Article from

Predictive Analytics and Futurism

December 2016

Issue 14

Guide to Deep Learning

By Syed Danish Ali

Up until the recent past, the artificial intelligence (AI) portion of data science was looked upon cautiously due to its history of booms and flops.[1] In the latest stream of events, major improvements have taken place in this field and now deep learning, the new leading front for AI, presents a promising prospect for overcoming problems of big data. Deep learning is a method of machine learning that undertakes calculations in a layered fashion starting from high level abstractions (vision, language and other AI related tasks) to more and more specific features[2]. Deep learning algorithms essentially attempt to model high-level abstractions of the data using architectures composed of multiple non-linear transformations. The machine is able to progressively learn as it digests more and more data, and its ability to transform abstract concepts into concrete realities has opened up a diverse plethora of areas where it can be utilized. Deep learning has various architectures such as deep neural networks, deep belief networks, Deep Boltzmann machines and so on that are able to handle and decode complex structures that have multiple non-linear features.[3]

Deep learning offers us considerable insight into the relatively unknown unstructured data which is 80 percent of the data that we generate as per IBM.[4] While traditional data analysis before 2005 focused on just the tip of the iceberg, the big data revolution sprang up and now deep learning offers us a better glimpse into the unconscious segment of data that we know exists, but is constrained in realizing its true potential. Deep learning helps us in both exploring the data and identifying connections in descriptive analytics for ratemaking, but these connections also help us in price forecasting what the result will likely be, given the particular combination as the machine learns from the data.

Deep learning has inputs, hidden layers where they are transformed by the weights/biases and output which is achieved through choice of activation function from various functions available (Softmax, sigmoid, hyperbolic tangent, rectified linear, maxout and so on). The weights/biases are learned by feeding training data to the particular deep learning architecture.

A de-mystified foundation of deep learning is that deep learning is mostly a way of using backpropagation with gradient descent and a larger number of hidden neural network layers, which is certainly not new. However, revival of deep learning was possible after 2010 and onwards due to drastically more computational power from GPUs, bigger datasets, and some key algorithm tweaks—mainly dropout and AdaGrad to increase accuracy rates. Moreover, the unique feature of deep learning is that it allows individual parts of the model to be trained independently of the other parts.[5]

Deep learning models can recognize human faces with more than 97 percent accuracy, as well as recognize arbitrary images and even moving videos. Deep learning systems now can process real-time video, interpret it, and provide a natural language description. It is becoming increasingly established that deep learning can perform exceptionally well on problems involving perceptual data like speech recognition image classification and text analytics.[6]

In a single formula, this is the formula for neural networks (for hyperbolic tangent activation function).[7]

$$p(x) = \beta_0 + \sum_{i=1}^{n_h} \beta_i \tanh\left(\alpha_{i,0} + \sum_{j=1}^{n} \alpha_{i,j} x_j\right).$$

So that essentially, p(x) = linear + non- linear.

Aside from exposures, the other side of ratemaking of general insurance is losses and loss trends. By building deep learning models we can analyze images to estimate repair costs. Also deep learning techniques can be applied to automatically categorize the severity of damage to vehicles involved in accidents. This will more quickly update us with more accurate severity data for modeling pure premiums.[8]

Deep learning is becoming the method of choice for its exceptional accuracy and capturing capacity for unstructured data. This is also emphasized ahead in the section machine learning-unstructured data mining and text analytics.[9]

One issue, however, with deep learning is trying to find the hyper-parameters that are optimal. The possible space for consideration is very large and it is difficult and computationally intensive to understand each hyper parameter in depth. One potential solution that the author of this report identifies is the possible use of a genetic algorithm to find optimal hyper parameters. Genetic algorithms are already used on GLMs on R "glmulti" packages to select optimum GLM equation as per a given criteria usually Akaike Information Criterion or Bayesian Information Criterion.

Moreover, another algorithm has been used to optimize both structure and weights of a neural network. ES HyperNEAT is Evolving Substrate Hyperbolic Neuroevolution Of Augmenting Topologies developed by Ken Stanley. It uses a genetic algorithm to optimize both the structure and weights of a neural network. Following from this, maybe the ES HyperNEAT framework can be extended to deep learning so that a genetic algorithm can optimize both the structure and weights of the neural networks in deep learning as well.[10]

Another problem is over-fitting. Machine unlearning can be used to solve this. We will try to explain machine unlearning in one sentence. Machine unlearning puts a new layer of a small number of summations between the training data and the learning algorithm so that the dependency between these two is eliminated. Now the learning algorithms depend only on the summations instead of the individual data from which over-fitting can arise more easily. No retraining or remodeling is required.[11]

Finally, there are huge numbers of variants of deep architectures as it's a fast developing field and so it helps to mention other leading algorithms. This list is intended to be comprehensive but not exhaustive since so many algorithms are being developed.[12, 13]

1. Deep High-order Neural Network with Structured Output (HNNSO)
2. Deep convex networks
3. Spectral networks
4. noBackTrack algorithm to solve the online training of RNN (recurrent neural networks) problem
5. Neural reasoner
6. Reccurrent Neural Networks
7. Long short term memory
8. Hidden Markov Models
9. Deep belief networks
10. Convolutional deep networks
11. LAMSTAR
12. a) Large memory storage and retrieval neural networks
13. b) Increasingly being used in medical and financial applications
14. Deep Q-network agent
15. a) Used by Google DeepMIND
16. b) Based on reinforcement learning, which is a major branch of psychology, aside from evolution ■

Syed Danish Ali is a senior consultant at SIR consultants, a leading actuarial consultancy in the Middle East and South Asia. He can be reached at *sd.ali90@ymail.com.*

ENDNOTES

1 Jack Clark (Feb. 3, 2015); Bloomberg Business, "I'll be back: The Return of Artificial Intelligence."

2 Will Knight (May 31, 2015); MIT Technology Review, "Deep Learning catches on in new industries, from fashion to finance."

3 Yoshua Bengio, University of Montreal, "Learning deep architectures for AI."

4 IBM Website Smarter Planet, "Improve decision making with content analytics."

5 Sean Lorenz (June 2016), Domino Datalab, "Deep learning with h20.ai."

6 PwC, March 2016, Top Issues, "AI in Insurance: Hype or reality?"

7 Dugas, et. al.; "Statistical Learning Algorithms Applied to Automobile Insurance Ratemaking."

8 PwC, March 2016, Top Issues, "AI in Insurance: Hype or reality?"

9 Ibid

10 Risi, S. and Stanley, K. University of Central Florida, The ES-HyperNEAT Users Page

11 Cao and Yang, 2015. IEEE symposium on security and privacy, pgs. 463-480. "Towards Making Systems Forget with Machine Unlearning."

12 Mayo M, Larochelle H, (Oct 2015) KDNuggets.com. "Top 5 arXiv Deep Learning Papers, Explained."

13 Mayo M, Larochelle H, (Jan 2016) KDNuggets.com. "5 More arXiv Deep Learning Papers, Explained."