BLUE WATERS

SUSTAINED PETASCALE COMPUTING

2/8/18

Machine Learning on Blue Waters Using TensorFlow with the Image Feature Detection Problem

Or: How I Learned to Stop Worrying And Love AI

Presented By: Dr. Aaron D. Saxton

NCSA NSF GREAT LAKES CONSORTIUM FOR PETASCALE COMPUTATION CRAY

Todays Topics

- Blue Waters overview
- TensorFlow Basics
- Statistics Review
- Neural Networks
- Convolutions
- Convolutional Neural Networks
- ImageNet
- Blue Waters TensorFlow Process
- Distributed Tensor Flow
- TensorBoard

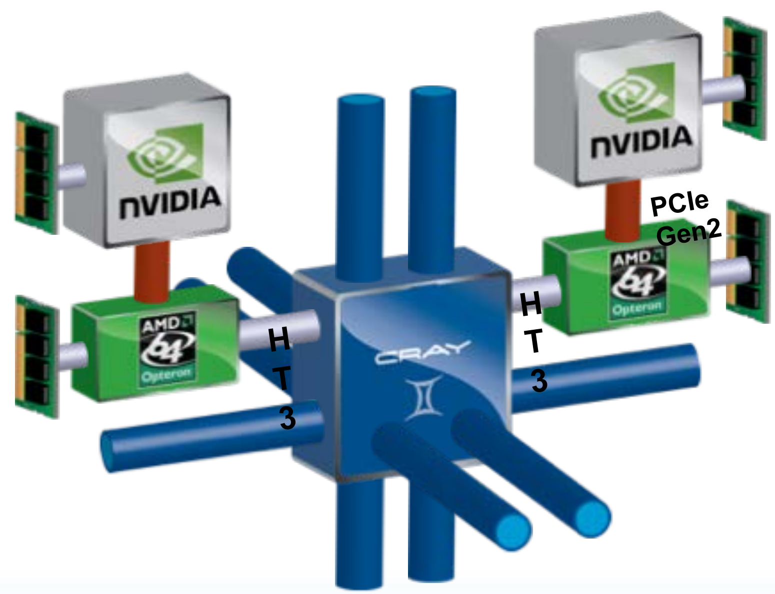
Blue Waters Overview

- Brief Summary
  - AMD Interlagos
  - NVIDIA Tesla
  - 22,636 XE Compute Nodes
  - 4,228 XK Compute Nodes
  - Cray Gemini Interconnect

Blue Waters Overview

2 XE nodes

2 XK nodes



Blue Waters Overview

Scuba Subsystem: Storage Configuration for **User Best Access**

10/40/100 Gb Ethernet Switch

External Servers

IB Switch

>1 TB/sec

100 GB/sec

400+ Gb/sec WAN

Spectra Logic: 200 usable PB

Sonexion: 26 usable PB

TensorFlow Basics

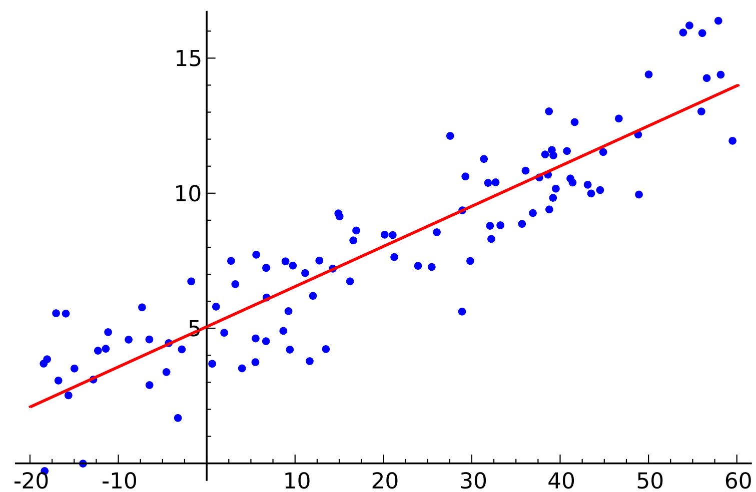
- Python API
  - C++ under the hood
- Mediator Design Pattern
  - Uses python context manager (`with`)
- Workflow
  - Construct operations
  - Assign to name scope and or device
  - Enter Session context
    - run()
- www.tensorflow.org/programmers_guide/low_level_intro

TensorFlow Basics

Demo

Statistics Review

- Simple $y = m \cdot x + b$ regression
  - Least Squares to find m,b
  - With data set $\{(x_i, y_i)\}_{i=1,\ldots,n}$
    - Very special, often hard to measure y_i
  - Let the error be
    - $R = \sum_{i=1}^{n} [(y_i - (m \cdot x_i + b)]^2$
  - Minimize Q with respect to m and b.
    - Simultaneously Solve
      - $R_m(m, b) = 0$
      - $R_b(m, b) = 0$
    - Linear System
- We will consider more general $y = f(x)$
  - $R_m(m, b) = 0$ and $R_b(m, b) = 0$ may not be linear



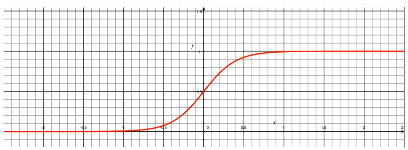
Statistics Review

- Regressions with parameterized sets of functions. e.g.
  - $y = ax^2 + bx + c$ (quadratic)
  - $y = \sum a_i \ x^i$ (polynomial)
  - $y = Ne^{rx}$(exponential)
  - $y = \frac{1}{1+e^{-(a+bx)}}$ (logistic)
- After optimal parameters found,
  - Use function for inference
  - Have x, compute y

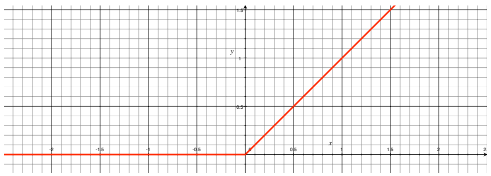
Neural Networks

- Activation functions

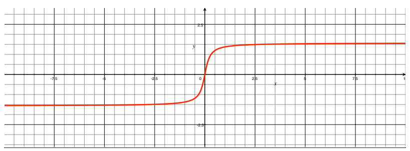
Logistic

$\sigma(x) =$ 

ReLU (Rectified Linear Unit)

$\sigma(x) =$ 

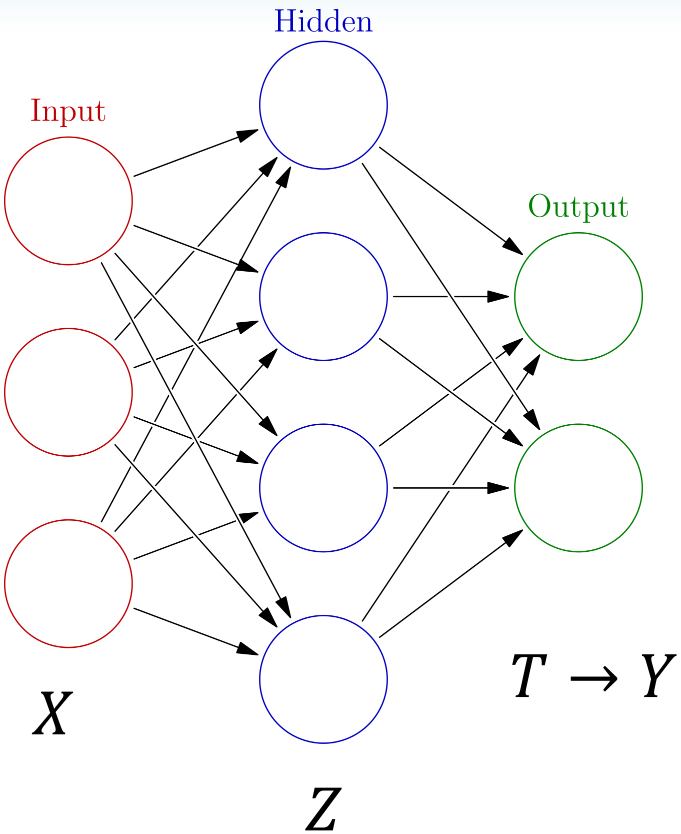
Arctan

$\sigma(x) =$ 

- Softmax
  - $g_k(x_1, x_2, \dots, x_N) = \dfrac{e^{x_k}}{\sum e^{x_i}}$

Neural Networks

- Parameterized function
  - $Z_M = \sigma(\alpha_{0m} + \alpha_m X)$
  - $T_K = \beta_{0k} + \beta_k Z$
  - $f_K(X) = g_k(T)$
- $\beta_{0i}, \beta_i, \alpha_{0m}, \alpha_m$
  - Weights to be optimized

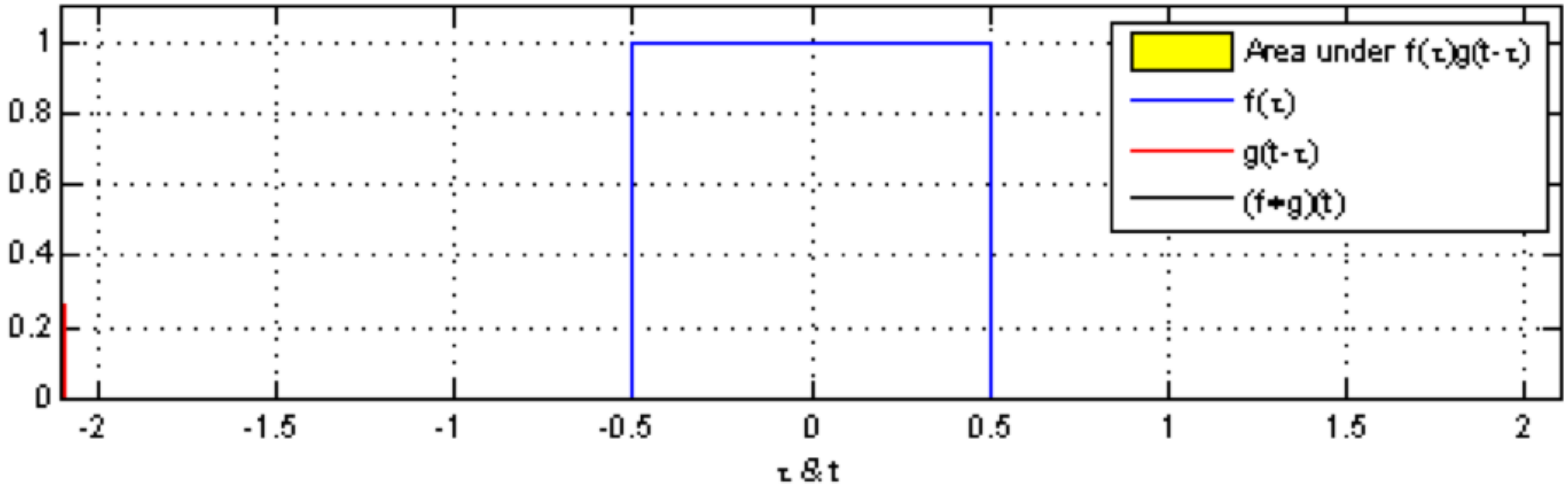


Neural Networks

- Finding Weights $\beta_{0i}, \beta_i, \alpha_{0m}, \alpha_m$
  - Back propagation
  - Nothing more than chain rule
  - Take partial derivative of error function R
  - This text is a good reference for nitty gritty details
    - The Elements of Statistical Learning, Second Eddition, by Trevor Hastie, Robert Tibshirani, Jerome Friedman
  - Back propagation give errors (or loss)
  - Gradient Decent tells you how to update weights

Convolutions

- For two functions, $f(x), g(x)$
  - $(f * g)(x) = \int_{-\infty}^{\infty} f(y)g(x-y)\,dy$

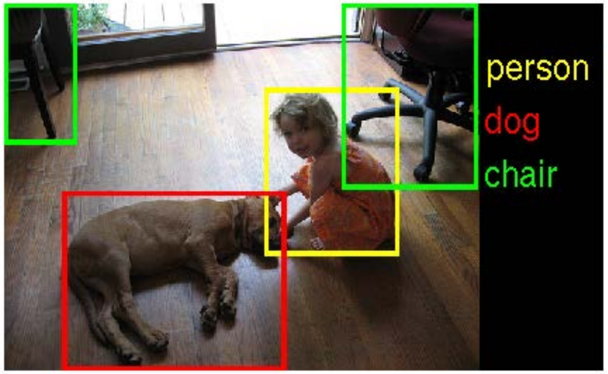


  - g is the kernel to f
  - Above is a rolling average
  - http://setosa.io/ev/image-kernels/

Convolutional Neural Networks

- https://adeshpande3.github.io/adeshpande3.github.io/The-9-Deep-Learning-Papers-You-Need-To-Know-About.html
- Highlights
  - AlexNet
  - VGG Net
  - GoogLeNet (Inception)
  - Microsoft ResNet

ImageNet



- www.image-net.org
- Large High Quality Dataset
  - 14,197,122 Images
  - 21841 synsets
- Runs the Large Scale Visual Recognition Challenge (ILSVRC)
- Annotated
  - Bounding Boxes
  - synset
    - WordNet (http://wordnet.princeton.edu)

ImageNet

- Blue Waters hosts copy of ImageNet
- Legal Term of Access
  - Create account on www.image-net.org
  - Navigate to Term of Access
  - Accept Term of Access
  - Take screen shot or print to PDF Term of Access with your name on it.
  - Email to saxton@illinois.edu
- After I receive your Term of Access I will give your Blue Waters user read permission to data

Term Of Access

You have been granted access for **non-commercial research/educational** use. By accessing the data, you have agreed to the following terms. Note: Our terms of access have changed. By continuing to download and/or access ImageNet data you agree to the new terms of access.

Aaron Saxton (the "Researcher") has requested permission to use **the ImageNet database** (the "Database") at **Princeton University and Stanford University**. In exchange for such permission, Researcher hereby agrees to the following terms and conditions:

1. Researcher shall use the Database only for non-commercial research and educational purposes.
2. Princeton University and Stanford University make no representations or warranties regarding the Database, including but not limited to warranties of non-infringement or fitness for a particular purpose.
3. Researcher accepts full responsibility for his or her use of the Database and shall defend and indemnify the ImageNet team, Princeton University, and Stanford University, including their employees, Trustees, officers and agents, against any and all claims arising from Researcher's use of the Database, including but not limited to Researcher's use of any copies of copyrighted images that he or she may create from the Database.
4. Researcher may provide research associates and colleagues with access to the Database provided that they first agree to be bound by these terms and conditions.
5. Princeton University and Stanford University reserve the right to terminate Researcher's access to the Database at any time.
6. If Researcher is employed by a for-profit, commercial entity, Researcher's employer shall also be bound by these terms and conditions, and Researcher hereby represents that he or she is fully authorized to enter into this agreement on behalf of such employer.
7. The law of the State of New Jersey shall apply to all disputes under this agreement.

ImageNet

Demo
(Archive Tour)

Blue Waters Tensorflow Process

- github.com/asaxton/ncsa-bluewaters-tensorflow
- Clone repo
- cd ncsa-bluewaters-tensorflow/datasets/imagenet
- qsub extract_data_from_archive.pbs
  - Wait for completion
- qsub build_imagenet_data.pbs
  - Wait for completion
- cd ncsa-bluewaters-tensorflow/run_scripts
- qsub distributed_tf_launch.pbs
  - Result will be in the directory checkpoint_dir

Blue Waters TensorFlow Process

Demo
(Code Tour)

Distributed TensorFlow

- Resoures
  - www.tensorflow.org/deploy/distributed
  - www.oreilly.com/ideas/distributed-tensorflow

Distributed TensorFlow

Demo
(Code Tour)

TensorBoard

- Grab your checkpoint

Demo