Neural Network Optimization based on Evolutionary Algorithm for Predicting Oil data

Baoyu Ma[1]
CECS of Australian National University
U6918231@anu.edu.au

Abstract. In neural networks, parameter adjustment plays an important role. This article focuses on using the EA algorithm (Evolutionary Algorithm) to adjust the weights and thresholds in the neural network, which significantly saves the time for manual parameter adjustment. In addition, this paper also uses the pruning method to optimize the model. This paper attempts to test the oil data set, and the results are roughly the same as the traditional manual adjustment results.

Keywords: Deep Neural Networks, Geology, Phi and logK, Evolutionary Algorithm, optimization, prunning

1 Introduction

With the advancement of technology, the computing power of computers is also increasing, which also makes more complex neural network models possible. For example, the latest GPT-3 model is a giant with 17 million parameters. In this case, manual adjustment All parameters are almost impossible. With the deepening of neural network research, people gradually understand the parameters from the learning rate to the weight adjustment between each layer, which also makes the calculation amount more complicated. Therefore, it is foreseeable that in the future, manually designing algorithms to let the computer find the optimal configuration of parameters will be the mainstream parameter adjustment method in the future.

In this paper, when the super-parameters such as the learning rate have been determined, the adjustment target is turned to weights and thresholds. In this paper, after the model based on determining the hyperparameters such as the learning rate, the target is adjusted to the weight and the threshold. The weight represents the mapping relationship between the nodes of different layers of the neural network, and the threshold represents an adjustment to the prediction result. Use the EA algorithm to find the optimal weights and thresholds, and then assign them to the neural network model for further training.

Finally, a pruning algorithm based on weights [1] is used to further optimize the model.

2 Method

2.1 Dataset and Pre-process

The dataset is the rock smples from three oil-wells located in the North West Shelf. Every rock have eleven parameters, thry are GR (Gamma Ray), RDEV (Deep Resistivity), RMEV (Shallow Resistivity), RXO (Flushed Zone Resistivity), RHOB (Bulk Density), NPHI (Neutron Porosity), PEF (Photoelectric Factor) and DT (Sonic Travel Time), PHI (Porosity), logK (Permeability) and FLAG (Frac, Good, and OK). [2]

1	GR	RDEV	RMEV	RXO	RHOB	NPHI	PEF	DT	Phi	logK	FLAG
2	0.61926064	0.4225968	0.49696303	0.48662867	0.6085044	0.50819672	0.10680123	0.64912059	0.11838059	0.46394969	Frac
3	0.56115747	0.17299829	0.1734992	0.14645057	0.86461388	0.54918033	0.14349428	0.6453528	0.0942772	0.57153068	Good
4	0.53899999	0.74878204	0.6660644	0.51023579	0.6945259	0.41734973	0.20846702	0.34106944	0.01958852	0.58966044	OK

Fig. 1. This is the representative of the data set sample

Fig.1 shows that the data set has been preprocessed, and the value of each parameter is between 0 and 1. At the same time, in order to ensure the reliability of the data, for the FRAC state data, although the statistics are included in the training model, the generalization ability of the model may be improved, but its side effects are far greater than the benefits it brings, so it is abandoned.

The purpose of this article is to predict PHI and logK without doing a classification study of rock quality. The first eight parameters are used as the input of the model, PHI and logK are used as the output of the model, and considering the rock as a whole, its parameters are related, so there is no need to predict with two models.

The total number of samples is 358, and the total number of available samples is 292. The data set adopts the cross-validation method, of which 30% is used as the test set, and in the remaining 70% the training set and the verification set are 4 to 1.

2.2 Model structure

This article has two model structures a three-layer neural network model and EA nodel, because the three-layer is conducive to the EA algorithm to calculate the extracted weight.

For neural network there are 8 neural units in the input layer, 16 neural units in the hidden layer, and 2 neural units in the output layer. Learning-rate is 0.01, iterations is 2000 times.
And for EA the bits if DNA is 178, population is 100, DNA crossover probability is 0.6 and mutation probability is 0.002, generation size is 100.

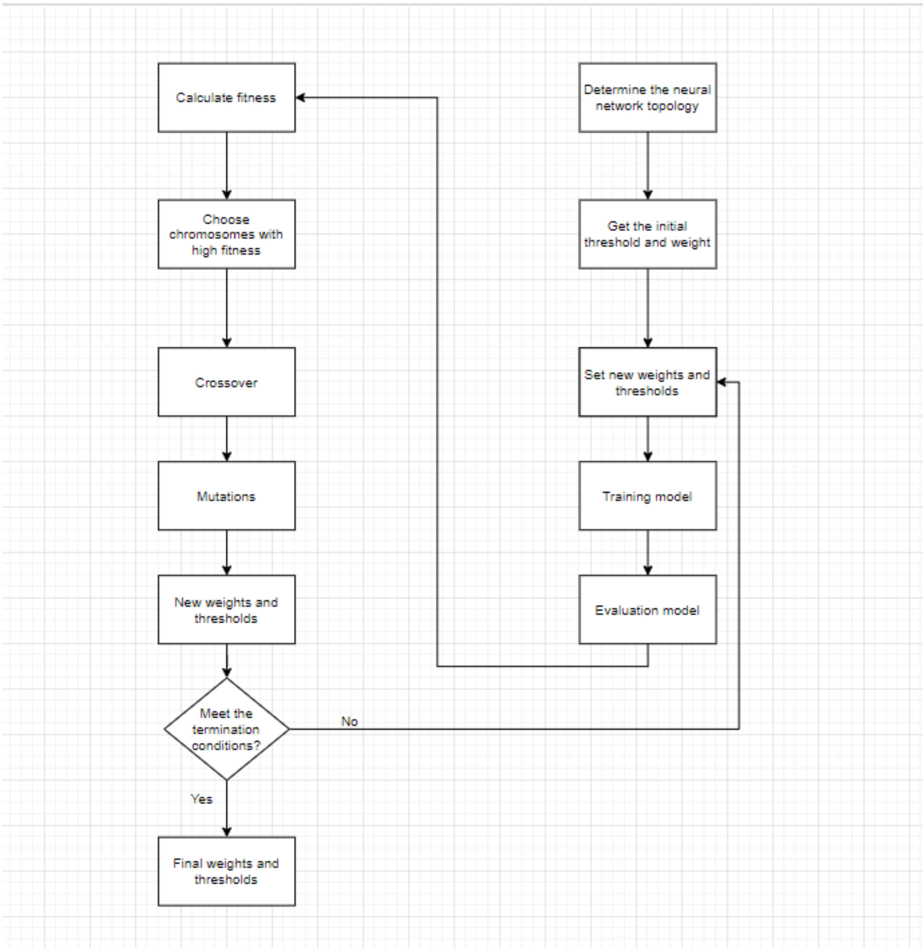

Fig. 2. This is the model flow chart

Fig 2 shows that how the model implements, firlstly, we initialize a neural network, extract the weights and thresholds, and input them as input to the EA model. The EA starts to perform cross-mutation iterations on this input variable and other operations, by bringing each configuration into the original neural network model for testing, Leaving the best test result (the lowest loss in this experiment) as the parent of the new generation, continue to repeat the above process until the termination condition is finally met, and then substitute the original neural network model for training to obtain the final result.

2.3 Evolutionary Algorithm

In artificial intelligence (AI), an evolutionary algorithm (EA) is a subset of evolutionary computation [3]. The EA algorithm has fixed algorithm steps and some parameters. In addition to the mutation rate mentioned above, there are also calculation functions including the target function and fitness.

In this experiment, the real number coding method is used, and a single unit of chromosome is a single weight or threshold of the neural network. The neural structure used in the article has a total of input layer neural units * hidden layer neural units + hidden layer neural units + hidden layer neural units * output layer neural units + output layer neural units which is 8 * 16 + 16 + 16 * 2 + 2 = 178. So the chromosome length is 178.

The target function is the test result loss of the neural network model. And for selecting the DNA, there are many mehods to seclect. The following two figures show the effect of different selection methods on the experimental results.

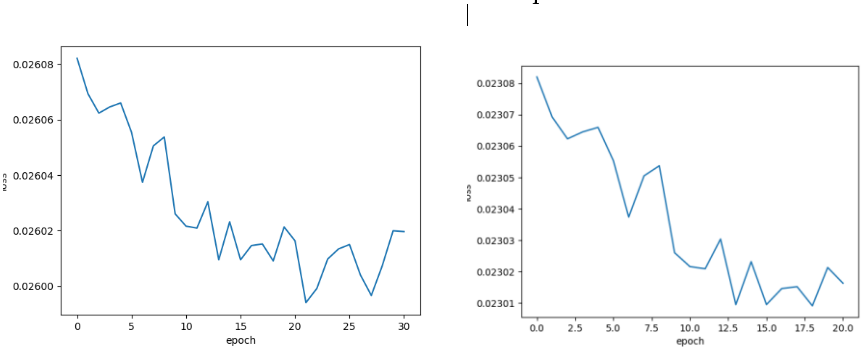

Fig. 2. The left patagraph uses the common mehod, and thr right is Boltzmann Selection, we could see that in the same epoch the right the picture on the right performs better.

2.4 Pruning

The more layers of the hidden layer, the better the data will be to the training set. However, when using the verification set to test, the error will increase, which is a typical overfitting phenomenon. Nevertheless, dropout solved the problem. The following three figures show the hidden layers of 1 and 2. Although the model with one hidden layer converges faster than the model with two hidden layers, the fluctuation after convergence is more extensive, so the model with two hidden layers is selected[3].

4 Result and Dicussion

The result without pruned is hovering around 0.23, after using pruning, the loss has no change and the pruned units' number is 0. In the last operation, the pruning rate was 23%, and the loss increased from 0.22 to 0.23. By comparison, we can see that under the premise of an angle of 15 degrees, the weights and thresholds obtained by the EA algorithm are already perfect, and close the results of manual parameter adjustment show that the EA algorithm has a strong adaptability to the adjustment of these hidden parameters such as weights and thresholds

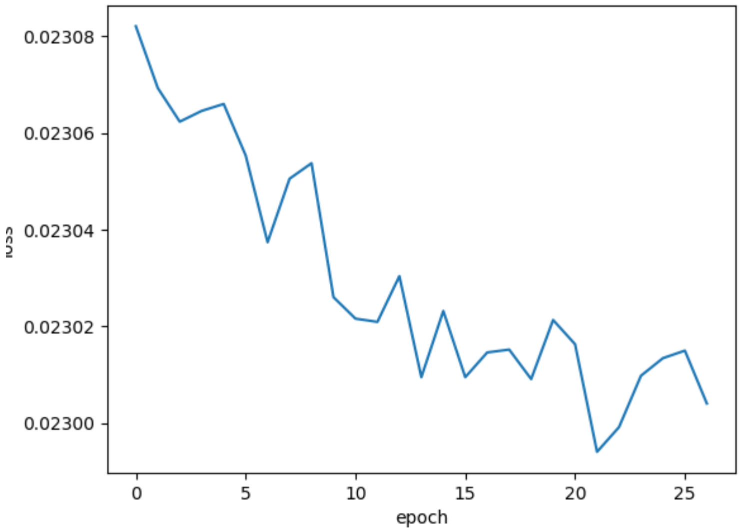

Fig. 4. This is the final results.

5 Conclusion

The results are excellent implies that the neural network can combine with other disciplines to create more achievements. Although the pruning technique does not play a big role in this model, it can become a standard for evaluating the model. The number of prunings of 0 reflects the good performance of the EA algorithm, and the results close to manual adjustment also indicate that in the future As the model becomes more complex and the computing power becomes more powerful, it is the future of artificial intelligence to let the computer adjust its participation. The EA algorithm will also have a greater role in the future. Although it may consume time, the calculation method similar to the evolution of species is likely to be a turning point for AI breakthroughs in the future.

References

[1][4]Gedeon, T.D. and Harris, D., 1991. Network reduction techniques. In Proceedings International Conference on Neural Networks Methodologies and Applications (Vol. 1, pp. 119-126).

[2] Kuo, H., Gedeon, T.D. and Wong, P.M., 1999, November. A clustering assisted method for fuzzy rule extraction and pattern classification. In ICONIP'99. ANZIIS'99 & ANNES'99 & ACNN'99. 6th International Conference on Neural Information Processing. Proceedings (Cat. No. 99EX378) (Vol. 2, pp. 679-684). IEEE.

[3] Ashlock, D. (2006), Evolutionary Computation for Modeling and Optimization, Springer, ISBN 0-387-22196-4.