DECRYPTING NEURAL NETWORK DATA: Are you really angry?

Xinyu He
Research School of Computer Science, Australian National University
u6003370@anu.edu.au

Abstract.
The goal of this paper was to use the technique of Decrypting Neural Network Data to analyses and process the Anger dataset and use an evolutionary algorithm to perform the feature selection to the processing Anger dataset. My research showed that input encoding can improve the performance of model. In addition to this, I used the evolutionary algorithm to select the useful feature in the Anger dataset, and the model is going to detect the veracity of anger emotions with accuracy above 99% after my pre-processing. This prediction accuracy is higher than the accuracy of 96% in paper [1].
Keywords: Emotion veracity, machine classification, pupillary response, evolutionary algorithm.

1 Introduction

In our daily life, the expression of emotion can be easier to influence the focus of human, and the displayer of emotion would have different internal mental state when they express their emotion. Therefore, to determine whether the expression of emotion is true or not in any situation has become an interesting topic to explore. In 2017 Are you really angry? Detecting emotion veracity as a proposed tool for interaction [1], Gedeon et al experimented that pupillary response patterns is an effect way to show the veracity of anger emotion rather than the verbal response of the human, and the ensembles of machine classifiers trained on pupillary responses can be used to predict the veracity of emotion. Another paper [3] also claims that actual expressions and false expressions of emotion can be classified after the research. The research indicates that the difference of pupillary is obviously when a participant watches at the expression of genuine or acted anger, so that the machine classification can predict the emotion veracity due to pupillary response with 95% mean accuracy.

In this paper, I explored whether a feedforward neural network (FNN) is useful to predict the veracity of anger emotion based on pupillary responses. During the time that we applied FNN, it is important that to decode the Anger data in order to reduce noise and the training size of data to improve the accuracy of model. I applied the technique of Decrypting Neural Network Data proposed by Gedeon et al [2] to analysis and process the Anger dataset. This technique increases the average test accuracy with 1.25%. In addition to this, I applied the Genetic algorithm to keep informative features and remove those redundant features in the Anger dataset.

2 Data description

The Anger dataset has 400 data and 8 features which include the following features in this paper:

- Video - 10 videos for the true angary emotion and another 10 videos for the false angary emotion.
- Mean – The mean of in pupillary response.
- Std - The standard deviation of in pupillary response.
- Diff1 – The change of left pupillary size after watching a video.
- Diff2 - The change of right pupillary size after watching a video.
- PCAd1- An orthogonal linear transformation with first principal component
- PCAd2- An orthogonal linear transformation with second principal component
- Label – The Genuine or Posed emotion.

3 Method

3.1 Feedforward neural network model design

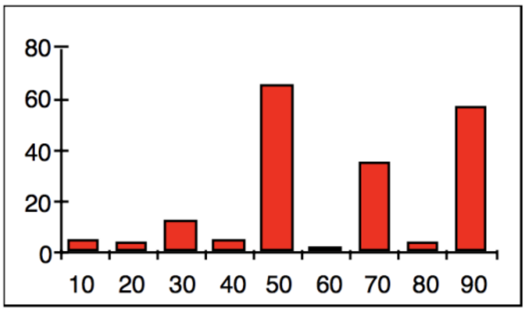
I developed a fully connected neural network with two layers which are one input layer with seven neurons corresponded to the seven anger input features and another output layer with one output neurons corresponded to the target label. Besides, there is one hidden layer with specified hidden units. I decided to separate the training data and the testing data by using the 5-folds cross validation. Since this is a classification model, so I use the sigmoid activation function [8] and a loss function called cross entropy error function [9]. Otherwise, Stochastic gradient descent (SGD) is currently a famous algorithm which can achieve great performance on a variety of machine learning tasks [10], so it is suitable for our task.

3.2 Decrypting Neural Network Data technique design and implementation

3.2.1 The introduction of technique

The Decrypting Neural Network Data technique usually predicts the output based on the available information, and the main idea is to invent some specific methods in order to deal with large amount of erratically data. Processing data before building a Neural Network is significant and critical step as researchers claim in the paper [4]. In the paper [2], as it describes the technique of Decrypting Neural Network Data, there is an example called GIS case study which explains how to predict the forest supra-type by using this technique. A common example is that the data analysis of Geology Descriptor in the GIS data as shown in figure 1[2]. According to this bar chart, the geology descriptor has three significant groups and it is not normally distributed. Besides, there is a problem that patterns with '50' output is similar to patterns with '70' output and this will lead to spurious '60' results. In the Encoding Decision part of this paper [2], it shows the solution that the categories will be represented by 4 inputs which have 3 popular types (output: 50, 70 and 90) and 1 rare type because all of rare outputs are very sparse, so that they can be represent by 1 type. This is how the Decrypting Neural Network Data technique works.

Geology descriptor:



(Figure 1. Geology descriptor data analysis)

3.2.2 The implementation of technique

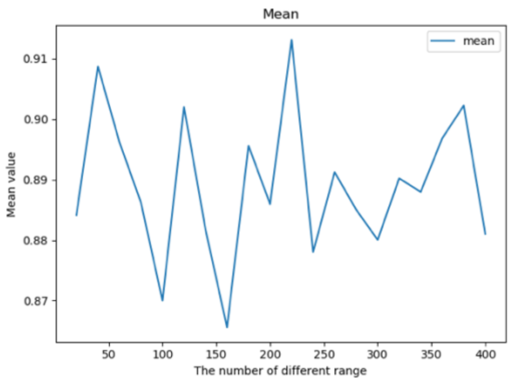
Label:
First of all, there are some features of text need to be analyzed and preprocessed before to apply the FNN, which are the feature 'Video' and 'Label'. Since this is a classification problem, so I decide to use two different numbers to represent the target Label, which 0 represents the Genuine and use 1 represents the Posed.

Video:
I use the one-hot encoding to process the feature 'Video' and separate Video to two features called Video1 and Video2. The objective of this preprocessing is to convert the categorical variable into a form that could be provided to FNN and also avoid the problem that label encoding always assumes higher the categorical value, better the category [7].

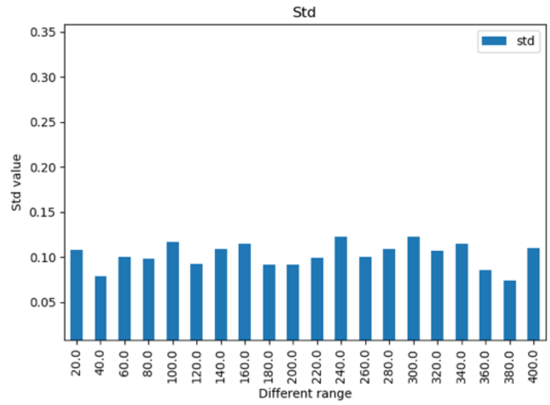
Mean & Std & Diff1 & Diff2 & PCAd1 & PCAd2:
For other normal features, I found that they do not have the evenly distributed at one same range. For example, all mean values are above 0.5 and the distribution shows in below Figure 2. In addition to this, values in Std feature have different ranges allocate at about 0.1 showing in the Figure 3 and Table 1. So, I normalized these data in order to allocate values evenly distributed for the range of 0 to 1. This processing would make each feature to have the same impact on weight update during the time of training the neural network.

Maximum Value	0.00799
Minimum Value	0.35837
Average Value	0.18524

(Table 1. Analysis table for Std)



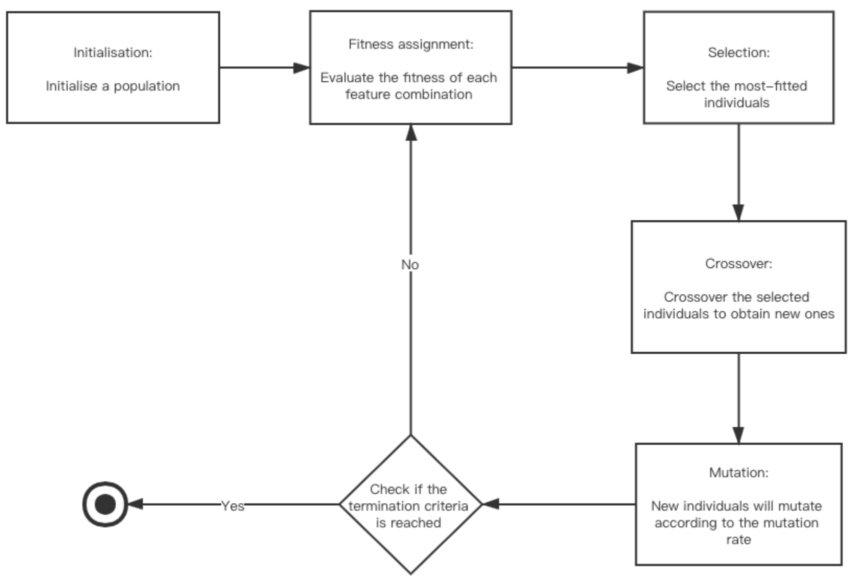
(Figure 2. Analysis Mean)



(Figure 3. Analysis Std)

3.3 Genetic algorithm design and implementation

Since there are some high related features in the Anger dataset, such as Diff1 and Diff 2, PCAd1 and PCAd2. Genetic algorithm will be useful to select best features in this dataset, so I applied GA to train the FNN after re-using the technique of Decrypting Neural Network Data to process data. Also, more configurations are in my consider like mutation and cross-over methods. During the time of implement genetic algorithm, I experimented two combinations which are Random cross over and Uniform mutation. Random cross over is that every parent network can pass its corresponding weight to its child network with the same chance. In addition to this, Uniform mutation is that setting a mutation probability firstly and then every child network would be mutated to a random value with the same mutation probability. Furthermore, I use the cross validation to split the dataset into 5 folds to increase the randomness. The procedure of genetic algorithm is introduced in Figure 4.



(Figure 4. The procedure of genetic algorithm)

4 Results and Discussion

Above all, after 500 epochs of training the Feedforward neural network model, the average test accuracy is 91% which means given a test data of anger emotion, the trained FNN could predict the emotion veracity on a Video for 91% accuracy. This accuracy is lower than the average test accuracy with 96% in paper [1]. In order to increase the test accuracy, I used the technique of Decrypting Neural Network Data. After I implement this technique, the average test accuracy is 92.60% which increases 1.60% to the original accuracy. This increasing of accuracy is because I use the technique to evaluate and decode all features.

	Testing accuracy
1-fold	90.00 %
2-fold	95.00 %
3-fold	93.75 %
4-fold	87.50 %
5-fold	88.75 %
Average	91.00%

(Table 2. Testing accuracy of FNN)

	Testing accuracy
1-fold	91.25 %
2-fold	96.25 %
3-fold	93.00 %
4-fold	88.75 %
5-fold	93.75 %
Average	92.60%

(Table 3. Testing accuracy after using technique)

Moreover, I set the data which used the technique as the input dataset for genetic algorithm in order to select best features. I found that Video is the most significant feature in the Anger dataset because even that I use one-hot encoding to make Video to be two features, the feature of Video mostly influenced the test accuracy. After implement of Genetic algorithm, the most fitted DNA [0 0 1 1 1 1 1 1], which means Video1 and Video2 randomly are dropped and other features as input to Neural Network, and this DNA has the worst performance which its testing accuracy is only has 77.01%, so that this is the reason why Video is the most important feature in the Anger dataset. In addition to this, PCAd2 is the most non-important feature in the Anger dataset because by comparing two DNA [0 0 1 0 1 1 1 1] and DNA [0 0 1 0 1 1 1 0], these two test accuracy is almost the same with about 98.24%, so dropped the feature of PCAd2 would not influence our test accuracy. This is because the PCAd1 and PCAd2 are highly related, so one of them may bring some noisy data and are not useful for us to improve our model. Overall, Genetic algorithm helped to increase the test accuracy by achieving feature selection in this situation that some features are highly related to each other.

5 Conclusion and Future Work

In summary, this investigation showed that the pupillary response is really useful for machine classification and the average test accuracy of pupillary response is higher than verbal response with 35% [1]. Also, technique of Decrypting Neural Network Data does have the effect to improve accuracy with 1.6% but it is not obvious. Genetic algorithm improved the test accuracy with 5.64% due to feature selection, but the limitation of computation influences our final result due to time reason. In the future work, I will adjust more parameters to get the most suitable model, such as the number of hidden neural, leaning rate and batch size.

References

1. Chen, Lu, et al. "Are you really angry? Detecting emotion veracity as a proposed tool for interaction." Proceedings of the 29th Australian Conference on Computer-Human Interaction. 2017.

2. Bustos, R. A., and T. D. Gedeon. "Decrypting neural network data: a GIS case study." Artificial Neural Nets and Genetic Algorithms. Springer, Vienna, 1995.

3. Kim, Chi Jung, and Min-Hyuk Chang. "Actual emotion and false emotion classification by physiological signal." 2015 8th International Conference on Signal Processing, Image Processing and Pattern Recognition (SIP). IEEE, 2015.

4. Yu, Lean, Shouyang Wang, and Kin Keung Lai. "An integrated data preparation scheme for neural network data analysis." IEEE Transactions on Knowledge and Data Engineering 18.2 (2005): 217-230.

5. Kim, Kwang In, Keechul Jung, and Hang Joon Kim. "Face recognition using kernel principal component analysis." IEEE signal processing letters 9.2 (2002): 40-42.

6. Mazurowski, Maciej A., et al. "Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance." Neural networks 21.2-3 (2008): 427-436.

7. Gedeon T D, Tamhane D, Lin T, et al. Use of linguistic petrographical descriptions to characterise core porosity: contrasting approaches[J]. Journal of Petroleum Science and Engineering, 2001, 31(2-4): 193-199.

8. MacLeod, C. (2010). An introduction to practical neural networks and genetic algorithms for engineers and scientists (ch.2 p.6). Scotland: The Robert Gordon University.

9. Bishop, C. M. (2006). Pattern Recognition and Machine Learning (Information Science and Statistics) (ch.5 Neural Network)

10. Recht B, Re C, Wright S, et al. Hogwild: A lock-free approach to parallelizing stochastic gradient descent[C]//Advances in neural information processing systems. 2011: 693-701.