Nikolas Remley

Notes and Using TensorFlow

**TensorFlow Basics**

* TensorFlow is a Machine Learning library developed by Google made specifically for Deep Learning, but still applicable to a wide variety of data analysis
  + Deep learning can be defined as “A class of machine learning techniques that exploit many layers of non-linear information processing for supervised or unsupervised feature extraction and transformation, and for pattern analysis and classification” (Deng, Yu)
* Data is stored and computed in a C++ backend with a Python interface
  + In python, one build a graph (how data moves and is computed) by linking together different objects called tensors
    - Tensors can be in a few different categories, but one should note the difference between placeholders (data the user inputs) and variables (weights/biases, the values being optimized)
  + Once the graph is built the Variables should be initialized with a Session
    - A session is what allows the data to flow through the graph. One can write a session into a script or use an InteractiveSession for on the fly analysis/debugging
    - [https://www.tensorflow.org/versions/r0.8/api\_docs/python/client.html#Session](https://www.tensorflow.org/versions/r0.8/api_docs/python/client.html" \l "Session)
  + Once a Session is activated, one let's the data flow through the graph with sess.run and feed dictionaries
    - The feed dictionaries fill any placeholder tensors with the data of your choosing as long as the shapes are equal.
* This overview is very generalized, for more information visit the TensorFlow website. It is filled with tutorials and information for getting started.
  + Important Note: TensorFlow requires a Linux or Mac OS

**Benefits of TensorFlow**

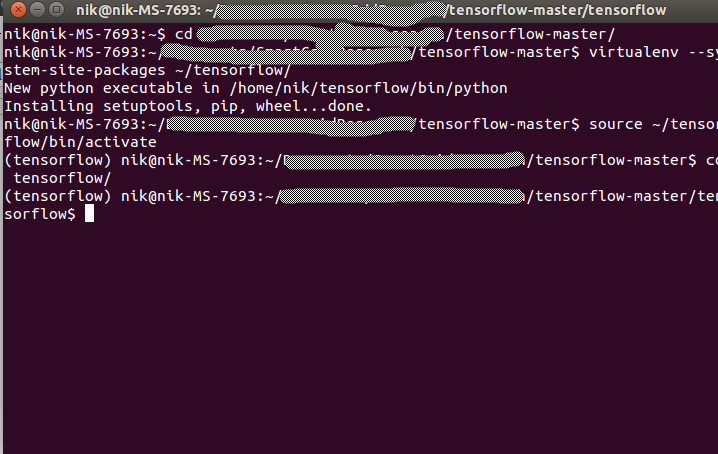
* TensorFlow is very powerful and is used in many applications, particularly Google Applications
  + A lot of TensorFlow use comes in the form of image categorization, which uses very large data sets
* TensorFlow is very fast using a C++ backend
* TensorFlow is very customizable, one can make many different types of Neural Networks and Machine Learning algorithms

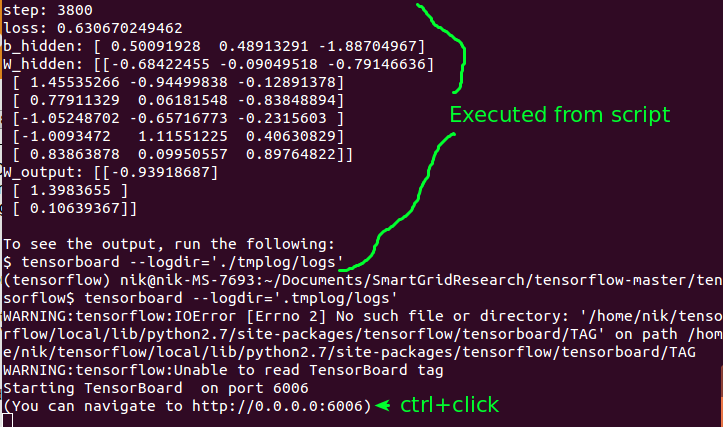
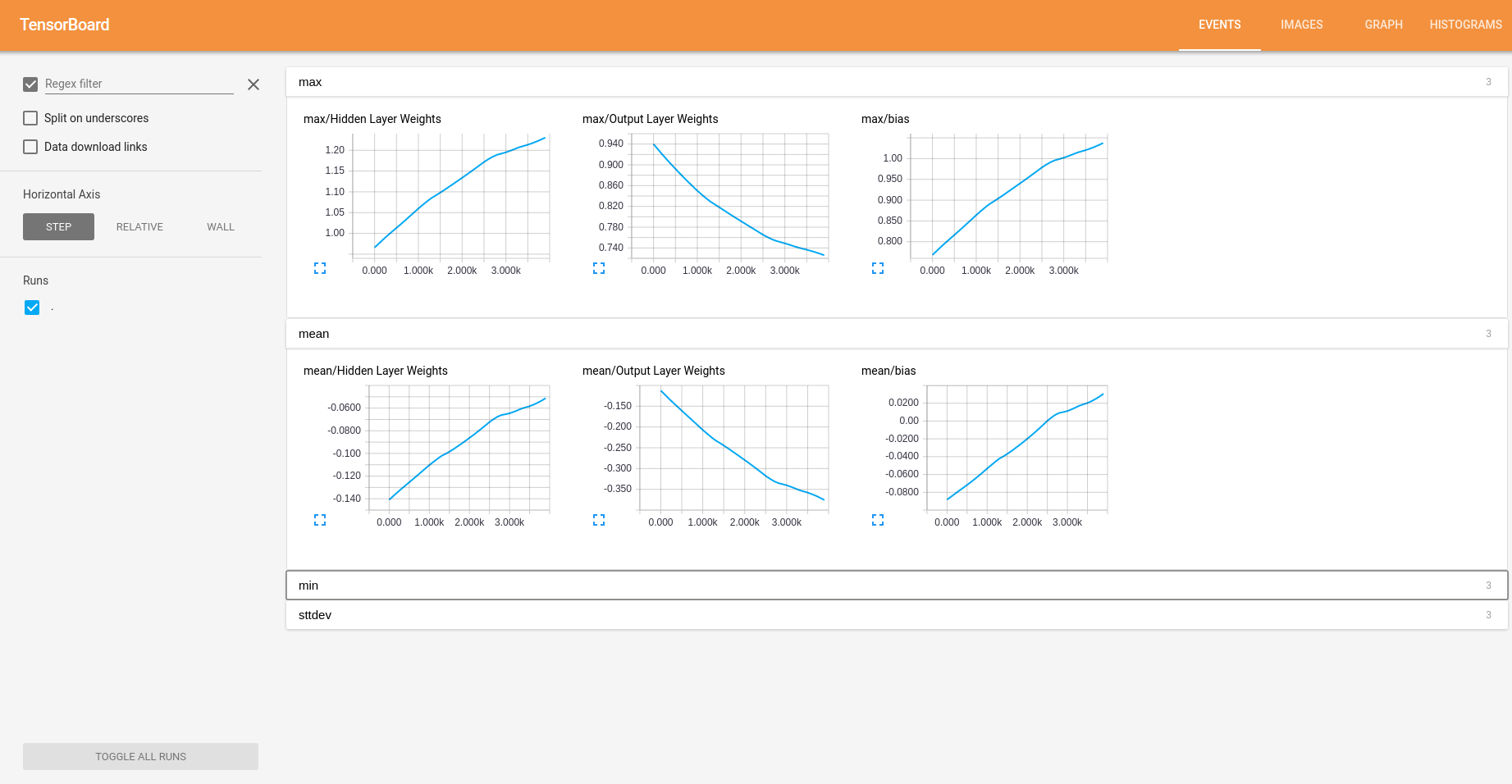
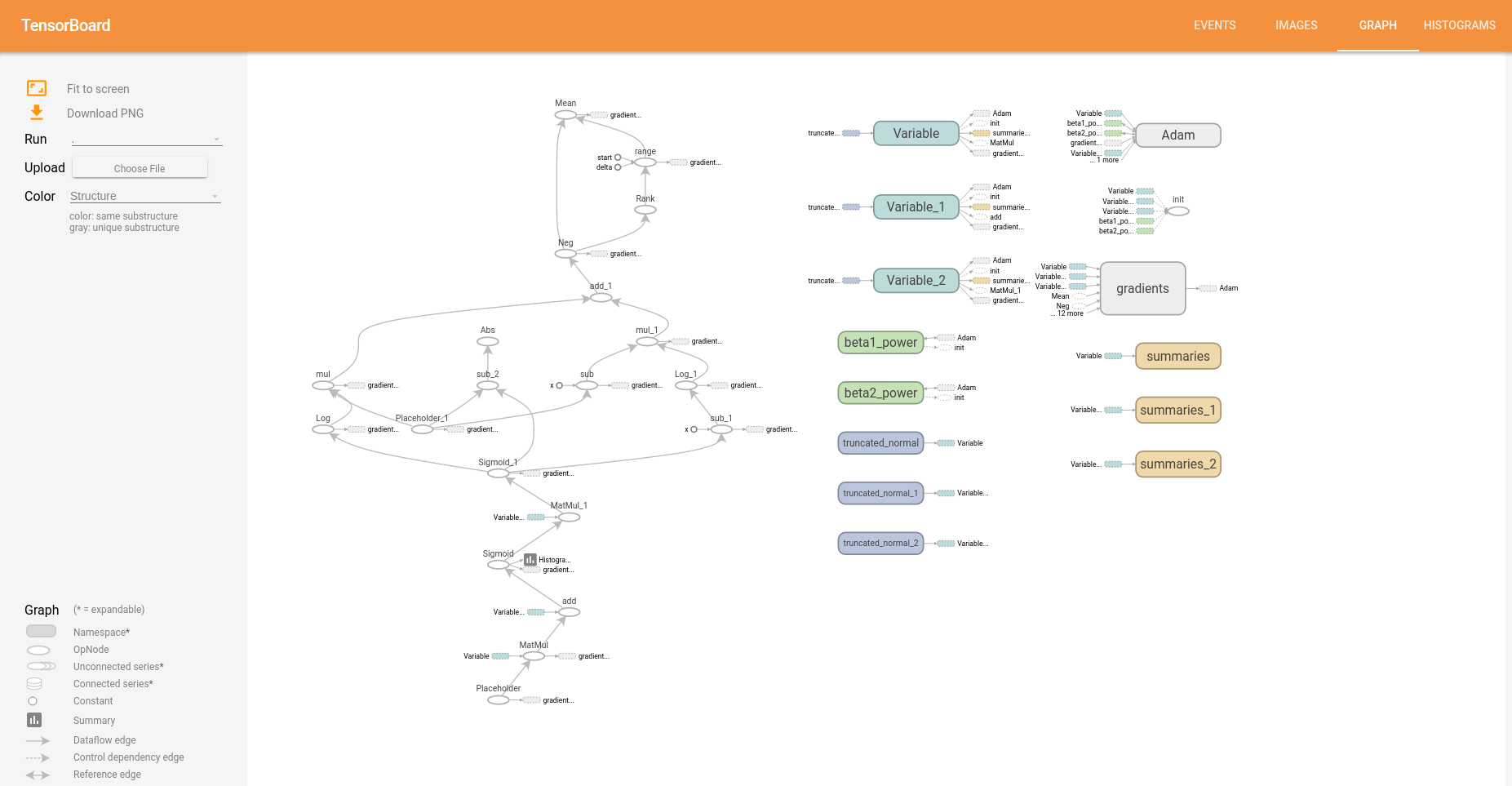
**Issues with TensorFlow**

* Location of the python scripts
  + Some people, including myself have run into issues with the placement of python scripts. A location that gives me no trouble is inside “tensorflow-master/tensorflow.” As a caution, there are some native files already present.
* Weight Initialization
  + When initializing weights some methods have bugs. “tf.truncated\_normal” has not given any issues
* Data Normalization
  + If the data is not normalized, one runs the risk of the weights converging to 0, null, or inaccurate values. Personally, I use “preprocessing.normalize” from the sklearn library. There are a few options, my preference being a “max” normalization. (<http://scikit-learn.org/stable/modules/preprocessing.html>)
* Importing Data
  + Be sure the data is in the correct shape when importing. If the output variable is supposed to be a column vector, be sure it is a column vector.
* Debugging
  + In general, debugging can be difficult since TensorFlow does a lot of it's work in c++. Tensor's (objects that a part of the TensorFlow graph) do not store values. One must be comfortable knowing how to get desired values (initializing variables, activating the session, and running the feed dictionary)
* TensorBoard
  + Known to be buggy, the TensorBoard sometimes does not grab it's event file correctly. If it is not, navigate to the “logdir” (will be within the /tensorflow-master/tensorflow) directory and be sure an event file is inside. If so, restart the browser and terminal and try again. If not, then there is an issue with the script.

**How to Run the Multilayer Perceptron script (Virtual Enviroment)**

* Installation
  + Follow all the instructions online for installation
    - TensorFlow requires a Mac or Linux OS. One needs a Virtual Machine if using Windows. All of this information is on the TensorFlow website.
  + <https://www.tensorflow.org/versions/r0.7/get_started/os_setup.html>
    - The rest of this tutorial focuses on the Virtualenv installation. This prevents tensorflow from overriding any pre-existing Python packages. Feel free to do any installation to your liking
  + The rest of this document will show how to use Tensorflow after those instructions have been followed
* Getting Started with a Virtual Enviroment
  + It is recommended to use Tensorflow in a Virtual Enviroment, the steps for installation of virtualenv are below the steps for the pip installation.
  + In the terminal, navigate to the directory containing the tensorflow folder. This should be directly in the “tensorflow-master” directory.
    - $ cd …/tensorflow-master (… is the directory containg the tfmaster folder)
  + Before starting, activate the virtualenv in the tensorflow directory
    - $ virtualenv --system-site-packages ~/tensorflow
    - $ source ~/tensorflow/bin/activate)
  + Now navigate into the directory containing your python script ($ cd tensorflow/) folder. Your terminal should look like this:



* + run the script ($ python Scipt.py)
    - The script should show a few different outputs related to the status and/or debugging
  + When one is done with the virtual enivroment use the command “$ deactivate”
* TensorBoard
  + The TensorBoard is a great way to visualize everything going on in script
  + In order to visualize certain elements with the TensorBoard, one needs to use SummaryWriters. The TensorFlow website has a great tutorial for making great TensorBoards
    - <https://www.tensorflow.org/versions/r0.8/how_tos/summaries_and_tensorboard/index.html>
  + To launch simply run your script enter ($ tensorboard –logdir=”./logdir'”). The script should prompt the user with the exact syntax
    - This will prompt to navigate to the localhost 6006, this is where you can look at the graphs and other ideas tensorflow has stored
    - The TAG warning is for debugging, not an issue
  + Here is the TensorBoard from the “TFMultilayerPerceptron.py” script
  + Ctrl+click on the link to open up the localhost on port 6006
  + If everything went smooth, your browser should look similar to this:
  + Check out the graph (the map for how data flows):
  + For a full interpretation of what everything means, visit: <https://www.tensorflow.org/versions/r0.8/how_tos/graph_viz/index.html>

Works Cited

1. Deng, L.; Yu, D. (2014). “Deep Learning: Methods and Applications”. *Foundations and Trends in Signal Processing* 7:3-4*.* pg. 199