Classroom Demonstration:

Deep Learning for Classification and Prediction, Introduction to GPU Computing

Prof. Eric A. Suess, Department of Statistics, CSU East Bay

July 30, 2018

Question:

Should GPU Computing be introduced and discussed in Statistics classes? No/Yes/Maybe

Should Neural Networks (NN) be introduced and discussed in Statistics classes? Yes/No/Maybe

Answers to first question:

No. The overhead of installing and maintaining a PC computer with the necessary video drivers for use with Machine Learning software (such as Tensorflow, Keras, MxNet, PyTorch, XGBoost) is considerable and it is beyond the average Statistics student. It can be done by the select few of interested students. The cloud offerings are also still a bit difficult to use and focus on Python.

Yes. The topic of GPUs is currently very popular in the field of ML and so all students who are interested in working in a company that is doing ML should become familiar with GPU and their use for training ML models. **Sorry this poster does not do it.**

Maybe. In the near future I hope that there will be a cloud computing platform that delivers R and RStudio, with the tensorflow and keras package installed, with access to CPU clusters and GPUs. The RStudio Cloud is almost there. Amazon, Google, IBM, and FloydHub are there, but for Python.

Answers to second question:

Yes. In my experience the use of Neural Networks/Deep Learning has changed my internet connected life considerably in recent years. In Data Products such as *Recommendation Engines* (Movies, Adds), *Facial Recognition* in photos (Facebook), *Translation software* (Text, Audio), etc. I believe it has changed the lives of my students or they are so used it that the use of neural networks is something they are not aware of the problems they solve.

No. Neural networks are not really statistics and should be taught in computer science or math classes.

Maybe. Neural networks are in statistics software (R, Python, SAS, JMP). While NN maybe not have fully theoretically understood statistical properties, they solve modeling problems that statistical models are applied to, but usually to very large datasets.

Small Data

Relating *data products* to *modeling problems*, discussing *Big Data*, and then relating the ideas to *statistical methods* applied to **sampled data** gives a nice basis in daily experience for the introduction to and the continued relevance of *Statistics* in the modern data world. The use of statistical methods for **personalized data collection** and **small group data analysis** is a point of discuss that is relevant to individuals.

Individuals care about their own **Small Data**!!!

Lots of problems of interest are related to Small Data. And as PC computers get more RAM, many Big Data problems can be run as Small Data problems. Or as elastic cloud computing become more available the distinction is becoming less important.

Abstract

We present examples of the use of basic **Artificial Neural Networks** (NNs) for introductory Statistics classes at the levels of undergraduate, major and first year graduate classes. Because of the available packages in R, NNs are easily included in the discussion of **Statistics classes** as alternative methods to linear regression and logistic regression.

With the increases in computational power (parallel computation on CPUs, parallel computation on GPUs, TPUs, and NPUs, and with increases in RAM) **Deep Learning** has become possible. With the newer packages in R to connect to *H2O*, *Tensorflow*, and *Keras*, implementing Deep Learning in class is possible.

We present **examples** for running NNs and Deep Learning in Statistics classes with discussion of the similarities and differences between traditional Statistical Methods and Deep Learning.

Examples of NNs in Statistics classes

Incorporate into existing classes, where appropriate.

- Undergraduate classes
  - Examples of Data Products that use NNs for classification.
- Majors in Statistics
  - Examples of NNs for numeric prediction as an alternative to Linear Regression.
- First year Graduate classes
  - Examples of numeric prediction as an alternative to Linear Regression.
  - Examples of NNs for classification as an alternative to Logistic Regression.
  - Examples of Deep Learning (DL) for multiclass prediction.
  - Examples of Deep Learning for image classification. Convolutional Neural Network (CNN).
- Beyond
  - Example of a Recurrent Neural Network (RNN).

Data Science

My Opinion: With the increasing interest in Data Science connecting Statistics instruction with Machine Learning and Deep Learning there is an opportunity to widen the interest in Statistics and to show the value of Statistical methods. While Deep Learning is being popularized by Google, Microsoft, Amazon, IBM, and other companies, with new software, many of the problems that Deep Learning has been applied to are related to Statistical methods and if only smaller data is available the traditional statistical methods may work better than applying Deep Learning.

By showing situations where Deep Learning is being used and showing that traditional statistical methods can also be applied to similar problems may excite students about learning regression, logistic regression and other statistical methods along with learning about Deep Learning.

This will well prepare students of Statistics with an understanding of how Statistics is connected to Machine Learning and Deep Learning.

Presenting NNs in Statistics classes.

My Questions:

1. Are NNs part of the Statistics curriculum?
2. Should NNs even be presented in Statistics classes?
3. How much detail should be presented?
4. Is it ok to introduce NNs as a black-box algorithm?
5. What is Deep Learning?

My current Answer:

1. Generally No
2. Yes
3. Very little at the undergraduate level, some at the undergraduate level, and more at the graduate level.
4. Yes
5. Acknowledging and connecting to the excitement of Deep Learning is *potentially good for attracting students to Statistics programs*.

The Problems with these suggestions

1. There is no extra time to present non-Statistics topics in the classes.
2. This is not a Statistical method.
3. We do not present black-box methods in Statistics classes.
4. The topic of NNs is a difficult topic that is too hard.
5. This is a Computer Science topic.

These problems are all true. I for one am still not certain if the inclusion of NNs in a perfect idea, however, most students have heard of Deep Learning at this point and connecting what they are studying in Statistics to these ideas *maybe be a good idea*.

My thinking

One main reason for making the connection to NNs and Deep Learning early in Statistics education is that Deep Learning seems to be most effective with Big Data and not all Data Analysis/Data Science work is related to Big Data, just Data or Small Data. Showing a few examples of the use of NNs with smaller data from Statistics classes would clarify when statistical methods are very useful and how NNs can be used as an alternative method for similar problems.

R has packages that can be used to implement NNs on the same data that linear regression, logistic regression, and time series methods can be used with.

The Statistical Machine Learning Framework

As everyone knows, not all of Statistics is about Hypothesis Testing.

This is something Statistics teachers should make clear in introductory statistics classes.

Prediction is important and taught in Regression courses.

Classification is also important and taught in Logistic Regression courses.

Both of these areas could be used in the context of **Exportatory Data Analysis** parts of Statistics courses.

The Statistical Machine Learning Framework

The framework within which NNs needs to be presented so students of Statistics can communicate with others in Data Science, Machine Learning, and Analytics.

- Holdout Method
- Training Data
- Test Data
- Validation Data
- Accuracy, Confusion Matrix
- Cross Validation

(I do think the basic ideas of Learning should be incorporated into discussions of Linear Regression presentations.)

R packages

Some R packages to try

- neuralnet
- nnet
- fpp2 for nnetar()
- nnfor
- RSNNS
- h2o
- tensorflow and keras

An aside,

- greta for Bayesian modeling uses tensorflow as a back-end.

Undergraduate classes

Point out to students that many internet service that they use on a daily basis now use Machine Learning and NNs.

Data Products using NNs for classification and prediction

- Facebook facial recognition NN
- Google photos, text search
- Amazon Alexa

Have students try them and think about them.

Statistics Major classes

After discussions of **Simple Linear Regression**, discuss the use of **Neural Networks** to predict a numeric outcome.

R code

In R a linear regression class we can be fit a line using the usual R model notation.

cars_train %>% lm(dist ~ speed, data = .)

Using the **neuralnet()** function from the neuralnet package a neural network can be fit using the usual R model notation.

library(neuralnet)

cars_train %>% neuralnet(formula = dist ~ speed, linear.output=TRUE))

First year Graduate classes

After discussions of **Multiple Linear Regression**, discuss the use of **Neural Networks** to *predict* a numeric outcome.

After discussions of **Logisitic Regression**, discuss the use of **Neural Networks** to *classify* outcomes.

R code

In R a linear regression class we can be fit a multiple linear regression using the usual R model notation.

Using the **neuralnet()** function from the neuralnet package a neural network can be fit again using the usual R model notation.

neuralnet(strength ~ cement + slag + ash + water + superplastic + coarseagg + fineagg + age, data = concrete_train, hidden = 5, act.fct = "tanh")

Using the **deeplearning()** function from the h2o package a multilayer neural network can be fit.

h2o.deeplearning(x=1:8, y="strength", training_frame=splits[[1]], activation = "Tanh", hidden = c(200,200), distribution = "gaussian")

Beyond

Discussion of Classification in Statistical Learning classes could/should include the discussion of Deep Learning.

RStudio has developed the tensorflow and keras packages, interfaces to these python packages. These can be run on CPUs. This is probably where students should begin to run these packages.

While they can be installed locally on a laptop, maintaining them has been a challenge.

GPU Computing! Wow this is much faster.

Beyond

R Studio Cloud can be used to run the h2o package with a multicore machine.

I have not found a good cloud based RStudio with the tensorflow and keras packages installed. There probably is a way to do this on AWS. (See the end of the Deep Learning with R book.) Maybe IBMs Data Science Experience?

Undergraduate Reading on Medium

Machine Learning is Fun!

- Part 1: The world's easiest introduction to Machine Learning
- Part 2: Using Machine Learning to generate Super Mario Maker levels
- Part 3: Deep Learning and Convolutional Neural Networks
- Part 4: Modern Face Recognition with Deep Learning
- Part 5: Language Translation with Deep Learning and the Magic of Sequences
- Part 6: How to do Speech Recognition with Deep Learning
- Part 7: Abusing Generative Adversarial Networks to Make 8-bit Pixel Art
- Part 8: How to Intentionally Trick Neural Networks

Statistics Major Reading

- neuralnet: Training of Neural Networks
- Deep Learning with R
- fast.ai
- Google Machine Learning Crash Course
- Tensorflow Playground

First year Statistics Graduate Student Reading

- neuralnet: Training of Neural Networks
- tensorflow R package from RStudio Examples
- [keras R package from RStudio Examples]
- Deep Learning with R
- Deep Learning with Python
- Deep Learning book
- Medium Deep Learning with Python
- Coursera Deeplearning.ai

Software: For these Python is used.

- Google Colab
- Google Colab Free GPU Tutorial
- Google Seedbank
- Amazon SageMaker
- FloydHub