Machine Intelligence at Google Scale: TensorFlow


Data is Everything.
How well you use your data can determine the degree of your success.

What is Machine Learning?

data

algorithm

insight


Coast
Water
Pool
Beach
Woman

Google Translate

Latin | English | Italian | English - detected

English | Latin | Italian | Translate

The internet has sufficient amounts of cat pictures!

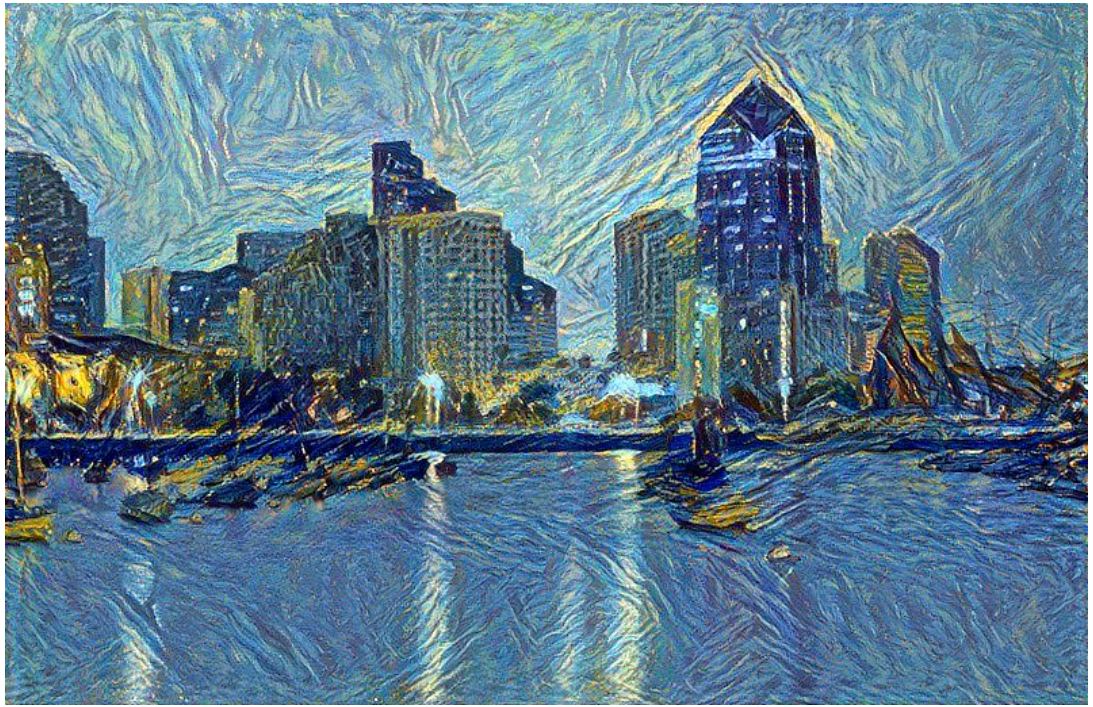
Internet ha una quantità sufficiente di foto di gatti!

Human:
Someone is using a small grill to melt his sandwich.

Neural Model:
A person is cooking some food on a grill.

Google

One example (of 100s): “Neural Art” in TensorFlow

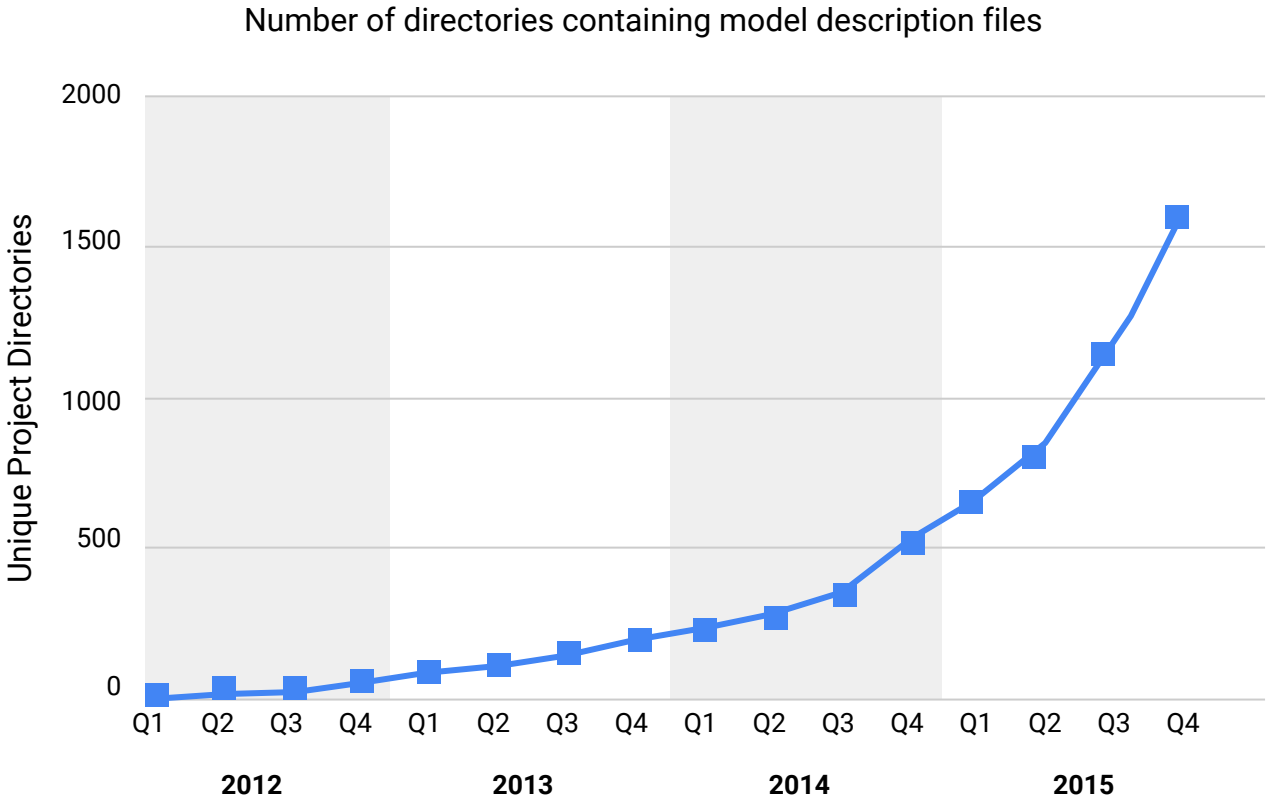


An implementation of "A neural algorithm of Artistic style" in TensorFlow, for

- Introductory, hackable demos for TensorFlow, and
- Demonstrating the use of importing various Caffe cnn models (VGG and illustration2vec) in TF.

github.com/woodrush/neural-art-tf

Growing Use of Deep Learning at Google

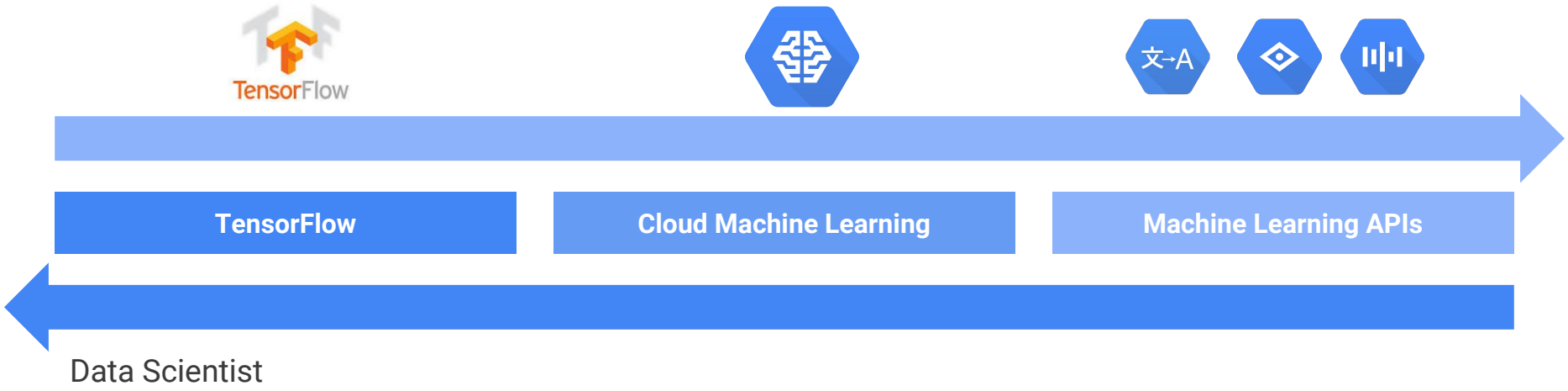


Across many products/areas

- Apps
- Maps
- Photos
- Gmail
- Speech
- Android
- YouTube
- Translation
- Robotics Research
- Image Understanding
- Natural Language Understanding
- Drug Discovery

Google

The Machine Learning Spectrum



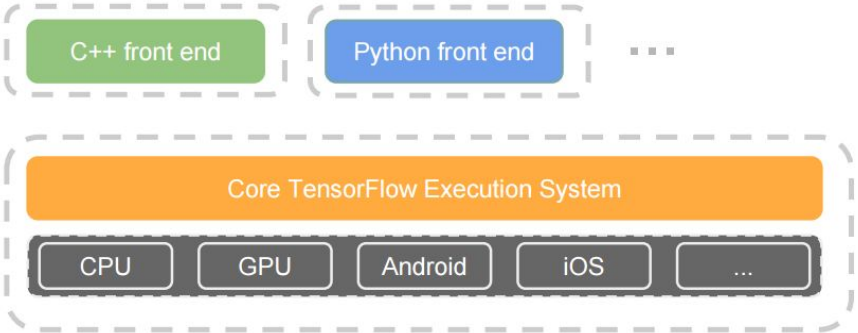
TensorFlow

What is TensorFlow?

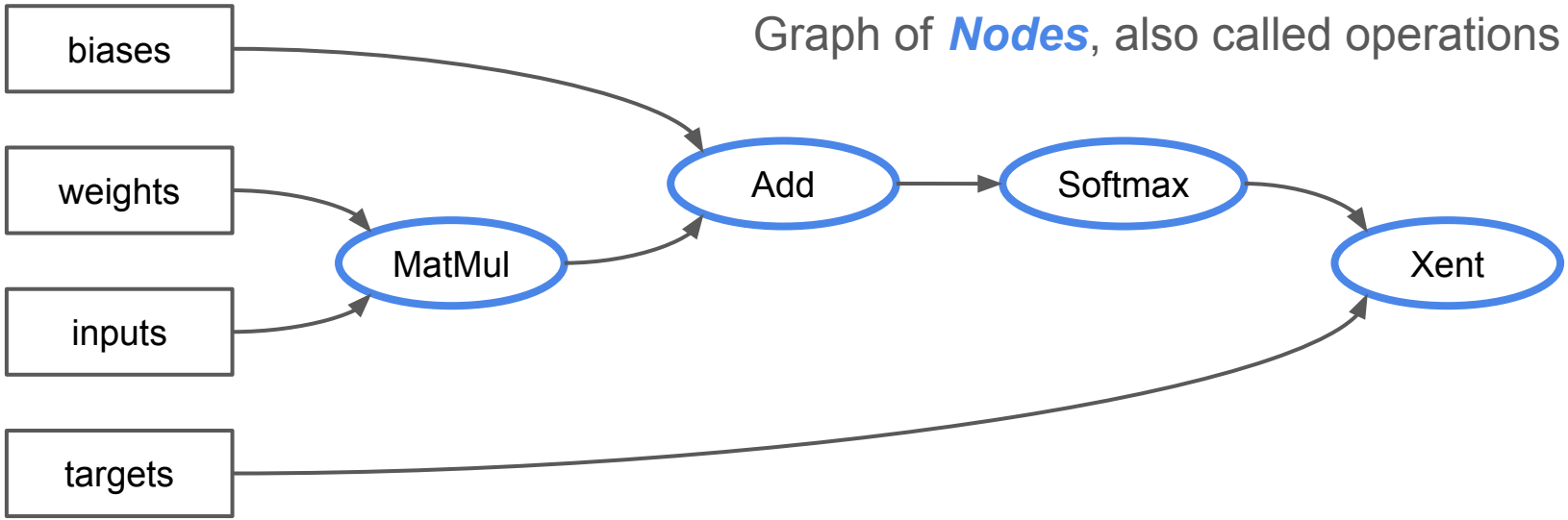
- Google's **open source** library for machine intelligence
- **Tensor**: N-dimensional array
- **Flow**: data flow computation framework (like MapReduce)
- **tensorflow.org** launched in Nov 2015
  - Most popular Machine Learning library
  - 20,000+ stars / 7000+ forks
  - 165 contributors / > 3000 commits

Portable

- Training on:
  - Data Center
  - CPUs, GPUs and etc
- Running on:
  - Mobile phones
  - IoT devices

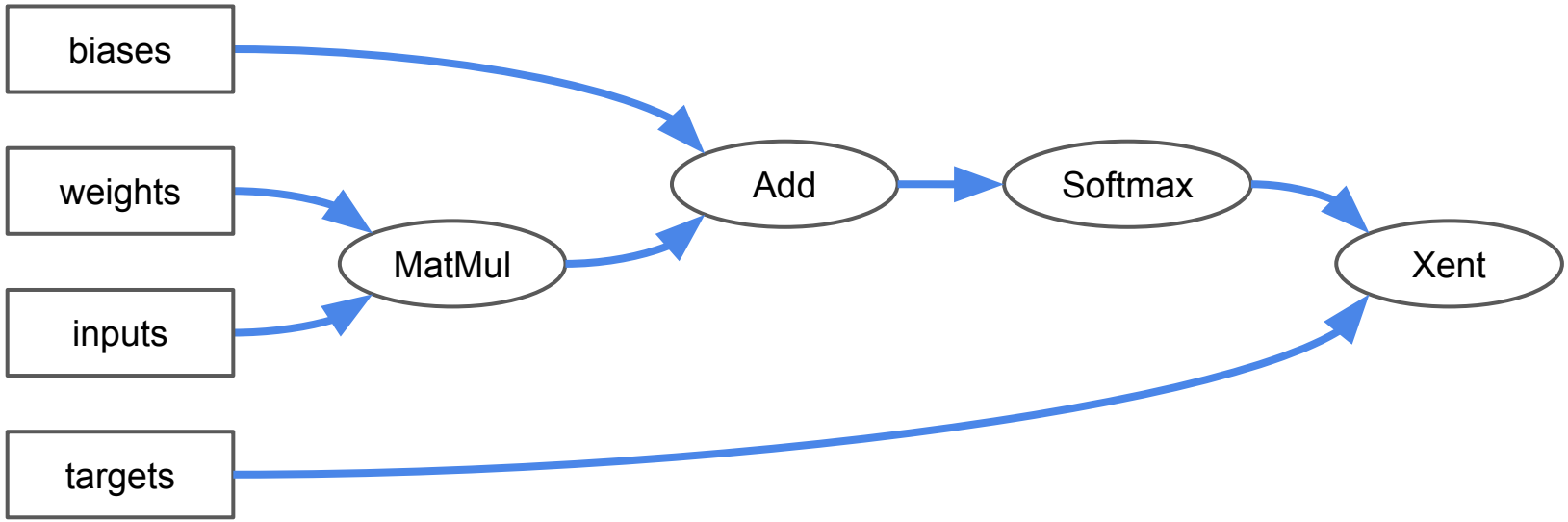


Computation as a Dataflow Graph



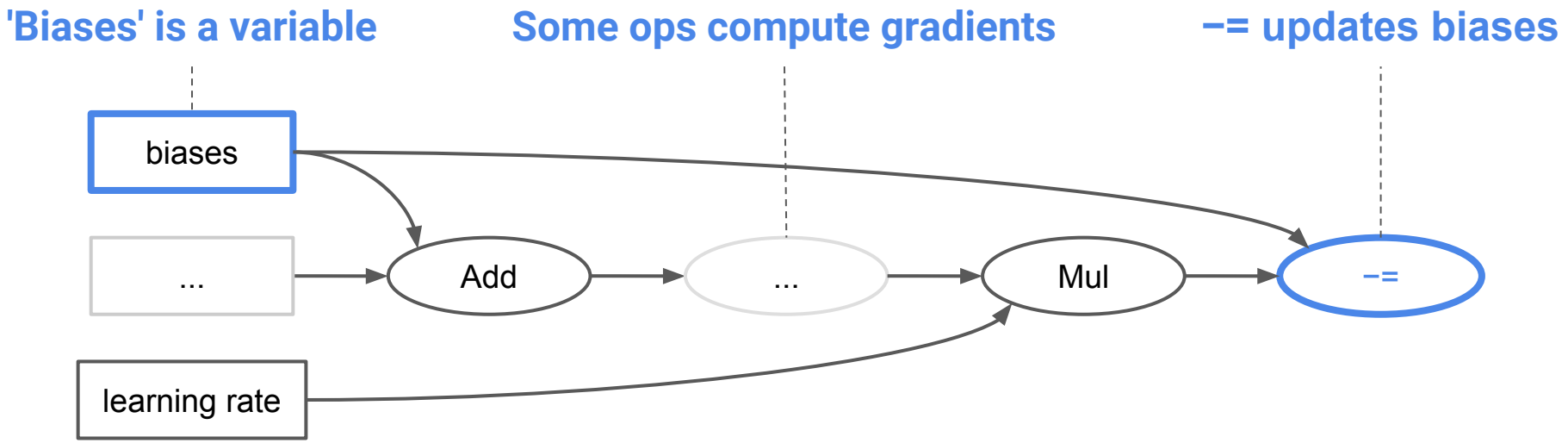
Dataflow Graph (forward)

Edges are N-dimensional arrays: ***Tensors***



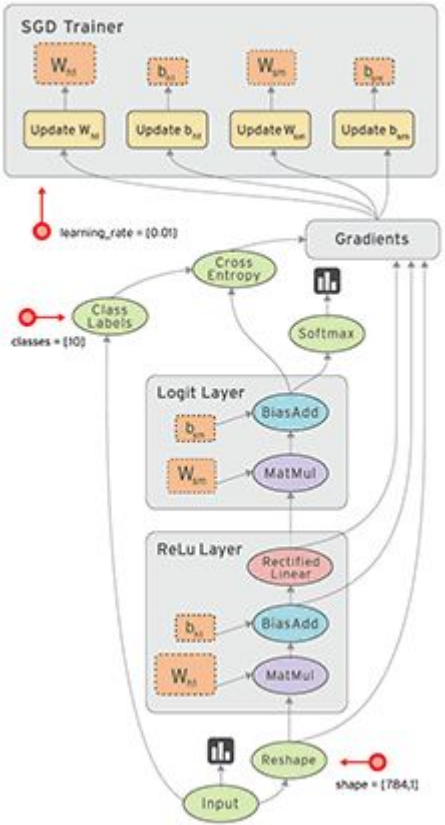
Dataflow Graph (backward graph and updates)

Backward graph and update are added automatically to graph



Dataflow Graph

- In practice often very complex with 100s or 1000s of nodes and edges
- “Inference” means to execute just the forward path of the graph

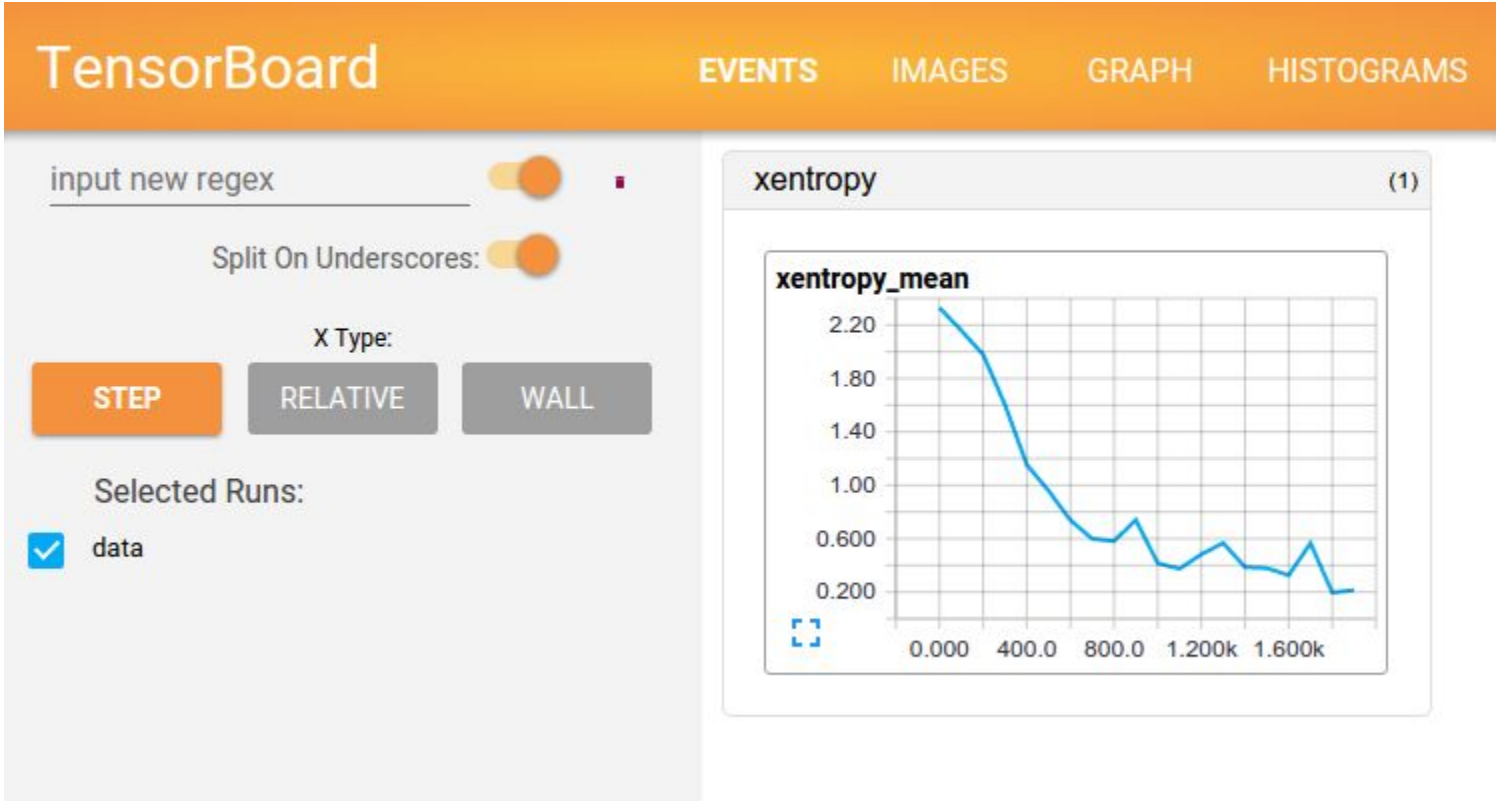


Example: Logistic Regression

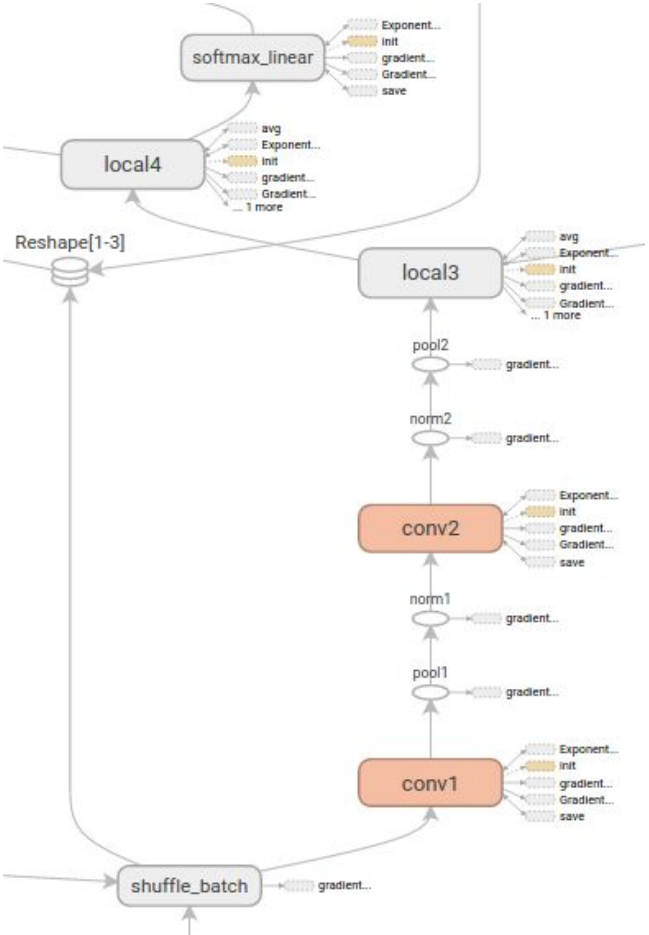
1. *N* examples of data
   - Input $\mathbf{x}_n$
   - Target $\mathbf{t}_n$
2. Model with parameters (weights **W,** biases **b**)
   - $\mathbf{y}_n = \text{Softmax}(\mathbf{W} * \mathbf{x}_n + \mathbf{b})$
3. CrossEntropy loss to minimize
   - ***loss*** $= -\text{Sum}(\mathbf{t}_n * \ln(\mathbf{y}_n))$
4. Minimization by Gradient Descent
   - requires gradients of ***loss*** with respect to parameters **W**, **b**

- TensorFlow does automatic differentiation
- Gradient calculation done by TF once model & loss defined

Visualizing learning

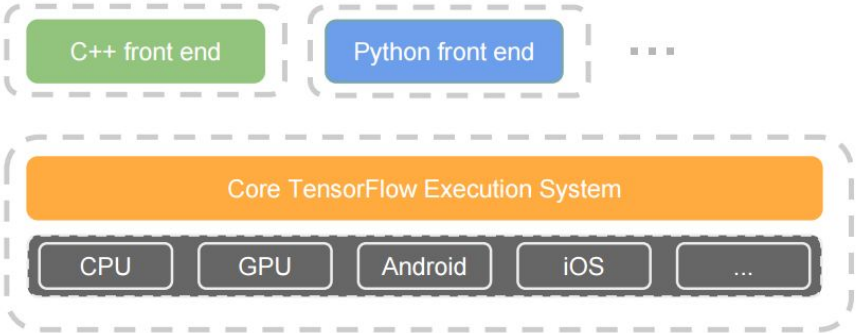


Visualizing graphs

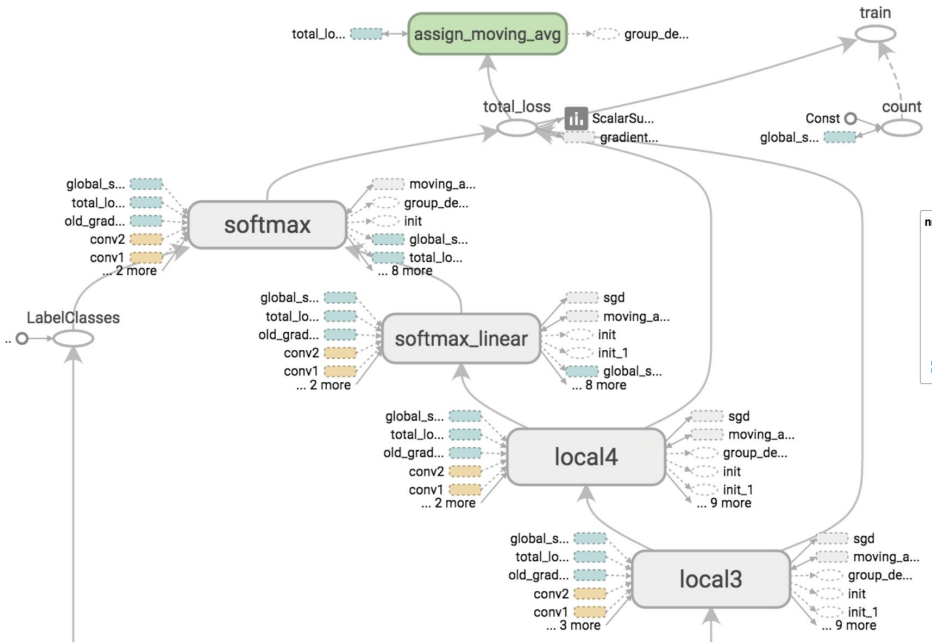


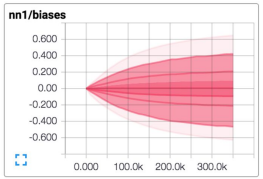
Portable

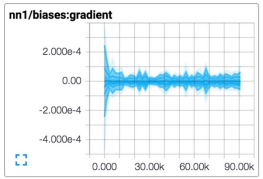
- Training on:
  - Data Center
  - CPUs, GPUs and etc
- Running on:
  - Mobile phones
  - IoT devices

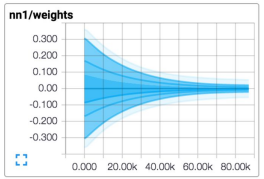


TensorBoard: visualization tool

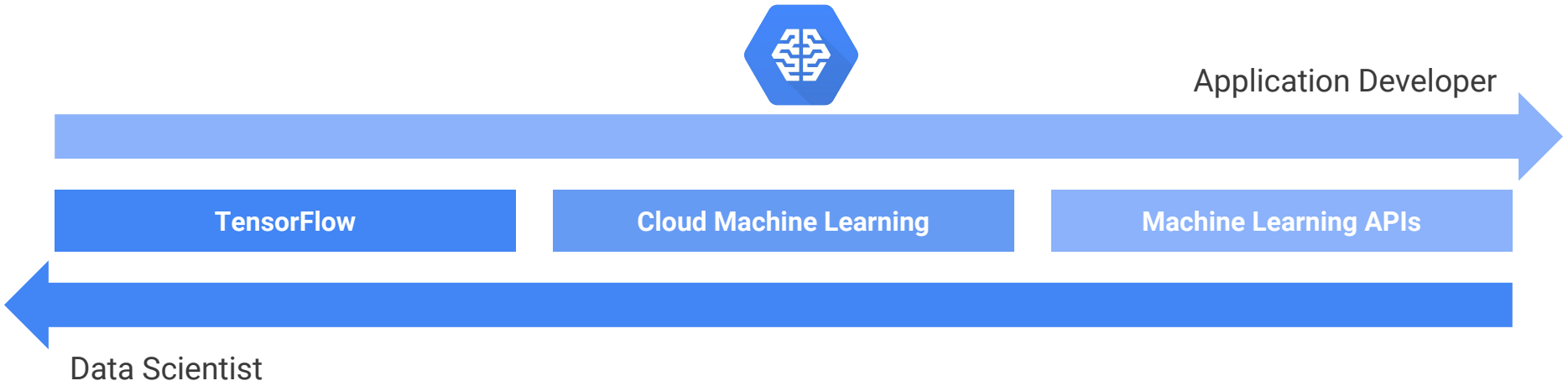








The Machine Learning Spectrum



Cloud Machine Learning Platform

Training

Data

Pre-Processing → **Training** → **Evaluation**

- Intelligent Text Processing
- Image Feature Processing

- Managed Infrastructure
- Scalable Algorithms
- HW Acceleration

- Detailed Model Metrics
- In-process model viz
- No-ops experience

Datalab

ML Authoring

ML Analysis

Prediction

input

Batch

Real-Time

output

Hosted Model

- Hosted Model experience
- AutoScaling
- Anytime access

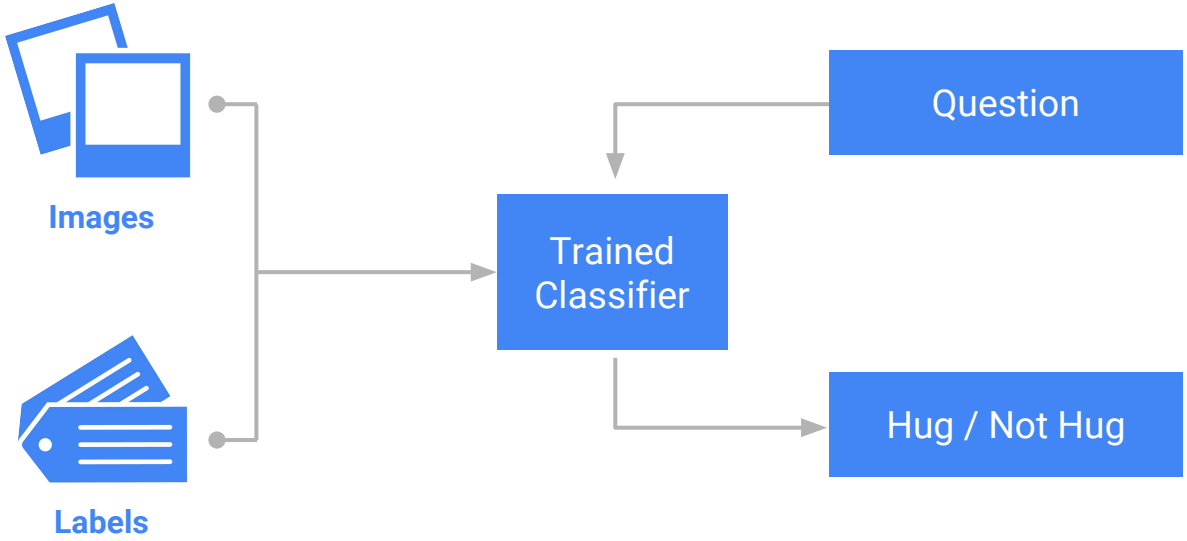
Console

Monitoring & Logging

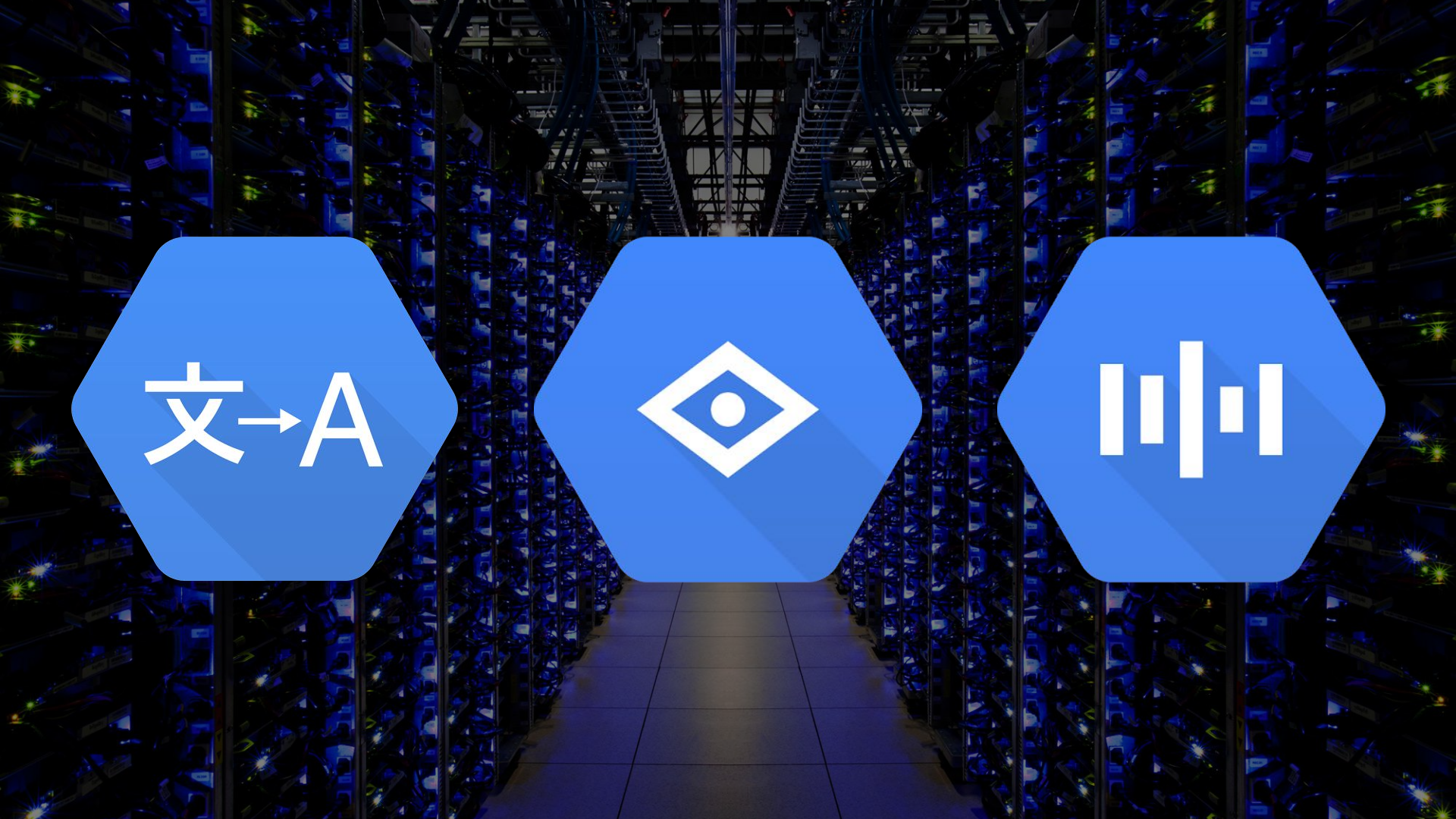
Model & Job Mgmt

Can I Hug That?

Can I Hug That?



Demo


文→A

Easy to use API: pass in an image, we give you insight:

Label Detection

OCR

Explicit Content Detection

Facial Detection

Landmark Detection

Logo Detection