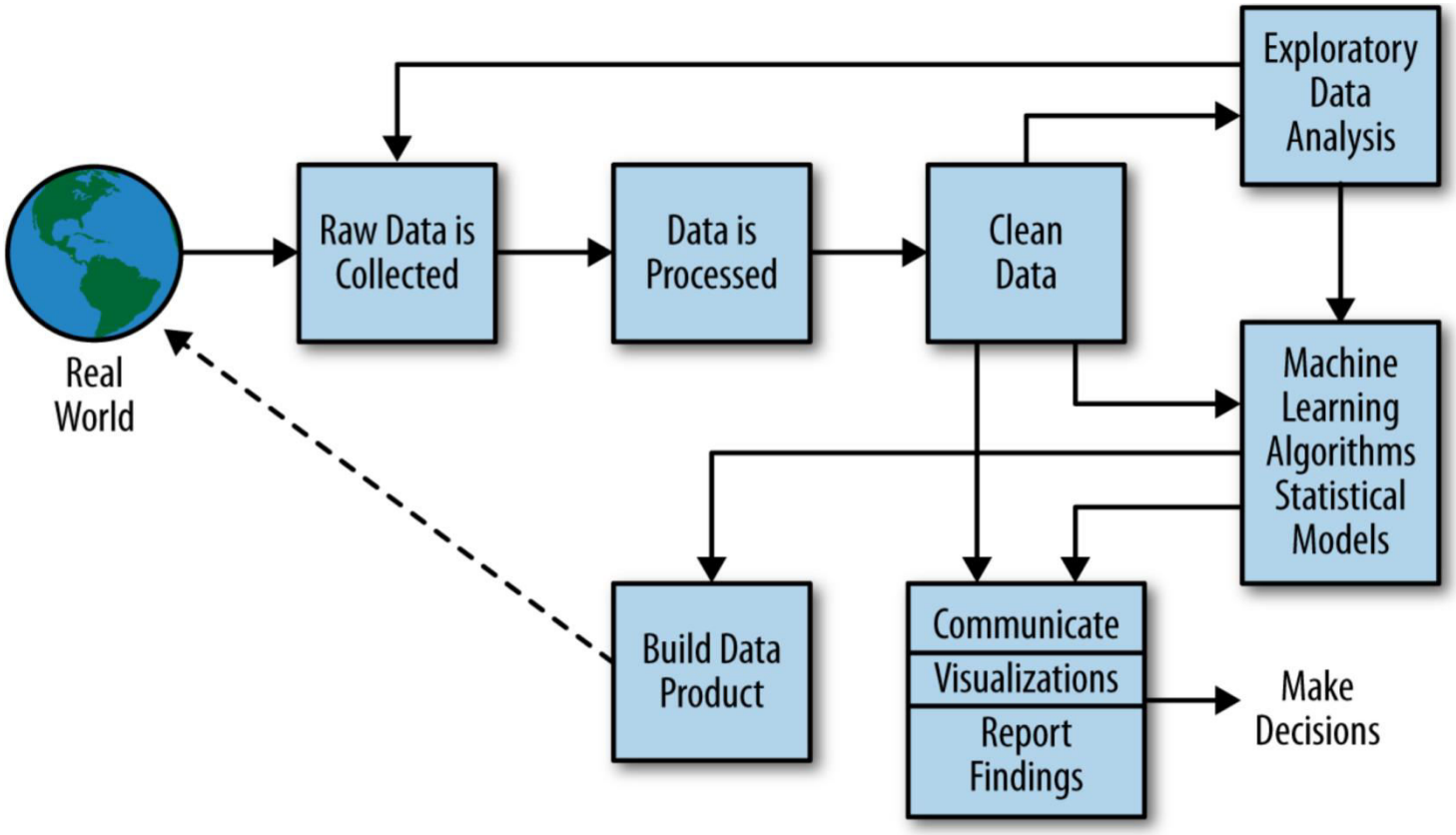
Deep Learning With TensorFlow Crash Course

Outline

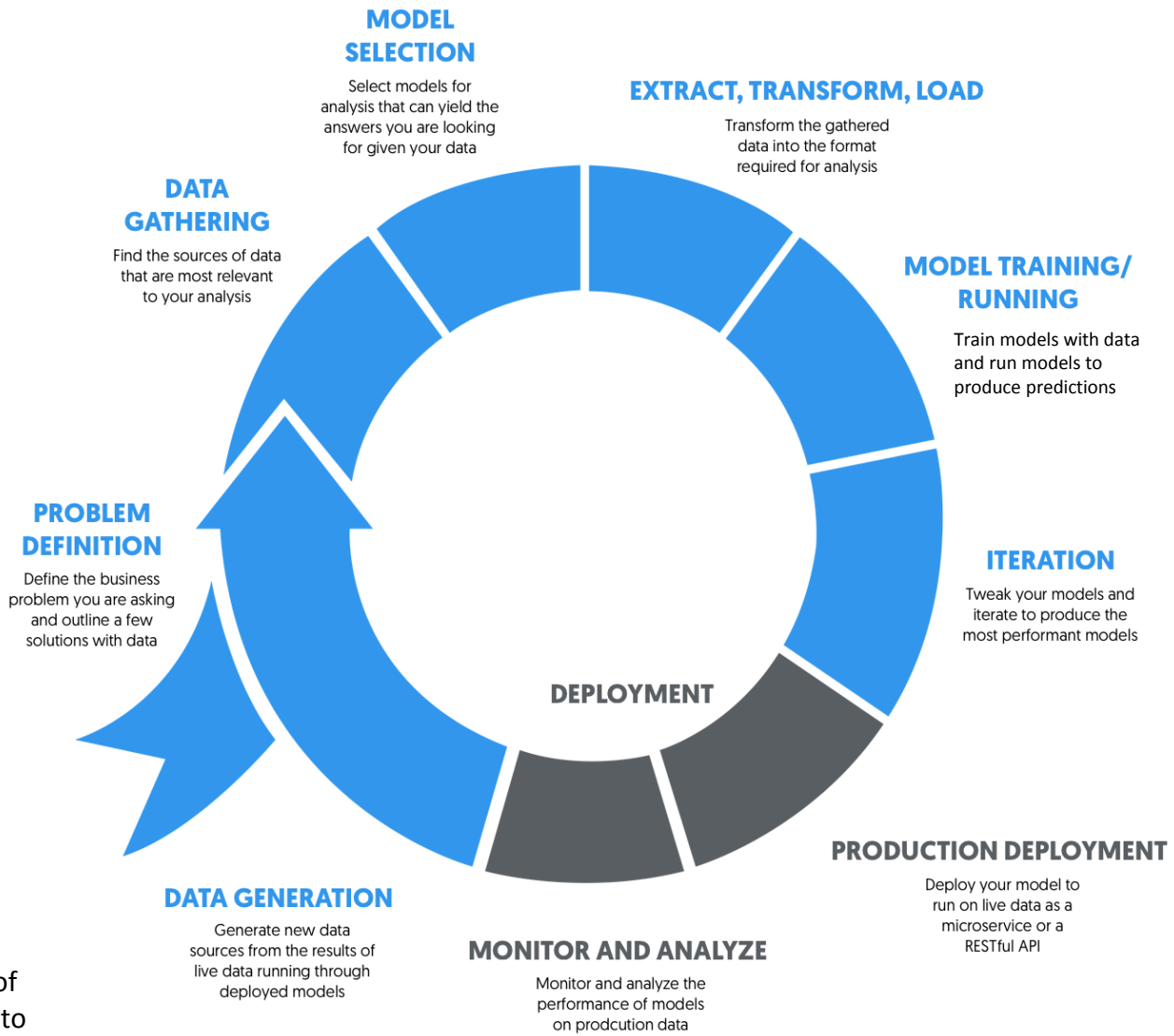
- Data science, machine learning and deep learning
- TensorFlow vs other deep learning frameworks
- TensorFlow basics
- Deep neural network
- Convolutional neural network
- TensorFlow in deployment
- Current trends
- Resources

The Data Science/Data Analytics Process



Doing Data Science, Schutt and O'Neil

Machine Learning Lifecycle



Overview: Deployment of Quantitative Workflows to Production by Anand Sampat

Typical Goal of Machine Learning

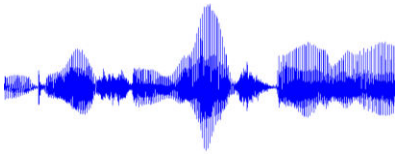
Input **Output**

Images /Video

ML

Label: “Motorcycle”
Caption: “Man with black helmet riding a black motorcycle”
Suggest tags
Image search
…

Audio



ML

Speech recognition
Music classification
Speaker recognition
Env. sound classification
…

Text

ML

Fake news classification
Anti-spam
Machine translation
Text search
…

Feature Representation

You see this

Camera sees this

194	210	201	212	199	213	215	195	178	158	182	209
180	189	190	221	209	205	191	167	147	115	129	163
114	126	140	188	176	165	152	140	170	106	78	88
87	103	115	154	143	142	149	153	173	101	57	57
102	112	106	131	122	138	152	147	128	84	58	66
94	95	79	104	105	124	129	113	107	87	69	67
68	71	69	98	89	92	98	95	89	88	76	67
41	56	68	99	63	45	60	82	58	76	75	65
20	43	69	75	56	41	51	73	55	70	63	44
50	50	57	69	75	75	73	74	53	68	59	37
72	59	53	66	84	92	84	74	57	72	63	42
67	61	58	65	75	78	76	73	59	75	69	50

Handlebars

Wheel

Feature Representation

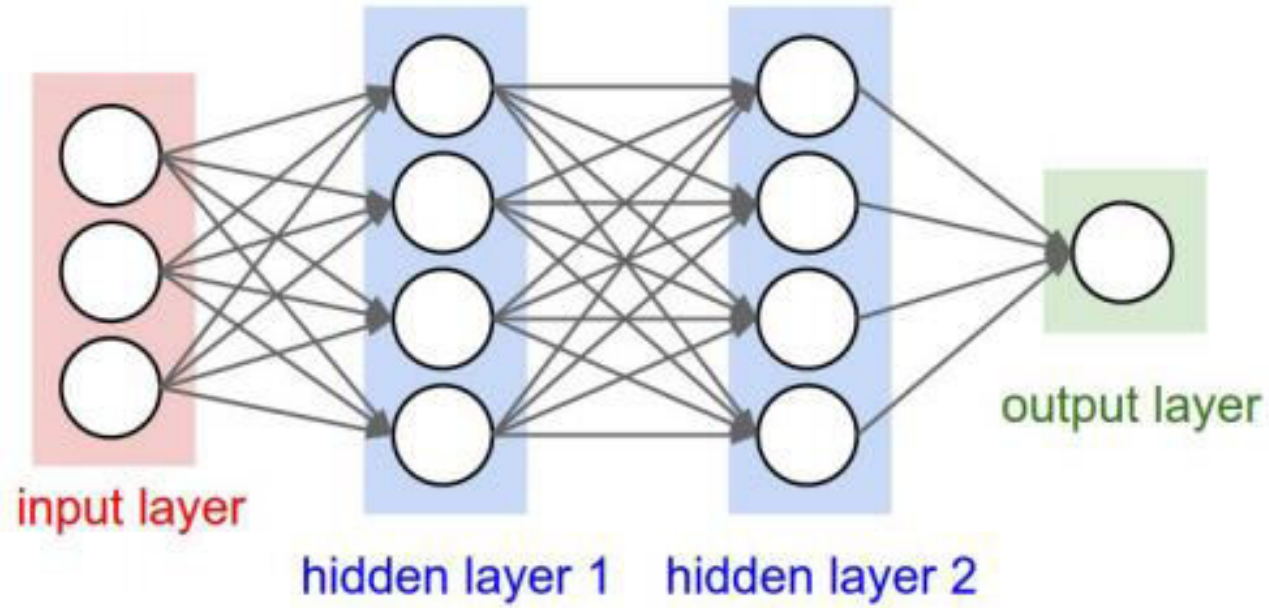
Learning Algorithm

E.g. Does it have handlebars? Wheels?

Expert knowledge required to come up with feature representations: e.g. SIFT, HoG, Textons, etc.

Deep Learning

- **Representation learning**: Learns feature representations from the data
- **Hierarchical** representation of features, i.e. deep

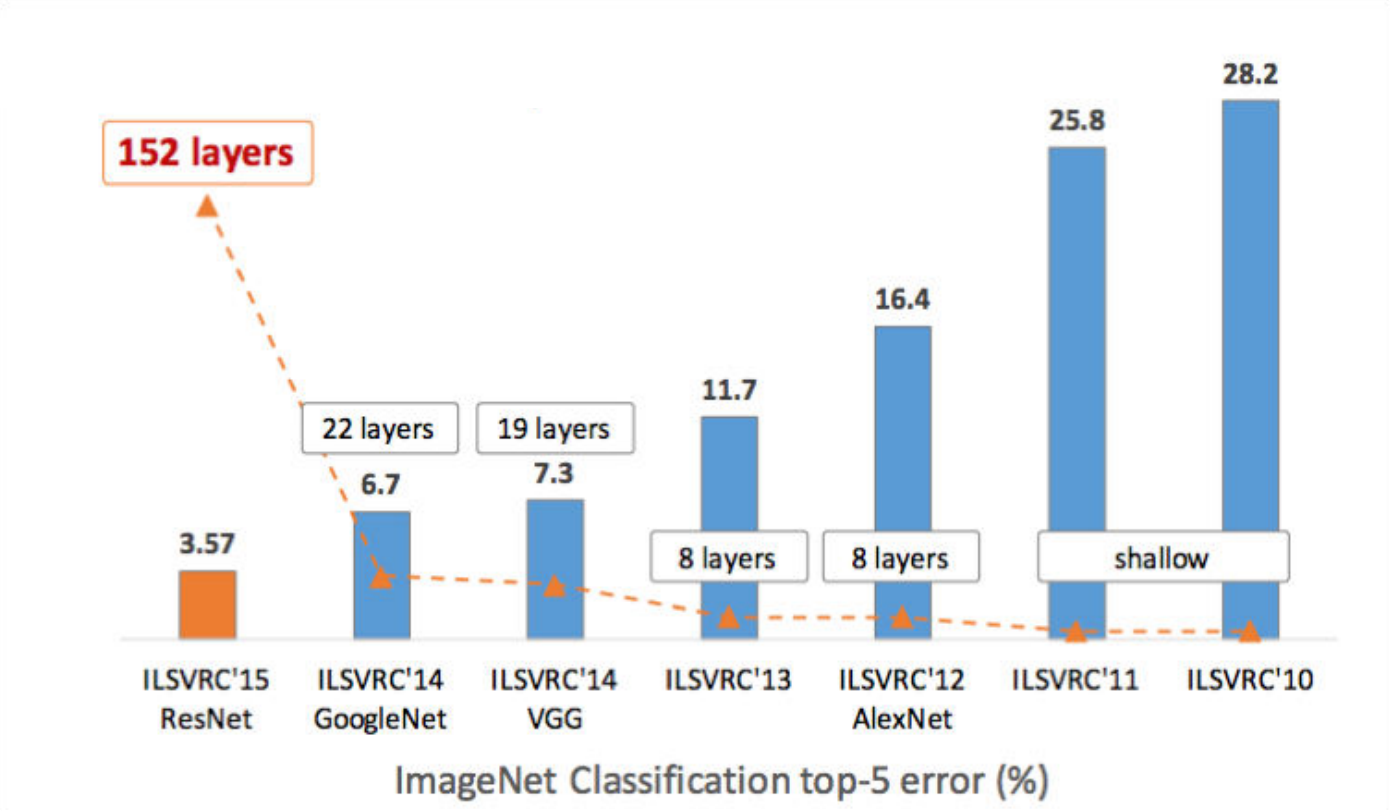


Deep Neural Network

Deep Learning Achievements

Object Recognition

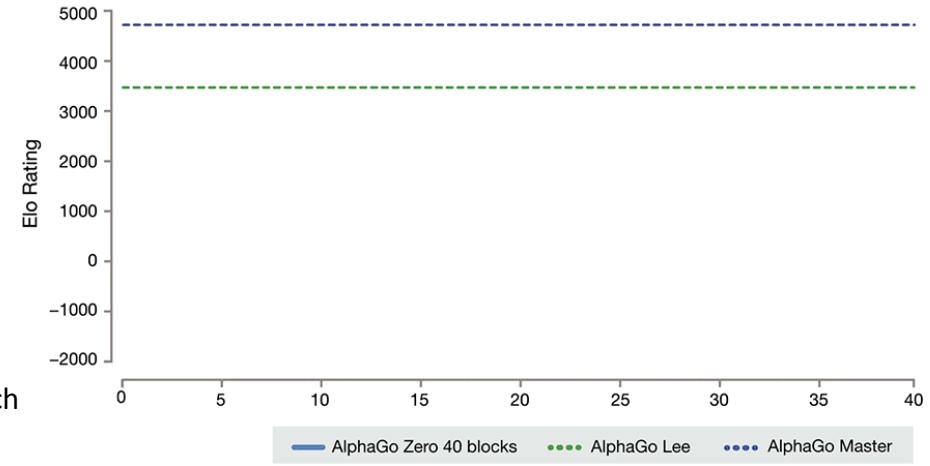
ImageNet Large Scale Visual Recognition Competition (ILSVRC)



Deep Learning Achievements

Playing Games

AlphaGo Zero



AlphaGo Zero: Learning from scratch by DeepMind

Deep Learning Achievements

Natural Conversation

Google I/O Restaurant Booking Demo

Deep Learning Frameworks

Framework	Distributed Execution	Architecture Optimisations	Visualisations	Community Support	Portability
TensorFlow	XX	XX	XX	XX	XX
PyTorch	XX	XX	XX	XX	XX
CNTK	XX	XX	X	–	XX
MXNet	X	XX	X	–	XX
Torch	–	XX	X	X	X
Caffe2	XX	XX	–	–	XX
Caffe	–	XX	X	X	X
Theano	–	XX	X	X	X

+ Support for Production

- Deploy as microservice
- Compress (quantize) your model
- Model versioning
- Serve multiple models
- Mobile/embedded support
- Lots of pre-trained models
- Integrated with Google Cloud ML

Choosing a Machine Learning Framework in 2018 by Abdelrahman Ahmed

About TensorFlow

- Open source library for numerical computation from Google Brain Team
- Version 1.8 recently released
- Written in C++ and CUDA
- Primarily for deep learning research
- Data flow graph-based
- Nodes represent operations
- Tensors flow through the nodes
- Leverages CPUs/Nvidia GPUs/Google TPUs on cloud
- Scalable across computer clusters
- Ecosystem for graph visualisations and serving production models, etc.

Youtube: TensorFlow 101 (Really Awesome Intro Into TensorFlow) by Brian Sletten in Devoxx Poland 2017

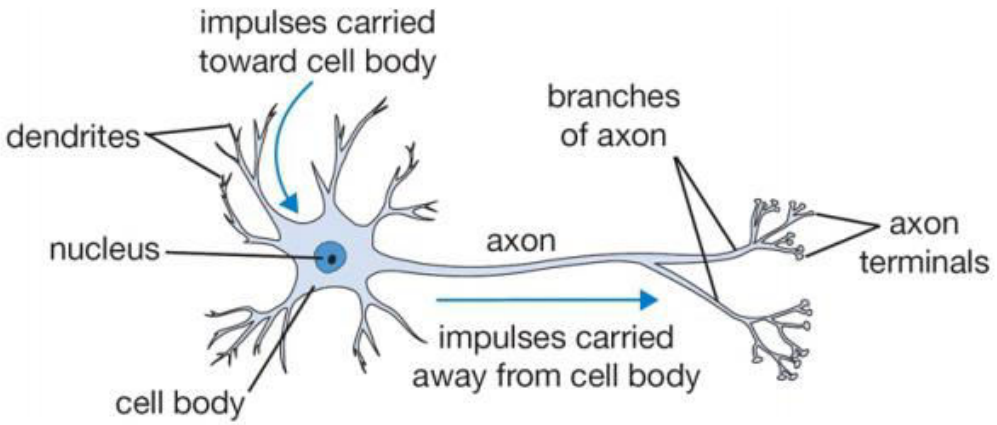
TensorFlow Installation

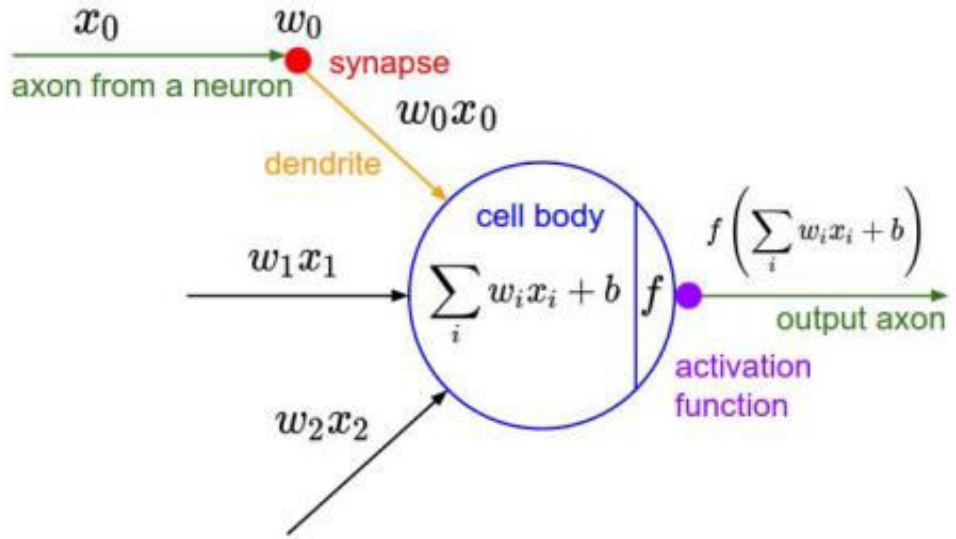
- Instructions: https://www.tensorflow.org/install/
- Nvidia GPUs need latest graphics driver, CUDA Toolkit and cuDNN
- Docker: nvidia-docker (see GitHub repo) for GPUs
- Options:
  - Python (for training/deployment):
    - Ubuntu, MacOS, Windows, From Source
    - CPU support only/with GPU support
    - Virtualenv, pip, Docker, Anaconda
  - Javascript (TensorFlow.js) (for training/deployment):
    - Script tag, NPM
    - CPU/GPU via WebGL
    - Web broswer, Node.js
  - Java (for deployment):
    - Ubuntu, macOS, Windows
    - Maven
  - Go (for deployment):
    - Linux, macOS
  - C/C++ (for deployment):
    - Linux, macOS

TensorFlow Basics

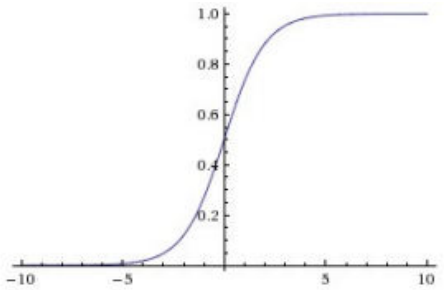
- What is a tensor?
- What is a computational graph?
- How to use TensorBoard?
- How to train a linear model?

Single Neuron



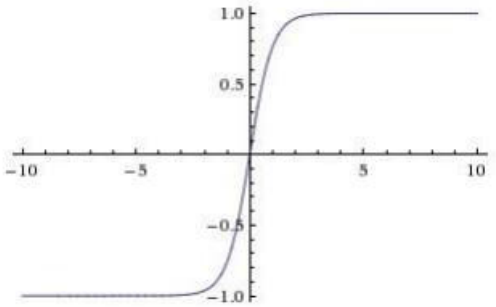


Activation functions

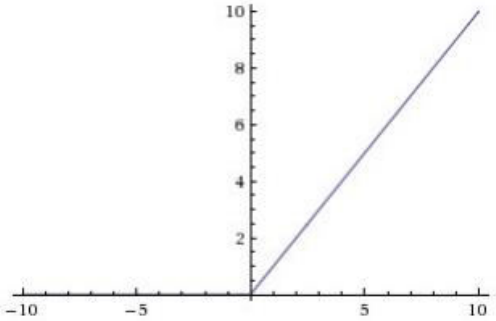


Sigmoid

$$\frac{1}{1+e^{-x}}$$

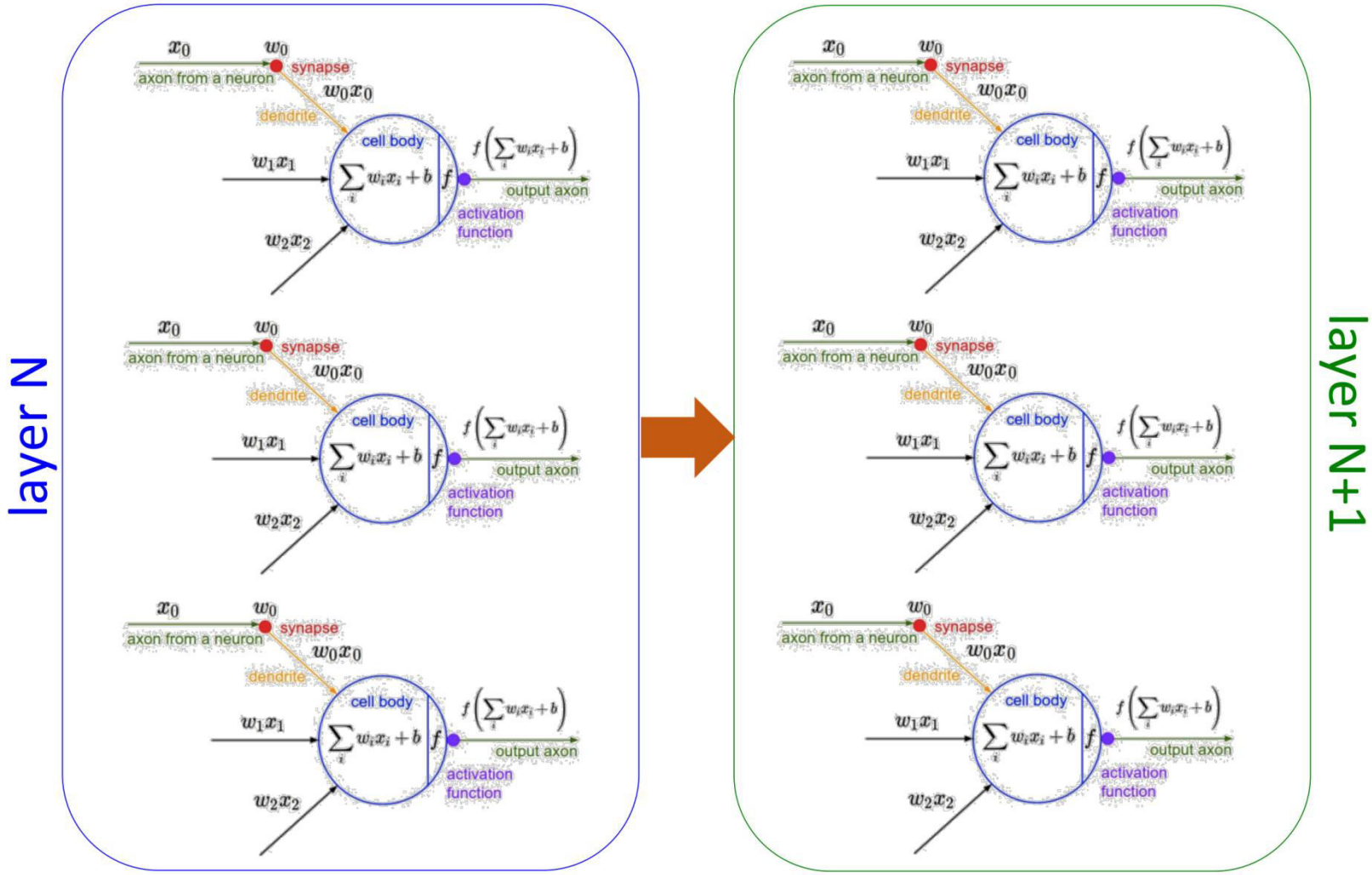


Hyperbolic Tangent
tanh(x)



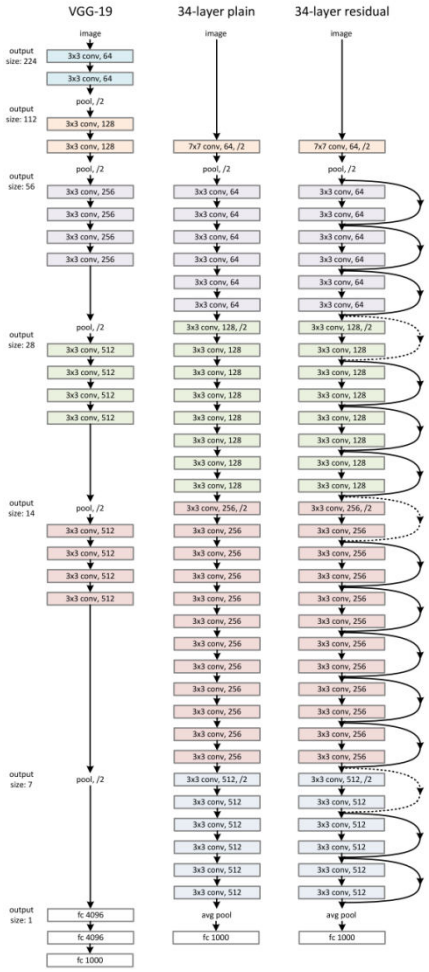
ReLU
max(0,x)

From Neuron to Neural Network

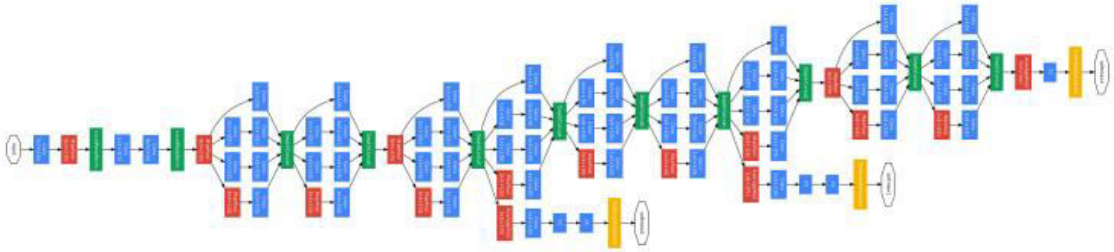


Stack Layers Like Lego Blocks

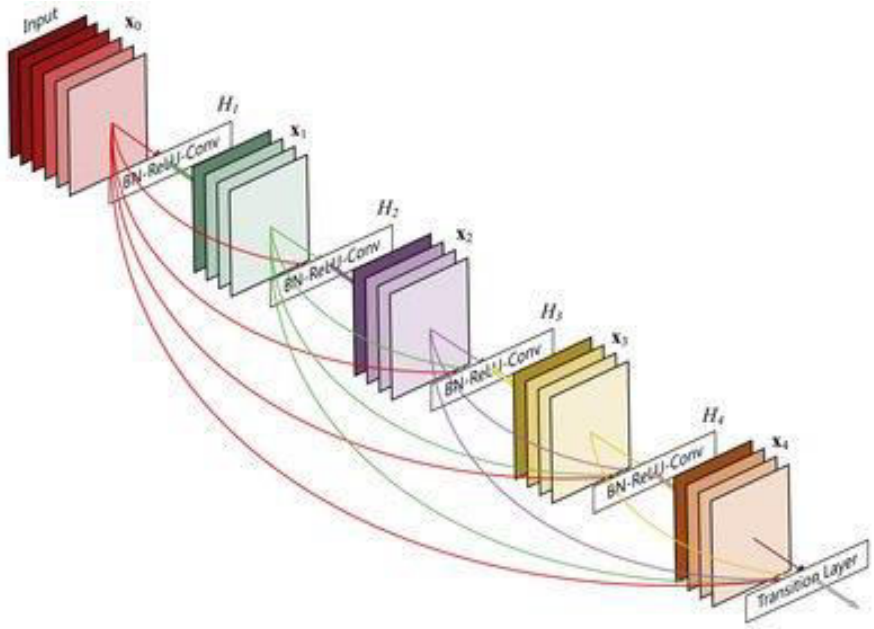
VGGNet/ResNet



GoogLeNet

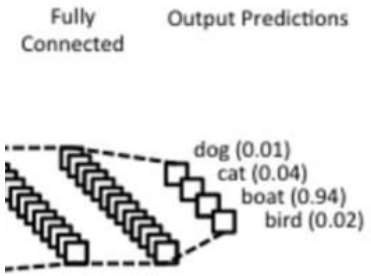


DenseNet



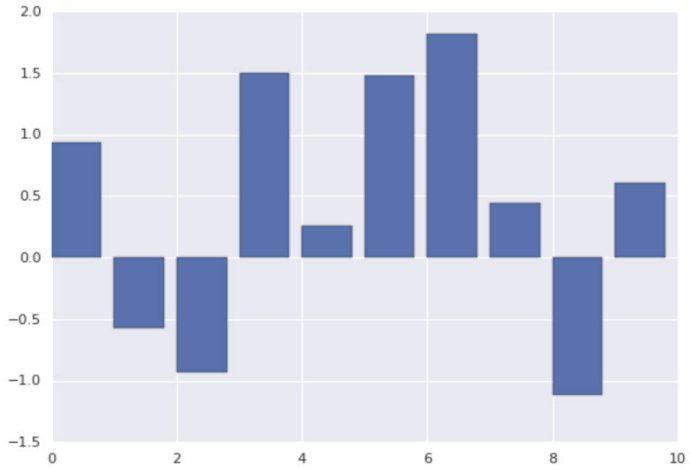
Output Predictions Using Softmax

- Softmax is a function that converts an array of values into an array of probabilities (0 - 1)

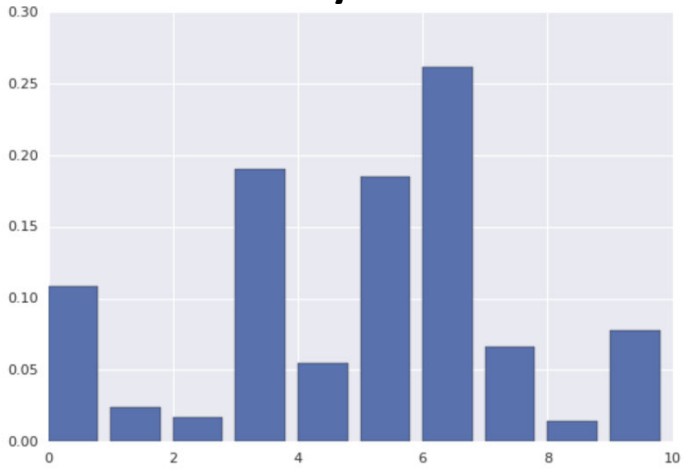


$$softmax(n) = \frac{\exp n_i}{\sum \exp n_i}$$

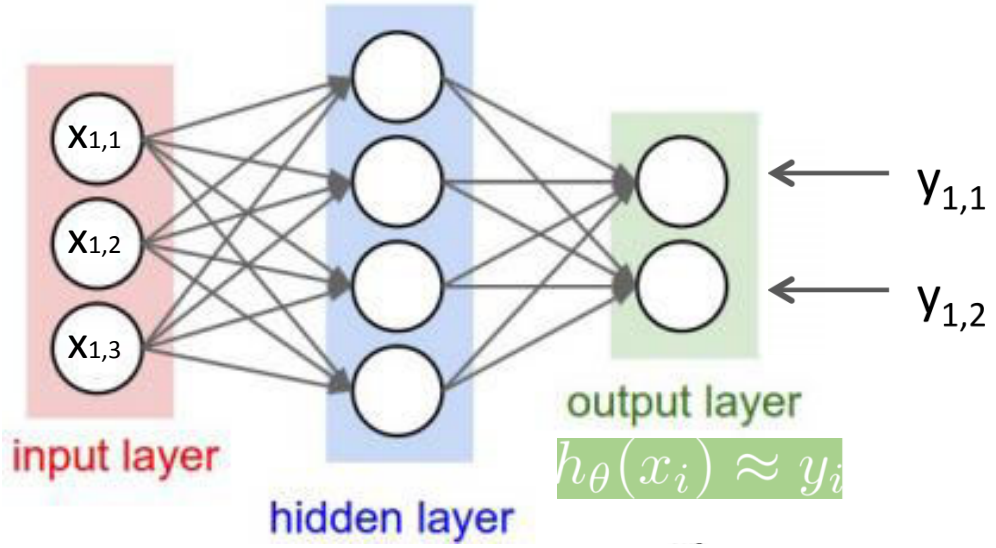
Logits



Probability Distribution



Training a Deep Neural Network



Given training set (x_1, y_1), (x_2, y_2), (x_3, y_3), ...

Adjust parameters θ (for every node) to make:

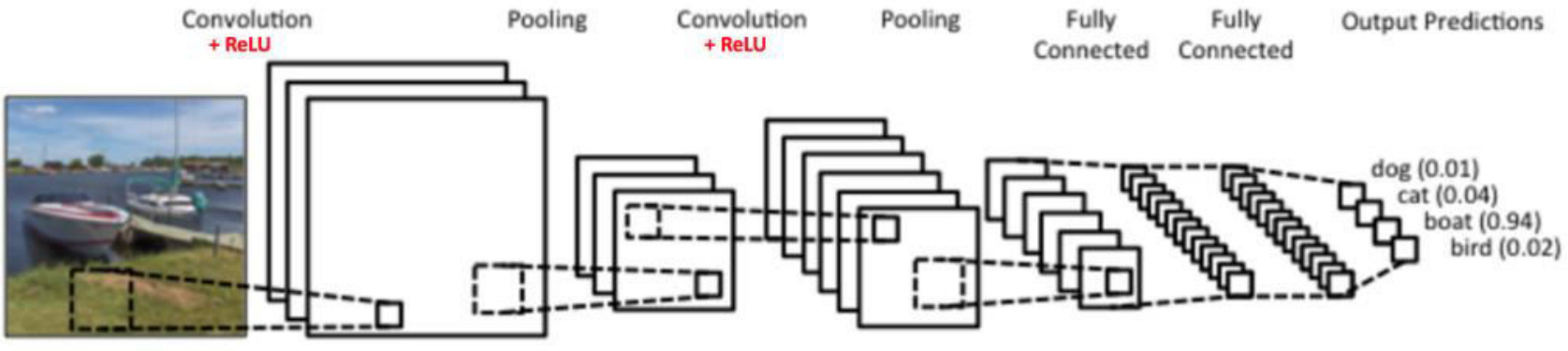
$$h_\theta(x_i) \approx y_i$$

(Use gradient descent - "Backpropagation" algorithm)

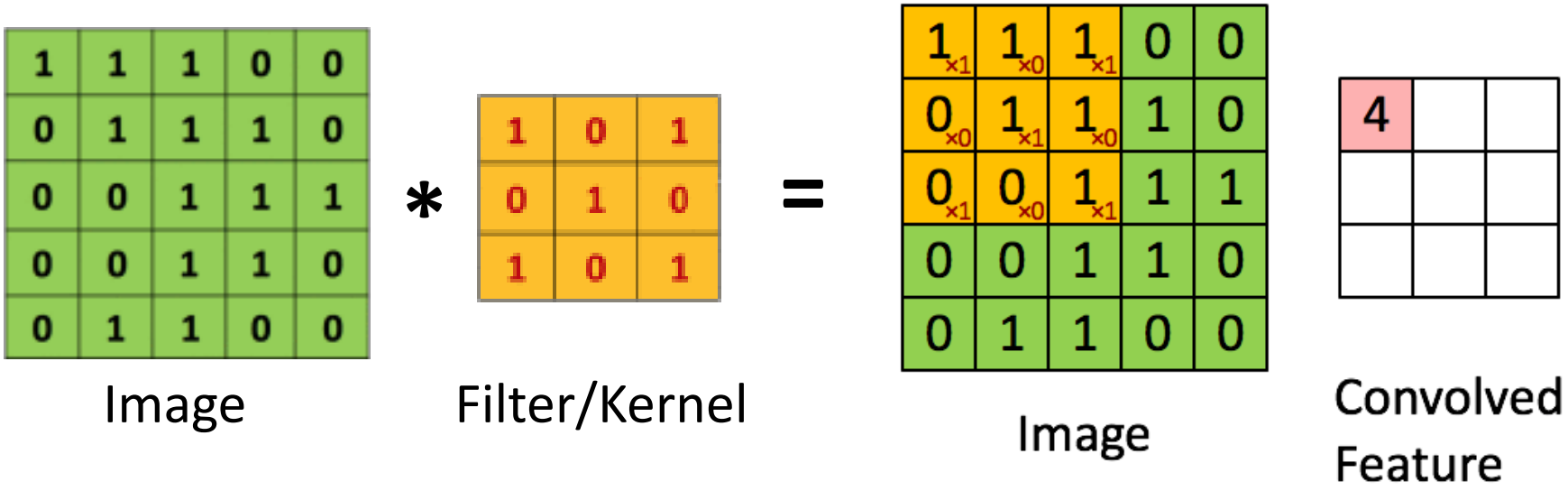
Deep Neural Network Demo

- A Neural Network Playground (https://playground.tensorflow.org/)
- Use TensorFlow to train a deep neural network and TensorBoard for monitoring

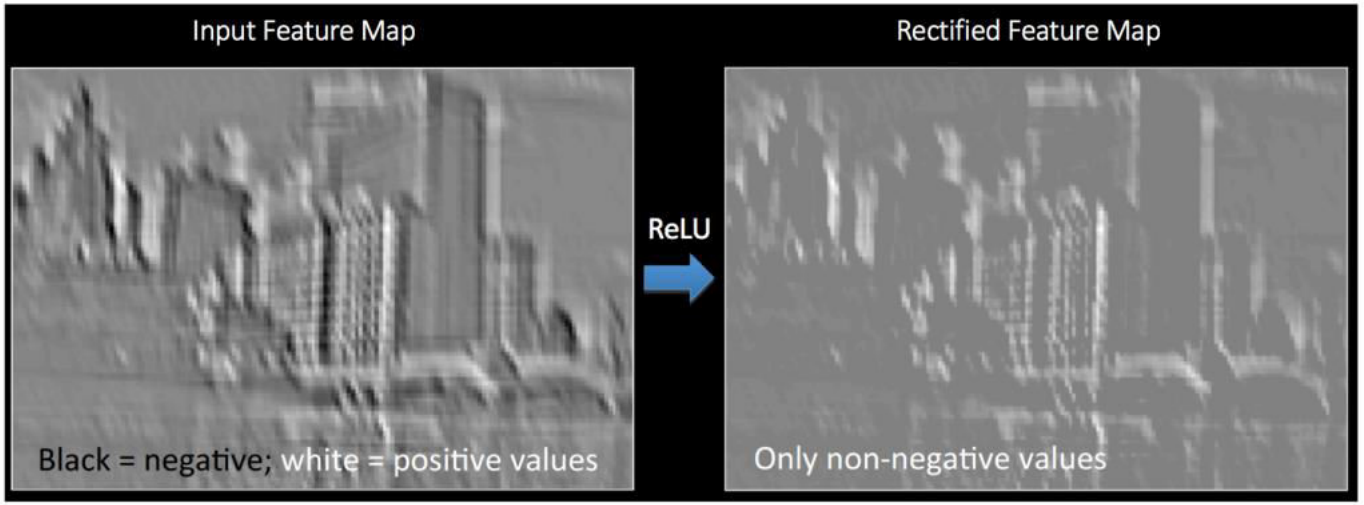
Convolutional Neural Network



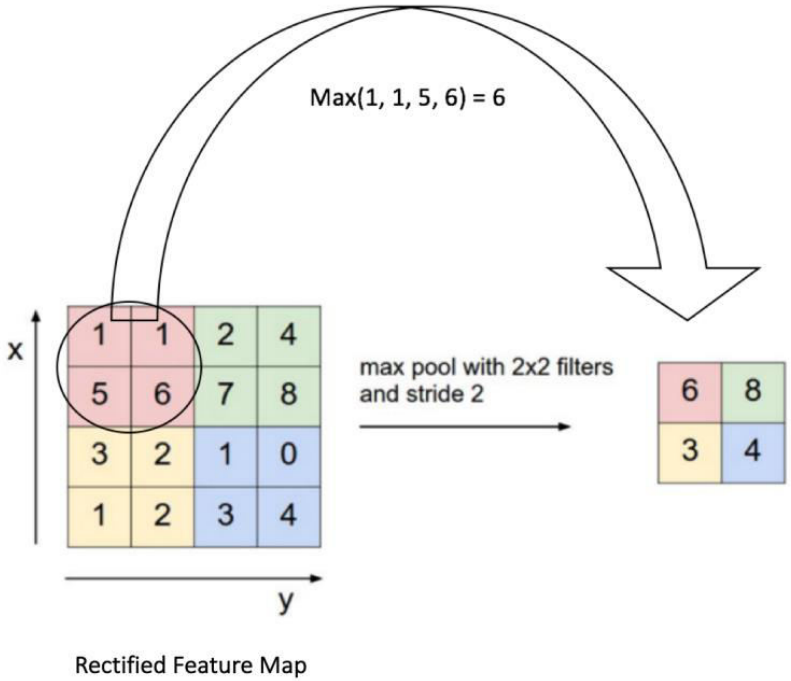
Convolution

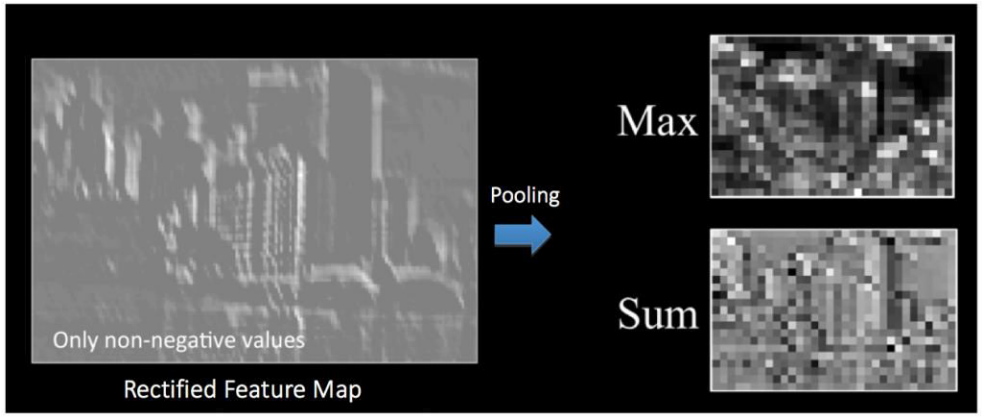


Rectified Linear Unit (ReLU)

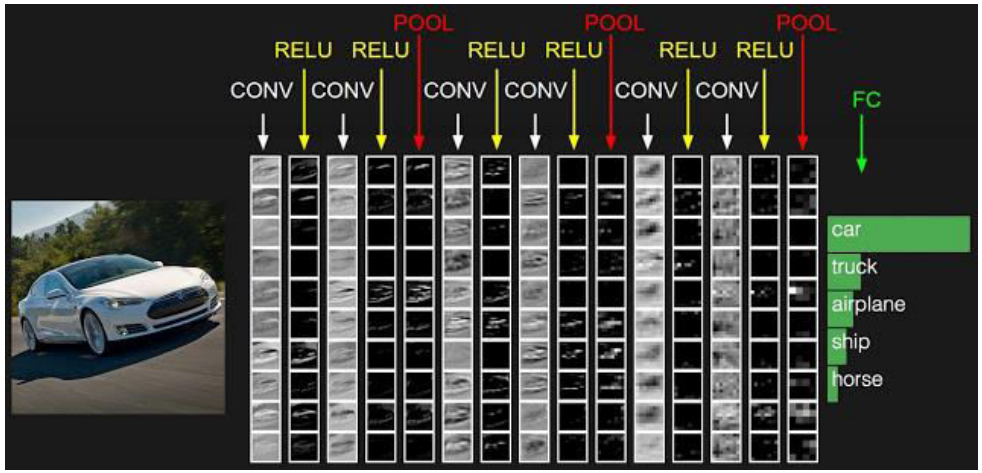


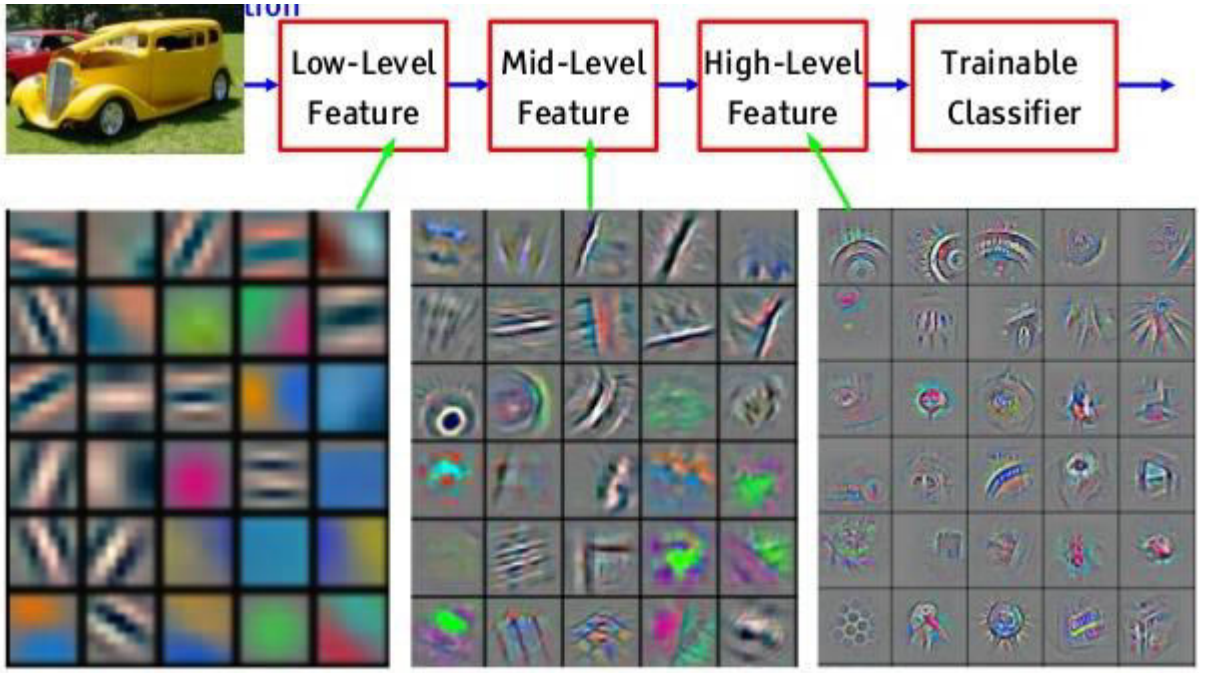
Pooling


Max(1, 1, 5, 6) = 6
x
y
1 1 2 4
5 6 7 8
3 2 1 0
1 2 3 4
max pool with 2x2 filters
and stride 2
6 8
3 4
Rectified Feature Map


Only non-negative values
Rectified Feature Map
Pooling
Max
Sum

Features From ConvNets



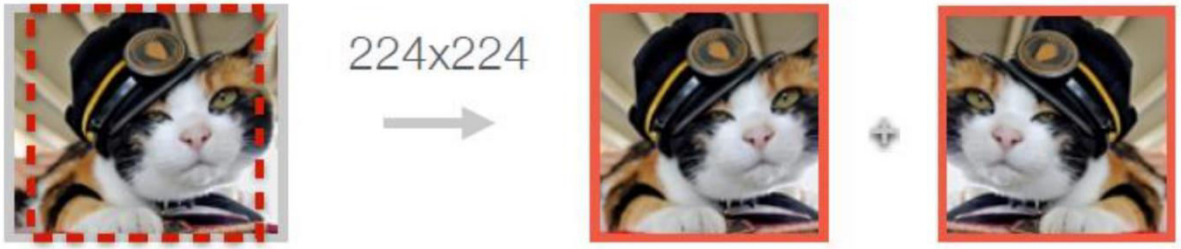


Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

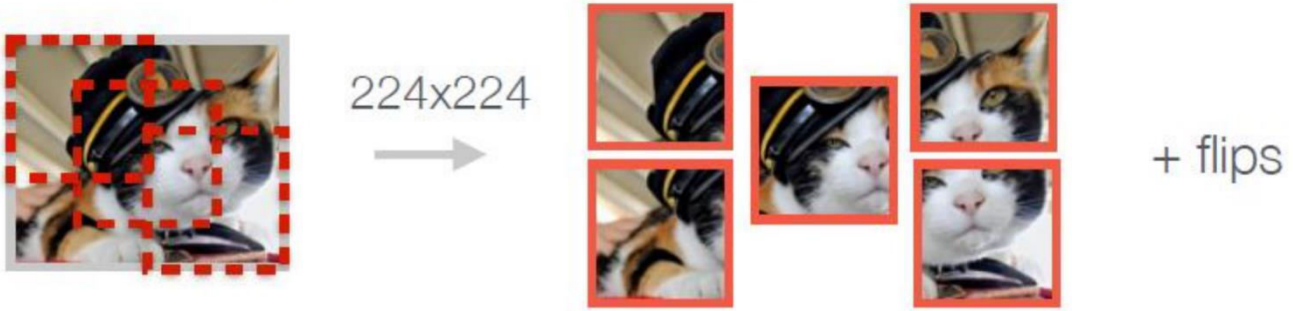
Data Augmentation

a. No augmentation (= 1 image)

b. Flip augmentation (= 2 images)



c. Crop+Flip augmentation (= 10 images)



Convolutional Neural Network Demo

- Use Keras to train a ConvNet/CNN 

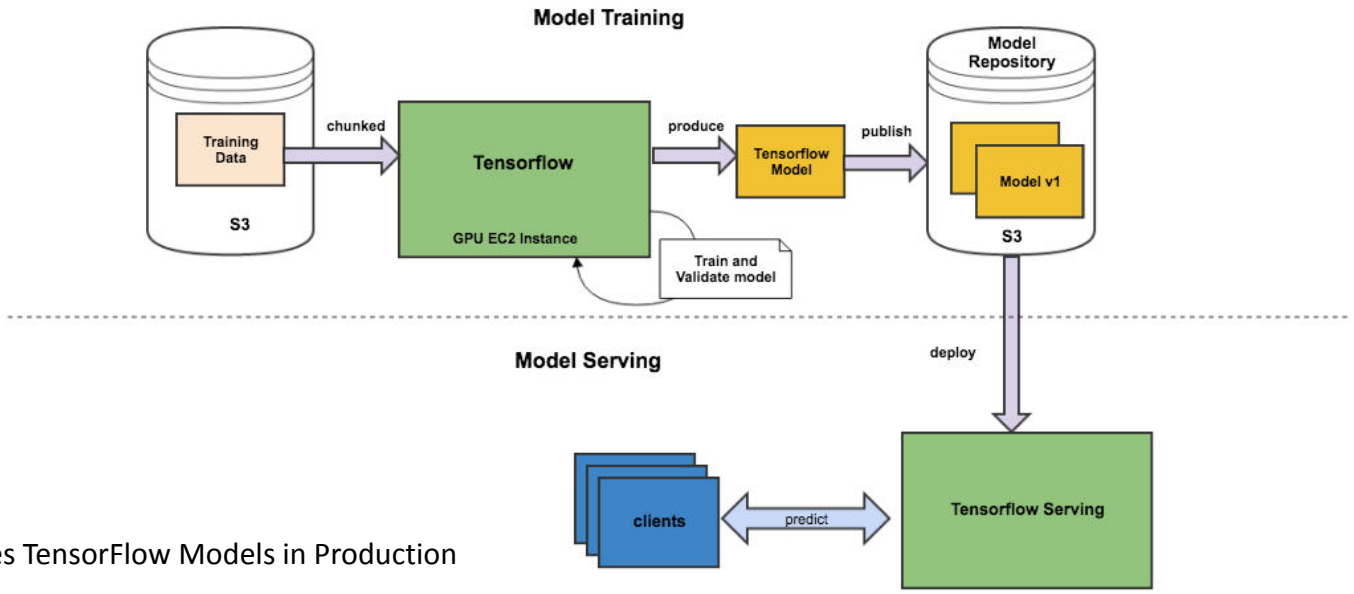
Not covered: Implementation using tf.data for managing data, tf.layers for creating the graph, tf.estimator or eager execution for training and evaluation

Deployment

- Load the saved model in Python, C++, Java, Go or Javascript applications
- Load the saved model using TensorFlow Mobile for Android apps and iOS apps
- **Upcoming:** Convert the saved model into a TensorFlow Lite model for Android apps, iOS apps (with CoreML support), or Raspberry Pi

Deployment

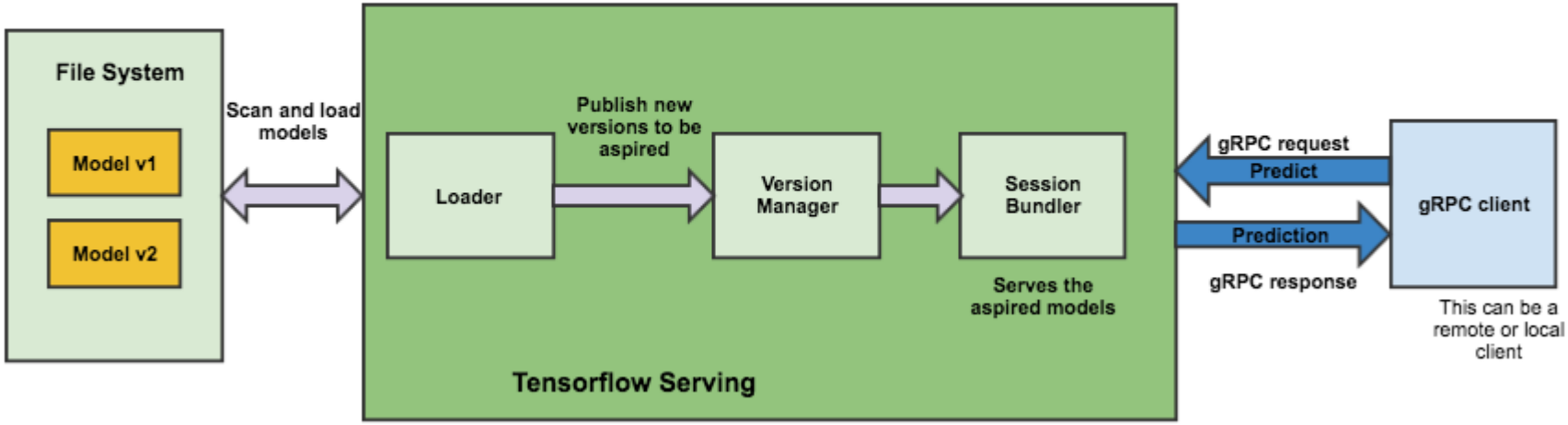
- TensorFlow Serving
  - Training to deployment workflow
    - Publish your saved model through the SavedModelBuilder module
      - ➢ Outputs:
        1. saved_model.pb (Protocol Buffer stores metadata: Signature, Version Number)
        2. variables



How Zendesk Serves TensorFlow Models in Production by Wai Chee Yau

Deployment

- TensorFlow Serving
  - Serve multiple models simultaneously
  - Serve multiple versions of the same model
  - Uses gRPC (Google Remote Procedure Call) to communicate with client

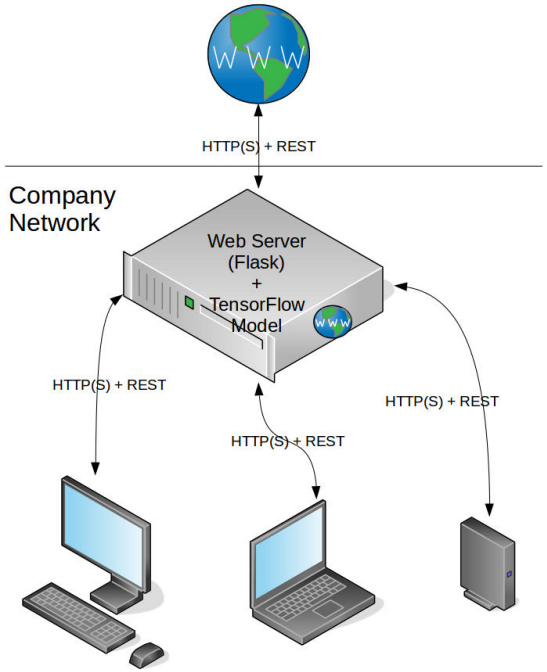


How Zendesk Serves TensorFlow Models in Production
by Wai Chee Yau

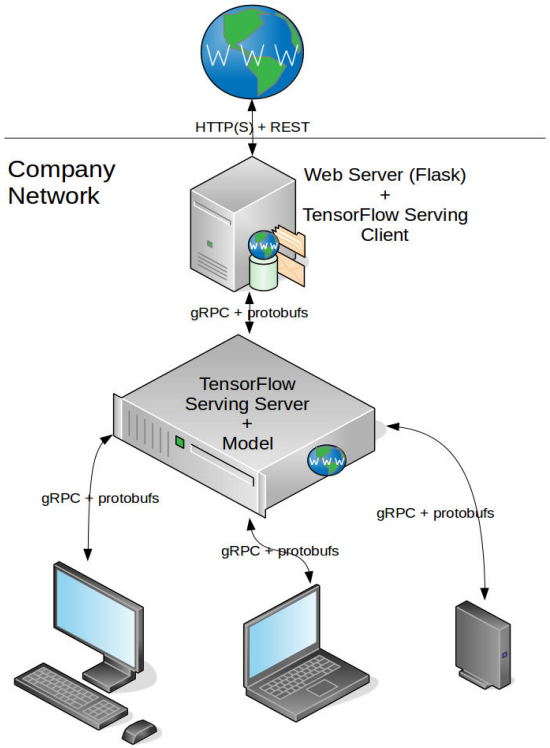
Deployment

- TensorFlow Serving
  - Serving with REST API adapter

Without TensorFlow Serving

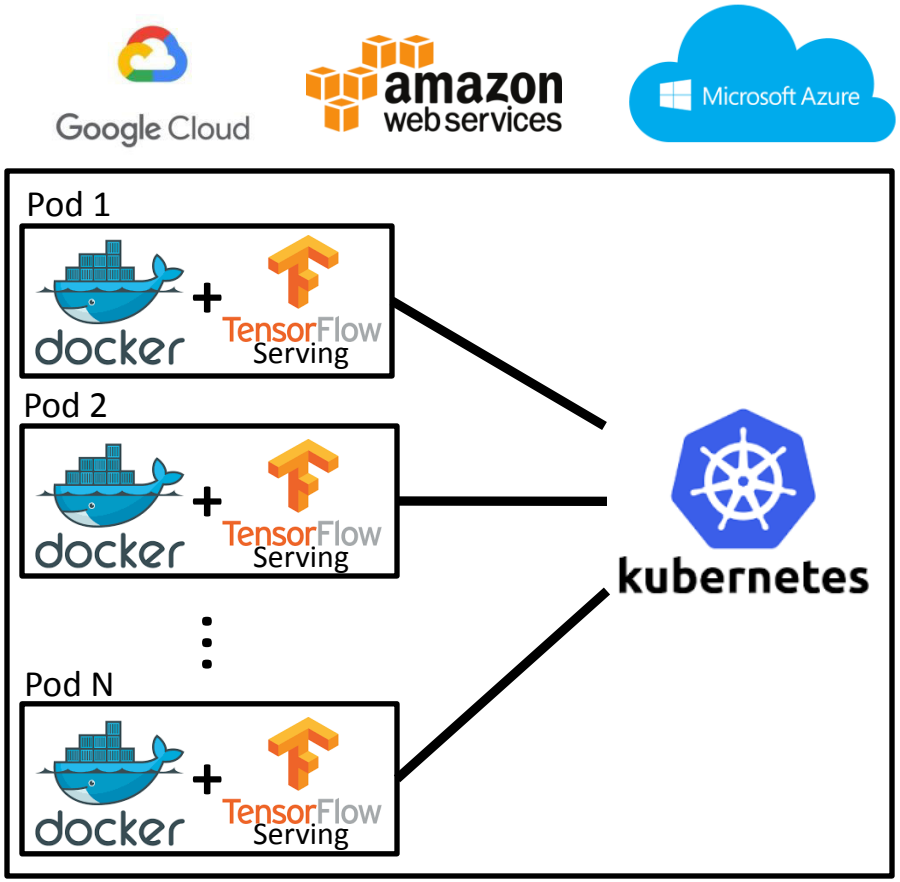


With TensorFlow Serving



Deployment

- On Docker and Kubernetes as a cloud service




```
python client.py
--server=<Service IP>:<Service Port>
--image=./image.jpg
--model_spec='obj_classifier'
--signature_name='model_signature'
--version='2'
```

Proto Request

Proto Response

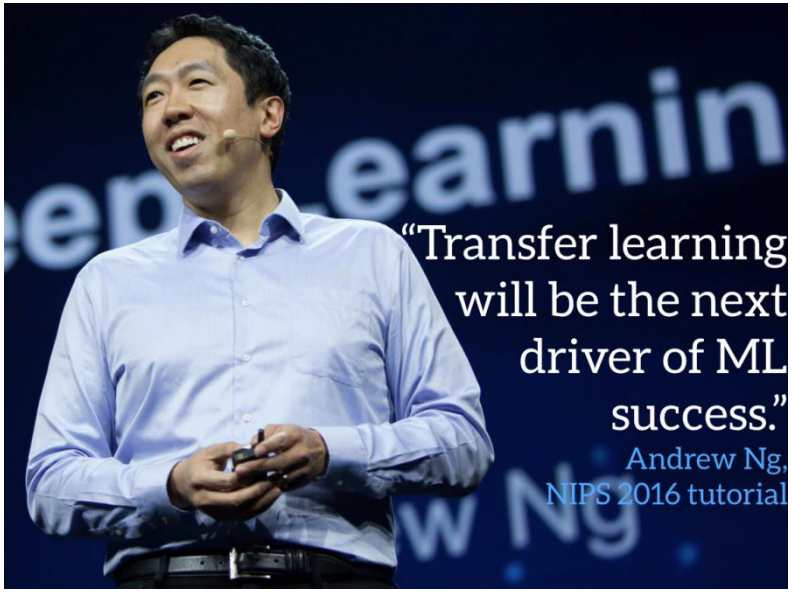
gRPC Client

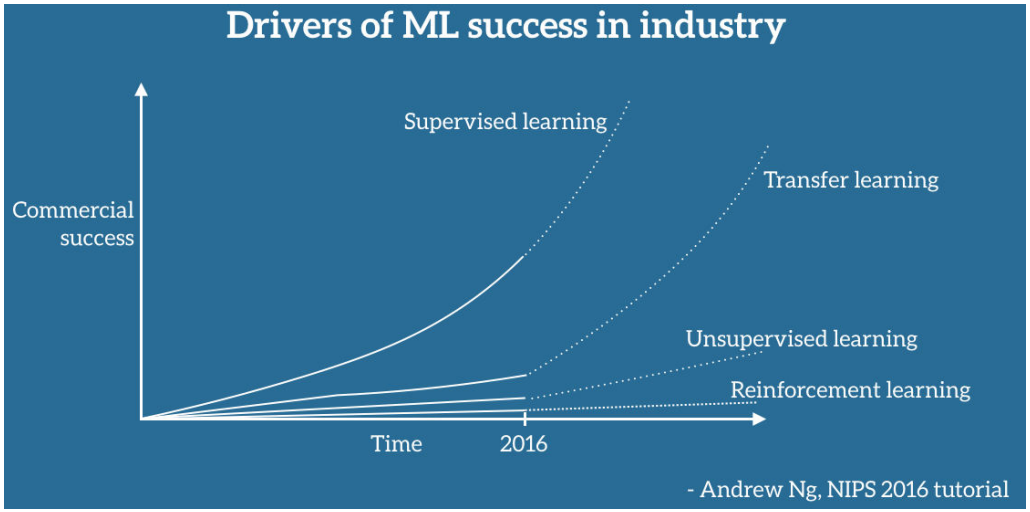
```
outputs {
  key: "scores"
  value {
    dtype: DT_FLOAT
    tensor_shape {
      dim {
        size: 1
      }
      dim {
        size: 10
      }
    }
    float_val: 8.630897802584857e-17
    float_val: 1.219293777054986e-09
    float_val: 6.613714575998131e-10
    float_val: 1.5203355241411032e-09
    float_val: 0.9999998807907104
    float_val: 9.070973139291283e-12
    float_val: 1.5690838628401593e-09
    float_val: 9.12262028080068e-17
    float_val: 1.0587883991775016e-07
    float_val: 1.0302327879685436e-08
  }
}
```

How to deploy Machine Learning models with TensorFlow
by Vitaly Bezgachev

Current Trends

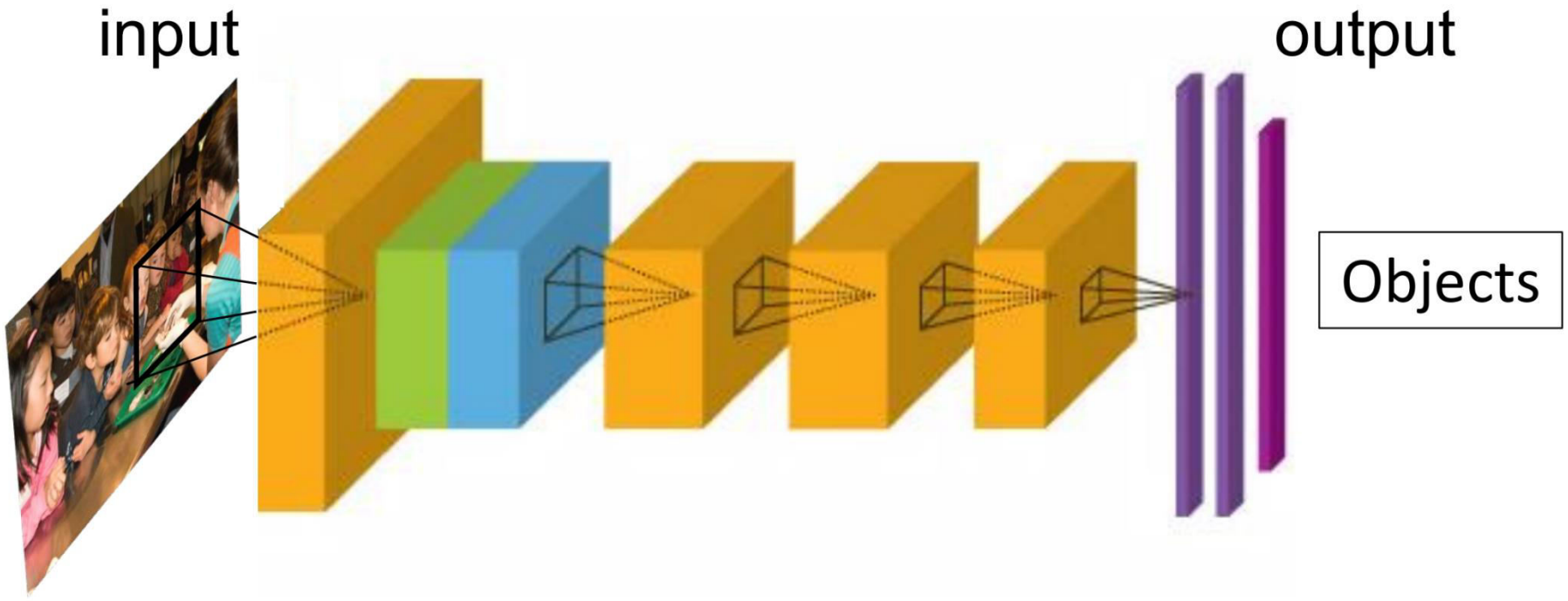
- Transfer Learning





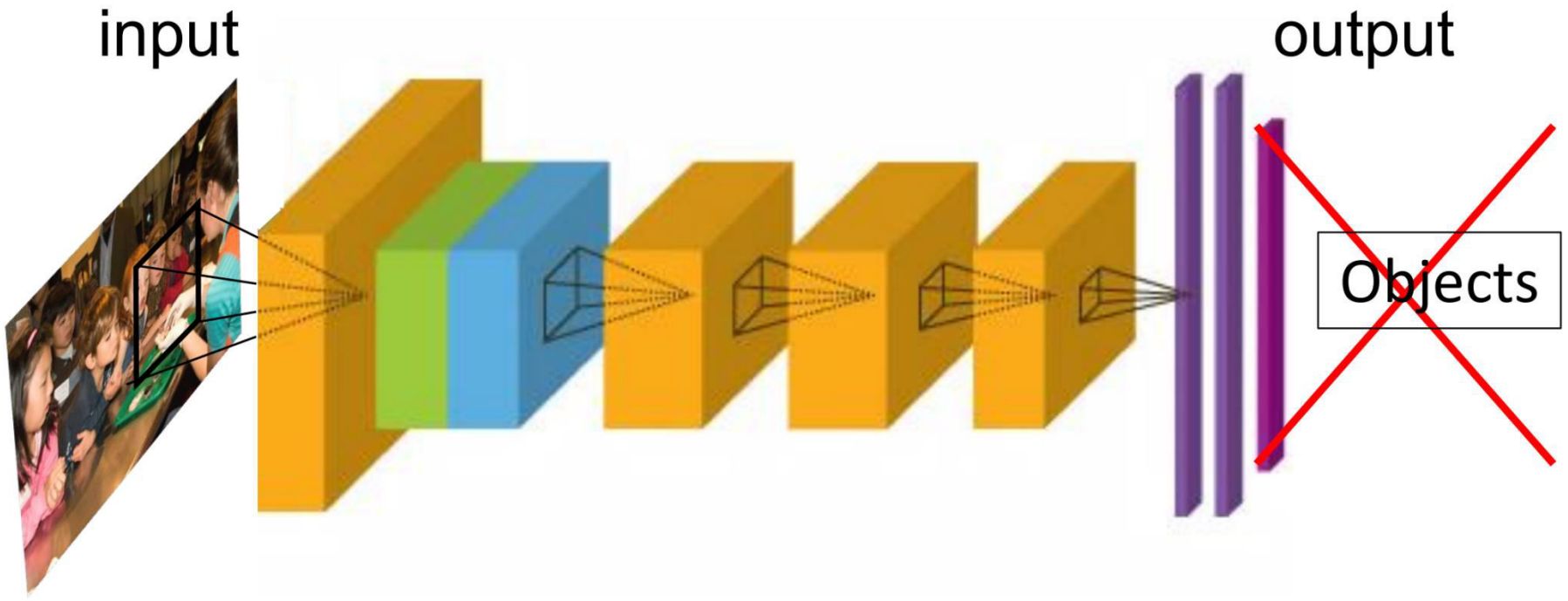
Current Trends

- Transfer Learning



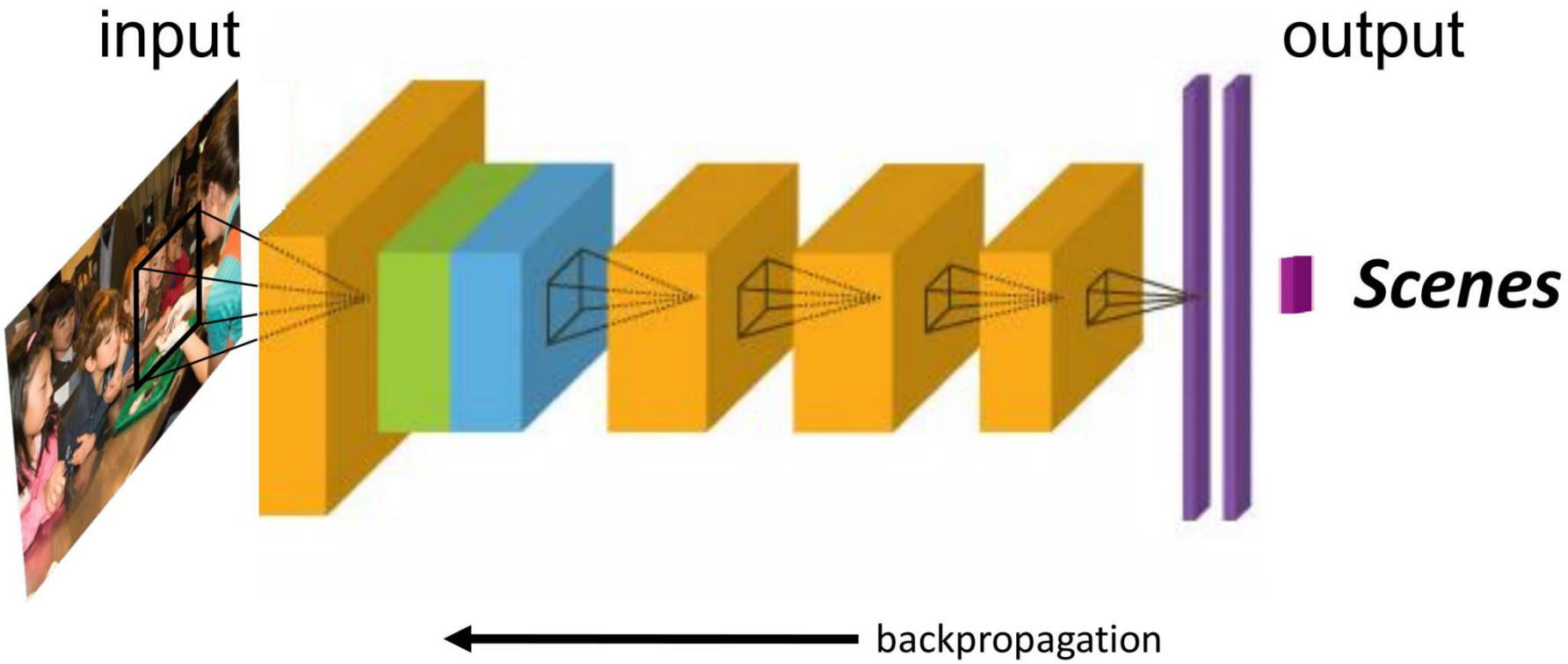
Current Trends

- Transfer Learning



Current Trends

- Transfer Learning

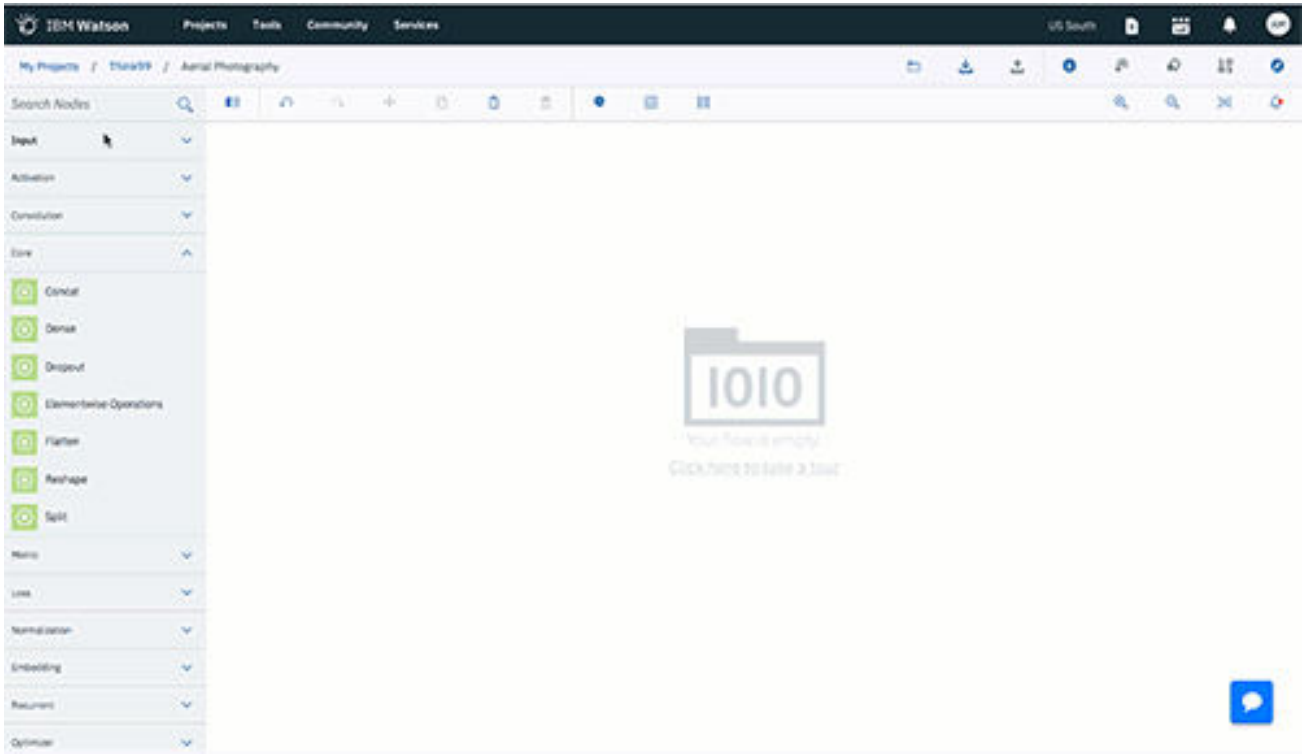


Transfer Learning Demo

- Use Keras to deploy a pretrained image recognition model trained on ImageNet data
- Use Keras to fine-tune a pretrained model

Current Trends

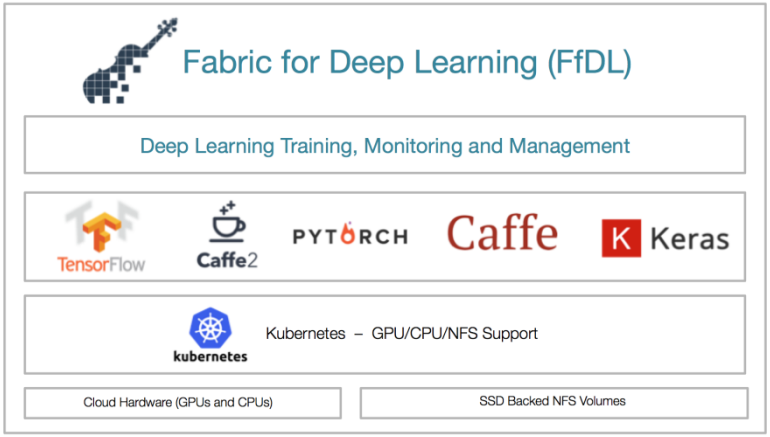
- Cloud Services
  - IBM Watson Studio
    - Deep Learning as a Service (DLaaS)



1. Deep Learning as a Service, IBM makes advanced AI more accessible for users everywhere
2. Deep Learning Advances from IBM Research

Current Trends

- Cloud Services
  - IBM Watson Studio
    - Open-source Fabric for Deep Learning (FfDL)



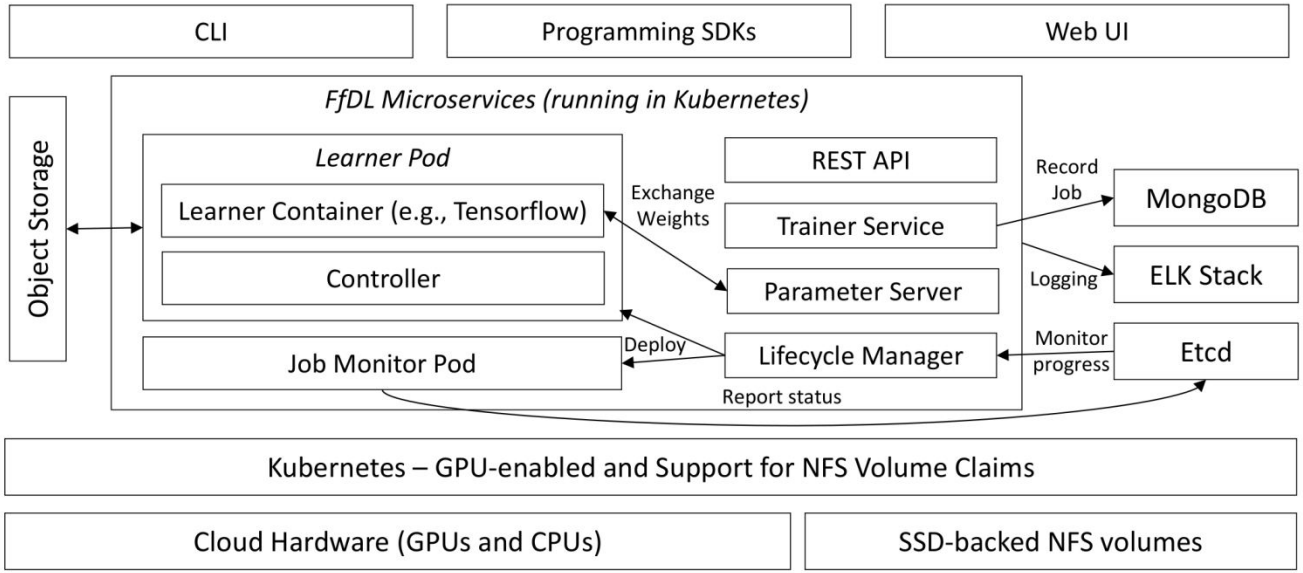


Figure 1: FfDL Architecture

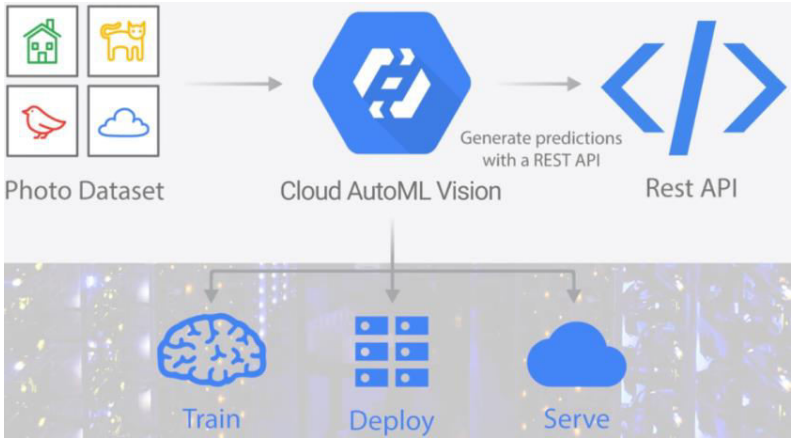
1. Deep Learning Advances from IBM Research
2. Scalable Multi-Framework Multi-Tenant Lifecycle Management of Deep Learning Training Jobs by Boag et al.

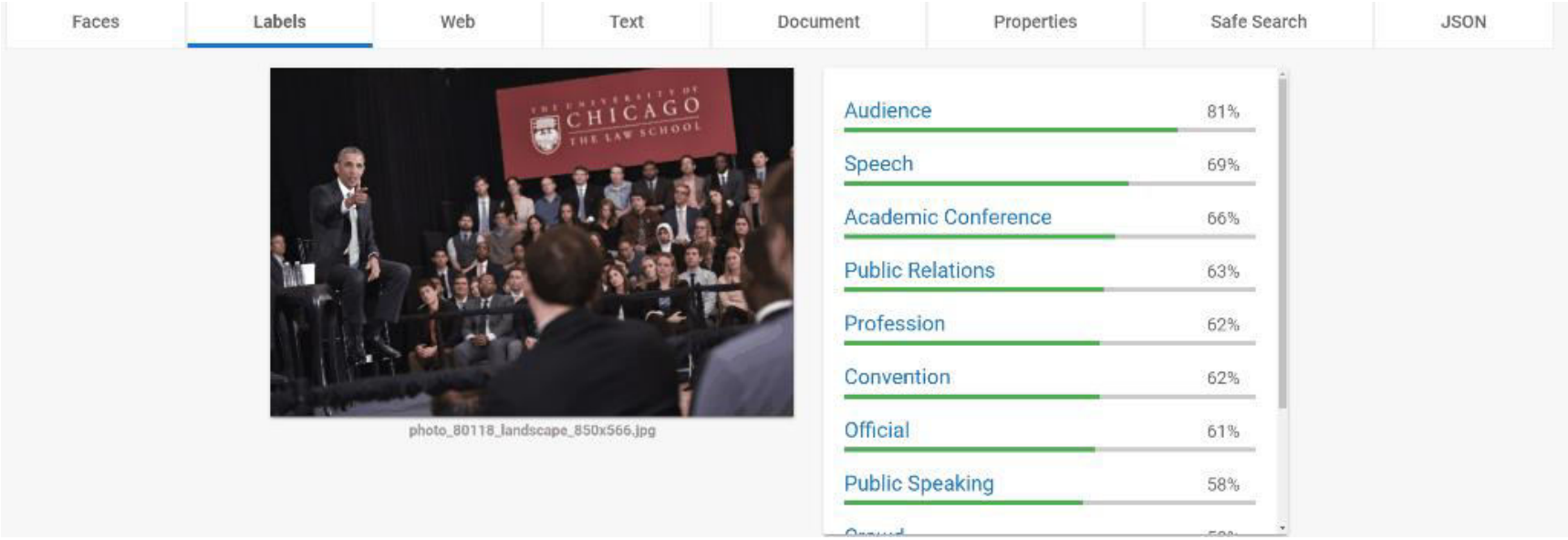
Current Trends

- Cloud Services
  - IBM Watson Studio Product Demo
    - Supermarket video object detection training using active learning to detect trolleys and persons
    - Tracking the time taken for a person to walk from an MRT gantry to the platform level
    - Counting people in a crowd
    - Car detection in a car park and illegal parking on a roadside

Current Trends

- Cloud Services
  - Google Cloud AutoML Vision
  - Google Cloud Vision API





Key Takeaways

- Democratisation of deep learning
- TensorFlow is built for production
- TensorFlow is just a toolkit: focus on getting clean, labelled data and try out transfer learning

Resources

Online courses with TensorFlow code exercises (materials are free):

- Coursera Deep Learning Specialization by Andrew Ng
- Udacity Deep Learning Nanodegree by Sebastian Thrun and Ian Goodfellow

Theory:

- Stanford CS231n Course Notes and Winter 2016 Lectures by Andrej Karpathy (you can watch Spring 2017 for later research developments but Karpathy explains it best)
- Deep Learning by Ian Goodfellow et al. (All the theoretical basics and good review of research until end 2016)
- Deep Learning Papers Reading Roadmap GitHub repo by floodsung

TensorFlow code examples:

- Official TensorFlow Tutorials (includes new features such as Eager Execution and Keras integration)
- TensorFlow Examples GitHub repo by aymericdamien
- Awesome TensorFlow GitHub repo by jtoy
- Deep Learning GitHub repo by Udacity

Inspirational:

- Siraj Rival on YouTube
- Two Minute Papers on YouTube
- Kaggle blog

Meetup Groups:

- TensorFlow and Deep Learning Singapore (Videos at engineers.sg)
- PyTorch and Deep Learning Singapore

Resources

Deployment:

- How to deploy an Object Detection Model with TensorFlow Serving by Gaurav Kaila
- How to deploy Machine Learning models with TensorFlow by Vitaly Bezgachev
- Creating REST API for TensorFlow models by Vitaly Bezgachev
- How Zendesk Serves TensorFlow Models in Production by Wai Chee Yau
- Introductory Guide To TensorFlow Serving by Weimin Wang
- TensorFlow Serving 101 by Stian Lind Petlund
- The Ultimate Guide on Deep Learning for web developers by zero to deep learning

Machine Learning:

- Coursera Machine Learning by Andrew Ng
- Stanford CS229: Machine Learning
- Kaggle.com

Backpropagation Explanation:

- Stanford Winter 2016 CS231n Lecture 4 on YouTube

Free GPU resource for 12 hours at a time to try things out:

- Google Colab