

Is the Algorithm of Artificial Neural Network a Deduction or Induction? Discussion between Natural Sciences, Mathematics and Philosophy

Jyh-Woei Lin

ABSTRACT

The algorithm of artificial neural network (ANN) has been defined as a supervised learning and heuristic algorithms. In training an ANN model, big data is necessary to use as training data to obtain perfectly accurate predicted data. However, big data really have no clear definition. Therefore, adding new training data to re-train an ANN model, by which can improve the predicted accuracy. This action of re-training this ANN model with added new training data is repeated to approach the laws of physics that is accessed to the principle of induction e.g., empirical formulas. However, accessing the principle of induction is limited. If the deduction is found using an ANN model, then approach of this ANN model with added new training data is also performed repeatedly to access the principle of deduction e.g., theory formulas. However, accessing the principle of deduction is also limited. It means the law cannot be easily deduced for an ANN model. Therefore, the algorithm of an ANN is not the canonical classical methods. On the other hand, the algorithm of an ANN does not belong to mathematical induction and deduction.

Keywords: Artificial neural network (ANN), induction, empirical formulas, deduction, theory formulas, canonical classical methods.

Published Online: September 30, 2021

ISSN: 2736-5492

DOI : 10.24018/ejcompute.2021.1.4.29

J.W. Lin*

Binjiang College, Nanjing University of Information Science & Technology, China.

Department of Electrical Engineering, Southern Taiwan University of Science and Technology, Taiwan.

(e-mail: pgjw11966@gmail.com)

**Corresponding Author*

I. INTRODUCTION

The algorithm of an artificial neural network (ANN), which is based on the multilayer perceptron (MLP) and an error back propagation (EBP) algorithm, is a supervised learning and heuristic algorithms [1], [2]. A large amount of training data is necessary to train an ANN model. A large amount of training data is called big data. Thereby, perfectly accurate predicted data of an ANN model can be obtained [3], [4]. However, the most common argument is to obtain real big data that is impossible because the big data have no clear definition. It claims that perfectly accurate predicted data is impossible to obtain.

Usually, some mathematical and physical problems are solved by the canonical classical methods, which are stated in next section. The problems of natural hazards such as seismic hazards are usually the non-linear problems [5]. However, to solve these complex non-linear problems, especially for the problems of environment and earth science e.g., Total electron content (TEC) precursors before the earthquakes [6], using the canonical classical methods, the calculated processing is too complicated. Simultaneously, an exact and perfect solution may or cannot be obtained [7], [8]. Therefore, an optimal solution can be given using a heuristic algorithm [9]. An optimal solution cannot access to have perfect accuracy.

The algorithm of an ANN belongs to the heuristic and

search algorithms [10], [11]. It is also referred to as metaheuristic algorithm [12]. The heuristic algorithm is a heuristic designed to find a heuristic algorithm such as random search algorithm [9], [11]. As stated previously, this algorithm can provide the sufficiently and optimization solutions to the complex non-linear problems. That can claim this is the advantage of an ANN model.

In training an ANN model, sometimes extra random values are added to obtain the stable computing efficiency [13]. Simultaneously, two types of datasets are used to train an ANN model; one is referred to as training data which are true values. Another is random values in a range referred to the initial weight and bias [14]. It means that the initial weight and bias are randomly selected in a range. Different ranges give rise to different training results of an ANN model. Different sizes of training data also result in different results of prediction. The learning rates and activation functions are used in training an ANN model. Different learning rates and activation functions also lead to different training results of an ANN model. Therefore, an ANN model has the complex relationships between inputs and outputs. In this respect, the algorithm of an ANN is not based on the empirical and theory formulas. Moreover, the algorithm of an ANN was performed only simulate some neurons of human brain and it is not parallel distributed processing (PDP) [15]. To simulate a complete and perfect ANN as the biological Neuron Network cannot access.

II. CANONICAL CLASSICAL METHOD

Many algorithms belong to the canonical classical methods [16] such as principal component analysis (PCA) [17], multivariate analysis, and regression analysis [4], [18], [19]. They have the empirical and theory formulas. The canonical classical methods include mathematical induction and deduction. Usually, they do not have the concept of network computing.

III. DISCUSSION

Physics is one of the subjects based on natural philosophy [20]. Typically, physicists express proposed phenomena using hypotheses and the laws of nature. If the hypotheses can be verified through rigorous experiments, they can be classified the laws of physics. However, like many other natural science theories, these laws cannot be proven, and their accuracy relies on repeated experiments or a large amount of information to confirm their reliability. A good example is the Mendelian inheritance of biological features, which is also referred to as Mendel's laws [21]. Mendel discovered that when purebred white flower and purple flower pea plants (P- generation) were crossed, the result was not a blend. Instead of the two, the offspring (F1 generation) was purple-flowered. When Mendel self-fertilized the F1 generation pea plants, the number of purple and white flowers in the F2 generation was 705 and 224, respectively, and their ratio was 3.15:1. However, this ratio should ideally be of 3:1 based on the Punnett square [22]. As F2 generation number increases, the ratio approaches 3:1. Therefore, this experiment must always have increased number of F2 generation near the theoretical value [23], [24].

As mentioned above, like many other natural scientific theories, these laws cannot be proofed. Only the accuracy of predicted values can be increased by repeatedly adding new training data to update an ANN model. Reference [25] built so called the Embedded Earthquake Richter Magnitude (ML) Prediction Backpropagation Neural Network (EEMPBPNN) model to approach more accurate predicted Richter Magnitude (ML) with added new earthquake catalogues as additional training data. Therefore, the EEMPBPNN model was re-trained to an updated better model with more accurate predicted values.

After adding new training data, this re-trained ANN model improves the accuracy of the predicted values. This is consistent with the point of view that when experiments are performed repeatedly to access the laws of physics e.g., empirical formulas, and degree of precision increases. An ANN training process is based on supervised learning to access the principle of induction. However, accessing to the principle of induction is limited. Therefore, it is not real mathematical induction [26]. If the deduction using ANN model can be found, then approach of this ANN model with new training data should be performed repeatedly to the principle of deduction e.g., theory formulas. However, the real law cannot be easily deduced for an ANN model. It is also not real mathematical deduction [27]. Therefore, an ANN model cannot become empirical and theory Formulas.

Most people want to use mathematical induction, e.g.,

inferring future events based on past events. For example, in the past, people observed that the Sun rises in the east. Thus, they inferred that the Sun would still rise in the east tomorrow. The disadvantage of this method is that, if in the past, the presence of certain conditions caused something to happen, and those conditions do not exist in the future, that thing will not happen again. This means that the people will see the things from a one-sided point of view, and this is euphemistically called induction. In fact, this representation evidence or observation target is greater, and people have more confidence in the conclusions of induction. However, it is difficult to obtain comprehensive information. A piece of information always comes, and such a piece of data compilation is just finished and hopes to learn a comprehensive case. However, it cannot be confirmed whether all the information has been obtained. People do not know if data process is complete and whether they have committed to the correct direction. They cannot tell the whole story. They can only try to not make mistakes and minimize the extent of error. Using past events to infer future events is based on the condition that the assumptions are unchanged, for the probability of an inference being correct. The existing knowledge of natural science also has some assumptions to deduce theory formulas. Therefore, only a relative accuracy, that is accurately than previously recognized and is closer to the right. However, the so-called relative truth (assuming it exists) has a large section of a great distance. An ANN model cannot tell the truth. The development of full artificial intelligence (AI) could spell the end of the human race, which is the statement by Prof Stephen Hawking (BBC New). It should not happen. His worry is not necessary.

IV. CONCLUSION

The canonical classical methods include mathematical induction and deduction. The algorithm of an ANN belongs to a supervised learning and heuristic algorithms. Training an ANN model, the training data are big data. However, big data have no clear definition. Therefore, re-training an ANN model with added new training data, can improve the predicted accuracy. The action of re-training this ANN model with new training data is repeated to approach the laws of physics and is accessed the principle of induction e.g., empirical formulas. However, accessing the principle of induction is limited. When training an ANN model for finding the deduction, then the approach of this ANN model with new training data is performed repeatedly to find the principle of deduction e.g., theory formulas. The law cannot be easily deduced for an ANN model. Therefore, the algorithm of an ANN is not the canonical classical methods. On the other hand, the algorithm of an ANN is not the canonical classical methods with the empirical and theory formulas.

APPENDIX

This article does not contain any studies with human participants or animals performed by any of the authors.

ACKNOWLEDGMENT

The author is grateful to his mother who was dead in 2016.

The author is also grateful the supporting of Prof. Dr Yuan Mei in Taiwan and all of my friends in Taiwan and China.

REFERENCES

- [1] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning representations by back-propagating errors," *Nature* 323, 533–536, 1986. doi: 10.1038/323533a0.
- [2] I. Hussain and D. M. Thounaojam, "SpiFoG: an efficient supervised learning algorithm for the network of spiking neurons," *Scientific Reports* 10, 13122, 2020. doi: 10.1038/s41598-020-70136-5.
- [3] M. Hilbert and P. López, "The World's Technological Capacity to Store, Communicate, and Compute Information," *Science* 332 (6025), 60–65, 2011. doi: 10.1126/science.1200970.
- [4] J. W. Lin, "An empirical correlation between the occurrence of earthquakes and typhoons in Taiwan: a statistical multivariate approach," *Natural Hazards* 65, 605–634, 2013. doi: 10.1007/s11069-012-0382-3.
- [5] S. Yaghmaei-Sabegh and R. Motallebzade, "An effective procedure for seismic hazard analysis including nonlinear soil response," *Natural Hazards* 64(2), 1731–1752, 2012. doi: 10.1007/s11069-012-0332-0.
- [6] J. W. Lin and J. S. Chiou, "Detecting Total Electron Content Precursors Before Earthquakes by Examining Total Electron Content Images Based on Butterworth Filter in Convolutional Neural Networks," *IEEE Access* 8, 110478 – 110494, 2020. doi: 10.1109/ACCESS.2020.3001337.
- [7] N. R. Jennings, "On agent-based software engineering," *Artificial Intelligence* 117 (2), 277–296, 2000. doi: 10.1016/S0004-3702(99)00107-1.
- [8] J. Stoer and R. Bulirsch, "Introduction to Numerical Analysis," 12, 746, *Springer Science & Business Media*, 2013.
- [9] Y. Wu, G. Peng, H. Wang and H. Zhang, "A Heuristic Algorithm for Optimal Service Composition in Complex Manufacturing Networks," *Complexity* 7819523, 2019. doi:10.1155/2019/7819523.
- [10] J. Kennedy, "Particle Swarm Optimization," *Encyclopedia of Machine Learning*, Springer, 760–766, 2010.
- [11] L. M. Almeida and T. Ludermir, "An improved method for automatically searching near-optimal artificial Neural Networks," *IEEE World Congress on Computational Intelligence* 2235–2242, 2008. doi: 10.1109/IJCNN.2008.4634107.
- [12] X. S. Yang, "Nature-inspired Metaheuristic Algorithms," *Luniver Press*, pp.148, 2010.
- [13] L. Bianchi, M. Dorigo, L. M. Gambardella and W. J. Gutjahr, "A survey on metaheuristics for stochastic combinatorial optimization," *Natural Computing* 8, 239–287, 2009. doi: 10.1007/s11047-008-9098-4.
- [14] T. M. Khoshgoftaar, J. V. Hulse and A. Napolitano, "Supervised Neural Network Modeling: An Empirical Investigation Into Learning From Imbalanced Data With Labeling Errors," *IEEE Transactions on Neural Networks* 21 (5), 813 – 830, 2010. doi: 10.1109/TNN.2010.2042730.
- [15] J. W. Lin, "Artificial Neural Network Related to Biological Neuron Network: A Review," *Advanced Studies in Medical Sciences* 5(1), 55–62, 2017. doi: 10.12988/asms.2017.753.
- [16] P. Legendre, and H. J. B. Birks, "From Classical to Canonical Ordination," in *Tracking Environmental Change Using Lake Sediments: Data Handling and Numerical Techniques*, 201–248, 2012. doi:10.1007/978-94-007-2745-8_8
- [17] J. W. Lin, "Is it possible to trace an impending earthquake's occurrence from seismo-ionospheric disturbance using principal component analysis? A study of Japan's Iwate-Miyagi Nairiku earthquake on 13 June 2008," *Computers & Geosciences*, 37 (7), 855–860, 2011. doi: 10.1016/j.cageo.2011.02.004.
- [18] I. Koch, "Analysis of multivariate and high-dimensional data," (Vol. 32). Cambridge University Press, 2013.
- [19] J. W. Lin, "Rainfall-triggered ordinary earthquakes in Taiwan: a statistical analysis," *Hydrological Sciences Journal* 59 (5), 1074–1080, 2014. doi: 10.1080/02626667.2014.900177.
- [20] D. P. Christopher, 2005, "Physics Made Simple: A Complete Introduction to the Basic Principles of This Fundamental Science," *The Crown Publishing Group* pp.208.
- [21] C. Correns, "G. Mendel's law concerning the behavior of progeny of varietal hybrids," *Genetics* 35(5), part 2, 33–41, 1950.
- [22] R. C. Punnett, "Early days of genetics," *Heredity* 4, 1–10, 1950. doi: 10.1038/hdy.1950.1.
- [23] I. Miko, "Gregor Mendel and the principles of inheritance," *Nature Education* 1(1):134, 2008.
- [24] E. C. Solomon, D. W. Martin, L. R. Berg, 2015, "Biology," 10th Edition, *Cengage Learning*, pp.1440, 2015.
- [25] J. W. Lin, C. T. Chao, and J. S. Chiou, "Determining Neuronal Number in Each Hidden Layer Using Earthquake Catalogues as Training Data in Training an Embedded Back Propagation Neural Network for Predicting Earthquake Magnitude," *IEEE Access* 6, 52582 – 52597, 2018. doi: 10.1109/ACCESS.2018.2870189.
- [26] M. Friend, N. B. Goethe, and V. S. Harizanov, "Induction, Algorithmic Learning Theory, and Philosophy," *Springer Netherlands*, pp.290, 2007. doi: 10.1007/978-1-4020-6127-1.
- [27] P. N. Johnson-Laird and R. M. J. Byrne, "Deduction (Essays in Cognitive Psychology)," *Psychology Press* pp.256, 1991.
- [28] J. Stoer and R. Bulirsch, "Introduction to Numerical Analysis," 12, 746, *Springer Science & Business Media*, 2013.



JYH-WOEI LIN received the B.Sc. degree from the Department of Physics, Chung Yuan Christian University, Chung Li City, Taiwan, in 1989, the M.Sc. degree from the Institute of Geophysics, National Central University, Chung Li City, in 1991, the Ph.D. degree from the Institut für Geophysik, Clausthal-Zellerfeld, Technische Universität Clausthal, Germany, in 2000, and the Ph.D. degree from the Department of Electrical Engineering, Southern Taiwan University of Science and Technology, Tainan, Taiwan, in 2019. Since 2019, he is a researcher in the Department of Electrical Engineering, Southern Taiwan University of Science and Technology. Until 2021, he has 71 SCI (SCIE) and EI papers include artificial intelligence, space physics, geophysics, medical sciences, and remote sensing; especially four books publications in Germany and two books in USA. After September 2020, he is as a Professor in Binjiang College, Nanjing University of Information Science & Technology, Wuxi, Jiangsu Province, China.