Pruning based Neural Networks with Genetic Algorithm of Depression Detection*

Shengqi Li

Research School of Computer Science, Australian National University, Australia
Shengqi.Li@anu.edu.au

Abstract. Depression detection is an important mission in our daily life. It is a challenging task to recognize depression from the human's physiological responses by using the neural network. More detailed data which record a lot of signal form human body must can improve the detection performance. However, if the input data has a large number of features. Neural networks are both computationally intensive and memory intensive. Thus, network should detect which weights are useful and which weight are useless. Using genetic algorithm to select feature, this paper introduces several pruning methods such as Random pruning, structured pruning, L1 Unstructured pruning which can improve the computation ability efficiently. Feature selection will be used in preprocess part which has been recognised as a powerful tool for network optimisation. Genetic algorithm will find the optimal hyperparameters of neural network.

Keywords: Pruning · Neural Network · Depression Detection · Physiological Signals · Genetic Algorithm.

1 Introduction

Depression, a common mental disorder[2], can lead people to a serious chronic health condition. Therefore, depression detection is an important mission. The etiology of depression is not very clear, but it is certain that biological, psychological and social factors are involved in the pathogenesis of depression. Thus doctors will try to analyze the different biological signals such as electrical signals from the skin to detect depression. In the previous work[6], they provide the several kinds of data such as Galvanic Skin Response, Skin Temperature and Pupillary Dilation. These data sets record the more than 20 indexes for a person. To deal with such a big data by using a simple neural network is a challenging work.

In this paper we introduce two different network structures(neural network with one layer which has 50 hidden neurons and 4 outputs, RNN [1]). Pruning[3][4] is an important technology in learning based method. Thus, based on these two structures we use the several pruning methods(Random pruning, structured pruning, L1 Unstructured pruning etc.) to reduce the weights and improve

* Supported by ANU.

both the accuracy and training speed. We also use the genetic algorithm to select features show as the figure 1.

And we give the comparison between different network structure with different pruning technology and genetic algorithm.

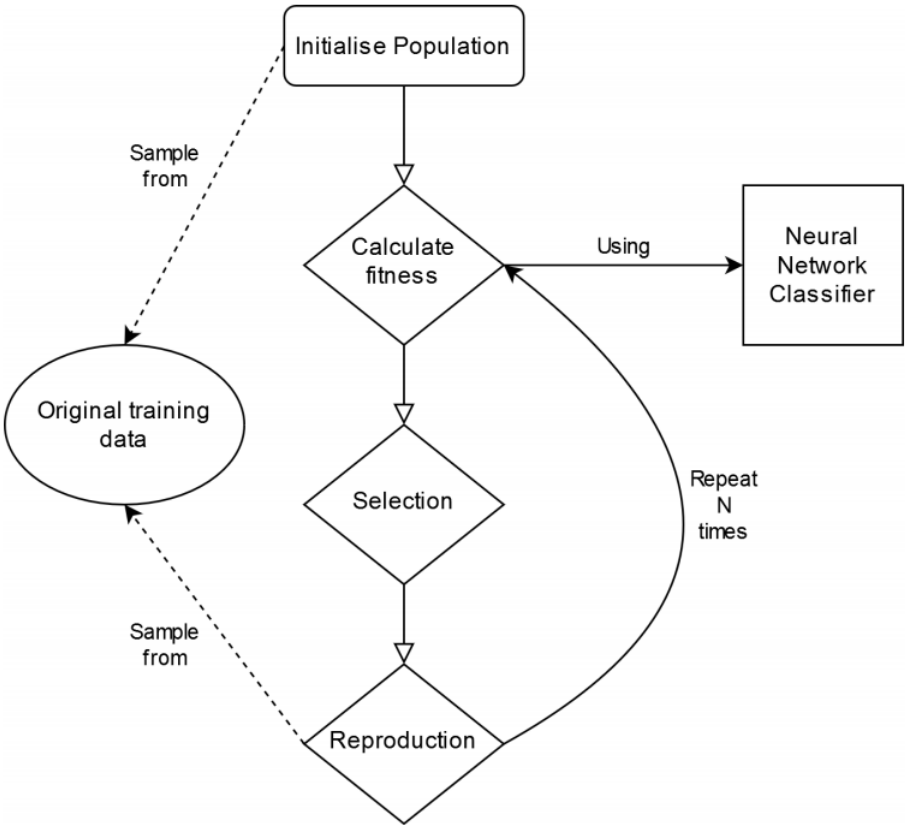

Fig. 1. Processing of genetic algorithm [7]

2 Methodology

2.1 Network Structures

Simple Structure Use a fully connected layer which has 23(has 23 different kinds of index to describe the condition of people) input neurons, 50 hidden neurons and 4 outputs. This 4 outputs is the predicted level of depression.

Recurrent Neural Network Convolutional neural network is same as human vision, but it hasn't memory ability, so it can only deal with a specific task, and can not deal with new tasks according to the previous memory.

Recurrent neural network is based on the idea of memory model. It is expected that the network can remember the previous features and infer the following results according to the features. Moreover, the overall network structure is constantly circulating, that is why named recurrent neural network

In this paper, Recurrent neural network has 23 input neurons, 50 hidden neurons and 4 output represent the level of depression,

2.2 Genetic Algorithm

Genetic algorithm is designed and put forward according to the evolution law of organisms in nature. It is a computational model simulating the natural selection and genetic mechanism of Darwin's theory of biological evolution. It is a method to search the optimal solution by simulating the natural evolution. The algorithm transforms the process of solving the problem into a process similar to the crossover and mutation of chromosome genes in biological evolution by means of mathematics and computer simulation. When solving complex combinatorial optimization problems, compared with some conventional optimization algorithms, they usually can get better optimization results quickly.

2.3 Pruning Method

We introduce 6 different pruning method, random unstructured pruning, L1 unstructured pruning, random structured pruning, Ln structured pruning, custom from mask pruning.

3 DataSet

These datasets [9] collect from 12 participants with no prior knowledge of depression identification. Participants will watching videos. Then These different dataset record different index of participants like the skin response or temperature.

3.1 Galvanic Skin Response

Galvanic Skin Response (GSR) is a physiological index of emotion. It represents the change of skin electrical conduction when the body is stimulated. It is generally expressed by resistance and its logarithm or conductance and its square root. Skin electrical response can only be used as a direct indicator of sympathetic nervous system function, and can also be used as an indirect indicator of brain arousal and alertness level, but it can not distinguish the nature and content of emotional response.

3.2 Skin Temperature

Skin Temperature (ST) reflect the change of body temperature. This fluctuates due to vasodilatation of peripheral blood vessels induced by increased activity of the sympathetic nervous system. Fluctuations are negatively correlated with unpleasant emotions, such as stress and fear.

3.3 Pupillary Dilation

Pupillary Dilation (PD) can't be controlled by human. The dilation and contraction of pupil truly reflect the complex and changeable psychology. For example, when a person feels happy, fond and excited, his pupils will expand.

4 Experiments

4.1 Environment and Setting

We use the python 3.6 and latest pytorch, with i7 cpu. Hyper parameters setting: epoch=500, hidden neurons=50, the initial ratio of importance matrix is 1/23, optimiser is Stochastic Gradient Descent[5][8], learning rate is 0.01. We divide the GRS and ST set into two parts as validation set and training set separately (ratio is 2:8), mutation rate is 0.2, chromosome will be encode into binary string with length 100. We use the crossover to change chromosome.

5 Conclusion and Future Works

Table 1. Accuracy of Simple Structure

pruning method/dataset	GSR	ST	GSR+GA	ST+GA
No pruning	70.92	71.33	75.22	76.32
RandomUnstructured	83.45	82.65	88.45	87.65
L1Unstructured	82.36	83.12	87.36	88.82
RandomStructured	85.27	85.81	90.20	89.87
LnStructured	84.52	83.51	89.76	88.36
CustomFromMask	86.32	87.63	91.02	91.81

5.1 Training Results and Conclusions

As the table 2 show that the accuracy of RNN is better than the simple structure, and both neural network has a better performance than results of depression prediction from observers' verbal response. And the network with pruning has a higher accuracy than the network without pruning. When we add the genetic

Table 2. Accuracy of RNN

pruning method/dataset	GSR	ST	GSR+GA	ST+GA
No pruning	72.11	74.03	69.22	74.82
RandomUnstructured	84.45	82.65	88.45	88.65
L1Unstructured	85.36	83.12	88.56	87.12
RandomStructured	84.24	87.81	90.20	89.87
LnStructured	88.32	83.51	89.88	89.36
CustomFromMask	88.91	87.63	91.72	92.01

algorithm to select important features(in this paper we choice 6 feature as input from 23 feature in the datasets), the accuracy also perform better than before.. At the same time, the converge speed of loss pruning network is faster(nearly converge in epoch 60),Fig. 2.

5.2 Future Works

In this paper we have compared two different structure with GA and pruning technology. However,there are still have some disadvantages. The number of selected feature is also set by a hyperparameter. Because some of features seems so important but it also can represent depression in a condition. Thus, in the future work, we will try the more deep structure and use a normalized method to let the feature selection process more clear.

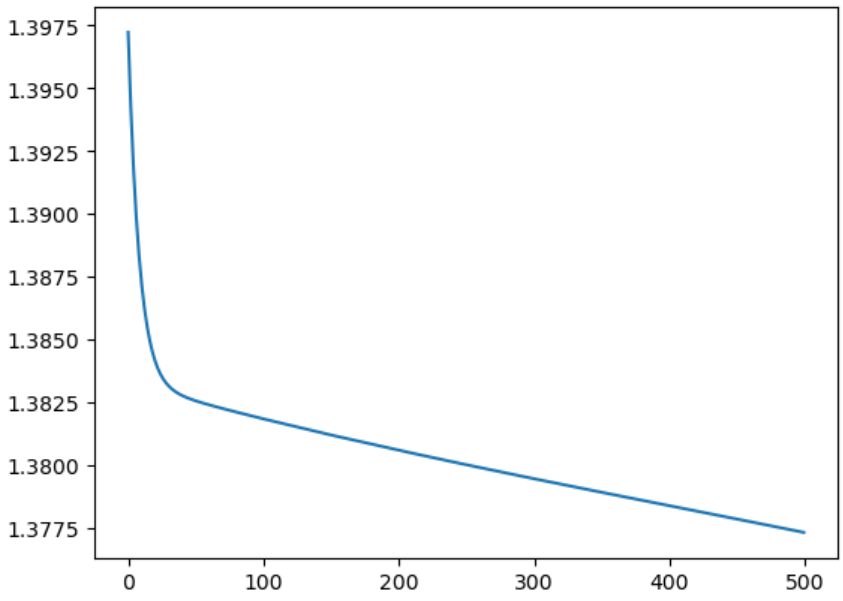

Fig. 2. Loss change of pruning network.

References

1. Zaremba, Wojciech and Sutskever, Ilya and Vinyals, Oriol:Recurrent neural network arXiv preprint arXiv:1409.2329,2014
2. N. Cummins, S. Scherer, J. Krajewski, S. Schnieder, J. Epps, and T. F. Quatieri, "A review of depression and suicide risk assessment using speech analysis," Speech Commun., vol. 71, pp. 10–49, 2015.
3. Karnin E D. A simple procedure for pruning back-propagation trained neural networks[J]. IEEE transactions on neural networks, 1990, 1(2): 239-242.
4. Luo J H, Wu J, Lin W. Thinet: A filter level pruning method for deep neural network compression[C]//Proceedings of the IEEE international conference on computer vision. 2017: 5058-5066.
5. Zinkevich M, Weimer M, Smola A J, et al. Parallelized Stochastic Gradient Descent[C]//NIPS. 2010, 4(1): 4.
6. Zhu X, Gedeon T, Caldwell S, Jones R. Detecting Emotional Reactions To Videos Of Depression.
7. Yuan L, Evolutionary Data Synthesis for Scarce Data Depression Classification, ABCs2020
8. Bottou L. Stochastic gradient descent tricks[M]//Neural networks: Tricks of the trade. Springer, Berlin, Heidelberg, 2012: 421-436.
9. Xuanying Zhu;Tom Gedeon;Sabrina Caldwell and Richard Jones. "Detecting emotional reactions to videos of depression". In: IEEE International Conference on Intelligent Intelligent Robots and Systems (2019).