Using Duo Output Neural Network to Solve Binary Classification Problems

PAWALAI KRAIPEERAPUN
Ramkhamhaeng University
Faculty of Science
Department of Computer Science
Bangkok, THAILAND
pawalai@yahoo.com

SOMKID AMORNSAMANKUL
Mahidol University
Faculty of Science
Department of Mathematics
Bangkok, THAILAND
and Centre of Excellence in Mathematics
CHE, Sriayudhaya Rd., Bangkok, Thailand
scsam@mahidol.ac.th

Abstract: This paper proposes an approach to solve binary classification problems using Duo Output Neural Network (DONN). DONN is a neural network trained to predict a pair of complementary outputs which are the truth and falsity values. In this paper, outputs obtained from two DONNs are aggregated and used to predict the classification result. The first DONN is trained to predict a pair of truth and falsity values. The second DONN is trained to predict a pair of falsity and truth values. The target outputs used to train the second network are organized in reverse order of the first network. The proposed approach has been tested with three benchmarking UCI data sets, which are ionosphere, pima, and liver. It is found that the proposed techniques improve the performance as compared to feedforward backprogation neural network and complementary neural network.

Key–Words: Feedforward Backpropagation Neural Network, Complementary Neural Network, Binary Classification, Duo Output Neural Network

1 Introduction

There are several methods used to solve binary classification problems. Some examples of these methods are neural network, support vector machine (SVM), decision tree, Naive Bayesian classifier, etc. It was found that neural networks provide better classification accuracy than traditional statistical methods in various areas of applications such as business, finance, health, medicine, engineering, marketing, geology, and Paleoceanography [2, 6, 10]. Even though SVM was found to give better classification accuracy than neural network in some applications [8], neural network was also found to perform better than SVM in various tasks such as document classification [7], exudate classification [9], bio-activity classification [11], and biological microscopic image classification [5].

However, it is found that neural network is one of the most widely used methods in binary classification problems. In order to classify the output obtained from neural network, one of the most well known techniques is applying a threshold value. The threshold value must be pre-set before making binary classification in order to examine whether the output exceed the threshold. This technique can cause vagueness in the classification since the output obtained from neural network is always uncertain. It is not exactly the truth output. Therefore, it would be better if we can deal with both truth and falsity output of a neural network for each input pattern. A neural network with multiple outputs can be applied to support this idea. Instead of dealing only with the truth output, a neural network trained to predict both truth and falsity outputs is considered. Instead of using only truth target values, complement of those target values called falsity target values are also considered. A neural network trained with both truth and falsity target values will provide both truth and falsity outputs. This method was created in [3] and named duo output neural network. Two duo output neural networks were created to solve a single output regression problem. One neural network was trained using reverse order of target values used in another neural network. The aggregation of outputs obtained from these two duo output neural networks was found to provide better performance when compared to backpropagation neural networks and support vector regression with linear, polynomial, and radial basis function kernels. From its success, this paper aims to apply duo output neural networks to solve binary classification problems. Three techniques used for binary classification are proposed. Furthermore, we can predict uncertainty in the classification based on truth and falsity outputs ob-

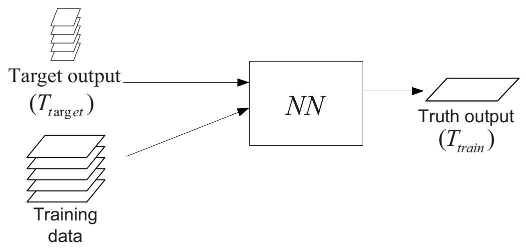

Figure 1: Neural network for binary classification. (Training Phase)

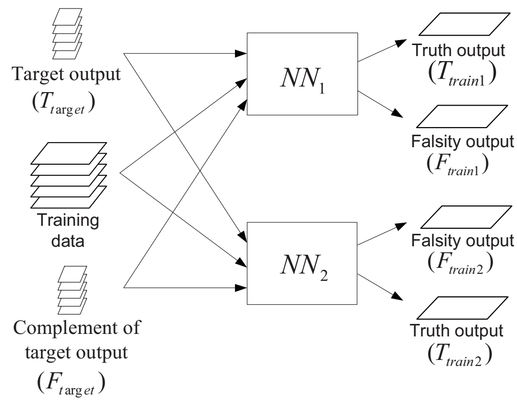

Figure 2: Duo output neural network for binary classification. (Training Phase)

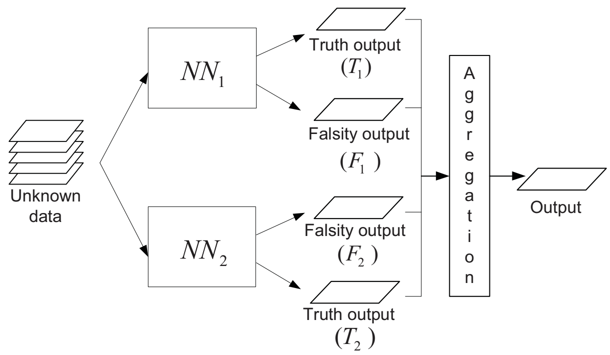

Figure 3: Duo output neural network for binary classification. (Testing Phase)

tained from both neural networks. These uncertainty can be used to identify level of confidence in order to belief that the input pattern is classified into the right class. We test our experiment based on three classical benchmark problems including ionosphere, pima, and liver from the UCI machine learning repository [1].

The rest of this paper is organized as follows. Section 2 describes the basic concept of duo output neural networks and the proposed classification techniques used to classify outputs from a pair of duo output neural networks. Section 3 describes data sets and results of our experiments. Conclusions and future works are presented in Section 4.

2 Duo Output Neural Network for Binary Classification

In traditional feedforward backpropagation neural network, only truth target values are used to train neural networks to predict truth outputs as shown in Fig. 1. However, the complement value of the truth which is the falsity value can occur in real world situations as well. Therefore, this paper also consider falsity target values together with truth target values in training process. In this case, we can create a feedforward backprogation neural network that is trained to provide two opposite outputs which are truth and falsity values.

Let $T_{target}(x_i)$ be the truth target value for the input pattern $x_i, i = 1, 2, 3, ..., m$ where m is the total number of training input patterns. Let $F_{target}(x_i)$ be the falsity target value for the input pattern x_i. The falsity target value is considered as the complement of the truth target value. The falsity target value can be computed as follows.

$$F_{target}(x_i) = 1 - T_{target}(x_i) \quad (1)$$

Duo output neural network can be created in two circumstances according to the sequence of truth and falsity target values used to train a neural network. In order to take advantages of both circumstances, a pair of duo output neural networks are created. Fig. 2 shows a pair of duo output neural networks in training phase. NN_1 is a feedforward backpropagation neural network trained to predict the truth output (T_{train1}) and the falsity output (F_{train1}). NN_2 is also a feedforward backpropagation neural network trained to predict the falsity output (F_{train2}) and the truth output (T_{train2}). Both networks are created based on the same architecture and parameters. Also, they apply the same input pattern data. However, they are trained using different order of truth and falsity target values.

In Fig. 3, the unknown input pattern y_j is assigned to a pair of duo output neural networks in the testing phase where $j = 1, 2, 3, ..., n$ and n is the total number of unknown input patterns. Let $T_1(y_j)$ and $F_1(y_j)$ be the truth and the falsity outputs for the unknown input pattern y_j of the first neural network (NN_1). These two outputs can be aggregated in two aspects: the average truth output ($T_a(y_j)$) and the average falsity output ($F_a(y_j)$). Both average outputs can be defined as follows.

$$T_a(y_j) = \frac{T_1(y_j) + (1 - F_1(y_j))}{2} \quad (2)$$

$$F_a(y_j) = \frac{F_1(y_j) + (1 - T_1(y_j))}{2} \quad (3)$$

Let $F_2(y_j)$ and $T_2(y_j)$ be the falsity and the truth outputs for the unknown input pattern y_j of the second neural network (NN_2). The average truth output ($T_b(y_j)$) and the average falsity output ($F_b(y_j)$) can be computed as follows.

$$T_b(y_j) = \frac{T_2(y_j) + (1 - F_2(y_j))}{2} \quad (4)$$

$$F_b(y_j) = \frac{F_2(y_j) + (1 - T_2(y_j))}{2} \quad (5)$$

Instead of using only the truth output for binary classification, both truth and falsity values can be applied. In this paper, we propose binary classification techniques based on the classification techniques used in [4]. In [4], a pair of feedforward backpropagation neural networks were created in which the first network is trained to predict only the truth output whereas the second network is trained to predict only the falsity output (see Fig. 4). This technique was named complementary neural network. In order to classify each input pattern, both truth and falsity outputs were compared. If the truth output is greater than the falsity output then the input pattern is classified as a value 1. Otherwise, it is classified as a value 0.

In this paper, the truth and falsity outputs which are T_a, T_b, F_a, and F_b can be used. Three classification techniques are proposed and described below.

1. For each input pattern y_j,
   if $T_a(y_j) > F_b(y_j)$ then

   the input pattern y_j is classified as a value 1

   else

   the input pattern y_j is classified as a value 0.

2. For each input pattern y_j,
   if $T_b(y_j) > F_a(y_j)$ then

   the input pattern y_j is classified as a value 1

   else

   the input pattern y_j is classified as a value 0.

3. For each input pattern y_j,
   if $\frac{T_a(y_j)+T_b(y_j)}{2} > 0.5$ then

   the input pattern y_j is classified as a value 1

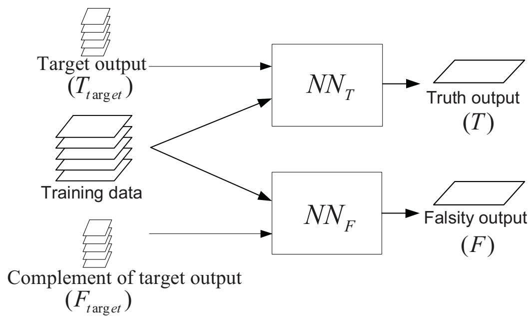

Figure 4: Complementary neural network for binary classification.

   else

   the input pattern y_j is classified as a value 0.

In [4], uncertainty in the classification of each input pattern were predicted based on truth and falsity outputs. This technique can be applied to our paper as well. Hence, the truth and falsity outputs obtained from a pair of duo output neural networks are used to quantify uncertainty in the classification. Let $U(y_j)$ be an uncertainty value in the classification of input pattern y_j. $U(y_j)$ can be computed as follows.

$$U(y_j) = 1 - |T_a(y_j) - F_b(y_j)| \quad (6)$$

or

$$U(y_j) = 1 - |T_b(y_j) - F_a(y_j)| \quad (7)$$

From our experiment in the next section, it is found that three proposed classification techniques provide the same results. These techniques can be used interchangeably. Uncertainty in the classification can be computed from equations (6) and (7) interchangeably as well. Degree of uncertainty can be used to identify level of confidence in order to belief that whether the input pattern is classified into the correct class. If the difference between the truth and falsity is high then the degree of uncertainty is low. On the other hand, if the difference is low then the degree of uncertainty is high. Uncertainty information obtained from the classification can be used to support users for selecting a proper classifier.

3 Experiments

3.1 Data Sets

Three benchmarking UCI data sets [1], which are ionosphere, liver, and pima are used in this experiment. The characteristics of these data sets can be shown in Table 1.

Table 1: UCI data sets used in this study

Name	Ionosphere	Pima	Liver
Feature type	numeric	numeric	numeric
No. of classes	2	2	2
No. of features	34	8	6
Sample set size	351	768	345
Training set size	200	576	276
Testing set size	151	192	69

3.2 Experimental Methodology and Results

In our experiment, each data set is applied to three types of neural network which are backpropogation neural network (BPNN), complementary neural network (CMTNN), and our proposed duo output neural network (DONN). In order to compare results obtained from those three methods, all environments are fixed for all methods. For BPNN, twenty feed-forward backpropagation neural networks are trained with twenty different randomized training sets for each data set. For CMTNN, twenty pair of feed-forward backpropagation neural networks are trained with the same twenty training sets used for BPNN for each data set. For our proposed neural network, twenty pair of DONNs are also trained with the same twenty training sets used for BPNN for each data set. All networks are having the same parameter values in terms of the network architecture and they are initialized with the same random weights. the number of input-nodes is equal to the number of input features, which is 34, 8, and 6 for ionosphere, pima, and liver data sets, respectively. They have one hidden layer constituting of $2n$ neurons where n is the number of input features. Hence, the number of neuron in the hidden layer for those data sets are 68, 16, and 12, respectively.

Table 2 shows the comparison among average classification accuracy obtained from twenty set of BPNN, CMTNN, and DONN for each test set of ionosphere, pima, and liver. It can be argued that the integration of both truth and falsity values in the prediction can provide better classification accuracy when compared to the prediction involved only in the truth values. It can be seen that both CMTNN and DONN can provide better performance than BPNN. Moreover, it is also found that a pair of neural network with multiple outputs (truth and falsity values) can provide better accuracy than a pair of neural network with single output dealing only with the truth values for one network and applying only falsity values to other network. Therefore, it is found that the results obtained from the proposed DONN outperform the results obtained from BPNN and CMTNN. Table 3 shows the percent improvement of the proposed DONN compared to BPNN and CMTNN.

Table 2: Average classification accuracy obtained from the test data set

Method	Ionosphere %correct	Pima %correct	Liver %correct
BPNN:			
$T > 0.5$	93.54	70.49	62.68
CMTNN:			
$T > F$	96.42	74.74	66.52
DONN:			
$T_a > F_b$	96.52	77.92	70.22
$T_b > F_a$	96.52	77.92	70.22
$\frac{T_a+T_b}{2} > 0.5$	96.52	77.92	70.22

Table 3: The percent improvement of the proposed DONN compared to BPNN and CMTNN.

Method	DONN (%improvement)		
	Ionosphere	Pima	Liver
BPNN:			
$T > 0.5$	3.19	10.53	12.56
CMTNN:			
$T > F$	0.10	4.25	5.56

From our experiment, it can be concluded that the proposed classification techniques using $T_a > F_b$, $T_b > F_a$ and $\frac{T_a+T_b}{2} > 0.5$ are found to provide the same results for each test set used in this study. Therefore, we can use these three techniques interchangeably. Moreover, uncertainty in the classification can be computed based on T_a and F_b using equation (6). Table 4 shows the ranges of uncertainty values in the classification of pima data set. One of twenty sets of pima used in the experiment is shown. Uncertainty values are grouped into three levels each with an equal range. The total number of correct and incorrect outputs in each range are shown together with their per-

Table 4: Uncertainty level obtained from a pair of DONN for the test set of pima data.

Uncertainty		Number of patterns		%correct
value	level	correct	incorrect	
0.68-0.99	High	23	23	50.00
0.36-0.67	Med	45	18	71.43
0.05-0.35	Low	77	6	92.77
Total		145	47	75.52

centage of correct classification. It can be seen that the proportion of the percentage of the correct outputs in each group conforming to the level of uncertainty in which the more correct outputs, the less uncertainty.

4 Conclusion

This paper has applied a pair of duo output neural networks to solve binary classification problems. The first duo output neural network is trained to predict a pair of truth and falsity outputs whereas the second duo output neural network is trained to predict a pair of falsity and truth outputs which are organized in reverse order of the first one. The proposed approached are tested based on three UCI data sets. It is found that the proposed approach provide better performance than traditional feedforward backpropagation neural networks and complementary neural networks. Uncertainty in the classification can also be quantified. These uncertainty values are very useful in order to enhance the decision making. In the future, we will apply our approach to multiclass classification problems.

Acknowledgements: This research project is supported by Faculty of Science, Mahidol University.

References:

[1] A. Asuncion and D.J. Newman. UCI machine learning repository, 2007.

[2] W. M. Brown, T. D. Gedeon, D. I. Groves, and R. G. Barnes. Artificial neural networks: A new method for mineral prospectivity mapping. *Australian Journal of Earth Sciences*, 47:757–770, 2000.

[3] Pawalai Kraipeerapun, Somkid Amornsamankul, Chun Che Fung, and Sathit Nakkrasae. Applying Duo Output Neural Networks to Solve Single Output Regression Problem. In C.S. Leung, M. Lee, and J.H. Chan, editors, *ICONIP (1)*, volume 5863 of *Lecture Notes in Computer Science*, pages 554–561. Springer, 2009.

[4] Pawalai Kraipeerapun and Chun Che Fung. Binary classification using ensemble neural networks and interval neutrosophic sets. *Neurocomputing*, 72(13-15):2845–2856, August 2009.

[5] Ilias Maglogiannis, Haralambos Sarimveis, Chris T. Kiranoudis, Aristotelis A. Chatziioannou, N. Oikonomou, and V. Aidinis. Radial Basis Function Neural Networks Classification for the Recognition of Idiopathic Pulmonary Fibrosis in Microscopic Images. *IEEE Transactions on Information Technology in Biomedicine*, 12(1):42–54, 2008.

[6] Björn A. Malmgren, Michal Kucera, Johan Nyberg, and Claire Waelbroeck. Comparison of Statistical and Artificial Neural Network Techniques for Estimating Past Sea Surface Temperatures from Planktonic Foraminifer Census Data. *Paleoceanography*, 16(5):520–530, 2001.

[7] Larry Manevitz and Malik Yousef. One-class document classification via Neural Networks. *Neurocomputing*, 70(7-9):1466–1481, 2007.

[8] David Meyer, Friedrich Leisch, and Kurt Hornik. The support vector machine under test. *Neurocomputing*, 55:169–186, 2003.

[9] Alireza Osareh, Majid Mirmehdi, Barry T. Thomas, and Richard Markham. Comparative Exudate Classification Using Support Vector Machines and Neural Networks. In *MICCAI '02: Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention-Part II*, pages 413–420, London, UK, 2002. Springer-Verlag.

[10] Mukta Paliwal and Usha A. Kumar. Neural networks and statistical techniques: A review of applications. *Expert Systems with Applications*, 36:2–17, 2009.

[11] Jehan Zeb Shah and Naomie Salim. Neural networks and support vector machines based bioactivity classification. In *Proceedings of the 1st International Conference on Natural Resources Engineering & Technology 2006*, pages 484–491, Putrajaya, Malaysia, July 2006.