Artificial Intelligence: The “New Electricity”

Marc Mailloux

Institute for Simulation and Training at the University of Central Florida

Table of Contents

Figures

Tables

Abstract

Jeff Bezos, CEO of Amazon, has recently said A.I. is currently in “a golden age”, where we are “solving problems with machine learning and artificial intelligence that were in the realm of science fiction for the last several decades. [“Jeff Bezos, Amazon and the final frontier”,2017]” Comparing A.I. science to a when it first begun back with Turing and McCarthy, huge transformations have happened. Todays A.I. are a result of the computational boom and have started to change the landscape of A.I. as it becomes smarter and more applicable. The notion of Deep Neural Networks have enabled humans to start making huge strides in A.I., especially in the combination of Neural Network models. CNN and LSTM models have been proven useful on their own but studies show that combining these models will make the resulting model more effective for learning. More effective learning leads to further progression in science as it continues to change the landscape in every industry. This new landscape will be powered through A.I., which is rapidly becoming the ‘new electricity’. To understand A.I.’s potential as the “new electricity”--literally changing the physical landscape and powering a brand new infrastructure through accessibility, affordability, and adaptability--this paper will examine a brief history of Neural Networks, analyze the operation of modern Deep Neural Networks, and look into a drastically different future full of new possibilities.

Keywords: Artificial Intelligence, Neural Network, Deep Neural Network, Convolution Neural Network, Bidirectional Neural Network, Long Short Term Memory, Backpropagation, Reinforcement Learning, Supervised Learning

Introduction

History has evidenced a huge potential for technological advancement, from the first instances of mechanical innovation to modern day breakthroughs in Artificial Intelligence and Intelligent Systems. Although each advancement represents a different turning point in history, influencing and progressing society in distinct ways, each also shares one commonality, the factors that characterize each as lasting movements: accessibility, affordability, and adaptability. Formerly complex and expensive systems are now widely available, user-friendly, and affordable, offering a strong framework for the establishment of future progress and application. To understand A.I.’s potential as the “new electricity”--literally changing the physical landscape and powering a brand new infrastructure through accessibility, affordability, and adaptability--this paper will examine a brief history of Neural Networks, and argue that our modern multi-method approach holds the key to a drastically different future. Despite opposing viewpoints, I will stand firm in my argument and provide counter evidence to address any concerns.

History

Early A.I. and Neural Networks

Prior to Intelligent Systems, the main focus was mechanical and electro-mechanical. But when Alan Turing created the first computer during WWII, an entirely new research tool emerged for modeling systems (B. & G., 2005). Using the computation power of computers, scientists developed algorithmic models and simulation-based verification processes. Even in these early phases, the notion of machines possessing the very human ability to “think” plagues Turing, leading him to create “The Imitation Game”, which probes the potential for human-like intelligence in I.S. and, ultimately, feasible solutions for our most complex problems (Byrnes,

2016) (Rajaraman, 2014). Years later in 1955, John McCarthy became known as the Father of Intelligence when he develops a new use for computers, coining the term “A.I.” with the development of a programming language called LISP (Rajaraman, 2014). McCarthy set A.I. into motion with LISP, which “pioneered many new ideas in computer science including list data structure, tree data structure, a common or same structure for program and data, [and] higher order functions” (Rajaraman, 2014). Then in 1956, after McCarthy successfully completed A.I. focused summer research at Dartmouth, A.I. begins to emerge in Europe, with “Machine Intelligence workshops” hosted in Edinburgh, Scotland (Sandewall, 2014).

With Artificial Intelligence on the rise, traditional input-output techniques begin to take on new complexities. In his book “The Organizational Behavior”, A.I. scientist Donald Hebb presents the concept of using biological neurons to model machine intelligence (Zahedi, 1991). This “Hebbian Learning” is now known as Unsupervised Learning, where the computer receives inputs but tries classify its own labels in much the same way our human brains classify and comprehend data (Ng, 2017). In contrast, Supervised Learning requires pre-labeled data inputs to perform the same “learned” outputs. These two distinct methods, when individually applied to their respective systems, form a new type of input-output system called Neural Networks (N.N.).

In 1962, Frank Rosenblatt, an American scientist of Cornell University was the first to propose N.N., a multi-layered perceptron comprised of neuron-like binary units (Garson, 1998). The three layers, input, hidden, and output, allow computer scientists to constrain the weights and biases of each neuron to produce certain outputs, essentially teaching the system to perform a desired function. To begin, each input neuron shares its data with each hidden layer neuron, which then takes the summation of all inputs to produce each output layer neuron, seen visually in Figure 1.

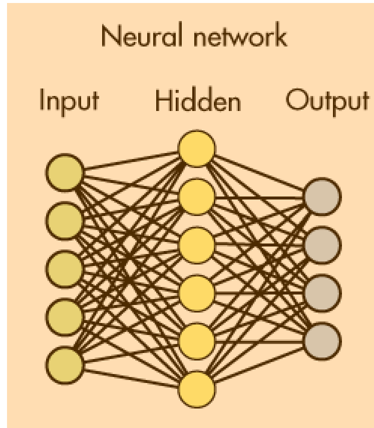

Figure 1: Diagram of a N.N. (Wong, 2017)

With this complex new computational method, A.I. begins to “[solve] problems that only intelligent people had previously been able to solve” (B. & G., 2005). But one problem persists. With countless input, hidden, and output neurons all sharing data in this web-like network, N.N. required massive amounts of computational power in a technologically limited landscape. It wasn’t until the modern development of Deep Neural Networks that A.I. began to take the form we recognize today.

Present

According to the CEO of Amazon, Jeff Bezos, A.I. is currently in “a golden age”, where we are “solving problems with machine learning and artificial intelligence that were in the realm of science fiction for the last several decades.” (“Jeff Bezos, Amazon and the Final Frontier”, 2017). As a result, this new technology is at once reshaping old industries and reimagining the most essential aspects of our futures: our jobs, our daily lives, and even our lifespans. Whereas only a few decades ago, the idea of using machines to solve everyday problems was a thing of the future; but now Amazon use A.I. for air drone delivery, the Echo Dot, and even predicting what customers want to buy (“Jeff Bezos, Amazon and the Final Frontier”, 2017). At this time, understanding A.I. and its capabilities is extremely important to the businesses of tomorrow.

In the present today, we use A.I. to develop models dependent on our inputs and hidden layers to obtain results. Deep Neural Networks are essentially standard Neural Networks with two or more levels of hidden layers, seen visually in Figure 2 (Wong, 2017). Whereas previously Neural Networks were subject to computational limitations, Deep Neural Networks are now creating a new, rich technological landscape--one that demands fast converging and precise algorithms with the potential to solve a whole new realm of problems (Chen & Liu, 2018). Of course, not all algorithms are programmed equally--the development of various models allow A.I. scientists to expand their knowledge of Neural Networks and conceive different applications for I.S.. The following sections will discuss the evolution of our modern Neural Networks, how they build upon each other, and how utilizing a combination of these methods is crucial for drastic technological progression in our society.

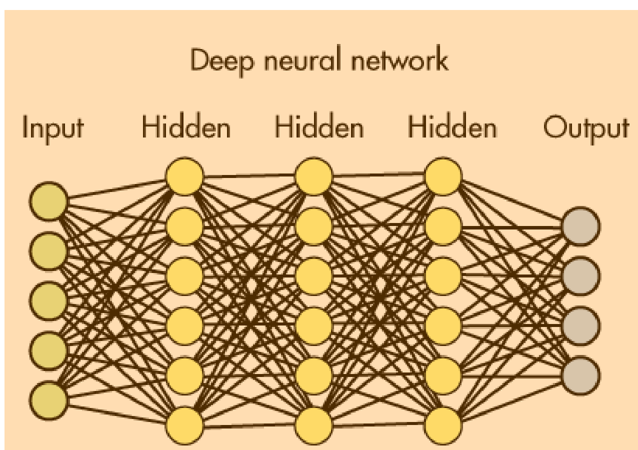

Figure 2: Diagram of a Deep N.N. (Wong, 2017)

Convolutional Neural Networks (CNN)

CNN is an example of an adaptable Neural Network that can be scaled into a Deep Neural Network with more computational layers. CNN operates by breaking up data sets to identify a multitude of features and, as a result, produces outputs with greater accuracy. In these early stages of A.I., CNN allows A.I. to be more accessible, representing a highly adaptable system of machine learning capable of solving a wide range of issues. For example, a CNN was used in the A.I. that defeat the top ranking GO player in his own game (Kohs,2017). “Many convolutional layers [were] employed to enhance the performance” of the A.I., named AlphaGo, which beat world champion Lee Sedol by a score of 4 to 1 (Li & Du, 2018). It is clear that A.I., when equipped with the appropriate level of computational layers, can effectively challenge even the smartest humans on Earth. As A.I. continues to progress and become more widespread, the average intelligence of the everyday person will consequently increase to “avoid becoming obsolete” (Kharpal, 2017) . Elon Musk, billionaire founder of Tesla and SpaceX, even ventures to argue that future “humans [will] need more processing power”(Kharpal, 2017) in order to survive in the drastically different A.I. landscape of the future .

But it is not CNN alone that will provide these huge leaps in A.I. technology. CNN, which operates through a re-iterative backpropagation method of machine learning, can be time consuming and inefficient compared to other models. Figure 3: The Error Function exemplifies the long, re-iterative process that a CNN may employ to reach an output, compare its output to the ideal, and adjust for greater accuracy. Researchers realize that A.I. will need to operate more like the human mind if it is meant to compete with our intelligence, and CNN alone cannot mimic the brain’s complex firing of every possible neuron layer. Rather, a combination of different modeling methods will together form the Artificial Intelligence that will ultimately reshape our society.

$$E(X, \theta) = \frac{(y_1 - f_\theta(\overrightarrow{x_1}))^2}{2} + \frac{(y_2 - f_\theta(\overrightarrow{x_2}))^2}{2}$$

Figure 3: The Error Function (Backpropagation, 2018)

Bidirectional Neural Networks + Long Short Term Memory (Bi-LSTM)

Other notions are necessary to improve certain aspects of the CNN or solve entirely different sorts of application problems. For example, time series problems like predicting stocks proved difficult to solve, producing unreliable outcomes because of the inability to store previous steps. Long Short Term Memory Neural Network models (LSTM) do allow the network to store information, and have thus “become the state of the art models for a variety of machine learning problems” (Greff, 2017). Bi-Directional LSTM (Bi-LSTM) makes the original algorithm “more effective” because it contains parallel layers that propagate in both directions, allowing information to be stored from both directions, forwards and backwards (Li, Yan, Wu, Zhou, 2017).

This model combined with the CNN model allows for the identification of features, and the ability to store information in a reliable, multi-directional way. With both methods working simultaneously, the system can work cyclically: LSTM receives live input data, CNN recognizes features and sorts, Bi-LSTM easily accesses information and updates CNN, and so on. The depth of computational power of these multi-method networks leave their possibilities up to the imagination, possessing the potential to transform our industries and create new ones that we have yet to imagine.

A Multi-Method Approach

The dataset in table 1 demonstrates the power of a multi-method approach to machine learning. In an experiment categorizing 25,000 movie reviews as positive or negative, CNN, LSTM, and Bi-LSTM alone produced significantly lower accuracy and recall averages than the CNN+Bi-LSTM model by a margin of nearly 2% after four iterations. These researchers concluded that the higher computational network CNN+Bi-LSTM “improves the accuracy compared to CNN, LSTM and Bi-LSTM model” (Li, Yan, Wu, Zhou, 2017). Where the technology is still in its infancy, these results show how rapidly the technology is progressing. The CNN+Bi-LSTM method and others like it are stretching our understanding of A.I. and how it could be adapted to further human interests.

Table 1: Movie Review Prediction Results (Li, Yan, Wu, Zhou, 2017)

	iterations	1	5	10	30	average
CNN	accuracy(%)	0.87804	0.89048	0.88768	0.88331	0.88487
	recall(%)	0.87661	0.87464	0.88297	0.87286	0.87677
LSTM	accuracy(%)	0.81036	0.83376	0.84648	0.83696	0.83189
	recall(%)	0.78497	0.89115	0.86034	0.83977	0.84405
Bi-LSTM	accuracy(%)	0.85368	0.82228	0.83208	0.81568	0.83093
	recall(%)	0.84838	0.75530	0.82638	0.80281	0.80821
CNN+Bi-LSTM	accuracy(%)	0.85991	0.89787	0.92843	0.91623	0.90061
	recall(%)	0.88525	0.90122	0.87243	0.94887	0.90194

To solidify the argument that a multi-method approach is essential for the evolution of A.I., consider the experimental platform in Figure 4, set to test the accuracy of different networks to predict the time of expected tool failure:

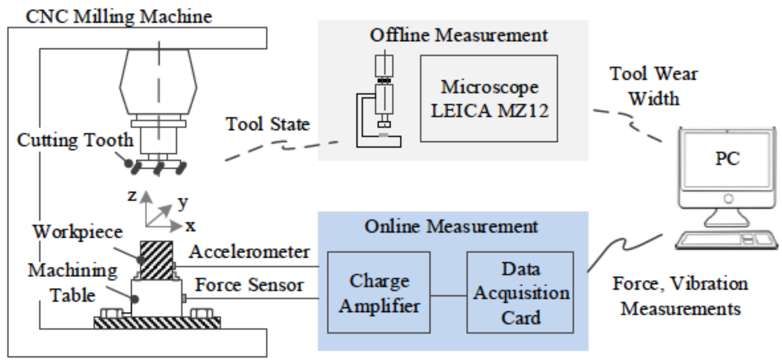

Figure 4: Operation Cycle of Experiment (Zhao, Yan, Wang, Mao, 2017)

This experiment compared a total of 10 different models with 3 different datasets to verify their model’s effectiveness. The results were compared using Mean Absolute Error and Root Mean Square Error to compare the predicted data to the actual data.

Tables 2 & 3: Mean Absolute Error and Root Mean Square Error Response (Zhao, Yan, Wang, Mao, 2017)

Category	Methods	Datasets		
		c1	*c4*	*c6*
Regression Models	LR	24.4	16.3	24.4
	SVR	15.6	17.0	24.9
	MLP	24.5	18.0	24.8
Recurrent Models	RNN	13.1	16.7	25.5
	Deep RNN	7.8	9.4	19.3
	LSTMs	19.6	15.6	25.3
	Deep LSTMs	8.3	8.7	15.2
	BLSTMs	9.9	10.8	15.7
	Deep BLSTMs	8.7	6.7	14.4
Our Model	CBLSTMs	**7.5**	**6.1**	**8.1**

Category	Methods	Datasets		
		c1	*c4*	*c6*
Regression Models	LR	31.1	19.3	30.9
	SVR	18.5	19.6	31.5
	MLP	31.2	20.0	31.4
Recurrent Models	RNN	15.6	19.7	32.9
	Deep RNN	12.5	11.8	22.9
	LSTMs	23.9	20.8	32.4
	Deep LSTMs	12.1	10.2	18.9
	BLSTMs	12.3	14.7	20.8
	Deep BLSTMs	11.5	9.1	18.9
Our Model	CBLSTMs	10.8	7.1	9.8

The experimental results show that the CBLSTM, or CNN+Bi-LSTM, has the lowest error for both error functions. The CNN+Bi-LSTM (CBLSTM) model uses a CNN layer to extract the favored features, and a Bi-LSTM layer to “encode the temporal information and capture long term dependencies in forward and backwards ways” (Zhao, Yan, Wang, Mao, 2017). This combination of methods gives the system a depth of computational power to solve extremely complex problems. At this point, A.I. has overcome formerly debilitating problems to reach this current peak. More accessible, affordable, and adaptable than ever, A.I. is now poised for widespread integration into the global market.

Future

As Artificial Intelligence pervades our everyday lives, as more complex systems become accessible, and as the systems adapt to meet society’s ever-changing needs, A.I. technology will continue to evolve at a rapid rate. Andrew Ng, Chief Scientist at Baidu University, agrees: “Just as electricity transformed almost everything 100 years ago, today I actually have a hard time thinking of an industry that I don’t think AI will transform in the next several years” (Ng, 2017). We can expect A.I. to power a brand new infrastructure, a literal new landscape, comparable in scale to the electric boom of the industrial revolution--global effects of electricity on the landscape appear in Figure 5& 6 (Mackenzie, 2012) (Locke, 2017). Where people were once afforded the luxury of indoor lights and longer hours of productivity, now Artificial Intelligence will bring a similar ease of living, freeing up time and resources for society to advance like never before, into industries we have yet to imagine.

Figure 5 & 6: Photo of Italy & America at Night (Mackenzie, 2012) (Locke, 2017)

The future of Artificial Intelligence in regards to Neural Networks is still yet to be uncovered. More research is necessary to discover the next phase of the multi-method approach, whether it incorporates more layers, or looks altogether different. The foreseeable next steps in this evolution is the application of our new methods--actually putting our research toward the solutions for our world's biggest problems. From here, the technology will continue to evolve through the side-by-side evolution of our ideas and creativity. To reference Ng (2017) once again: where our current reliance on electricity has changed the way humans perceive our reality, so too will A.I. literally change the landscape, giving us a different view of how the world could and should operate. Just as the internet can be accessed anywhere through smartphones, people will experience similar accessibility with new A.I. technology. With A.I. being more applicable, people will be forced to adapt, leaving present notions of reality behind to help aid the betterment of self-preservation and society. It is inevitable for jobs to become automated as some of these tasks are optimally performed by the A.I. compared to humans. Through our symbiotic relationship with A.I., our society will push to adapt and strive for greater goals.

Conclusion

A.I. has made considerable leaps since John McCarthy coined the term in 1955. We now see a widespread acceptance of Artificial Intelligence pervading every facet of our lives, from Amazon’s Echo Dot, to their Air Delivery Drones. But even in the early phases of A.I., we still wondered if machines could ever possess the ability to “think”, and if they could, how would it be possible? This challenge led to the development of perceptron’s, Neural Networks that could be compounded and optimized for an endless number of intended applications. With this level of computational power at our disposal, the application of A.I. is only limited by our imagination and the data needed to train the network. It is hard to predict how our understanding and application of A.I. will manifest in the future, but A.I.’s foothold on future industry is undeniable. As the “new electricity,”: A.I. will illuminate the way for industries to advance like never before, changing the face of technology and ultimately the landscape as a whole.

References

B., & G., B. (2005, December 22). A (Very) Brief History of Artificial Intelligence. *AI Magazine*.

Backpropagation. (n.d.). Retrieved March 26, 2018, from https://brilliant.org/wiki/backpropagation/

Burger, C. (2017, February 06). Google DeepMind's AlphaGo: How it works. Retrieved March 25, 2018, from https://www.tastehit.com/blog/google-deepmind-alphago-how-it-works/

Byrnes, N. (2016, April 04). Intelligent software and machines are showing up in new places. Retrieved March 25, 2018, from https://www.technologyreview.com/s/600986/ai-hits-the-mainstream/?set=600969

Chen, C. L., & Liu, Z. (2018). Broad Learning System: An Effective and Efficient Incremental Learning System Without the Need for Deep Architecture. *IEEE Transactions on Neural Networks and Learning Systems*,*29*(1), 10-24. doi:10.1109/tnnls.2017.2716952

Garson, G. D. (1998). *Neural networks: An introductory guide for social scientists*. London: Sage.

Greff, K., Srivastava, R., Koutnik, J., Steunebrink, B., & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. *IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*, *28*(10).

He, T., Zhang, Y., Droppo, J., & Yu, K. (2016). On training bi-directional neural network language model with noise contrastive estimation [Abstract]. *2016 10th International Symposium on Chinese Spoken Language Processing (ISCSLP)*. doi:10.1109/iscslp.2016.7918423

Jeff Bezos has big plans for Amazon. (2017, November 24). Retrieved March 25, 2018, from https://www.intheblack.com/articles/2017/11/01/amazon-power-invention

Kharpal, A. (2017, February 13). Elon Musk: Humans must merge with machines or become irrelevant in AI age. Retrieved March 25, 2018, from https://www.cnbc.com/2017/02/13/elon-musk-humans-merge-machines-cyborg-artificial-intelligence-robots.html

Gary Krieg, Josh Rosen, Kevin Proudfoot, Greg Kohs. 2017. *AlphaGo*

Li, F., & Du, Y. (2018). From AlphaGo to Power System AI: What Engineers Can Learn from Solving the Most Complex Board Game. *IEEE Power and Energy Magazine*,*16*(2), 76-84. doi:10.1109/mpe.2017.2779554

Li, S., Yan, Z., Wu, X., Li, A., & Zhou, B. (2017). A Method of Emotional Analysis of Movie Based on Convolution Neural Network and Bi-directional LSTM RNN [Abstract]. *2017 IEEE Second International Conference on Data Science in Cyberspace (DSC)*. doi:10.1109/dsc.2017.15

Locke, C. (2017, July 12). Dazzling Photos Let You Orbit Earth Aboard the Space Station. Retrieved March 25, 2018, from https://www.wired.com/2017/06/tim-peake-iss-photos/

Mackenzie, C. (2012, July 09). From shimmering cities to the dark side of the world: Stunning satellite images which show how we use light and energy. Retrieved March 25, 2018, from http://www.dailymail.co.uk/news/article-2170422/Revealed-The-stunning-images-Europe-wastes-energy-pollutes-light-Africa-South-America-Asia-darkness.html

Ng, A. (2017, February 02). Andrew Ng: Artificial Intelligence is the New Electricity. Retrieved 2018, from https://www.youtube.com/watch?v=21EiKfQYZXc&t=258s

Rajaraman, V. (2014). JohnMcCarthy — Father of artificial intelligence [Abstract]. *Resonance,19*(3), 198-207. doi:10.1007/s12045-014-0027-9

Sandewall, Erik (2014). A perspective on the early history of artificial intelligence in Europe

Wong, W. (2017, July 18). What's the Difference Between Machine Learning Techniques? Retrieved 2018, from http://www.electronicdesign.com/automotive/what-s-difference-between-machine-learning-techniques

Zahedi, F. (1991). An Introduction to Neural Networks and a Comparison with ... Retrieved 2018.

Zhao, R., Yan, R., Wang, J., & Mao, K. (2017). Learning to Monitor Machine Health with Convolutional Bi-Directional LSTM Networks [Abstract]. *Sensors,17*(2), 273. doi:10.3390/s17020273