Application of Machine Learning Algorithm Based on Neural Network Technology

Yiqiang Lai*
South China Business College, Guangdong University of Foreign Studies, Guangzhou 510545, China
*Corresponding author e-mail: laiyq1982@foxmail.com

Abstract. Neural networks have strong characteristics for processing data and information. At the same time, the current computer technology is also very advanced, and many kinds of very powerful information technologies have been developed under the promotion and promotion of modern science and technology. Therefore, relevant personnel will carry out advanced technology and neural network structure methods. Fusion, and then an artificial neural network was established on this basis. In a broad sense, machine learning refers to how to enable a machine to acquire relevant knowledge through autonomous learning, and the purpose is to enable the machine to have relevant skills similar to what people need to acquire knowledge. The research in this article aims to explore the machine learning algorithms based on neural network technology, and through literature research methods, case analysis methods, etc., to have an in-depth understanding of machine learning algorithms in neural network technology, and then through the analysis of machine learning algorithms in neural network technology The learning advantage and its influencing factors are designed based on the machine learning algorithm in the neural network and experimented. Experimental results show that LSTM performs well in replication tasks, and the performance of LSTM even far exceeds that of NTM, but the performance of LSTM in addition and multiplication tasks is much lower than that of NTM. Although the accuracy of NTM on the test set is higher than that of LSTM and RNN, the performance of the model is still relatively poor.

Keywords: Neural Network, Machine Learning, Algorithm Learning, Writing Mechanism

1. Introduction
The main classification of neural networks can be roughly divided into two types: neural networks in neurobiology and artificial neural networks [1-2]. Biological neural network, as the name implies, is a network composed of the nervous system of a living body. With the continuous development of science and technology and the deepening of bionic technology, the eyes of artificial intelligence have gradually turned to this field [3-4]. The so-called artificial neural network is to compare the various information transmission and processing methods of a biological neural network with other neural
networks. In an artificial neural network system, it simulates a biological brain and nervous system. Neurons are connected to each other to form a complex computing network system similar to biological neural networks [5-6]. Artificial neural network has many advantages, such as parallel processing, autonomous learning, strong nonlinear ability and high fault tolerance. It is currently widely used in image recognition, machine learning, pattern recognition and other fields [7-8].

Artificial intelligence has made remarkable achievements in recent years and has gained a high degree of popularity. An important reason behind it is the rise of deep learning technology based on neural networks, and the subject of machine learning to which deep learning belongs. The content of the research is actually to train and adjust the parameters of the model on a large data set [9]. Due to the large amount of data processed by machine learning, the amount of calculation required is extremely huge. The problem that is caused is that the time required for calculation is often very long. Therefore, people have looked for various ways to speed up the calculation, one of which is It is with the help of quantum computing, which has brought a new cross-discipline—quantum machine learning, which is a very cutting-edge problem studied in the academic circles in recent years[10].

This article introduces NTM-based algorithm learning. First, it introduces in detail how to apply NTM to algorithm learning tasks, including transforming algorithm learning into sequence-to-sequence problems, and the formal definition of NTM. Then it analyzes the influence of the size of the memory module on the performance of the model, and finally compares the performance differences of NTM, RNN, GRU and LSTM on algorithm learning tasks.

2. Research on Machine Learning Algorithms Based on Neural Network Technology

2.1 Advantages of Neural Network Technology
(1) Speed advantage: Although using software to simulate neural networks on a computer processor has certain flexibility and does not require the use of hardware, the speed of software simulation of neural networks is very slow and usually not suitable for all occasions, so hardware is used Computational processing of convolutional neural networks is a practical and effective method to solve the practical application of neural networks. Hardware implementation is in parallel processing and distributed computing, and its acceleration effect can even reach several orders of magnitude improvement, so hardware implementation has very high computational efficiency.

(2) Fault tolerance: Because serial implementation applications have certain instabilities, when any problem occurs in any link, the entire processor system will lose function, and the main reason for this problem is that the single processor is not enough Redundancy. Even if it is a multi-core processor, as far as it is concerned, is there an effective fault-tolerant mechanism? However, compared to hardware, when an exception occurs in a certain part, due to the use of a parallel structure and a distributed structure, the overall performance can also be reduced, so that the application can operate normally. Therefore, compared with the software-based neural network, the realization of the digital hardware circuit of the neural network has incomparable advantages.

As a feedforward network structure, neural network has a high degree of independence between layers. Each layer of neural network is calculated independently, and there is no feedback data between layers. Therefore, the neural network is a highly parallel artificial neural network structure.

2.2 Influencing Factors
(1) Data set
The sample variables in the research work mainly have the following two characteristics. First, one big data sample should be selected as much as possible, so as to promote its confidence time to meet the research requirements. Second, the samples should be dispersed as much as possible, and they should not have too strong mutual relations.

(2) Initialization
 Initialization mainly includes two components, namely weight and threshold. These two factors will have a certain impact on the operation process in training, so that every node on the surface is
moved, and the maximum characteristic of the surface. The problem is that they are not balanced enough, so we need to change their convergence goal through the initialization value, so that it can be further optimized on the original basis. Under normal circumstances, the initial value has strong randomness, which makes the difference between the various networks too large, which leads to the optimization can only be limited to some areas, which cannot directly affect the final convergence goal, which will inevitably be. It directly affects the stability of the training results, making it more difficult to control.

(3) Training speed
And the speed of training may also be affected by many factors. In addition to the structure or algorithm, the basic parameters may also cause the speed of training to slow down. In order to limit the training time to a minimum range, when selecting the hidden layer, you must carefully check the number of their nodes. At the same time, you must carefully consider the structural issues, otherwise the structure of the sample may no longer be possessive effective. In addition, in order to prevent the training period from being too long, it is necessary to choose the most appropriate node count as far as possible to avoid the number of nodes exceeding the required matching area.

2.3 Algorithm Analysis
The write mechanism combines the input gate and forget gate in LSTM. Therefore, the writing mechanism of NTM is divided into two steps: the first step is to erase the memory, and the second step is to write the memory.

At time t, the write head outputs a weight vector \( w^{(t)} \), an E-dimensional elimination vector \( e^{(t)} \) and an E-dimensional addition vector \( a^{(t)} \). Each element of \( e^{(t)} \) belongs to the interval \([0, 1]\). Then the value of the memory matrix can be calculated according to formula (1) to formula (2):

\[
e^{(t)} = \sigma(W^e h^{(t)} + b^e) \quad (1)
\]

\[
a^{(t)} = W^a h^{(t)} + b^a \quad (2)
\]

\[
M^{(t)} = M^{(t-1)} \odot [1 - W^e (e^{(t)})] + W^a (a^{(t)})^T \quad (3)
\]

Among them, 1 in formula (3) represents a matrix of all ones with a size of \( N \times E \), \( h^{(t)} \) represents the output of the controller at time t, and \( W^e, b^e \) are the weights and biases corresponding to the elimination vector and the addition vector, respectively.

3. Experimental Research on Machine Learning Algorithms Based on Neural Network Technology
The program code in this experiment is based on the deep learning framework Tensorflow. The machine running the program uses a cloud server with an 8-core CPU, 64G memory, 100G solid-state hard drive and 1-core Tesla M60 GPU.

3.1 Evaluation Index
(1) The accuracy of the model refers to the accuracy of the model on the test set after the training. The accuracy of the model on the test set refers to the ratio of the number of samples that the model predicts correctly to the number of all test samples. It should be noted that for a test example, the model will output a prediction sequence. Only when the entire sequence is correct can the prediction of this test sample be considered correct. As long as the prediction for one time step is incorrect, then the prediction is wrong.

(2) The training speed can be measured by the number of training times, the number of samples used for training each time, and the training time. In this experiment, the number of training times and the number of training samples for each task are the same, so the training time can be measured. To measure training speed. For the same model and the same task, the longer the training time, the slower the training speed; the shorter the training time, the faster the training speed.
3.2 Network Structure
The experiments in this chapter will involve four neural network structures, including:

1. Ordinary RNN. It contains a hidden layer with 128 memory cells.
2. LSTM. It contains a hidden layer with 128 memory cells.
3. GRU. Its setting is similar to LSTM, including a hidden layer, the hidden layer has 128 memory units.
4. NTM.

3.3 Network Hyperparameters and Their Tuning
For each task, the hyperparameters related to the task need to be adjusted. The values of the desired parameters are shown in Table 1. The sequence length represents the length range of the input sequence. For example, in a copy task, the sequence length ranges from 2 to 11. It should be noted that the input sequence includes the terminator ",", the plus sign "+" and the multiplication sign "×". Therefore, the actual number of characters to be memorized by the model is smaller than the sequence length. The minimum length of the sequence to be copied by the copy algorithm is 1 and the maximum is 10; in addition, the length of each addend ranges from 1 to 10; in multiplication, the length of each multiplier ranges from 1 to 10.

| task   | Sequence length | Memory size | Training times |
|--------|-----------------|-------------|----------------|
| copy   | [2,11]          | 10×8        | 50000          |
| addition | [4,22]      | 21×10       | 150000         |
| multiplication | [4,22]   | 30×10       | 300000         |

4. Experimental Analysis of Machine Learning Algorithms Based on Neural Network Technology
The experiments presented in this article compare the performance of different models on algorithm learning tasks, including RNN, LSTM, and NTM. The experimental tasks are copy, addition and multiplication. Each task trains a model independently. In the replication task, the model was trained with 50,000 batches; in the addition task, the model was trained with 150,000 batches; in the multiplication task, the model was trained with 300,000 batches. In order to compare the performance differences of different models, this article draws Figure 1, Figure 2 and Figure 3.
Fig. 1, Fig. 2 and Fig. 3 show the graphs of the accuracy of the model as a function of the number of model training. From top to bottom are copy, addition, and multiplication. The vertical axis represents the accuracy of the model on the test set, and the horizontal axis represents the number of training sessions. The blue solid line represents NTM, the green solid line represents LSTM, and the red solid line represents RNN. It can be seen from the figure that LSTM performs well in replication tasks, and the performance of LSTM even exceeds NTM, but the performance of LSTM in addition and multiplication tasks is much lower than that of NTM.

Although the accuracy of NTM on the test set is higher than that of LSTM and RNN, the performance of the model is still relatively poor. The specific performance is that the accuracy of NTM in algorithm learning is not high. Although the accuracy of NTM on algorithm learning tasks is
higher than the other two. For example: obetter than LSTM 18% higher, 45% higher than RNN. But the accuracy of the model is below 90%. Esn the addition task, after 150,000 batch training, the accuracy of NTM is 17% higher than that of LSTM and 47% higher than RNN; on the multiplication task, after 300,000 batch training, the accuracy of NTM is pecially in the addition task, after NTM has trained 150,000 batches, the accuracy of the model is only 77%, and from the trend of the curve, the rising speed of the curve is gradually slowing down. This shows that the learning speed of the model is slowing down, and the accuracy of the model is difficult to improve.

5. Conclusions
Neural network has a high amount of processing data and information. At the same time, the current computer technology is also very advanced, and many kinds of very powerful information technologies have been developed under the promotion and promotion of modern science and technology. Therefore, we will now carry out advanced technology and computer neural network structure methods. A holistic combination, and on this basis, it is concluded that too large a memory module will increase the training time of the model, and a too small memory module will reduce the accuracy of the model. In addition and multiplication tasks, although the accuracy of NTM is higher than the other three recurrent neural network models, the accuracy is below 90%.

Authors
Yiqiang Lai, he Received a master degree in computer technology from Jinan University, China, in 2010. He is now working in the faculty of information science technology, South China Business College, Guangdong University of Foreign Studies. His research interests include database technology, image processing, machine learning, etc.

References
[1] Liang S, Gan F. Research and Application of Balance Control of two-wheeled Robots Based on Machine Learning. Computer Measurement & Control, 2011, 19(12):2972-2974.
[2] Mirzal A. Nonparametric Tikhonov Regularized NMF and Its Application in Cancer Clustering. IEEE/ACM Transactions on Computational Biology & Bioinformatics, 2014, 11(6):1208-1217.
[3] Song S, Bing L, Li K, et al. A Novel PSO Algorithm Model Based on Population Migration Strategy and its Application. Journal of Computers, 2011, 6(2):280-287.
[4] Xu Z, Hu T, Peng Z. Research on classification of spiral galaxies and stars based on convolutional neural network. Journal of Physics: Conference Series, 2020, 1684(1):012032 (5pp).
[5] Peker M, En B, Delen D. Computer-Aided Diagnosis of Parkinson's Disease Using Complex-Valued Neural Networks and mRMR Feature Selection Algorithm. Journal of Healthcare Engineering, 2015, 6(3):281-302.
[6] Zhang F, Liu Y, Chen C, et al. Fault diagnosis of rotating machinery based on kernel density estimation and Kullback-Leibler divergence. Journal of Mechanical Science and Technology, 2014, 28(11):4441-4454.
[7] Zhu X, Zhang Y, Zhu Y. Intelligent fault diagnosis of rolling bearing based on kernel neighborhood rough sets and statistical features. Journal of Mechanical Science & Technology, 2012, 26(9):2649-2657.
[8] Li L, Wu K, Zhou D W. Extraction algorithm of mining subsidence information on water area based on support vector machine. Environmental Earth Sciences, 2014, 72(10):3991-4000.
[9] Bing S. Comprehension and Thinking of Complex Network Construction and Clustering Based on Neural Network Algorithm. MATEC Web of Conferences, 2018, 173(6638):03040.
[10] Zeng S C, Chen Y, Guo Z H, et al. The Practice Research of Power System Operation Mode Calculation Based On Machine Learning. Beijing Ligong Daxue Xuebao/Transaction of Beijing Institute of Technology, 2017, 37:193-197.