

# Technical Basic Guide to Deep Learning

Syed Danish Ali

## Guide

Up till recent past, the artificial intelligence portion of data science was looked upon cautiously due to its history of booms and flops.<sup>1</sup> In the latest stream of events, major improvements have taken place in this field and now deep learning, the new leading front for Artificial Intelligence, presents 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 Artificial Intelligence related tasks) to more and more specific features<sup>2</sup>. 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.<sup>3</sup>

Deep learning offers us considerable insight into the relatively unknown unstructured data which is 80% of the data that we generate as per IBM.<sup>4</sup> 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. Deep learning is different from neural networks as it has multiple hidden layers whereas neural network only has one.<sup>5</sup>

A de-mystified the foundation of 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

---

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

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

<sup>3</sup> Yoshua Bengio, University of Montreal; "Learning deep architectures for AI".

<sup>4</sup> IBM Website; Smarter Planet; improve decision making with content analytics

<sup>5</sup> Jeff Heaton, SOA Predictive Analytics and Futurism Newsletter; Issue 9, 2014. An Introduction to Deep Learning

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.<sup>6</sup>

Deep learning models can recognize human faces with over 97% accuracy, as well as recognize arbitrary images and even moving videos. Deep learning systems now can process real-time video, interpret them, 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.<sup>7</sup>

In a single formula, this is the formula for neural networks (for hyperbolic tangent activation function)<sup>8</sup>

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

So that essentially,  $p(x) = \text{linear} + \text{non-linear}$ .

Aside from exposures, the other side of ratemaking 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 with us more accurate severity data for modeling pure premiums.<sup>9</sup>

Deep learning is becoming the method of choice for its exceptional accuracy and capturing capacity for unstructured data. This is also emphasized ahead in section machine learning-unstructured data mining and text analytics.<sup>10</sup>

One issue however with deep learning is trying to find the hyper-parameters that are optimum. 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 which the author of this report identifies is the possible use of genetic algorithm to find optimal hyper parameters. Genetic algorithms are already used on GLMs on R 'glmulti' package 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 ES HyperNEAT framework can be

---

<sup>6</sup> Sean Lorenz (June 2016); Domino Datalab; Deep learning with h20.ai

<sup>7</sup> PwC; March 2016; Top Issues: AI in Insurance; hype or reality?

<sup>8</sup> Dugas et al; Statistical Learning Algorithms Applied to Automobile Insurance Ratemaking

<sup>9</sup> PwC; March 2016; Top Issues: AI in Insurance; hype or reality?

<sup>10</sup> Ibid

extended to deep learning so that genetic algorithm can optimize both the structure and weights of the neural networks in deep learning as well.<sup>11</sup>

Another problem is over fitting. Machine unlearning can be used to solve this. Explain machine unlearning in one sentence. Machine unlearning puts a new layer of 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.<sup>12</sup>

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. The list is intended to be comprehensive but not exhaustive since so many algorithms are being developed.<sup>13,14</sup>

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

---

<sup>11</sup> Risi, S. and Stanley, K. University of Central Florida; The ES-HyperNEAT Users Page

<sup>12</sup> Cao and Yang, 2015. IEEE symposium on security and privacy pgs 463-480. Towards making systems forget with machine unlearning.

<sup>13</sup> Mayo M, Larochelle H (Oct 2015) KD Nuggets.com. Top 5 arXiv Deep Learning Papers explained.

<sup>14</sup> Mayo M, Larochelle H (Jan 2016) KD Nuggets.com. 5 more arXiv Deep Learning Papers explained.

## The Impact of Deep learning on Investments

Deep learning offers us considerable insight into the relatively unknown unstructured data which is 80% of the data that we generate as per IBM<sup>15</sup>. 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 are constrained in realizing its true potential. Deep learning helps us in both exploring the data and identifying connections in descriptive analytics but these connections also help us in forecasting what the result will likely be, given the particular combination as the machine learns from the data.

Deep learning, in collaboration with other machine learning tools is making headways in possible applications. All major giants like Google, IBM, Baidu are aggressively expanding in this direction but startups are providing the most vivid applications so far. [Kensho](#) is a startup that aims to use software to perform tasks in minutes that would take analysts weeks or months. Just like searching via Google, the analyst can write their questions in the Kensho's search engine. The cloud based software, as per Forbes reporter Steven Bertoni, can find targeted answers to more than 65 million combinations in the flick of a second by scanning over 90,000 actions which are as myriad as political events, new laws, economic reports, approval of drugs etc and their impact on nearly any financial instrument in the world<sup>16</sup>. Another startup, [Ufora](#) is set to automate a large part of quantitative finance work undertaken by quants, especially on the stochastic modeling front. Even some hedge funds like [Renaissance Technologies](#) are proactively working on machine learning and deep learning algorithms to better see patterns in the financial data to exploit opportunities (which stocks are overrated or underrated, market is going strong on fundamentals or approaching the bubble stage and so on) to guide their investment strategies<sup>17</sup>.

On the other hand, Firms like [Narrative Science](#) and [Automated Insights](#) working on text analytics are utilizing deep learning to create lively and interactive narrative reports out of data and numbers. This essentially means report written by a machine that reads like it is almost written by a human author. To elaborate this feature, Narrative Science's Quill platform undertakes statistical analysis of applying time series, regression etc and then the semantic engine evaluates the important data signal from the unimportant noise as per the needs of the particular audience in question like different reasoning if it is related for a quant or a trader of investments. The patterns are spotted and made sense out of in a holistic manner. Particular fuzzy attention is given to anomalies and elements of results that are deviant from the main normal body of the results to ascertain their impact and proper interpretation. It remembers previous reports made so it doesn't become repetitive. Natural Language Generation is applied with a surgeon's precision and expertise in forming such a dynamic semantic engine.

This is indeed a leap forward as report writing consumes a lot of human time and efforts and because machines making such reports was before unheard of practically. Deep learning allows us not just to explore and understand the data better, but also to perform forecasts better. For predictive analytics part, the startup [MetaMind](#) is working to help financial firms assess chances of selling of stocks by going through corporate financial disclosures<sup>18</sup>. It identifies from previous experiences when a particular

---

<sup>15</sup> IBM Website; Smarter Planet; improve decision making with content analytics

<sup>16</sup> Forbes, Steven Bertoni (Nov 11, 2014); Goldman Sachs leads \$15 million investment in tech start up Kensho

<sup>17</sup> Bloomberg Business, Kelly Bit (May 20, 2015): The \$10 Hedge Fund Supercomputer that's sweeping Wall Street".

<sup>18</sup> MetaMind company's website

combination of actions lead to a particular result to assess chances of the same result happening in the future.

Extrapolating this trend into the future, it is my opinion that such analytics might soon find their way into Mergers & Acquisitions (M & A) and will be able to come up with probability of some key event happening and the consequences of it when involved in a high stake M & A. Another application can be to apply deep learning applications to help us for one of the most vexing problems, i.e, financial crises. Economists, financial experts and social scientists have elaborated on a lot of key issues that lead to financial crises in general as well as specifically for a particular meltdown. These can form the modeling methodology for the deep learning machine to analyze the cosmic scale of data available on any and every platform that it can garner. Such evaluation can perhaps help us to see patterns that we could have missed otherwise as well as to allow us to understand more accurately the sequential movements and mechanisms involved in a particular financial contagion and crisis. There is no guarantee that this will work. But perhaps it can shed some light inside the 'quantum black box' of financial crises. This seems to be the need of the hour with recurring financial hemorrhages such as EU crisis on Greek Debt as well as the recent massive and escalating falls in Chinese stock exchanges reminding us of the bitter past we faced in Wall Street Crisis of 2008-09.

Given all these developments, there are still a myriad of issues that need clarification with not just deep learning in specific but also with big data generally. Automation of such unprecedented scale and intensity raises the possibility of mass redundancies in labor force across the economy. Are we comfortable with giving up our controls to such applications without knowing the full implications of such a move? Not every innovation brings positive results or sustains in the long run. Technology is progressing rapidly at an unstoppable pace but can we manage the social consequences and make it sustainable in the long term? Human efforts are seemingly being diverted from other fields into IT which consequently can imply a concentration of power in one overlord field to the potential detriment of others. Are we ready for this? From a consumer point of view how ethical is it that marketing personnel know you so well that it makes rational optimization very difficult on the part of the consumer?

These are all good questions and should be adequately and mutually tackled and addressed by all the stakeholders involved such as the data scientists, government, professions and consumers so as to be able to reach a mutual policy that can better alleviate such concerns. The core aim of the policy has to be to sustain technology for the benefit of our societies, to lead to value creation, to reduce scarcity and reduce fragility of our systems as well as to generate more resources for our prosperity instead of creating the monster of Frankenstein, as Terminator and other doomsday movies will have us believe.