SBP-BRIMS 2016 Tutorial: Deep Learning Applied

Alexander G. Ororbia II
College of Information Sciences & Technology
The Pennsylvania State University, University Park
ago109@psu.edu

Proposed Tutorial Content   Recent advances in artificial neural network research have greatly impacted domains ranging from computer vision to bio-informatics to natural language processing. Furthermore, these improvements have led to a seemingly explosive growth in the number of intelligent applications, including image-based search, image-captioning, automated language translation, and basic conversation modeling. Beyond their utility, deep architectures have the potential to serve as predictive models trained on rich, large datasets in the social sciences, where relationships among variables can be highly non-linear.

However, while these deep neural architectures have outperformed shallow, traditional statistical learning models in a variety of settings, getting the great generalization performance as reported in the literature can be difficult in practice. The process of training neural models has been labeled a "dark art", as it requires expertise in the theoretical and engineering underpinnings of neural-based learning. The intent of this tutorial is to guide the audience through the process of properly training state-of-the-art deep neural networks. As human annotation of data in the social sciences is costly, we emphasize semi-supervised learning. Concepts such as the choice of loss functions in regression and classification settings, neural architecture design, approaches to selecting and tuning various hyper-parameters, basics in data pre-processing for neural models, and gradient-descent-based parameter optimization will be presented.

The audience will follow along by modifying code samples that use our own Scala-based deep learning library, Yet Another Deep Learning Library (YADLL), written with convenient Java and Python interfaces. Specifically, we will develop a semi-supervised model capable of basic automated content coding of text, a task which is often critical in social science research and theory-building. By the end of the tutorial, the audience should have a basic grasp of how to properly train a deep neural network and methods for identifying undesired behavior and bad performance, familiarity with the application of these models to a real-world data-setting, and pointers to key literature that provide further in-depth information for practically applying neural models to data.

Expected Audience   This tutorial is intended for a more general audience, primarily social scientists interested in applying state-of-the-art neural models to their own work. No knowledge of neural networks will be assumed and only basic familiarity with machine learning is required. Knowledge in calculus and linear algebra will certainly help, but any necessary mathematics will be explained as needed in the context of successfully applying these models to data.[1]

About the Presenter   Alexander G. Ororbia II is a doctoral graduate candidate at the Pennsylvania State University, advised by C. Lee Giles and David Reitter. His research is centered on developing novel neural architectures capable of semi-supervised, lifelong learning from non-stationary distributions, primarily from text data. The presenter will require funding for conference travel and lodging.

---

[1]Plenty of references will be provided throughout the tutorial and made publicly available for those interested in building deeper knowledge on the top of the foundation provided by this tutorial.