AI Education: Deep Neural Network Learning Resources

Todd W. Neller (Gettysburg College; tneller@gettysburg.edu)
DOI: 10.1145/3137574.3137580

Introduction

In this column, we focus on resources for learning and teaching deep neural network learning. Many exciting advances have been made in this area of late, and so many resources have become available online that the flood of relevant concepts and techniques can be overwhelming. Here, we hope to provide a sampling of high-quality resources to guide the newcomer into this booming field.

Textbooks and Papers

Deep Learning (Goodfellow, Bengio, & Courville, 2016) is a popular recent textbook that seeks to briefly introduce background mathematical topics (e.g. linear algebra, probability and information theory, numerical computation) as well as machine learning basics. The second part of *Deep Learning* treats core material of deep learning practice (e.g. deep feedforward networks, regularization, convolutional networks, recurrent networks, etc.), whereas the third part covers topics of modern research interest in deep learning.

This textbook is available in HTML form on the authors' Deep Learning Book website and it is not difficult to find other e-book formats online that have been built from these HTML pages. However, this text alone is not the easiest introduction to the field. We would recommend Andrew Ng's Machine Learning MOOC, *An Introduction to Statistical Learning with Applications in R* (James, Witten, Hastie, & Tibshirani, 2014), and other resources listed in this AI Education Matters column in Volume 3, Number 2 as starting points for background material relevant to all machine learning.

Also recommended as a gentler introduction is Michael Nielson's *Neural Networks and Deep Learning* online book.

Web Resources

One of the best news feeds for following Deep Learning research developments and learning resources is Waikit Lau and Arthur Chan's *Artificial Intelligence and Deep Learning (AIDL)* Facebook group. At the time of this writing, it has over 30,000 members and features an active and steady flow of research results, tutorials, announcements, and Q&A discussion relevant to deep learning. Recommendations much like these can be found in questions 2-4 of the AIDL FAQ.

Other good sites for suggested starting points for learning about DL is *A Guide to Deep Learning* by YerevaNN Labs, Piotr Migdał's *Learning Deep Learning with Keras*, a16z team's reference links, Stanford's CS 231n *Convolutional Networks* course website, and, of course, various Wikipedia pages concerning artificial neural networks.

MOOCs

In April 2017, David Venturi collected an impressive list of Deep Learning online courses along with ratings data. In August 2016, Arthur Chan listed his top 5 lists. Concurring with these bloggers, we found Geoffrey Hinton's *Neural Networks for Machine Learning* course lectures to be a good high-level introduction to the field. However, this course is not oriented towards the beginner.

We recommend taking Andrew Ng's Machine Learning MOOC for background coverage and then supplementing Hinton's MOOC with applied tutorial exercises found elsewhere.[1] Kaggle is a data science competition website that features tutorials, datasets, and challenges that offer practical experiential learning opportunities. Beyond Hinton's general introduction, Arthur Chan also recommends Hugo Larochelle's graduate-level online Neural Network course.

[1] At time of writing, Andrew Ng has announced a new Coursera specialization in deep learning.

Software

There are several popular software frameworks that facilitate rapid prototyping of deep learning systems. Most popular is Google's TensorFlow. An even higher-level layer that has become popular is Keras, which can run as a layer on top of TensorFlow, Microsoft Cognitive Toolkit (CNTK), or Theano. To grasp how high-level Keras is, consider these small MNIST digit recognition training examples featuring multi-class logistic regression, single-hidden-layer neural network training, and convolutional network training implemented in under ten lines each.

Other popular software for deep learning includes Torch (and PyTorch), Caffe, MXNet, and DeepLearning4J. While Python appears to be the most popular language for deep learning development, support exists for other programming languages.

Hardware

Researchers seem to obtain hardware with GPUs that support fast deep learning in three main ways. One expensive route is to buy machines marketed directly for deep learning that typically have high-end GPU specifications and come preinstalled with popular deep learning software. However, there are numerous DIY tutorials for buying the necessary parts and putting together an inexpensive deep learning machine. These options are at the extremes of the high-cost/low-effort and low-cost/high-effort spectrum.

We would recommend a middle-ground approach for time-strapped faculty with tight budgets: Buy a high-end gaming machine (e.g. Dell's Alienware desktops) with good GPUs and install the necessary software as needed. Tim Dettmers has shared results of a recent GPU comparison study, and Ved's d4datascience blog entry describes the process of installing CUDA libraries and TensorFlow in detail.

Your Favorite Resources?

These are but a few good starting points for learning about deep neural network learning. If there are other resources you would recommend, we invite you to register with our wiki and add them to our collection at http://cs.gettysburg.edu/ai-matters/index.php/Resources.

References

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. (http://www.deeplearningbook.org)

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2014). *An introduction to statistical learning: With applications in r*. Springer Publishing Company, Incorporated. (http://www-bcf.usc.edu/~gareth/ISL/)

Todd W. Neller is a Professor of Computer Science at Gettysburg College. A game enthusiast, Neller researches game AI techniques and their uses in undergraduate education.