The Application of Bidirectional Neural Network and Genetic Algorithm in Neural Network

Hualin Bai

College of Engineering & Computer Science
Australia National University
u5908811@anu.edu.au

Abstract. The objective of this paper is applied Bidirectional Neural Network as well as Genetic Algorithm for feature selection to evaluate the performance of a neural network. Based on the basic neural network model, Bidirectional Neural Network will be implemented to investigate the change of performance for the neural network. Genetic Algorithm will be further implied on the neural network to investigate relationship between feature selection with the network accuracy. Based on the results, the applied of Bidirectional Neural Network slight decreased the performance of a neural network for stress prediction from 52.6% to 50.5% while the application of genetic algorithm for feature selection increased the accuracy of the neural network from 52.6% to 56%-58%.

Keywords: Neural Network, Bidirectional Neural Network, Genetic Algorithm

1 Introduction

During the past decades, with the development of artificial neural network, there have been numbers of studies focus on the performance of the networks. At current stage, variety of methods have been applied to improve the performance of a neural network. One of these possible methods is applied feature selection. Genetic Algorithms inspired by John Holland during the 1970s, is one of the most popular advanced algorithms for feature selection [1]. From the other hand, Bidirectional Neural Network is another method for enhance the structure of a neural network and improve the performance of the network.

The main aim of this study is to investigate the different effects brought by Bidirectional Neural Network and Genetic Algorithm to the performance of an artificial neural network. At the beginning of this paper, the dataset pre-processing will be discussed. After ensuring the input and output of the network, a simple fully connection neural network will be constructed. Then, bidirectional neural network and genetic algorithm will be applied based on the simple neural network to compare the performance of the neural networks.

2 Method

2.1 Data Pre-processing

Considering about the data set we used in this project, there are 12 columns in total. The first column is the subject ID of the data while the second column is labels which represent the stress. The other 10 columns are features data collected via two types of cameras which include RGB camera as well as thermal camera.

Feature data will be used as the input of our neural network. Considering about that all features have been represented as the numerical form which is not need any further pre-processing. The output of the network should be the label which represent the stress. It can be seen from the dataset that there are two labels in this column which include calm and stressful. Labels are given in the string value form which should be transferred into numerical form. In our project, label stressful will be represented by number 0 while calm will be represented by 1.

Data will be randomly split into two sets for training and testing process respectively. 67% of the data will be selected randomly from the whole dataset as training data while the rest 33% of the data will be selected as data for testing. After split the data, labels are visualized by plot as shown below. It can be seen from the plot that two labels are distributed evenly which makes the further training process more reliable. Besides, it can be seen from the dataset that there is no missing data which need no further data pre-processing for data missing.

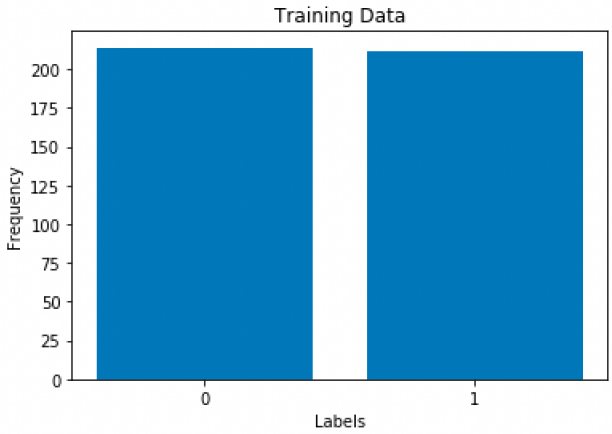

Figure 1Data distribution of the labels

2.2 Network Architecture

The first network used in this project can be seen as one fully connection basic artificial neural network which include 3 layers [2]. This neural network used features from the dataset as inputs. Considering about there are 10 features in the dataset, there are 10 nodes in the input layer of this neural network. Due to the size of the dataset as well as the function of this network, there is 1 hidden layer in this network. Based on the performance of experiments the number of neurons in the hidden layer is 10. There are two labels in the data, thus, there are 2 output neurons in the output layer.

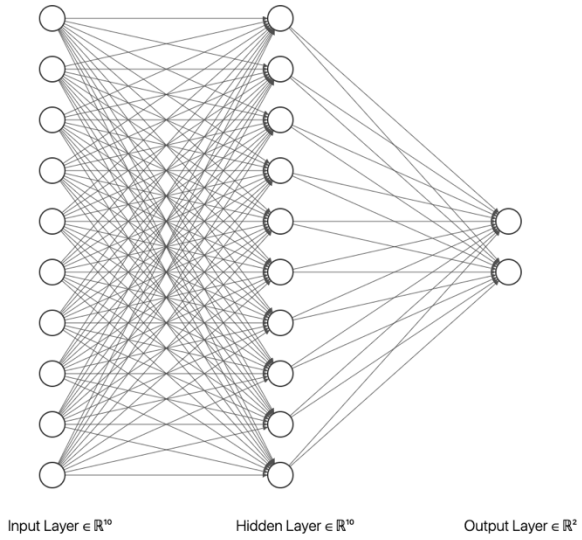

Figure 2 Neural Network Architecture

Considering about the aim of the project which is classify and predict the stress labels based on the given feature data, Cross Entropy Loss Function are going to be applied in the neural network for evaluation [3]. The mathematical formular of Cross Entropy Loss Function for 2 classes classification is shown below:

$$L = \frac{1}{N}\sum_{i} L_i = \frac{1}{N}\sum_{i} -[y_i \cdot log(p_i) + (1 - y_i) \cdot log(1 - p_i)] \quad \textbf{(1)}$$

As it is necessary for this project to predict the probability of the stress labels as one of the outputs, Sigmoid function which has the range between 0 to 1 is demonstrated in the hidden layer as activation function [4]. The Sigmoid function can be represented as shown below:

$$\phi(z) = \frac{1}{1 + e^{-z}} \tag{2}$$

Optimizer used in this project is Stochastic Gradient Descent (SGD) with momentum. The momentum of this project has been setting to 0.9 for the better performance of the whole network.

$$V_t = \beta V_{t-1} + (1 - \beta) S_t, \beta \in [0,1] \tag{3}$$

2.3 Evaluation Method

To evaluate the performance of the neural network, accuracy of both training and testing dataset are calculated based on the function shown below. Besides, the confusion matrix is used for analysis of the performance. The column of the confusion matrix represents the predicts number while the row represents the actual number. Loss and accuracy of the neural network are also been visualized by plots.

$$Accuracy = \frac{number\ of\ correct\ classified\ patterns}{total\ number\ of\ patterns} \tag{4}$$

2.4 Bidirectional Neural Network (BDNN)

In this project, bidirectional neural network is constructed based on the fully connection neural networks as describe above. However, considering about the performance of the whole network as well as the properties, some settings, parameters, and functions are slightly changed.

Firstly, the main difference between simple neural network and bidirectional neural network is that in BDNN, after some of the training, the direction will be reversed. In the other words, the output will be used as input for weight modified [5]. Figure 3 shows an architecture of BDNN. As can be seen from figure 3 that the neural network can work in two directions.

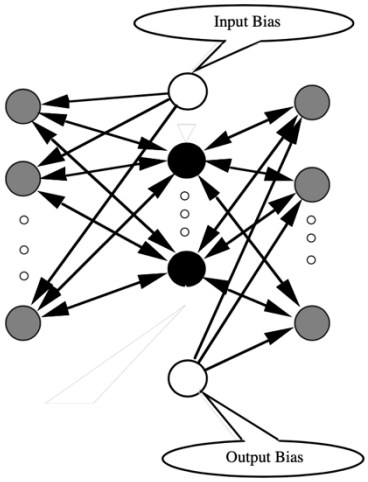

Figure 3 BDNN Architecture

Based on the architecture, the amount of input neurons in the input layer remains the same which is 10. The number of neurons in the output layer is 1. This mainly because for both labels which include stressful and calm can be seen as 1 feature when the neural network reverse the direction. Considering about the data size, the number of hidden neurons is set to 50 for better performance of the whole network.

As the network reverses the direction while training process, the loss function used in this network changed to Mean Squared Error loss function which shown as function below:

$$MSE = \frac{1}{n}\sum_{i}^{n}(\hat{y}_i - y_i)^2 \quad \textbf{(5)}$$

Other functions which include SGD with momentum and the sigmoid function remains the same with previous simple neural network which are used as optimizer and activation function respectively.

2.5 Genetic Algorithm (GA)

Genetic Algorithm is a search heuristic that is inspired by theory of natural evolution which reflects the process of natural selection where the fittest individuals are selected for reproduction in order to produce offspring of the next generation [6]. In GA, the process of survival of the fittest is mimicked to evolve the solution.

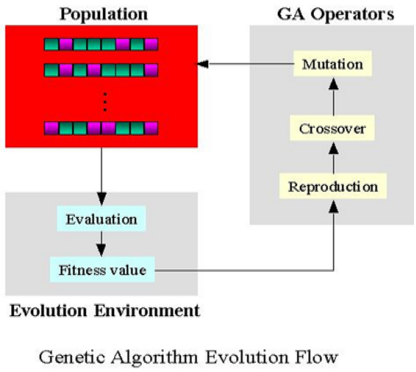

Figure 4 Genetic Algorithm working flow

As the step of GA, initial population of candidate solutions is generated in initialization. Randomly initialized binary strings will be generated in this step. Followed by initialization, chromosome is represented by binary which use 0 to represent removed features while 1 represent features kept.

Besides, in this project, fitness of this project is going to be evaluated via two criteria which include the accuracy and the cost [7]. To be more specific, it will be calculated by fitness function as shown below. In this function, fitness(x) is the feature subset represented by x while accuracy(x) is the test accuracy [7]. $cost_{max}$ is the upper bound on the costs of candidate solutions.

$$fitness(x) = accuracy(x) - \frac{cost(x)}{accuracy(x) + 1} + cost_{max} \quad \textbf{(6)}$$

3 Results and Discussion

Table 1 shows the accuracy of training and testing process based on simple artificial neural network as well as bidirectional neural network (BDNN):

Table 1Accuracy of Simple Neural Network and Bidirectional Neural Network

	Training accuracy	Testing accuracy
Simple Network	52.59%	52.59%
BDNN	50.48%	50.48%

It can be seen from the figure above that both training accuracy and testing accuracy slightly decreased in the Bidirectional Neural Network. In figure 5 the loss (left) and accuracy (right) of training process with simple neural network have been visualized by plots. The loss of training keeps decreasing at the beginning and kept stable after that. While the accuracy, which fluctuated slightly at the beginning, keep increasing rapidly and remaining stable at around 52.5% after the increasing.

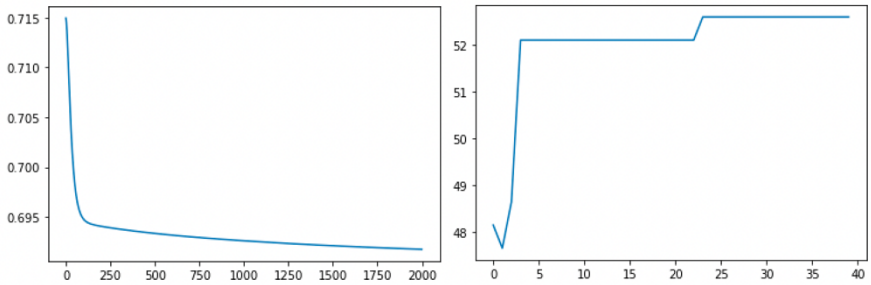

Figure 5 Loss (Left) and Accuracy (Right) of Simple Neural Network

Figure 6 below shows the loss of BDNN model. As the loss calculated by the sum of feed forward direction as well as the reserved direction, thus, in terms of magnitude, the total loss for BDNN is significantly larger than feed forward neural network. In terms of trend, the loss of BDNN keeps decreasing at the very beginning and remains the stable value since about 50 epochs to the end.

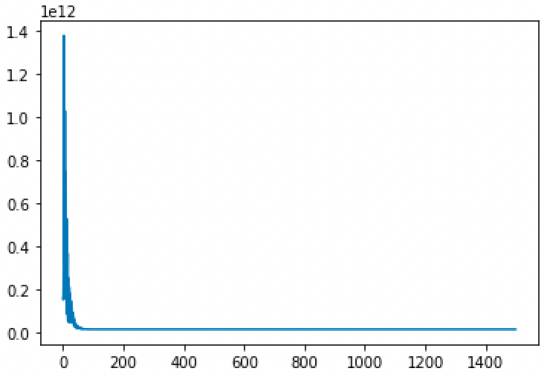

Figure 6 Loss plot for BDNN

Figure 7 shows the fitness of GA. It can be seen from the figure that the fitness mainly concentrated at the range of [0.037,0.043]. The fitness value frustrated at the first beginning then become more stable. After tests, the cross rate is set as 0.8, mutation rate of this GA is 0.02 while the population size is 500.

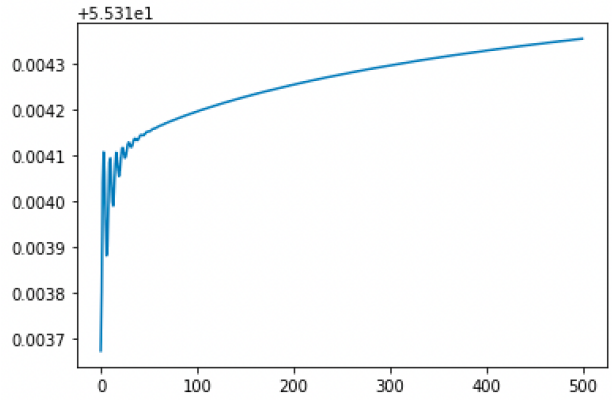

Figure 7 Fitness of Genetic Algorithm

Table 2 Testing Accuracy of Three Methods

Method	Testing Accuracy(%)
Neural Network	52.59
Bidirectional Neural Network	50.48
Genetic Algorithm	59.18

Table 2 shows the testing accuracy of all three models. Overall, compare with the simple feed forward neural network the implement of Genetic Algorithm for feature selection successfully increases the performance of the neural network for about 7%. The Bidirectional neural network model decreases the accuracy of the network for around 2% in testing accuracy. Thus, in terms of testing accuracy, neural network which with genetic algorithm for feature selection perform the best among the three methods.

4 Conclusion and Further Work

In the project, thermal video related dataset is used for stress classification and prediction. Based on the data collected by two types of camera, individuals are divided into two stress related labels which include stressful and calm. At the beginning of this project, data pre-processing is implemented. Then, a simple feed forward neural network with 1 hidden layer is constructed for classification and prediction. This simple feed forward neural network achieves the accuracy by around 52.59%.

Then, another bidirectional neural network is constructed. During the training process of this network, the direction will be reversed for further weight evaluating. However, as limited by the size of original dataset, the bidirectional neural network decreases the accuracy to around 50.48%. The reason lead this happened may include for reversed

direction, there are only 1 input feature with 2 different values fed into the network as input while 10 features are expected can be predicted as outputs.

Followed by that, Genetic Algorithm are implemented based on the feed forward neural network which constructed at the first stage of the project. Genetic algorithm provide feature selection to the network input features which successfully improved the accuracy of the network from 52.59% to 59.18%.

Finally, several problems can be further researched and discussed in the future. First of all, the size of the dataset used in current stage is quite limited. Issues like overfitting and underfitting may occur. Besides, the size of data also limit the implement of bidirectional neural network as there are only two labels as output. Moreover, the neural network used in this project is quite basic. In the future, more complex network can be constructed based on CNN or RNN method.

Reference

[1] S. Sayed, M. Nassef, A. Badr and I. Farag, "A Nested Genetic Algorithm for feature selection in high-dimensional cancer Microarray datasets," *Expert Systems With Applications,* vol. 121, pp. 233-243, 2018.

[2] E. Guresen and G. Kayakutlu, "Definition of artificial neural networks with comparison to other networks," *Procedia Computer Science,* vol. 3, pp. 426-433, 2011.

[3] J. Otander, "What is the different between MSE error and Cross-entropy error in NN," 5 September 2017. [Online]. Available: https://susanqq.github.io/tmp_post/2017-09-05-crossentropyvsmes/. [Accessed 20 May 2021].

[4] S. SHARMA, "Activation Functions in Neural Networks," 6 September 2017. [Online]. Available: https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6. [Accessed 20 May 2021].

[5] T. D. Gedeon, J. Catalan and J. Jin, "Image Compression using Shared Weights and Bidirectional Networks," The University of New South Wales, Sydney.

[6] V. Mallawaarachchi, "Introduction to Genetic Algorithms — Including Example Code," 8 July 2017. [Online]. Available: https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3. [Accessed 23 May 2021].

[7] J. Yang and V. Honavar, "Feature Subset Selection Using A Genetic Algorithm," Iowa Sate University, Iowa, 1997.