Statistical Perspectives in Teaching Deep Learning from Fundamentals to Applications

Nathan Kim & David Han, Ph.D. (david.han@utsa.edu)

The University of Texas at San Antonio, TX 78249

Abstract

The use of Artificial Intelligence (A.I.), Machine Learning (ML) and Deep Learning (DL) have gained a lot of media attention and become increasingly popular in many areas of scientific and business applications. Historically ML and its theory had strong connections to statistics; however, the current DL context is mostly in computer science perspectives and lacks statistical perspectives. In this work, we address this research gap and discuss about the critical components in teaching DL to the next generation of statisticians and data scientists.

Keywords: artificial intelligence, curriculum development, deep learning, machine learning, artificial neural network, statistical education
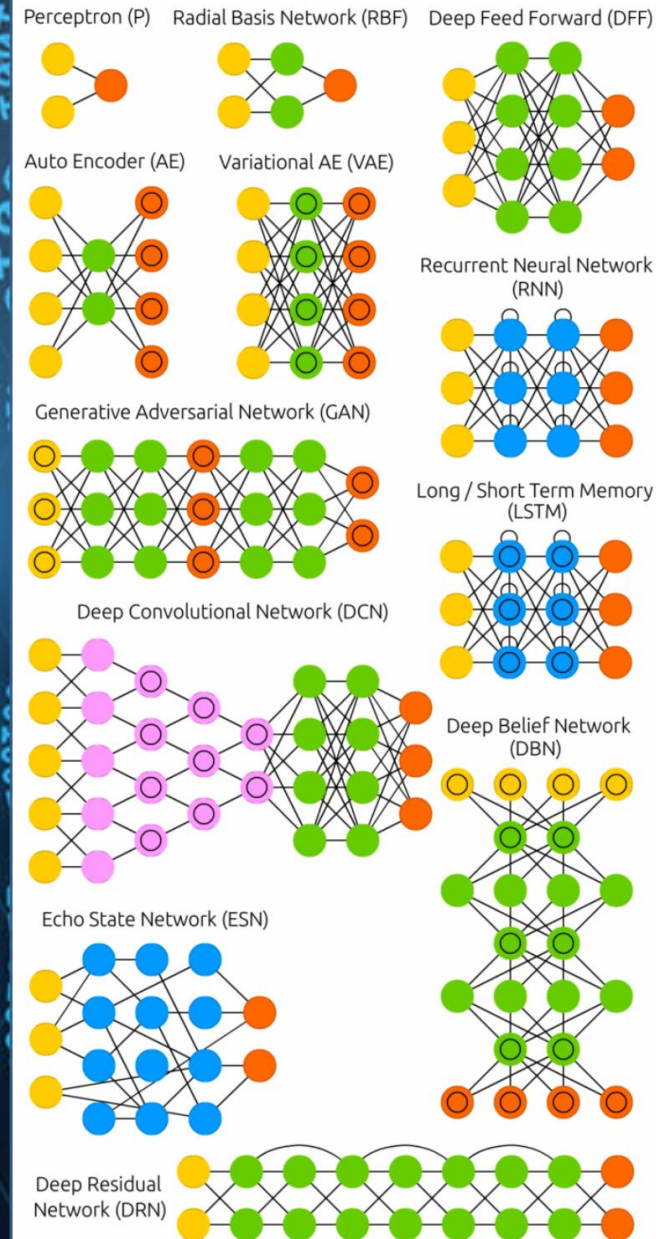
Practical Challenges

- practice-oriented curricula development, rather than proof-based curricula
  - Students' appetite and drive is to do DL, rather than to prove theorems about DL.
- textbook selection, other than Deep Learning by Goodfellow et al. (2016)
- software management due to diverse programming environments and available open-source tools
  - Google TensorFlow, Keras, PyTorch, Theano, Caffe, Apache MxNet, MS CNTK, DL4J, Chainer, …
  - must consider the learning curve, stability and developer/community support, tool ecosystem.
- cognitive burden due to inconsistent terminology and jargon between statistics and computer science
  - *e.g.*, bias term, noise robustness, normalization, hyperparameter, etc.

Course Contents

- basic concept of perceptron, (deep) feed-forward neural network, back-propagation, non-convex optimization, empirical risk minimization, regularization and dropout, feature selection and dimension reduction, interpretable deep networks, convolutional neural network (CNN), recurrent neural network (RNN) and long short-term memory (LSTM), variational auto-encoders (VAE), generative adversarial network (GAN),
representation learning, reinforcement learning, uncertainty quantification (UQ) via deep Gaussian process (DGP) and Bayesian variational inference, …
- should cover other ML techniques such as support vector machines (SVM), random forests, boosting and bagging, natural language processing (NLP), etc.
- *MUST* discuss how all these can be viewed as statistical techniques.
  - connect to the fundamental statistical theory or methods.
- address model assumptions and potential drawbacks.
  e.g., overfitting and poorly approximated UQ, sub-optimal network architectures due to trial and error or based on abstract

Pedagogical Approach

- curriculum should be dynamical, adjustable, and adaptable.
- prepare lectures in interactive and dynamic platform.
  e.g., Jupyter Notebooks
- utilize MOOC in a complementary manner.
  i.e., bootcamp, preparatory and supplementary
- consider audience level: beginner or advanced
  - (senior) undergraduate? graduate (M.S./Ph.D.)?

Perceptron (P) — Radial Basis Network (RBF) — Deep Feed Forward (DFF) — Auto Encoder (AE) — Variational AE (VAE) — Generative Adversarial Network (GAN) — Deep Convolutional Network (DCN) — Echo State Network (ESN) — Deep Residual Network (DRN) — Recurrent Neural Network (RNN) — Long / Short Term Memory (LSTM) — Deep Belief Network (DBN)

Programming Skill

- significant prerequisite, and should be enhanced over the course.
- Python as the main programming language, not R
  - must be adaptable as software tools can/will change constantly.
- secure computational resources and infra (*e.g.*, HPC servers, cloud computing) as most algorithms are computation-intensive.

Homework & Projects

- balance between theory and practice
- must be prescriptive and application-oriented.
- start with classical datasets to get familiar.
  e.g., MNIST, Cifar-10, SVHN, etc.
- use large open datasets (*e.g.*, Kaggle) for implement-ation and practices w/ Big Data
- introduce domain knowledge specific to each problem from various fields
  - help building cross-/inter-/multi-disciplinary collaborations with domain experts
- encourage/incorporate team-based learning

References

- Geron, A. (2017). *Hands-on Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. Sebastopol, CA: O'Reilly Media, Inc.
- Goodfellow, I., Bengio, Y., and Courville, Y. (2016). *Deep Learning*. Cambridge, UK: MIT Press.
- Sutton, R.S. and Barto, A.G. (2018). *Reinforcement Learning*. Cambridge, UK: MIT Press.
- van Veen, F. and Leijnen, S. (2019). A mostly complete chart of neural networks (asimovinstitute.org)