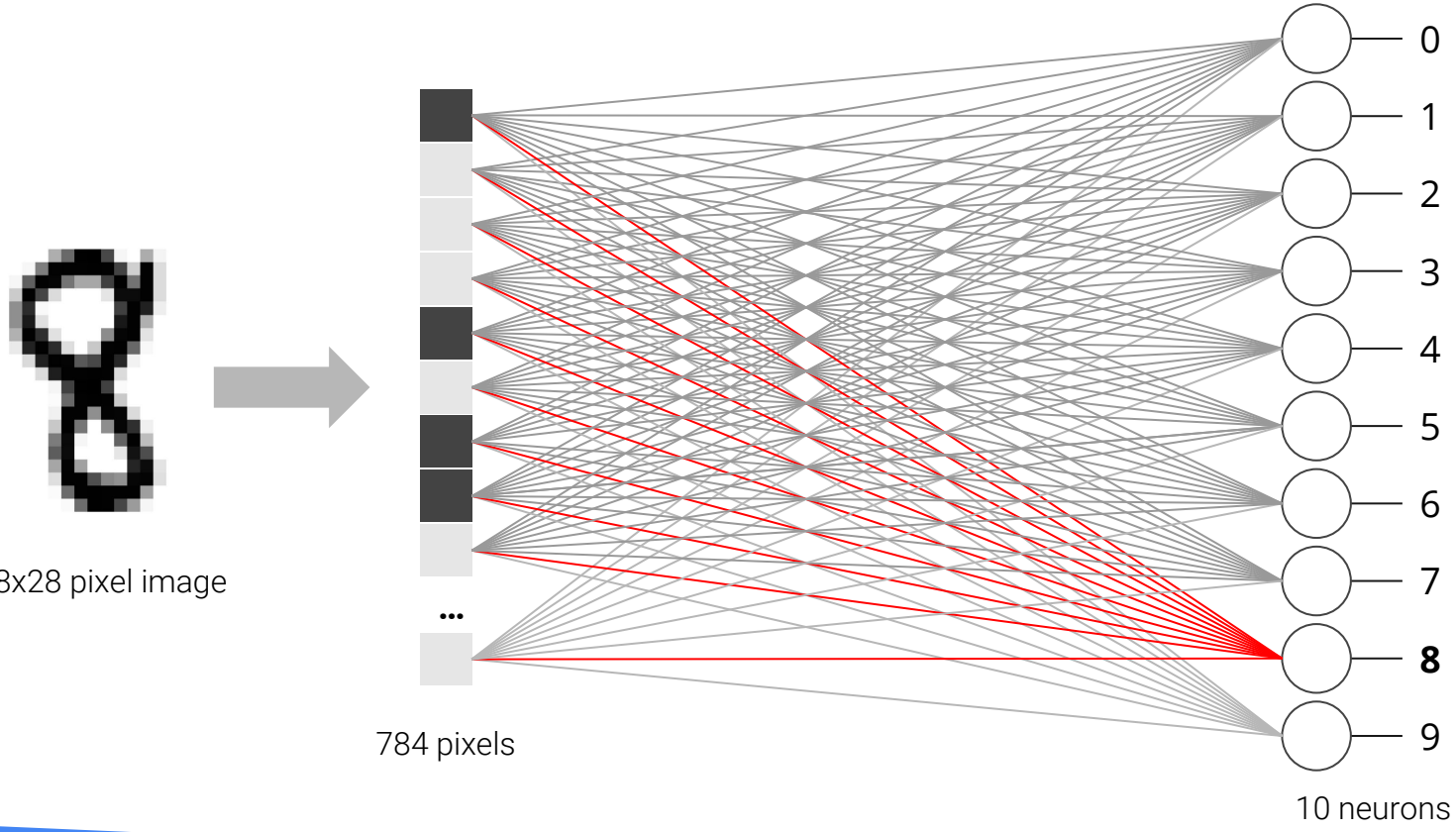
Mimic the behavior of human neurons with simple math



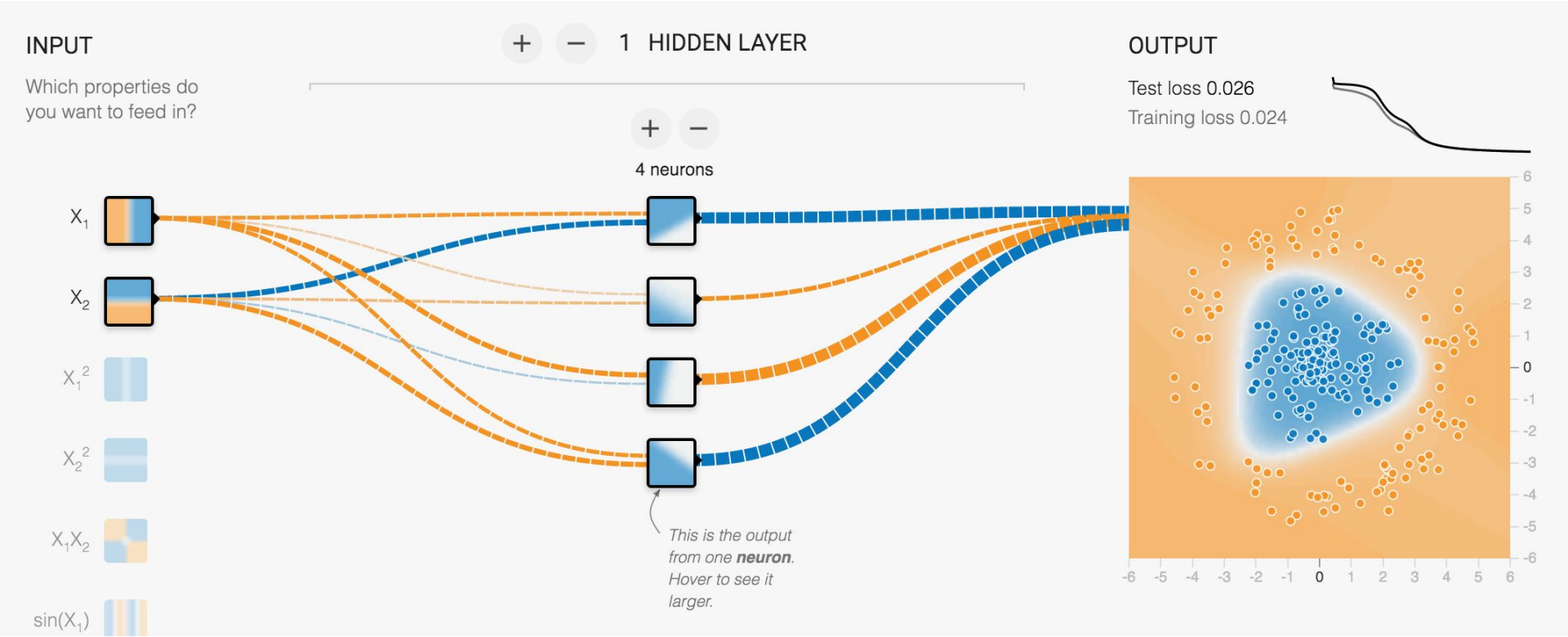
Let’s see how the neuron solves the problem



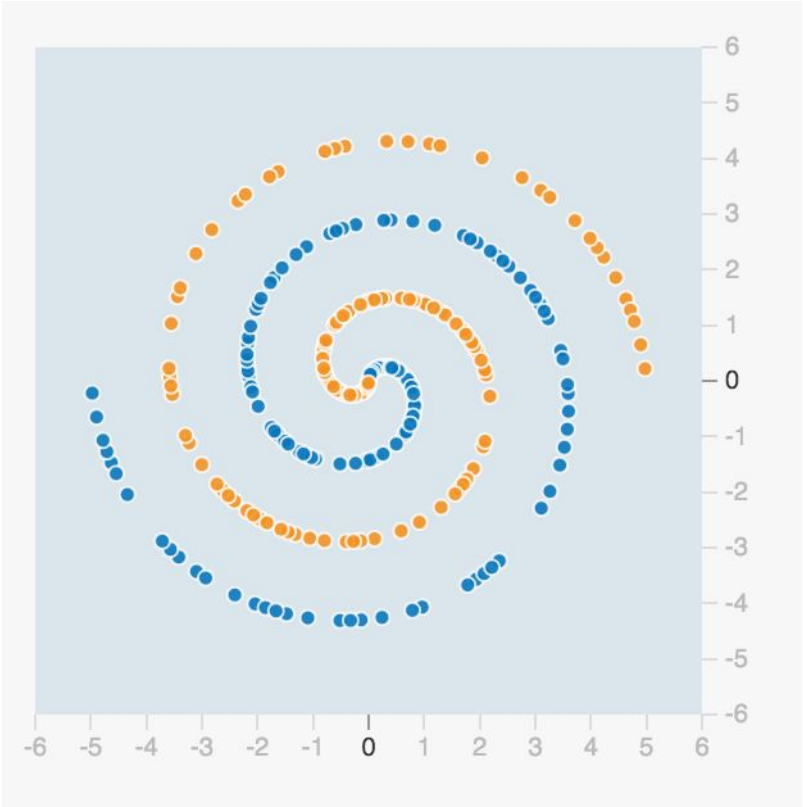
Even from hundreds or thousands of features in big data



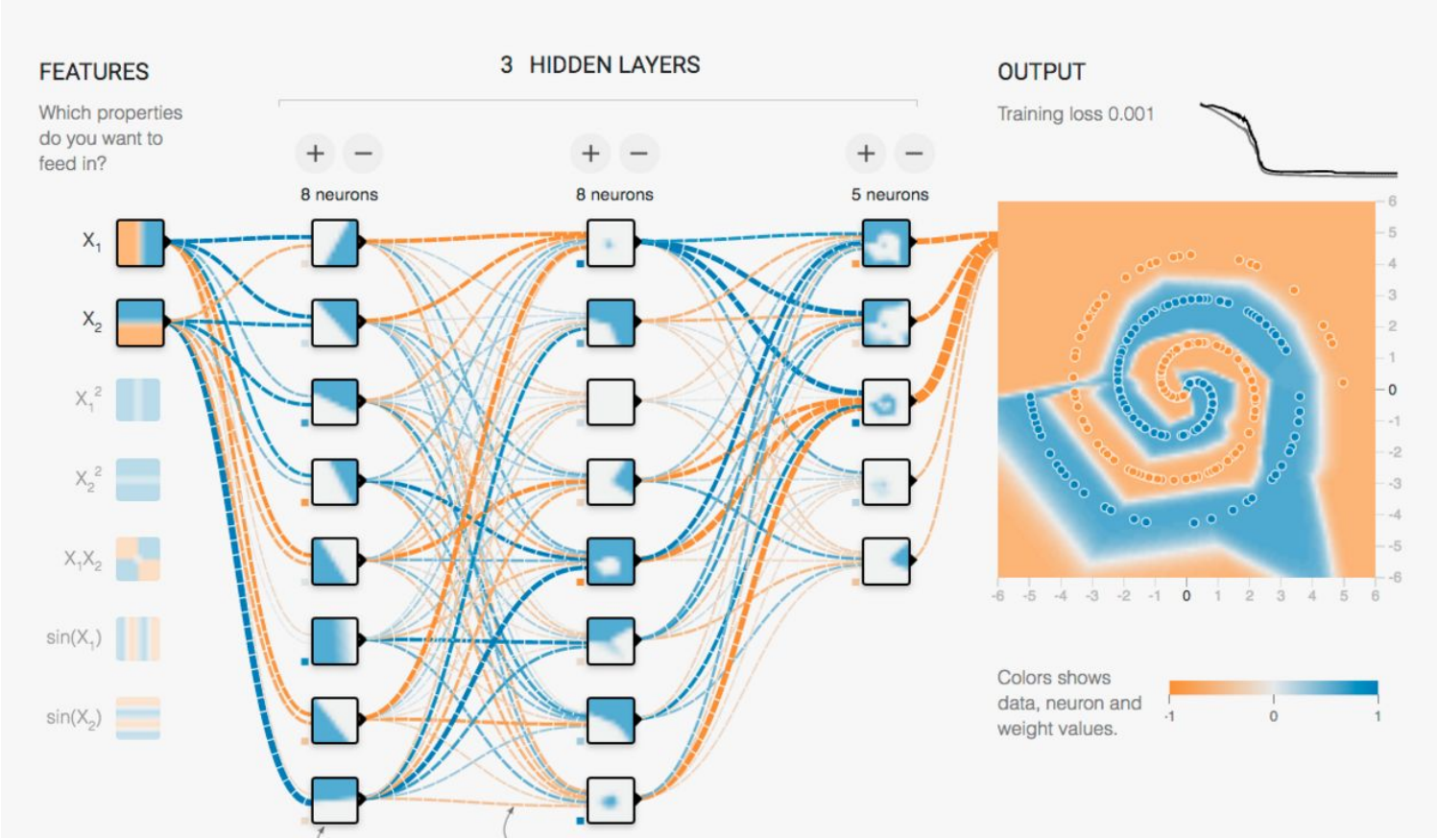
What is Deep Learning



More **neurons** = More **features** to extract

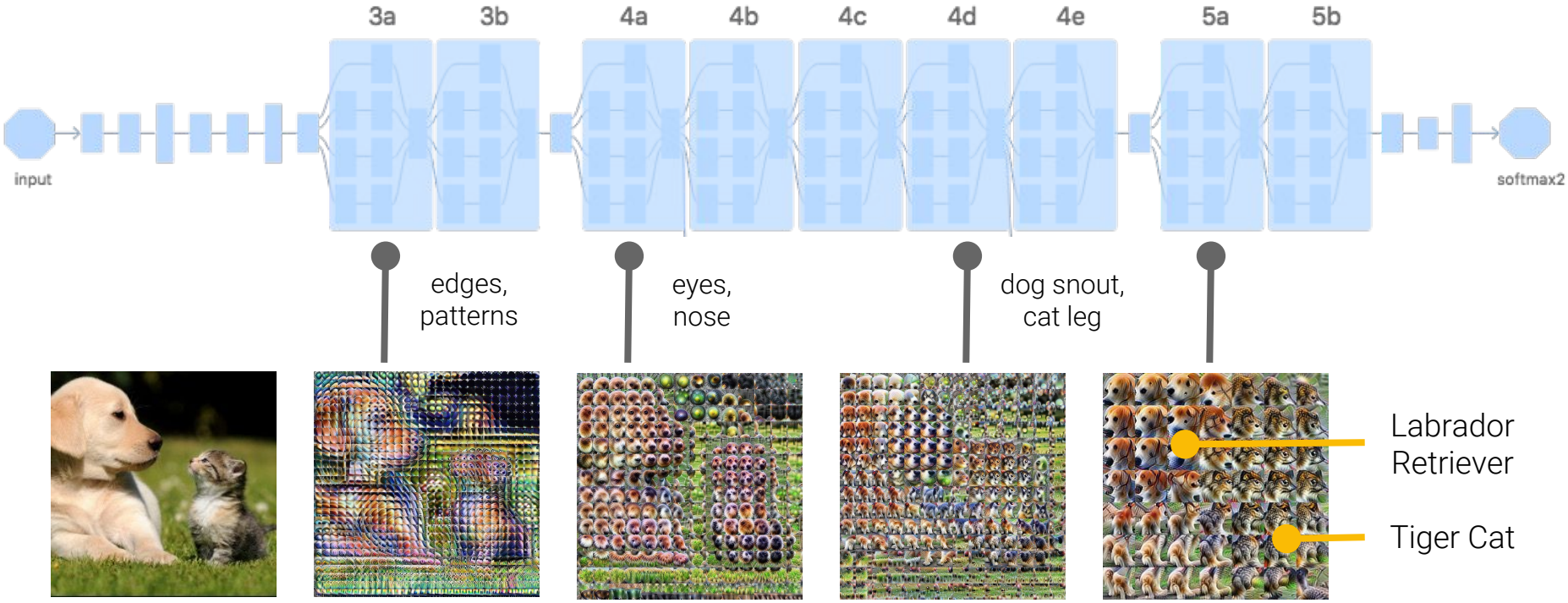


How about this?



More **hidden layers** = More **hierarchies** of features

More layers for deeper insights = **Deep Learning**

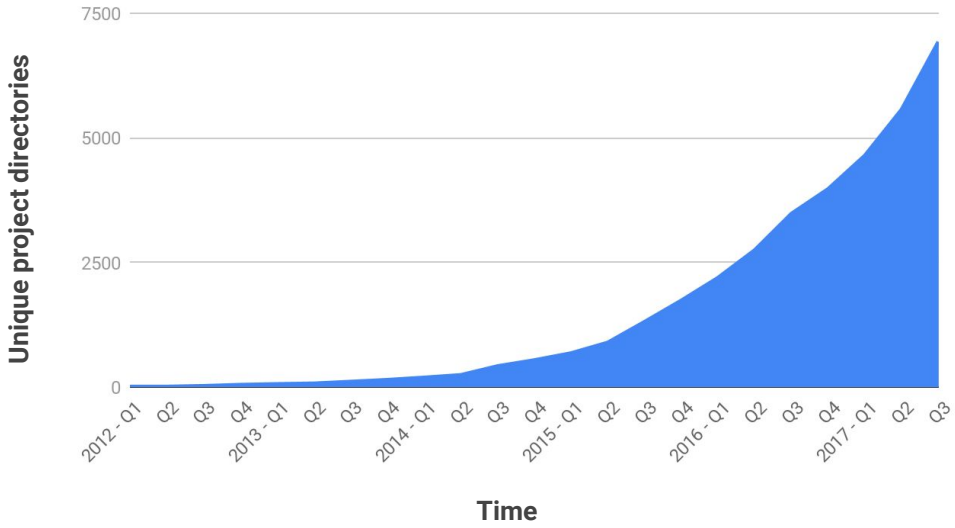


Source: The Building Blocks of Interpretability, Chris Olah et. al.

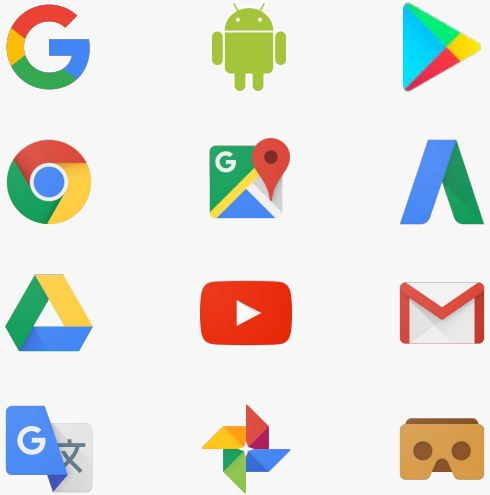
ML use cases in Google services

Google is an AI company

Google 3 directories containing Brain Model

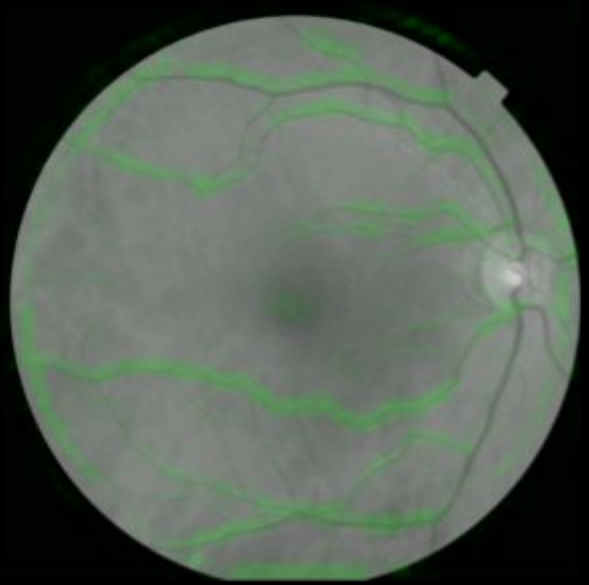


Used across products:

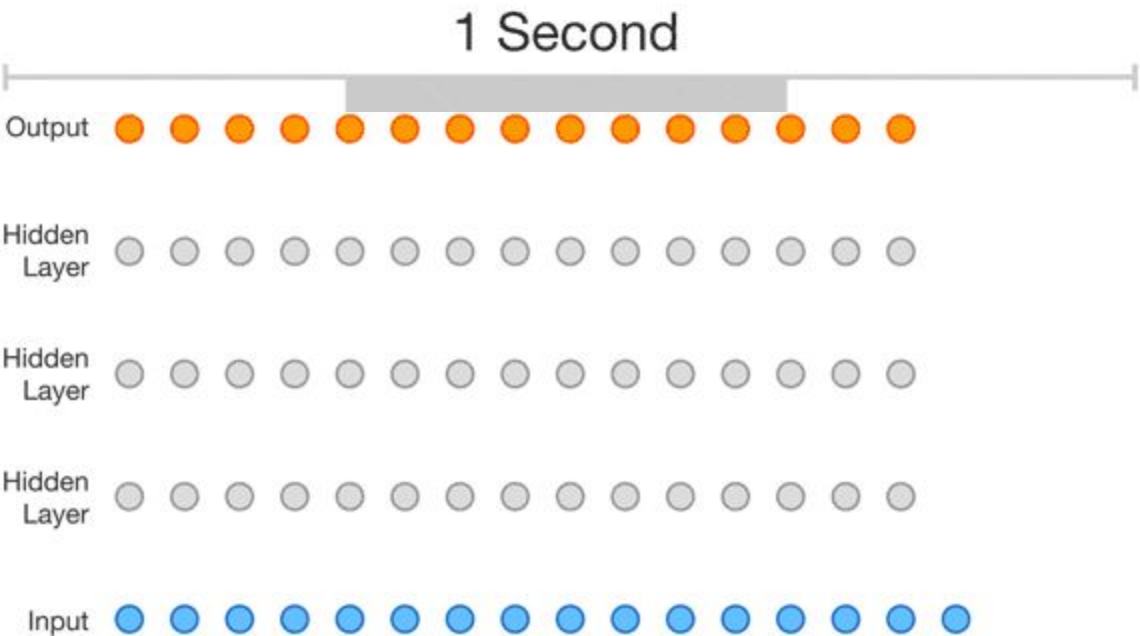



Google Photos
“boats”
Google

Image of retina

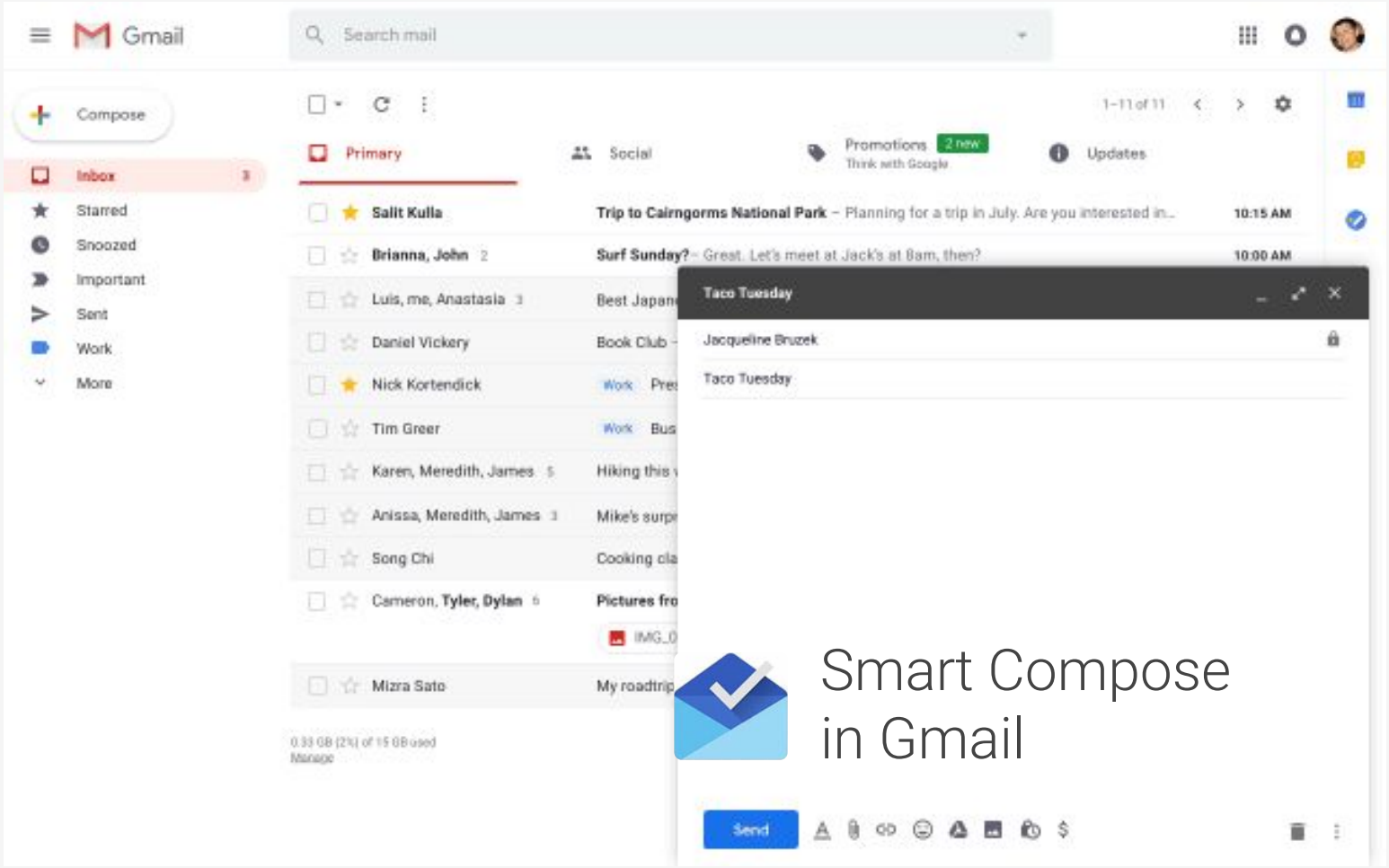


Blood pressure predictions focus on blood vessels



WaveNet by DeepMind


2018
Google I/O
Keynote


Gmail
Search mail
Compose
Inbox 3
Starred
Snoozed
Important
Sent
Work
More
Primary
Social
Promotions 2 new
Think with Google
Updates
1–11 of 11
Salit Kulla
Trip to Cairngorms National Park – Planning for a trip in July. Are you interested in...
10:15 AM
Brianna, John 2
Surf Sunday? – Great. Let's meet at Jack's at 8am, then?
10:00 AM
Luis, me, Anastasia 3
Daniel Vickery
Nick Kortendick
Work
Tim Greer
Work
Karen, Meredith, James 5
Anissa, Meredith, James 3
Song Chi
Cameron, Tyler, Dylan 6
Mizra Sato
0.33 GB (2%) of 15 GB used
Manage
Taco Tuesday
Jacqueline Bruzek
Taco Tuesday
Send
Smart Compose
in Gmail

Google Translate

ML services from Google Cloud

Sharing the power of ML

ML services from Google Cloud

AI Building blocks

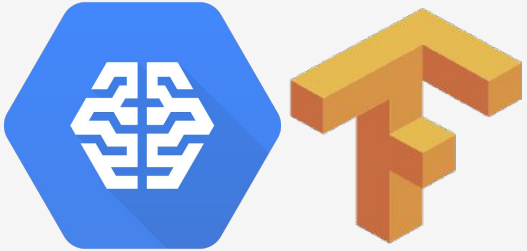
Pre-trained,
easy to use

Cloud AutoML

Customizable,
easy to use

Cloud ML Engine /TensorFlow

Scalable and flexible ML
for data scientists



Cloud AI building blocks

SIGHT

Cloud Vision API
Image recognition and classification.

Cloud Video Intelligence API
Scene-level video annotation.

AutoML Vision BETA
Custom image classification models.

LANGUAGE

Cloud Translation API
Language detection and translation.

Cloud Natural Language API
Text parsing and analysis.

AutoML Translation BETA
Custom domain-specific translation.

AutoML Natural Language BETA
Custom text classification models.

CONVERSATION

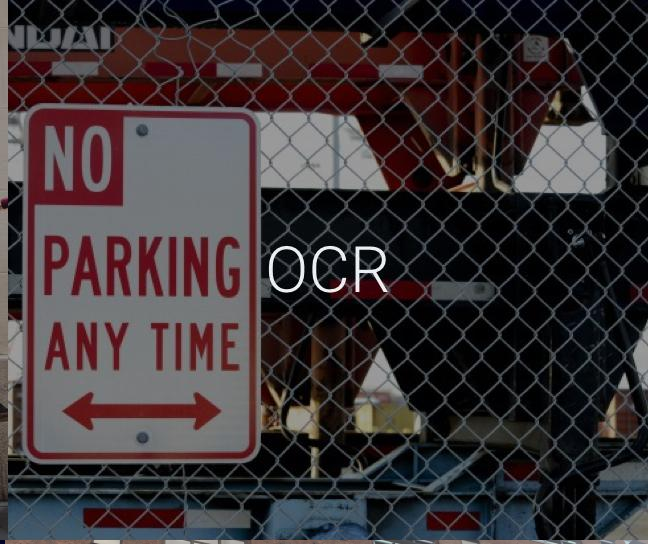
Dialogflow Enterprise Edition
Build conversational interfaces.

Cloud Text-to-Speech API
Convert text to speech.

Cloud Speech-to-Text API
Convert speech to text.


Label
Detection


Face
Detection


NO
PARKING
ANY TIME
OCR


Explicit Content
Detection
Next


Landmark
Detection


Logo Detection
Google
Google Cloud

Vision API: Web annotations


```
{
  "entityId": "/m/0gff2yr",
  "score": 5.92256,
  "description": "ArtScience Museum"
}
```

```
{
  "entityId": "/m/016ms7",
  "score": 1.44038,
  "description": "Ford Anglia"
}
```

```
{
  "entityId": "/m/0h898pd",
  "score": 7.4162,
  "description": "Harry Potter (Literary Series)"
}
```


Google Cloud Platform

Search

Why Google | Products | Solutions | Launches | Pricing | Customers | Documentation | Su

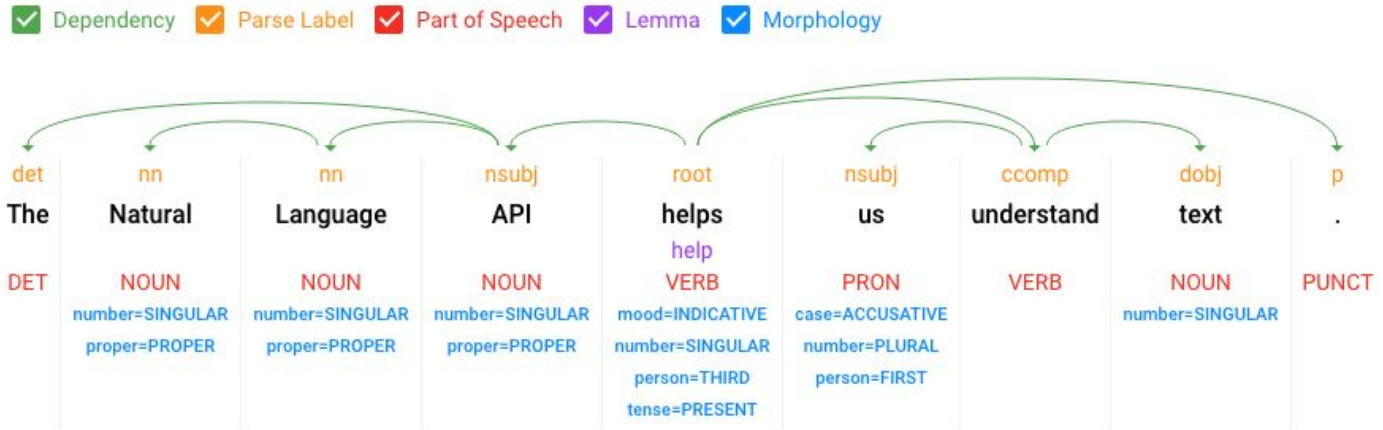
CONTACT SALES

Labels | Web | Properties | Safe Search | JSON

Dog	96%
Dog Like Mammal	96%
Mammal	92%
Dog Breed	78%
Carnivoran	74%
Komondor	74%
Fur	65%
Goldendoodle	60%

Natural Language API: Entity and Syntactic analysis

***Joanne "Jo" Rowling**, pen names **J. K. Rowling** and **Robert Galbraith**, is a **British** novelist, screenwriter and film producer best known as the author of the **Harry Potter** fantasy series*



Google Cloud Platform

Search

CONSOLE

Why Google | Products | Solutions | Launcher | Pricing | Customers | Documentation | Su

CONTACT SALES

CLOUD NATURAL LANGUAGE API

Derive insights from unstructured text using Google machine learning

VIEW DOCUMENTATION

VIEW CONSOLE

Powerful Text Analysis

Google Cloud Natural Language API reveals the structure and meaning of text by offering powerful machine learning models in an easy to use REST API. You can use it to **extract information** about people, places, events and much more, mentioned in text documents, news articles or blog posts. You can use it to **understand sentiment** about your product on social media or **parse intent** from customer conversations

Cloud AutoML

Automating the craft of data science

Google Cloud

State of the Industry: Complex & Time Intensive

Large computational resource . Machine learning expertise . Manual data labeling

Google Cloud

Introducing Cloud AutoML

A technology that can automatically create a Machine Learning Model

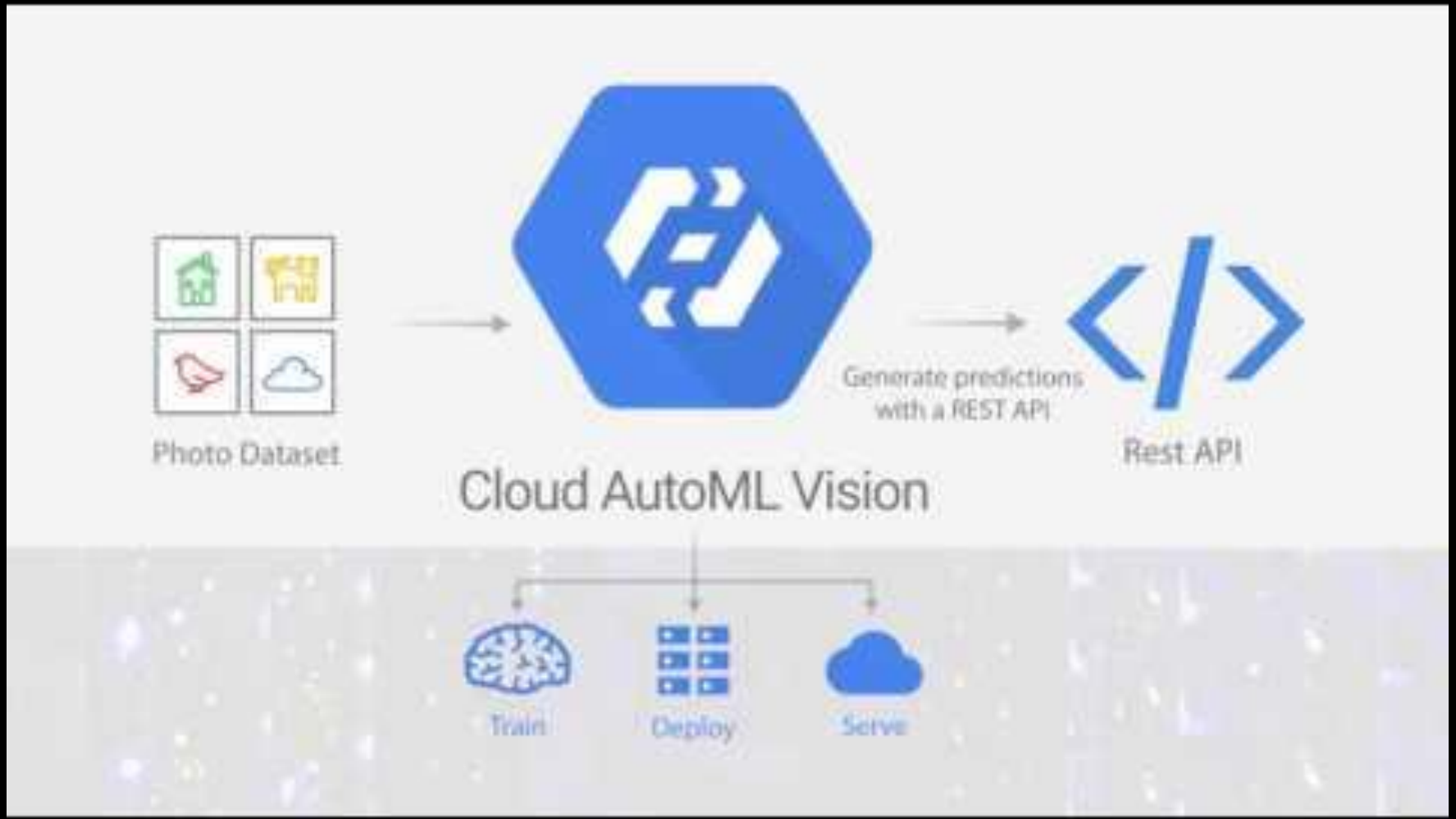
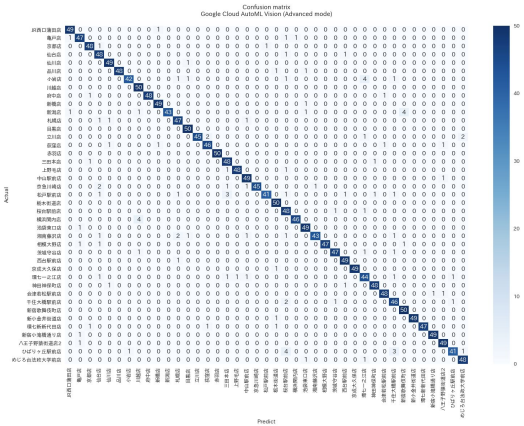


Photo Dataset
Cloud AutoML Vision
Generate predictions with a REST API
Rest API
Train
Deploy
Serve

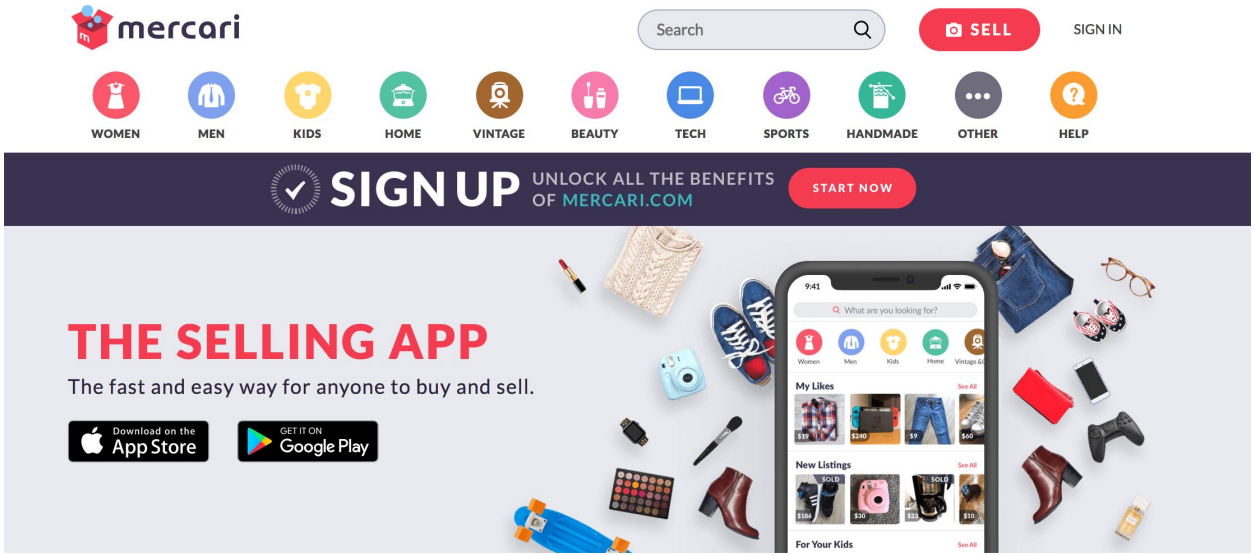
Predicting Ramen shop at 95% accuracy

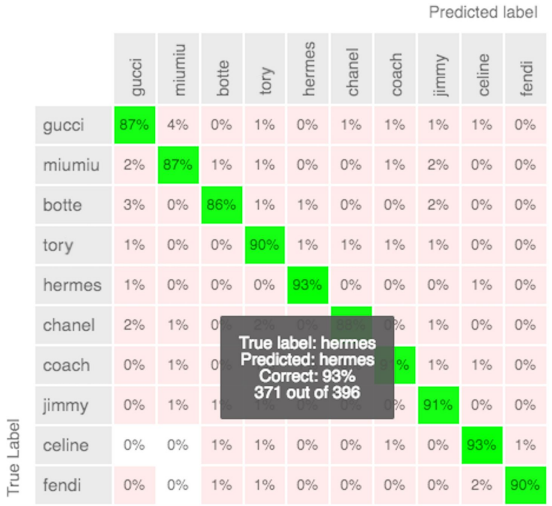
from 41 Ramen Jiro branches



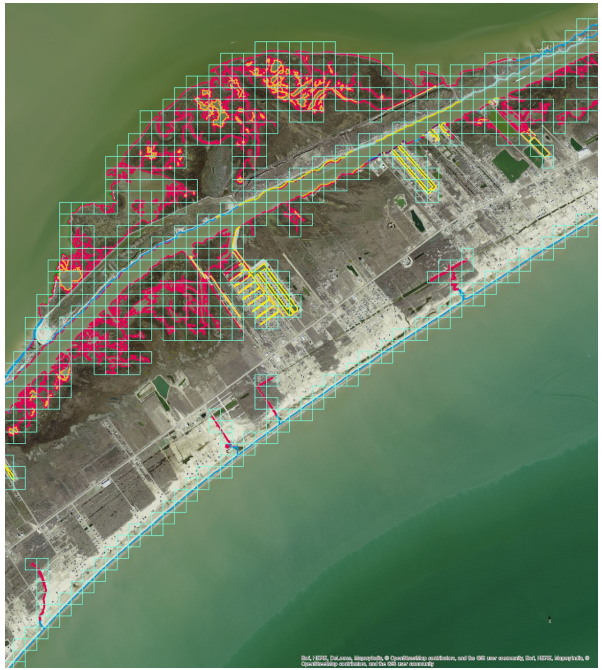
Predicting Brand names at 90% accuracy

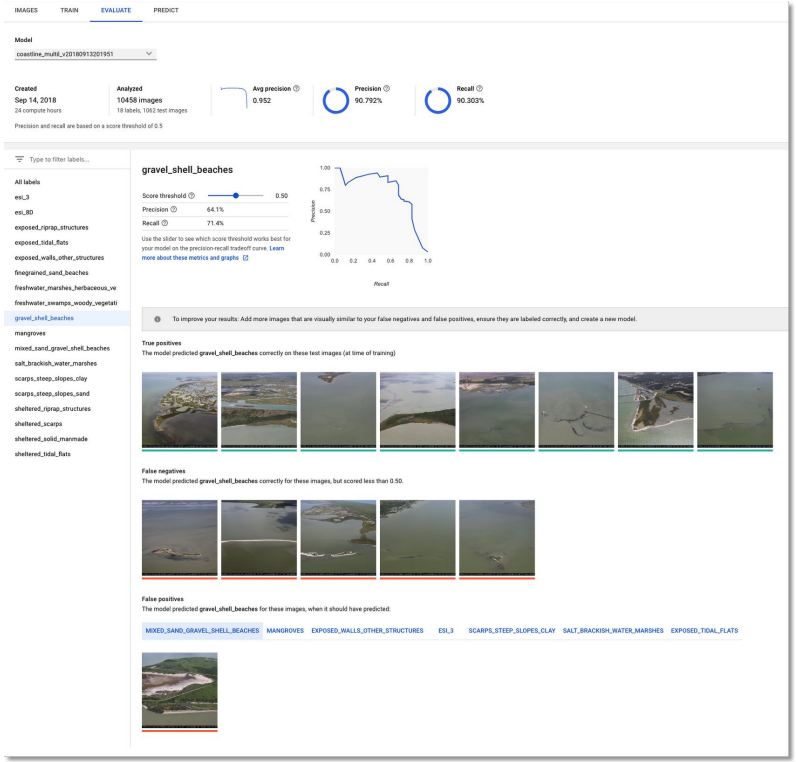
from 12 popular brands at Mercari





Tracking environmental changes at the Gulf of Mexico





TensorFlow

The open source software
for ML development

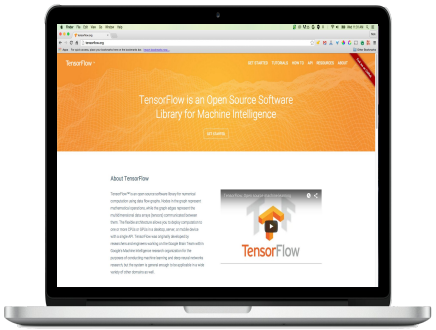
What is TensorFlow?

Open source software for ML development

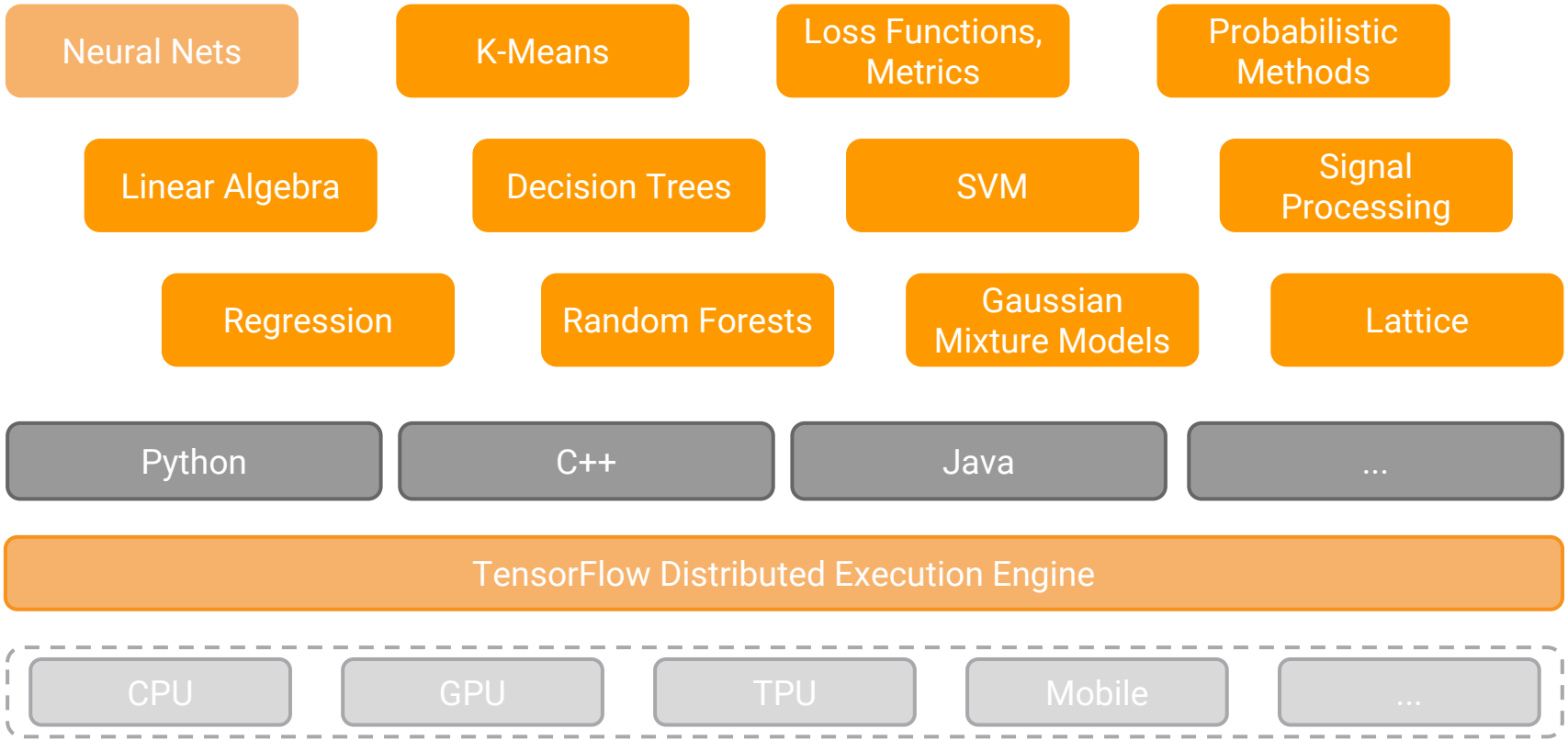
tensorflow.org launched in Nov 2015

Created by Google Brain team

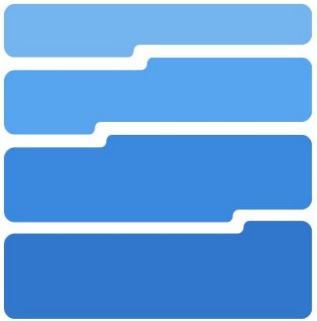
The standard tool for ML and AI development at Google



A comprehensive ML toolkit

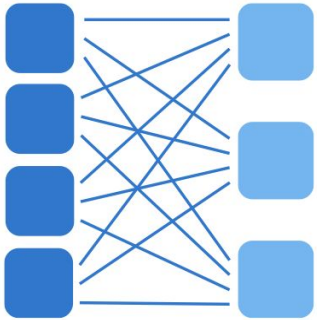


Bucketing



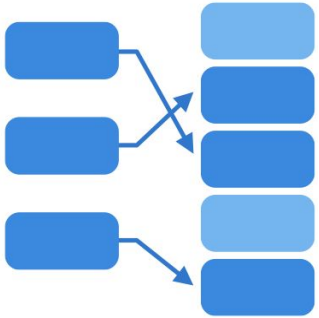
Partition by range

Crossing



Create new combinations

Hashing



Limit size

Embedding

Learn a new representation

https://www.youtube.com/watch?v=d12ra3b_M-0 — https://www.tensorflow.org/get_started/feature_columns

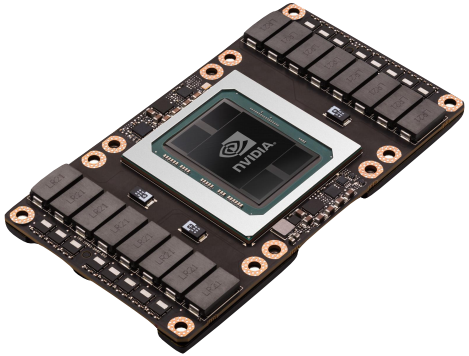
Scalable and portable training and prediction

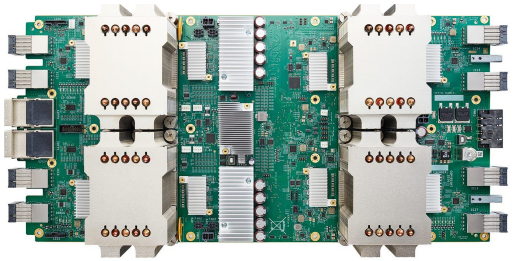
Training model on:

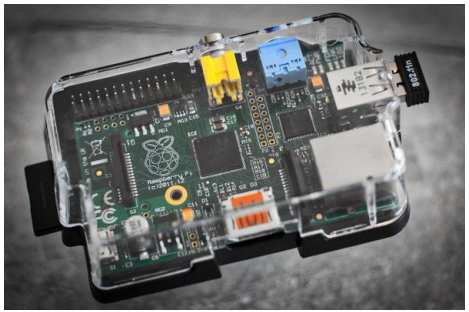
- Mac/Windows
- Single GPU
- Multiple GPUs and TPUs

Use (prediction) model on:

- Android and iOS
- RasPi and TPU

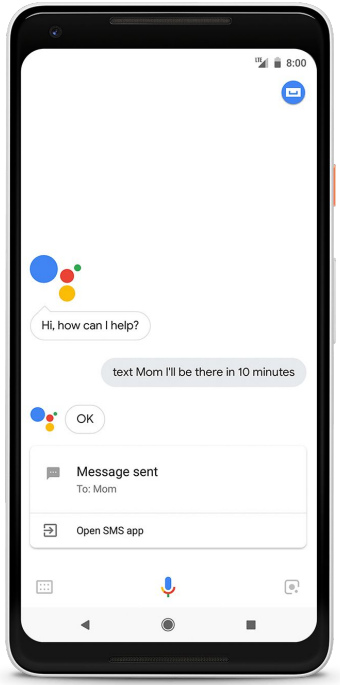






Google apps using TensorFlow Lite

Portrait Mode on Android Camera



‘Hey Google’ in Google Assistant

SmartReply on Google WearOS

Using TensorFlow Lite

Convert to TensorFlow Lite Format

```
$ toco --savedmodel_directory=awesome_model/
--output_file=awesome_model.tflite
```

```
import org.tensorflow.lite.Interpreter;
try (Interpreter tflite = new Interpreter("awesome_model.tflite")) {
  // Create inputs and outputs.
  ...
  // Invoke the interpreter.
  tflite.run(inputs, outputs);
}
```

Invoke TensorFlow Lite Interpreter

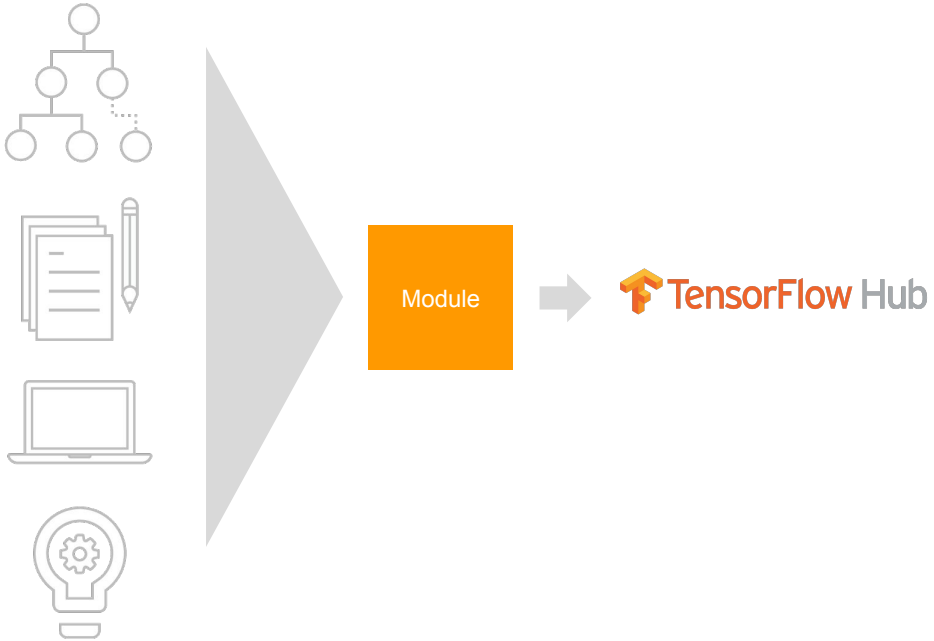

TensorFlowLite
82ms
scissors: 0.88
tensorflowlogo: 0.03
pen: 0.02

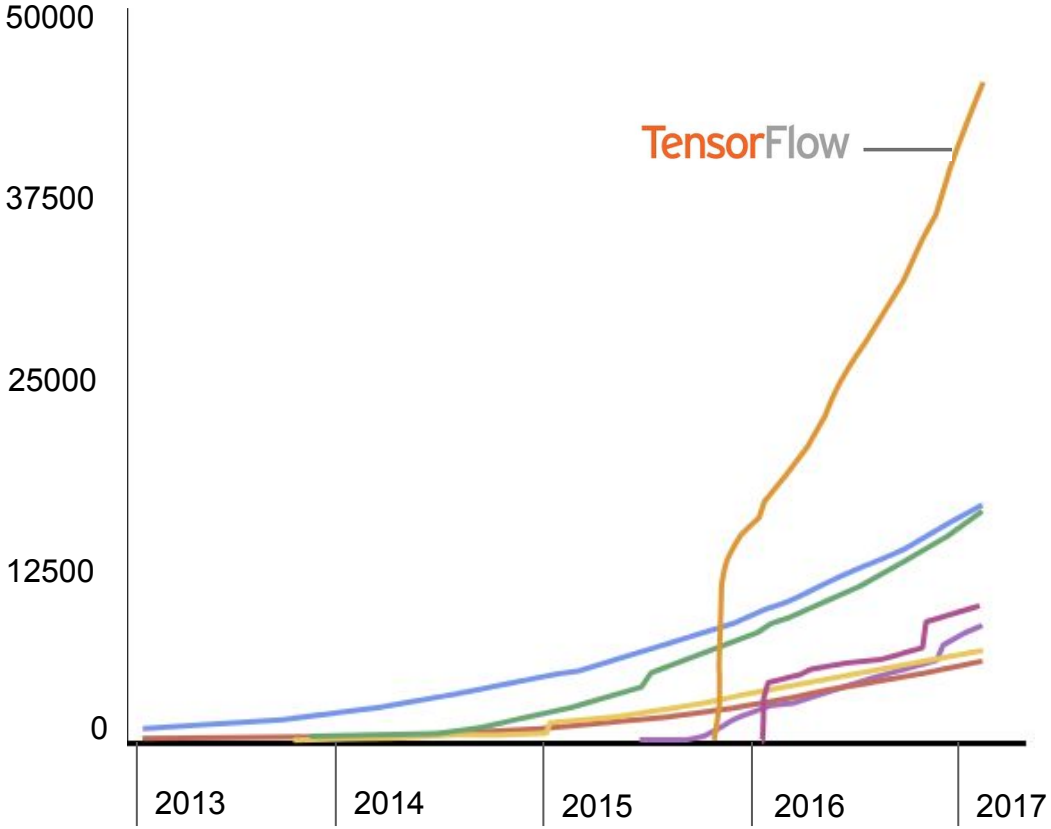
TensorFlow Hub

A repository of pre-trained model components, packaged for one-line reuse

Easiest way to share and use best approaches for your task

tensorflow.org/hub





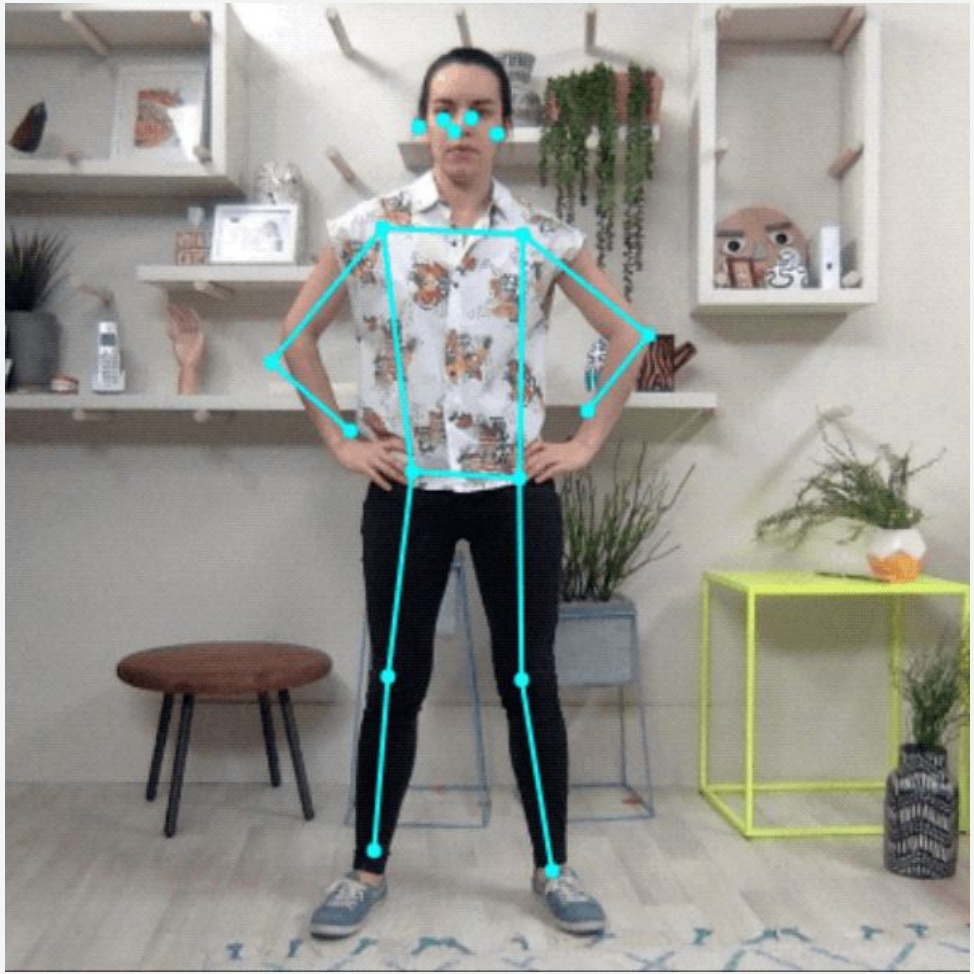
Framework	GitHub Star Count
TensorFlow	44508
scikit-learn	16191
Caffe	15690
CNTK	9383
MXNet	7896
Torch	6285
Theano	5568

Companies using TensorFlow

Use Cases:
Image Recognition

Real Time Pose Estimation

In-the-Browser with JavaScript




きゅうり農家 + AI = ?
FPS:3.35

AUCNET: Real-time car auction

5M cars, 30K car dealers

20 images per each car

Labeling takes **15 mins**

- Identifying parts and car make

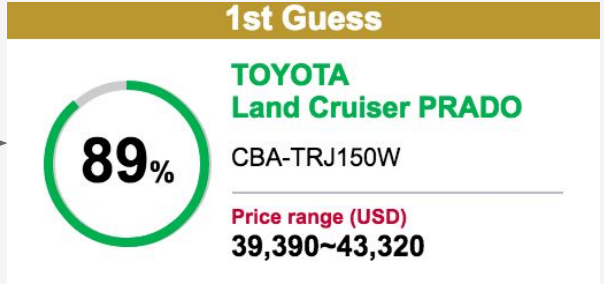
Car Image Classifier with TensorFlow/ML Engine

Inception v3 + Transfer Learning

200 images per car model, 150K images total

ML Engine: increased training performance for **86X** faster

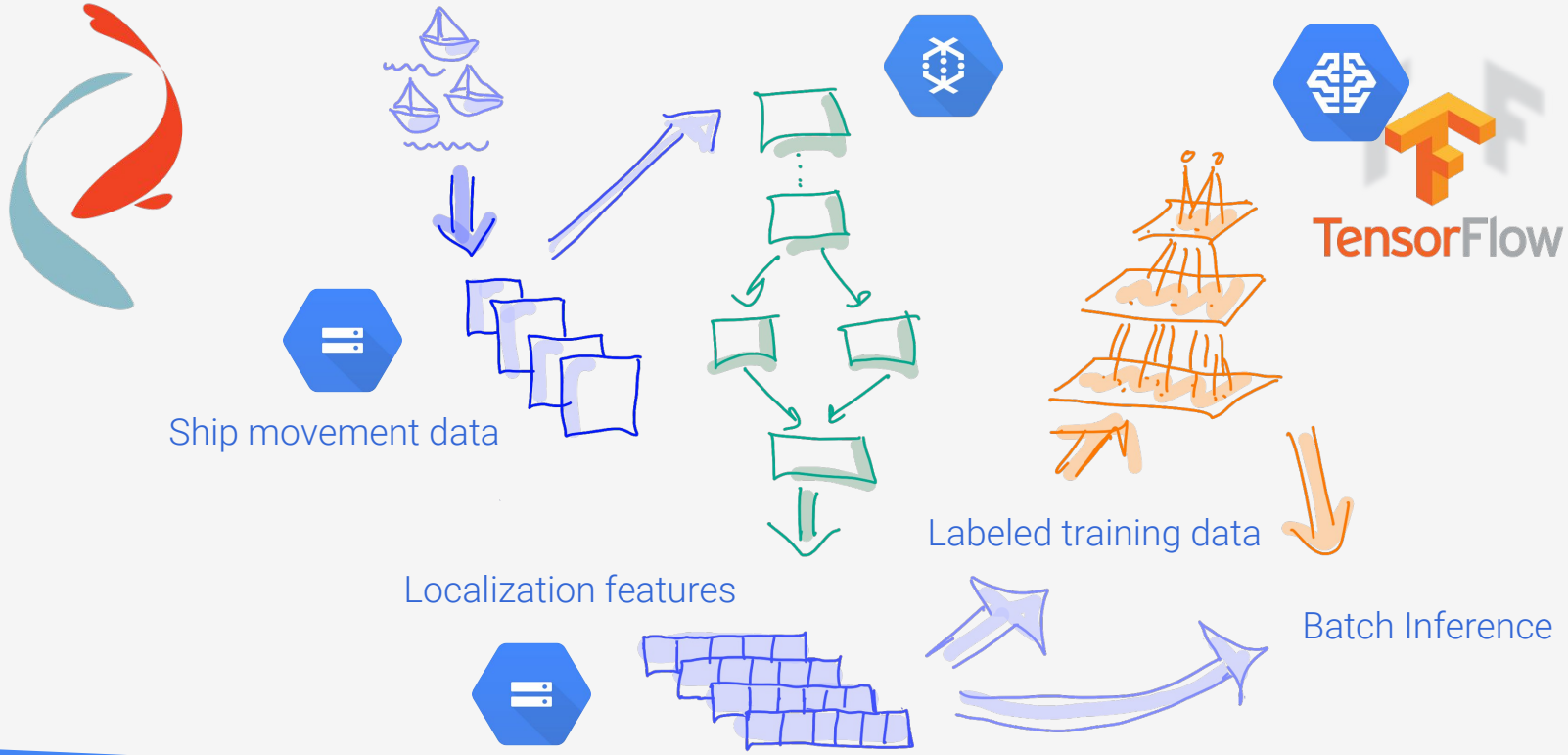
Operation time reduced from **15 min to 3 min**

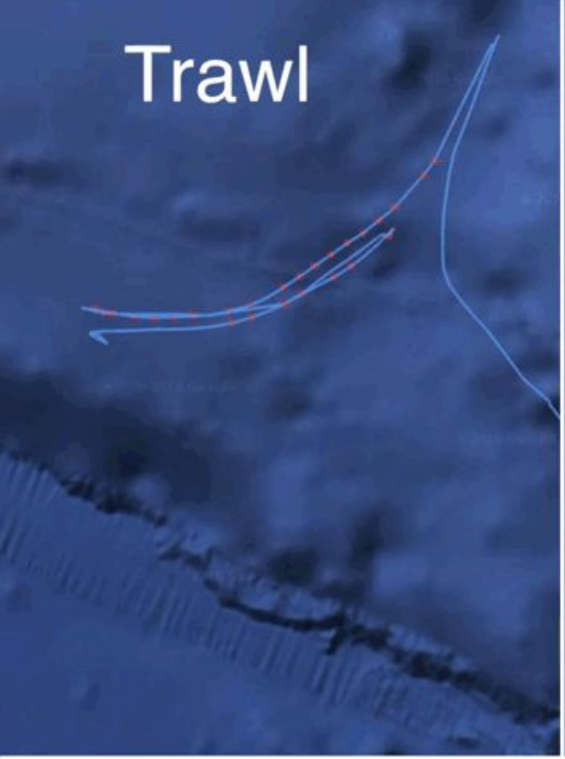


Use Cases: Internet of Things (IoT)

GLOBAL FISHING WATCH BETA
Madagascar
Mauritius
Saint-Denis
Reunion
Swaziland
Lesotho
Pretoria
Maputo
Durban
Richards Bay
East London
12 NOV 2013
10 FEB 2014
NOV
DEC
JAN
FEB
MAR
2014
90 days
-39.6226, 27.9053
OCEANA
SKYTRUTH
Google
Map data ©2016 Google, INEGI Imagery ©2016 NASA, TerraMetrics

Global Fishing Watch: preventing overfishing






Ivonne Pellworm Marinephotobank

Chris Howell, Shipspotter


Ulrich Karlowski 2008_Marine Photobank

Google DC Ops

Applying ML lead to 40% reduction in cooling energy in Google datacenters.

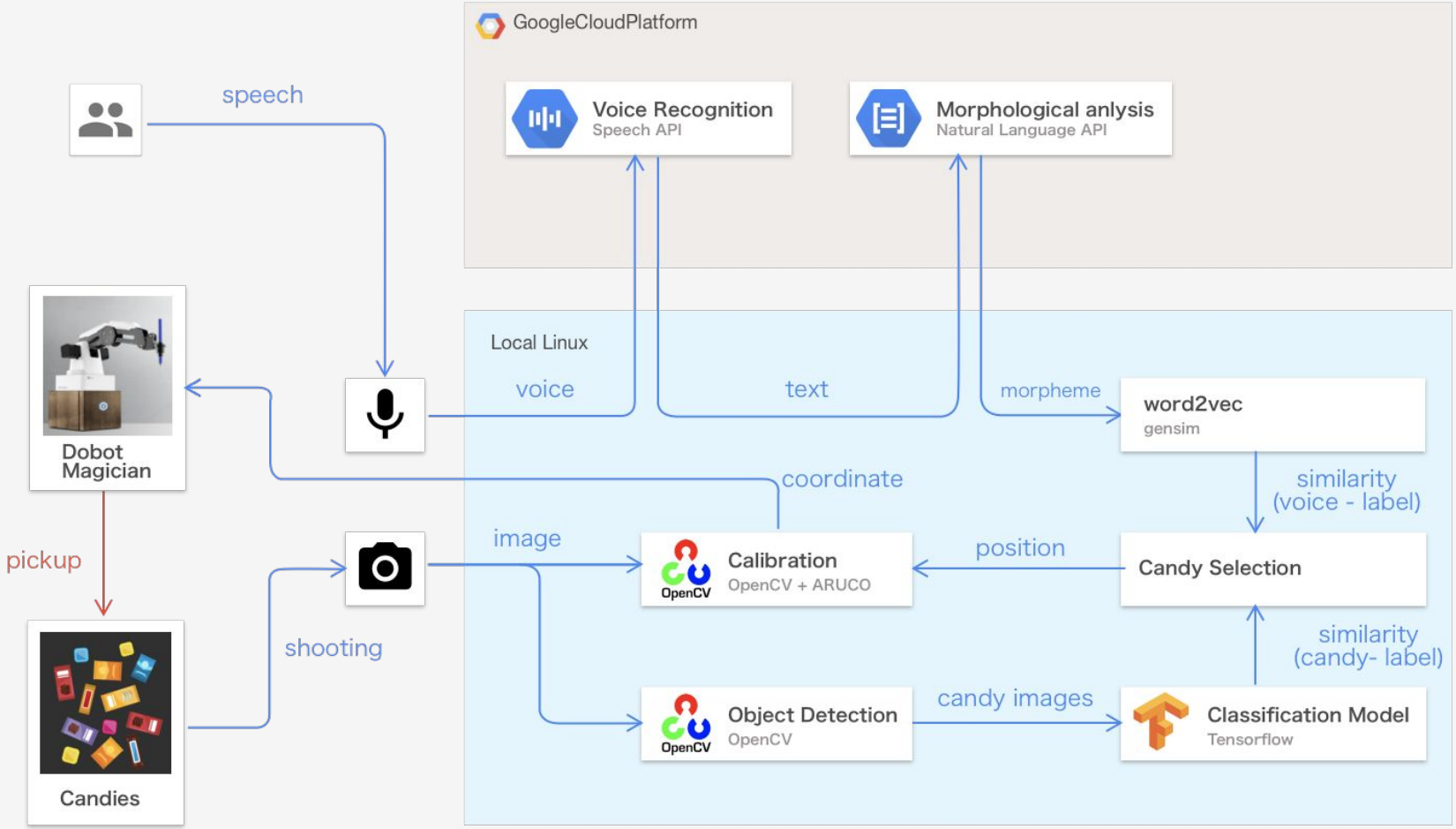
"Find Your Candy" Demo

Robot arm picks a candy

Speech API and **NL API** for voice command control

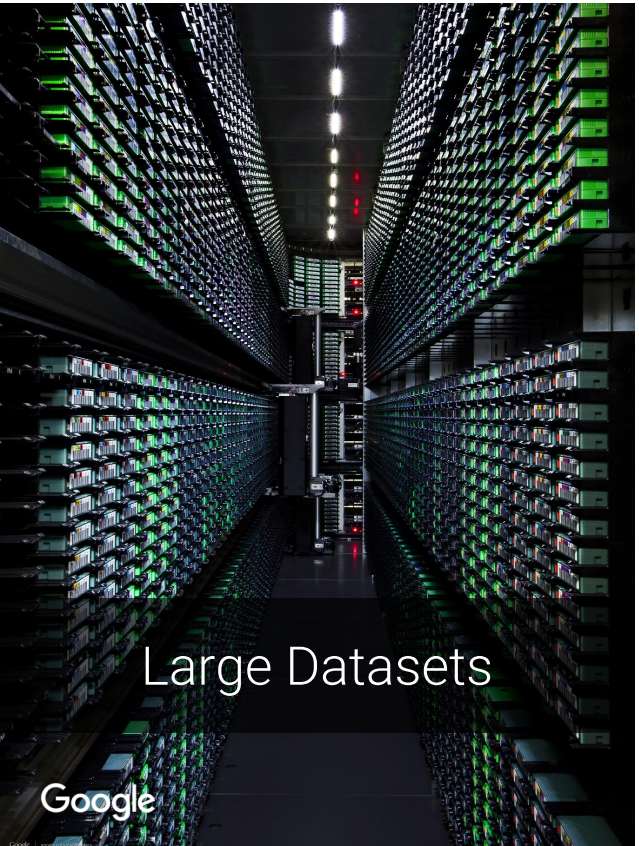
TensorFlow and **ML Engine** for object recognition


17
Cloud TPU

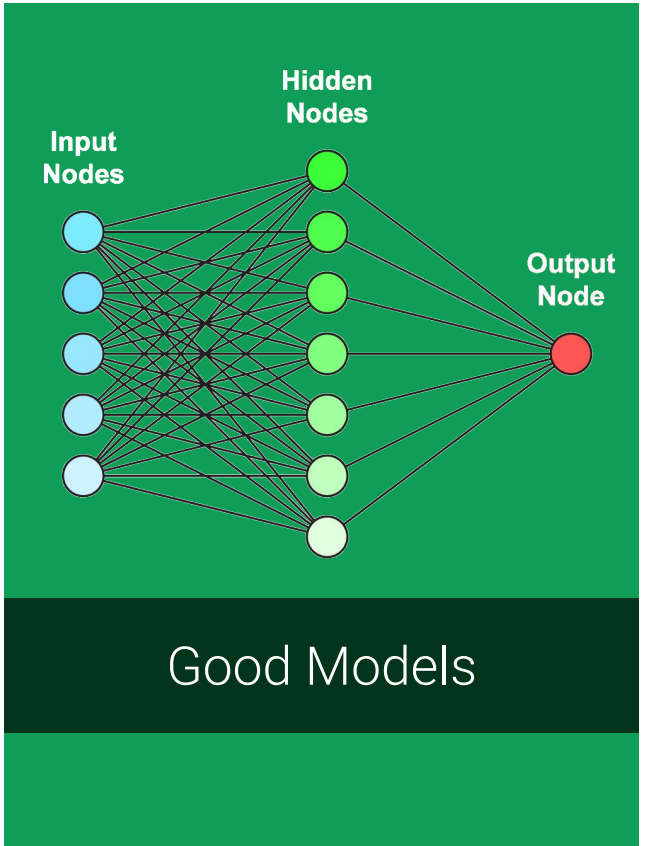

GoogleCloudPlatform
Voice Recognition
Speech API
Morphological anlysis
Natural Language API
speech
Local Linux
voice
text
morpheme
word2vec
gensim
Dobot
Magician
coordinate
similarity
(voice - label)
pickup
image
position
Calibration
OpenCV + ARUCO
Candy Selection
shooting
similarity
(candy- label)
Object Detection
OpenCV
candy images
Classification Model
Tensorflow
Candies

ML dev tools

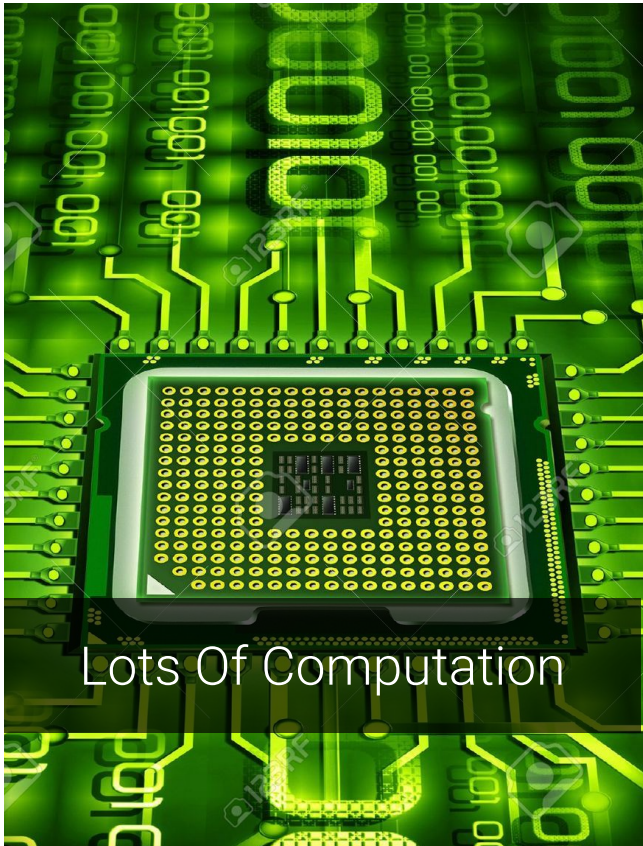
The Challenges of ML



Large Datasets

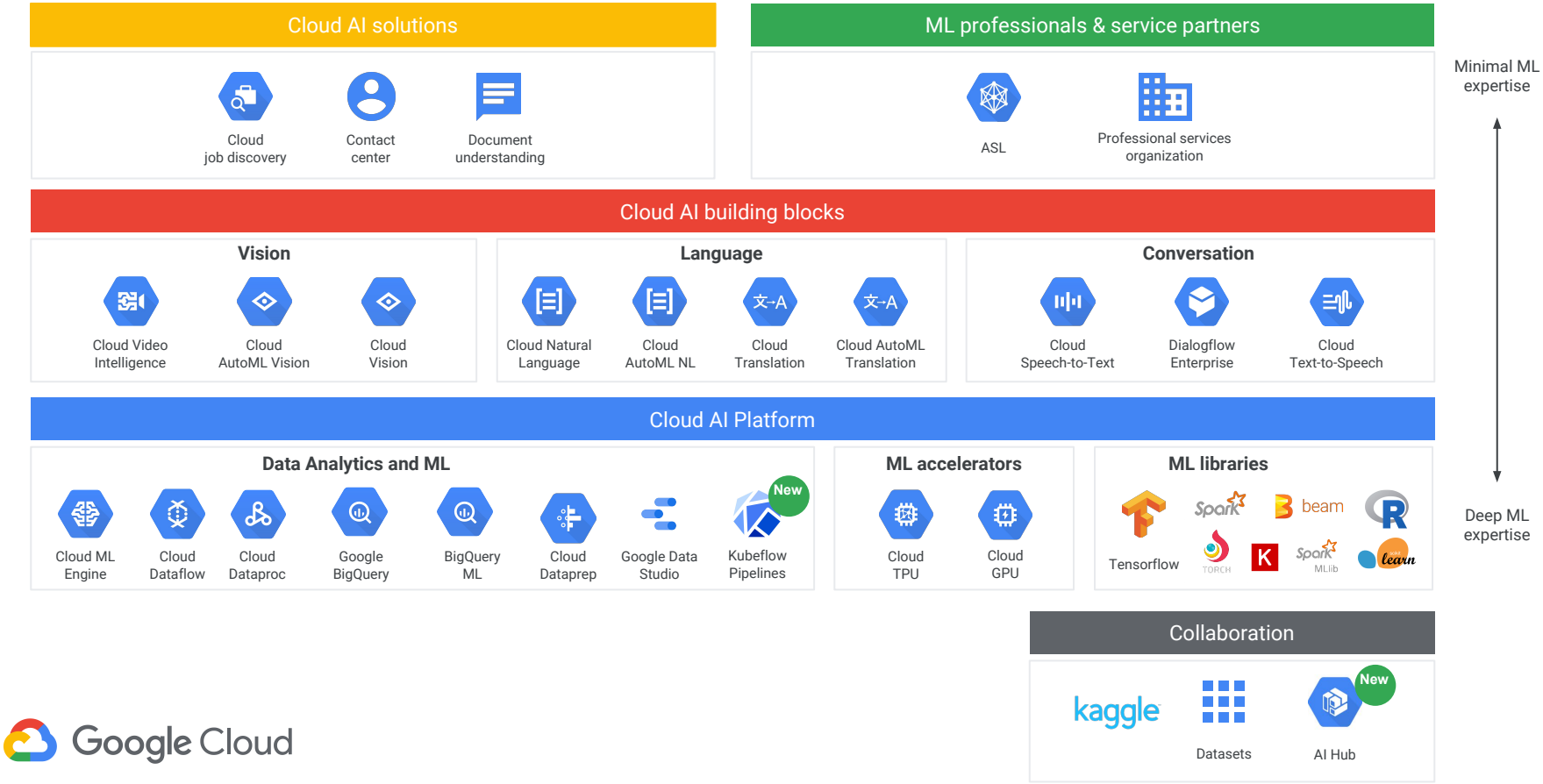


Good Models

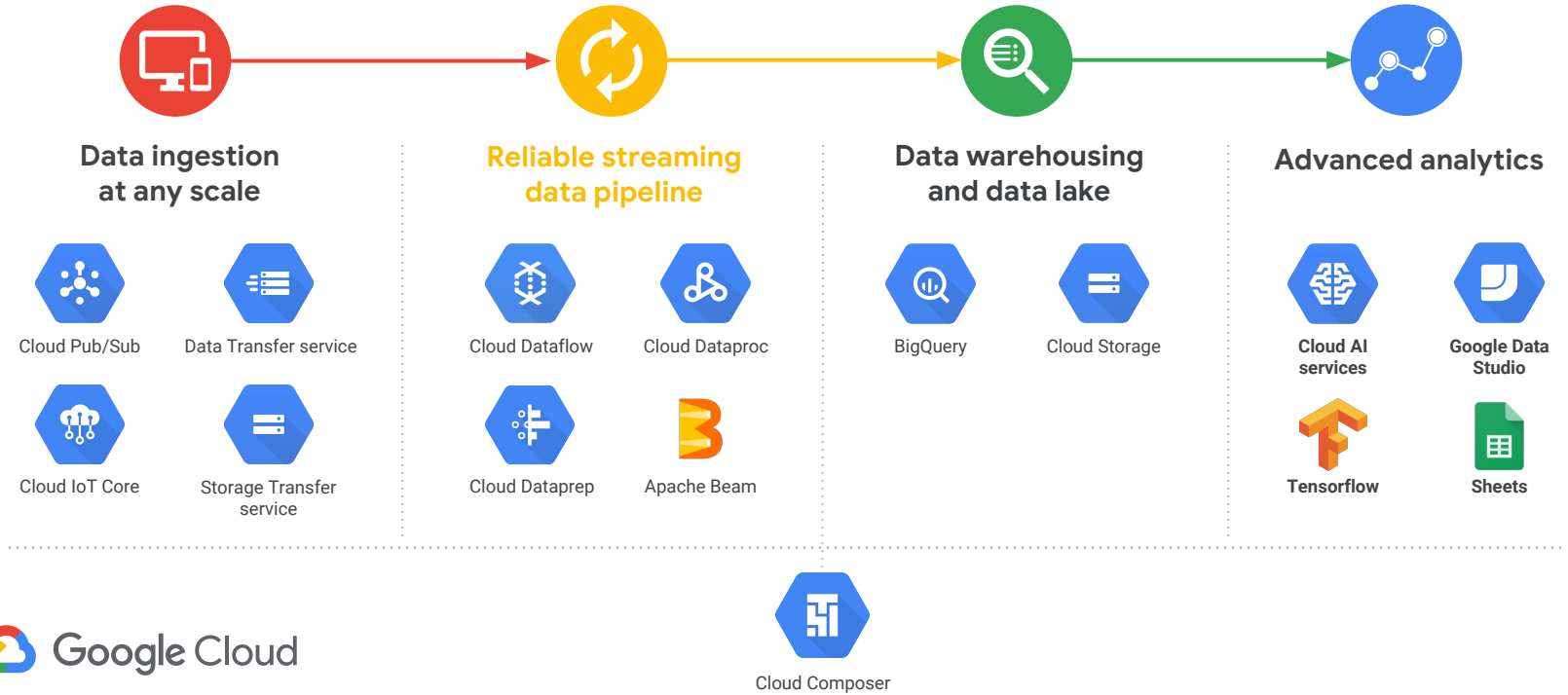


Lots Of Computation

Solution:
The Datacenter as a Computer


Cloud AI solutions
Cloud job discovery
Contact center
Document understanding
ML professionals & service partners
ASL
Professional services organization
Minimal ML expertise
Deep ML expertise
Cloud AI building blocks
Vision
Cloud Video Intelligence
Cloud AutoML Vision
Cloud Vision
Language
Cloud Natural Language
Cloud AutoML NL
Cloud Translation
Cloud AutoML Translation
Conversation
Cloud Speech-to-Text
Dialogflow Enterprise
Cloud Text-to-Speech
Cloud AI Platform
Data Analytics and ML
Cloud ML Engine
Cloud Dataflow
Cloud Dataproc
Google BigQuery
BigQuery ML
Cloud Dataprep
Google Data Studio
New
Kubeflow Pipelines
ML accelerators
Cloud TPU
Cloud GPU
ML libraries
Tensorflow
Spark
beam
TORCH
Spark MLlib
learn
Collaboration
kaggle
Datasets
New
AI Hub
Google Cloud

Serverless data analytics and AI



Google Cloud

Colaboratory

Run pre-built ML models and check its behavior for free

Colaboratory is being developed in Google Research Project

1. **Free of charge**

   Check and validate model behavior for free

2. **Interactive experiments**

   Various ML codes are shared in colab format. So you can quickly check how it works.

3. **Access to GPU and Cloud TPU**

   Can access to GPU and Cloud TPU for FREE!

Cloud Deep Learning VM Image beta

VM Images works with your GCE instances

TensorFlow PyTorch Chainer dmlc XGBoost

You can quickly use above frameworks by specifying VM images that is prepared for each framework.

Fast prototyping

Popular ML frameworks are pre-installed, and you can quickly start prototyping your custom models.

Train your models faster

Leverage multi-GPUs and Cloud TPU to make your training faster.

Interactive experiments

Pre-installed JupyterLab enables you to experiment ML models interactively. In addition you can see how many GPUs are utilized using web based console.

Fast Prototyping

Pre-installed ML packages make you start prototyping immediately

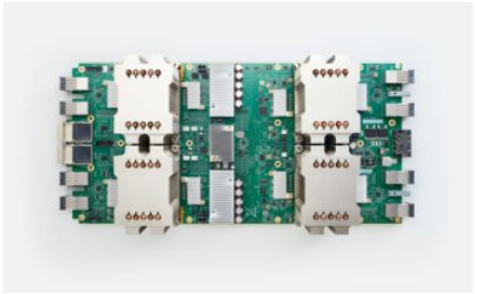
- Python 2.7 and 3.5
- Deep Learning Frameworks (TensorFlow, PyTorch, Chainer and XGBoost)
- GPU / DNN Libraries (CUDA / CuDNN / NCCL)
- Development environment for ML models (JupyterLab)
- Scientific Libraries (numpy / scipy)
- Data Analysis Libraries (pandas / matplotlib)
- Machine Learning Library (sklearn)
- Natural Language Processing Library (nltk)
- Image Processing Libraries (Pillow / scikit-image / Opencv-python)

Train your model faster

You can leverage GPU and Cloud TPU

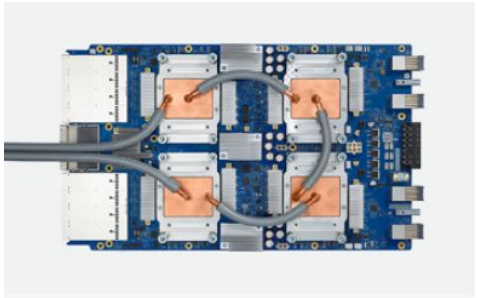
NVIDIA TESLA
V100 for NVLINK

125 TFlops / 16 GB HBM2
$2.48 USD per GPU (US)



Cloud TPU v2

180 TFlops / 64 GB HBM
$4.50 USD per TPU v2
$1.35 USD per preemptible TPU v2



Cloud TPU v3 beta

420 TFlops / 128 GB HBM
$8.00 USD per TPU v3
$2.40 USD per preemptible TPU v3

Prices are based on US region

AI Hub ALPHA

1. **One stop AI catalog**

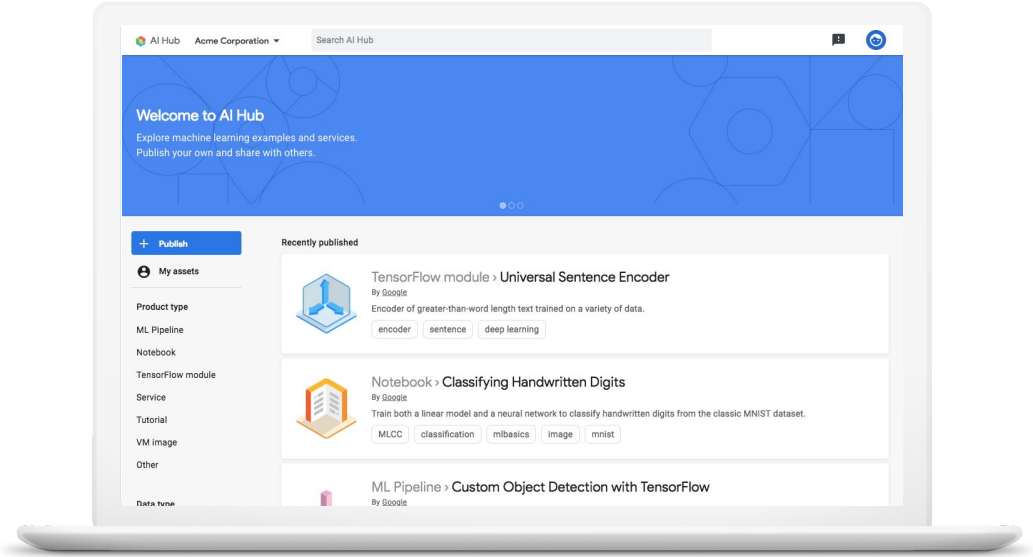
   Easily discover plug & play pipelines & other content built by Google AI and partners.

2. **Enterprise-grade sharing controls**

   Host pipelines and ML content with private sharing controls within an enterprise to foster reuse within organizations.

3. **Easy deployment on GCP and hybrid**

   Deploy pipelines via Kubeflow on GCP and on premise.



Google Cloud

AI Hub alpha

Enables ML model discovery and reuse for everyone

AI Hub

Public Contents

By Google

Unique AI assets
by Google

AutoML, TPUs, kaggle
Cloud AI Platform, etc.

Research at Google

DeepMind

By partners

Created, shared &
monetized by anyone.

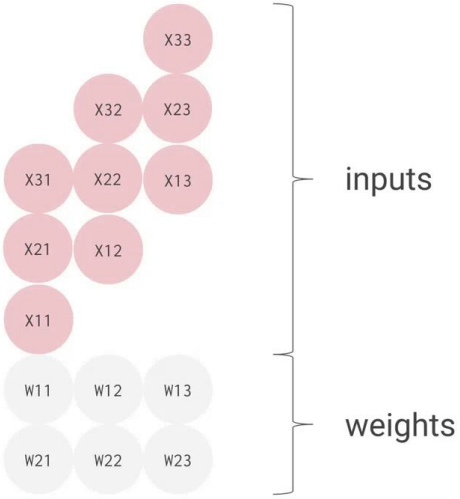
+ Private Contents

By customers

Content shared
Securely within and
with other organizations.

ML at Google scale

Large scale training and prediction
with Cloud ML Engine



TPU v2

ASIC for TensorFlow

Designed by Google

180 Tflops / chip

TPU v2 Pod: 11.6 PFLOPS

Cloud TPU Performance

ResNet-50 on TPU v2 half-pod

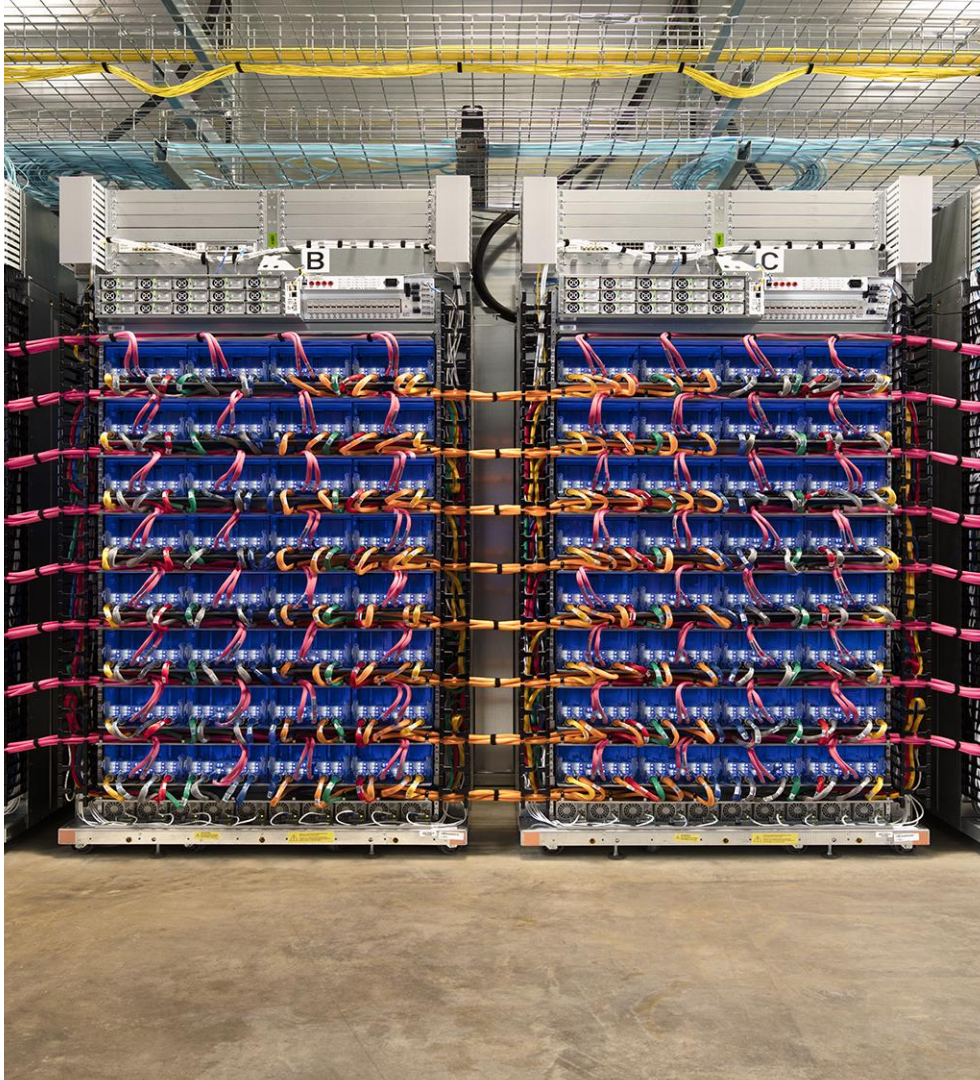
Real data: 77,392 images/sec

Final accuracy: 93%

Training time: 30 min

#1 training time on DAWNBench

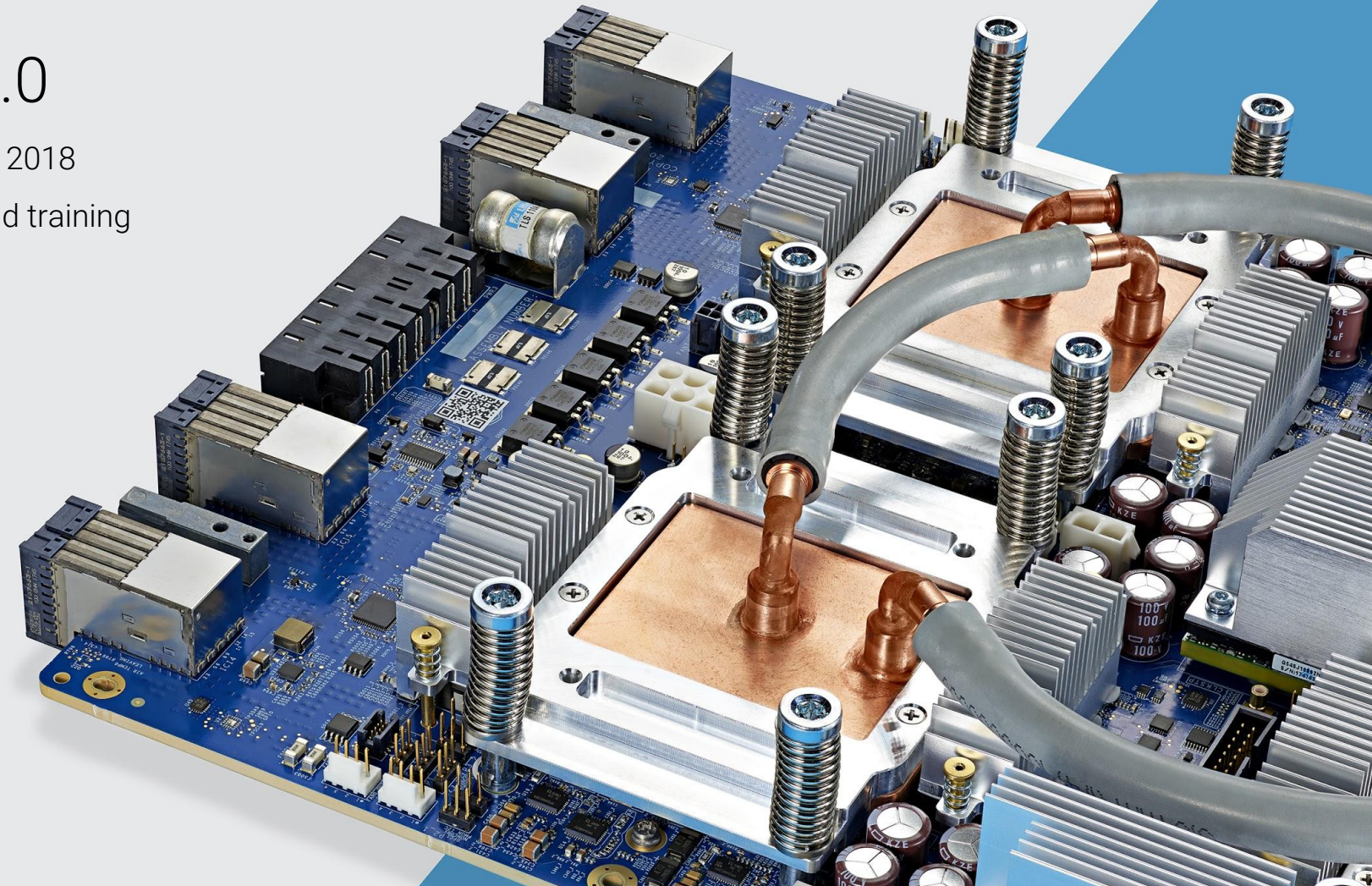
Google Cloud

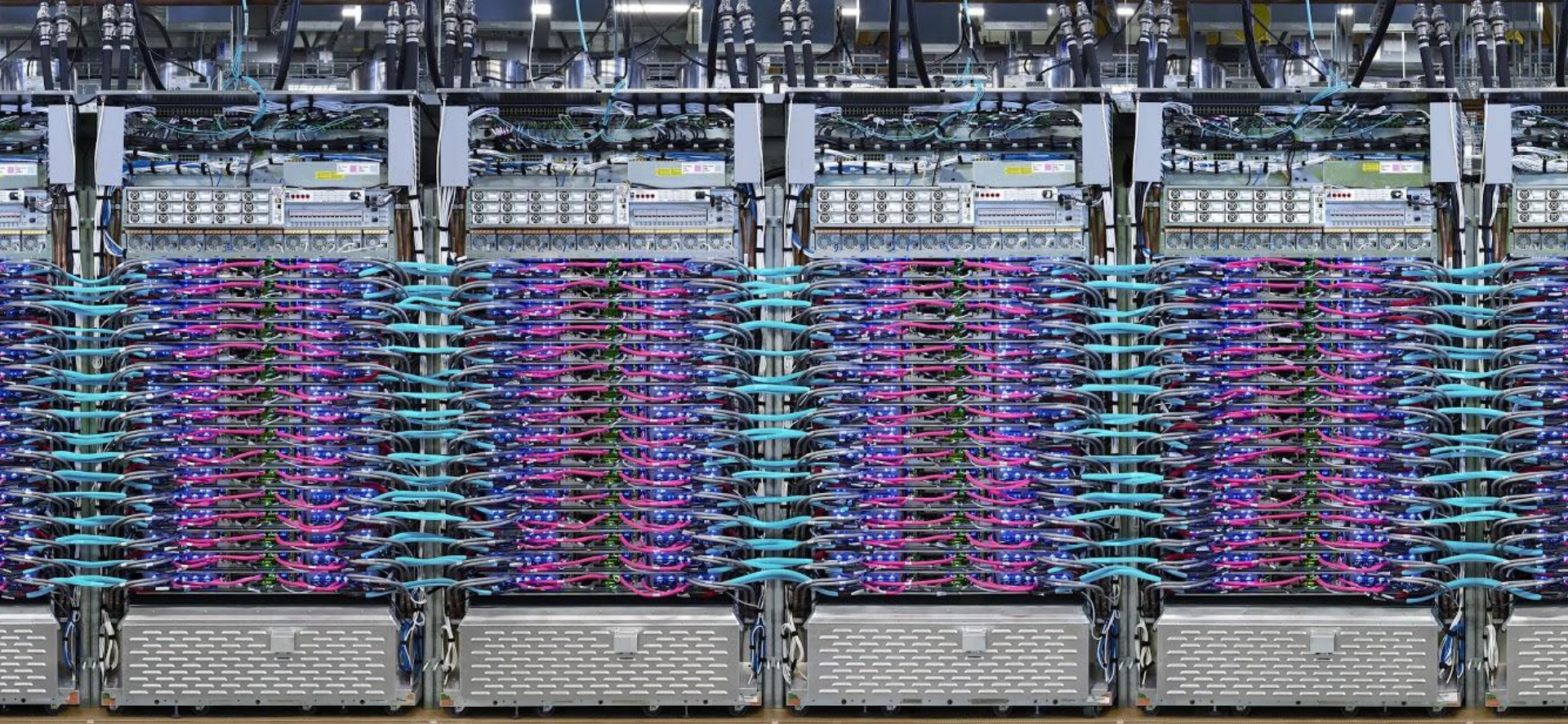


TPU 3.0

Launched in 2018

Inference and training





TPU 3.0 Pod: 100 PFLOPS (8X faster than v2)

Google Cloud
Summit

cloud.google.com

for more info and demo

Thank You!