

## An Online Chronic Disease Prediction System Based on Incremental Deep Neural Network

Bin Yang<sup>1,\*</sup>, Lingyun Xiang<sup>2</sup>, Xianyi Chen<sup>3</sup> and Wenjing Jia<sup>4</sup>

<sup>1</sup>School of Design, Jiangnan University, Wuxi, 214122, China

<sup>2</sup>School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha, 410114, China

<sup>3</sup>Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information Science and Technology, Nanjing, 210044, China

<sup>4</sup>School of Electrical and Data Engineering, University of Technology Sydney, Sydney, 2000, Australia

\*Corresponding Author: Bin Yang. Email: Yangbin@jiangnan.edu.cn

Received: 21 October 2020; Accepted: 12 November 2020

**Abstract:** Many chronic disease prediction methods have been proposed to predict or evaluate diabetes through artificial neural network. However, due to the complexity of the human body, there are still many challenges to face in that process. One of them is how to make the neural network prediction model continuously adapt and learn disease data of different patients, online. This paper presents a novel chronic disease prediction system based on an incremental deep neural network. The propensity of users suffering from chronic diseases can continuously be evaluated in an incremental manner. With time, the system can predict diabetes more and more accurately by processing the feedback information. Many diabetes prediction studies are based on a common dataset, the Pima Indians diabetes dataset, which has only eight input attributes. In order to determine the correlation between the pathological characteristics of diabetic patients and their daily living resources, we have established an in-depth cooperation with a hospital. A Chinese diabetes dataset with 575 diabetics was created. Users' data collected by different sensors were used to train the network model. We evaluated our system using a real-world diabetes dataset to confirm its effectiveness. The experimental results show that the proposed system can not only continuously monitor the users, but also give early warning of physiological data that may indicate future diabetic ailments.

**Keywords:** Deep learning; incremental learning; network architecture design; chronic disease prediction

### 1 Introduction

Chronic disease is a common physiological health problem among humans across gender, race and age. In this work, we take type 2 diabetes as an example to verify the effectiveness of the proposed method. It has been shown that early detection of high-risk population can prevent



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

or delay 80% of type 2 diabetes complications. The medical decision-making process for early detection of diabetes is an urgent need. Due to the low efficiency of traditional manual data analysis, computer-based analysis becomes essential [1]. To this aim, many computer data analysis methods have been considered and examined. Data mining is a big advance in the type of analysis tools. It has been proved that the introduction of data mining into disease prediction can improve the accuracy of diagnosis, reduce costs and save human resources [2].

A number of disease predictive systems using various classification techniques (such as Naive Bayes (NB) [3], Multi-domain learning [4], dual-chaining [5] and nonlinear activation [6]) have been proposed. However, one of the main challenges of neural networks is that when users define the topology of neural networks, the number of hidden layers, the number of neurons, the number of times and the learning rate must be optimized. These parameters have to be confirmed before training. To automatically adjust the network parameters, AutoMLP [7] was proposed. However, the accuracy of this method still unsatisfactory in real environment.

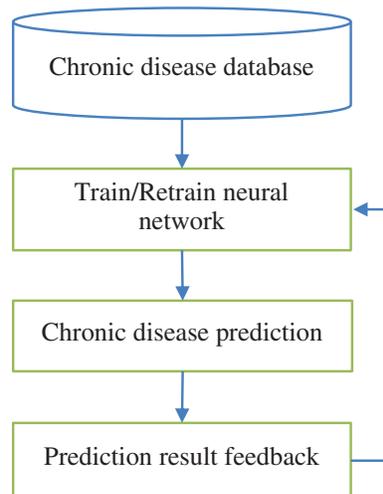
Deep learning [8–11] has developed rapidly in recent years. They are widely used in medical field, including medical prognosis and cell event detection. Phan et al. [12] found that due to the lack of annotated data, supervised deep learning methods have inherent limitations. They proposed a novel unsupervised two-path input neural network architecture to capture the irregular changes in cell appearance and motion. Unsupervised deep learning methods show an advantage in general (non-cell) video because they can learn the visual appearance and motion of events that occur periodically. A problem of poor performance in the prognosis of diabetes is data overfitting. Ashiqzaman et al. [13] developed a prediction system for the disease of diabetes. The overfitting problem was minimized by using dropout method. They tested their system on Pima Indians Diabetes Dataset (PIDD), and obtained a remarkable performance. However, the lack of data and training attributes of PIDD limits its application.

The parameters of network model need to be constantly updated due to the complexity and unpredictability of human body and disease, as well as the irregularity in of biological information [14]. In this paper, a novel network model is proposed to adapt to the changing human physiological data through incremental learning. With the increase of user's time, the intelligent system can predict diabetes more and more accurately according to feedback information. The overall architecture is shown in Fig. 1.

Compared with the current state-of-the-art of prediction methods of diabetics, the contribution of our work can be summarized as follows:

- The proposed network model is able to predict the risk of diabetes for patient in real time.
- The proposed system can adapt to the constantly changing physiological data of users by using incremental learning technique.
- When misjudgment occurs, the error samples will be fed back to the server for network model adjustment. Thus, after a period of incremental learning, the network model will be more and more adaptable to user.

The rest of the paper is organized as follows. Section 2 presents the related work of incremental learning algorithms. The proposed method is presented in Section 3. Section 4 presents the experimental results and discussions. Finally, the concluding remarks are given in Section 5.



**Figure 1:** Overview of our proposed diabetes prediction system

## 2 Related Work

Machine learning can be divided into supervised learning, unsupervised learning [15] and semi-supervised learning [16]. In supervised learning, training samples need to be labeled, while unsupervised learning can train without labeled samples. Most samples have no labels in the real world, so unsupervised learning is easier to apply than supervised learning. As an important type of unsupervised learning, clustering has been widely used in data classification when training data are not available. Some clustering algorithms, such as k-means [17], Expectation-Maximization (EM) [18], and their variations [19–24] have been exploited for classification applications. BP network model has been widely used in diabetes prediction system. However, there is a disadvantage that the neural network model must be retrained when new detection data is generated. In addition, when the number of online users increases, the remote server may not be able to complete the training task in time. Reference [25] discovered that BP neural network models are not suitable for on-line prediction system. Incremental learning [26,27] was proposed to overcome this weakness. In this section, a brief background on classical incremental learning algorithms are presented.

### 2.1 Self-Organizing Map

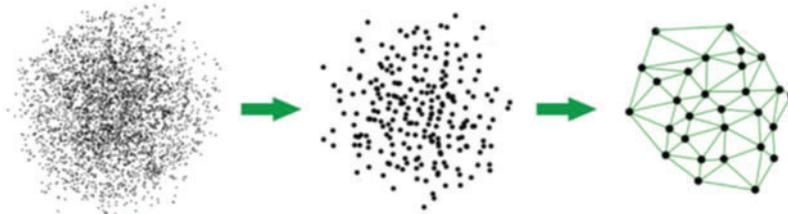
Self-organizing Map (SOM) [28] is one of the famous incremental learning models. It provides a guaranteed topological mapping from a high dimensional space to a mapped neuron. SOM usually contains a two-layer neural network based on competitive learning, which can be used for online clustering and topological representation without prior knowledge. Moreover, it is robust to noise data. There are many improved versions of SOM that have similar learning mechanisms.

A certain number of neurons are randomly distributed in a certain space in SOM models. Their connections are initialized to determine the initial topology of the network. In training process, all neurons compete with each other for the response power to the current input. The winning neurons update their parameters to adapt to the new input. This competition mechanism eventually makes neurons in different regions more sensitive to different input patterns. Therefore, competitive neural network is a model for pattern recognition. However, it is difficult to obtain stable learning results while maintaining plasticity. We need an adaptive learning system which

can adapt to the changing environment in real time. If the system is too stable, it cannot adapt to the fast-changing environment. On the contrary, if the system is too sensitive to external stimuli, it is difficult to stably save previously learned knowledge or even converge to a stable state [29]. When there is no prior knowledge and the external input mode changes with time, the self-organization of the network and the incremental of the algorithm are the key points of the learning system. One of the most representative self-organization models is Growing Neural Gas (GNG) [30] network. Its neurons can dynamically increase with the input data. GNG network has no parameters that change with time and can learn continuously. It can continuously update the network model by adding nodes and connections until it reaches the performance standard. Thus, GNG can adapt to the change of input mode and is more dynamic than SON network.

## 2.2 Self-Organizing Incremental Neural Network

In order to further improve the plasticity of the network, Self-organizing Incremental Neural Network (SOINN) based on SOM and GNG was proposed in [31]. SOINN uses two-layer neural network to represent the topological structure of unsupervised on-line data. By using a similarity threshold-based and a local error-based insertion criterion, the network is able to grow incrementally. Its working process is shown in Fig. 2.



**Figure 2:** Two stages of competitive learning in SOINN

SOINN finds the closest node (winner) and the second closest node (second winner) of the input vector. Similarity threshold is used to determine whether the input vector belongs to the same cluster or the second winner. If node  $I$  has a neighbor node, the maximum distance between node  $I$  and its neighboring nodes is used to calculate the similarity threshold  $T_i$ .

$$T_i = \max_{j \in N_i} \|W_i - W_j\| \quad (1)$$

where  $N_i$  is the set of neighbor nodes of node  $i$  and  $W_i$  is the weight vector of node  $i$ . If node  $i$  has no neighbor node, similarity threshold  $T_i$  is defined as the minimum distance between node  $i$  and other nodes in the network.

$$T_i = \min_{j \in N \setminus \{i\}} \|W_i - W_j\| \quad (2)$$

where  $N$  is the set of all nodes. If the input vector  $V$  is defined to belong to the cluster of first winner  $S_1$  or second winner  $S_2$ , then an edge will connect with  $S_1$  and  $S_2$ . The 'age' of the edge is set to '0'.

$$age(S_1, S_2) = 0 \quad (3)$$

The age of all edges linked to the winner are then increased by ‘1’. The weight vector of the winner and its neighboring nodes is updated as follow:

$$age(S_1, i) = age(S_1, i) + 1 \quad (4)$$

To find the winner  $S_1$ , and second winner  $S_2$ , the nodes set is searched by the following:

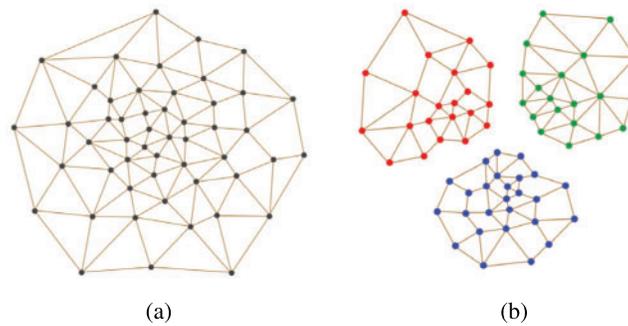
$$\begin{aligned} S_1 &= \arg \min_{C \in A} \|X - W_C\| \\ S_2 &= \arg \min_{C \in A \setminus \{S_1\}} \|X - W_C\| \end{aligned} \quad (5)$$

where connection set  $C$  is initialized to empty set  $C \subset A \times A$ . The input vector  $V$  is defined as a new node and is added into  $A$ , if  $\|X - W_{S_1}\| > T_{S_1}$  or  $\|X - W_{S_2}\| > T_{S_2}$ . If  $S_1$  and  $S_2$  are not connected, these closest nodes are connected with ‘age’ of ‘0’.

The change  $\Delta W_i$  to the weight of winner and change  $\Delta W_j$  to the weight of the neighbor node  $j \in N_i$  of  $i$  are defined as  $\Delta W_i = \frac{1}{M_i} (W_S - W_i)$  and  $\Delta W_j = \frac{1}{100M_i} (W_S - W_i)$ , where  $W_S$  is the weight of the input vector. If an edge is older than the predefined parameter,  $agemax$ , the edge is removed.

The SOINN network would add the new node into the right position where the accumulating error is extremely large after  $\lambda$  learning iterations ( $\lambda$  is a timer). And the insertion would be canceled if the insertion cannot decrease the error. SOINN can find new clusters in data flow and learn without affecting the architecture of previous results by using incremental way. SOINN can be adjusted accordingly, which is suitable for robot intelligence, computer vision, expert system, anomaly detection and other fields.

Although SOINN has shown excellent classification ability in some applications, it still has some disadvantages. User should determine the stop time of the first level learning and the start time of the second level learning. In addition, if there is a high-density overlap between clusters, the clusters in the network will be linked together to form a new cluster. If the learning results of the first level changed, the second level must be completely retrained. Therefore, the second layer of SOINN is not suitable for online incremental learning. If a prototype  $x_i$  is the nearest neighbor or the second nearest neighbor of the given sample  $n$ , the threshold will change. This means that new information can be learned without destroying previous knowledge. Thus, the second layer does not need to be fully retrained. Furao et al. [32] proposed an Enhanced Self-organizing Incremental Neural Network (ESOINN) to accomplish online unsupervised learning tasks. ESOINN is proofed to be superior to SOINN in the following respects: (1) it adopts a single-layer network to take the place of the two-layer network structure of SOINN; (2) it separates clusters with high-density overlap; (3) it uses fewer parameters than SOINN; and (4) it is more stable than SOINN [32]. Fig. 3 presents a comparison of SOINN and ESOINN in same dataset.



**Figure 3:** The comparison of SOINN and ESOINN in same dataset. (a) SOINN (b) ESOINN

### 3 Chronic Disease Prediction

First of all, we cooperated with a hospital to obtain pathological diagnosis data of diabetes on a certain scale. Before providing the pathological data of these diabetic patients, name, ID and other information had been deleted to protect the privacy of patients. Second, we input these hospital patients' pathological data into the proposed neural network for training. The pathology of diabetes was closely related to some factors, such as the changes of daily life habits, clinical symptoms, standard values, high-risk groups and so on. Finally, an intelligent prediction system based on incremental deep learning was established, which can accurately predict the incidence of diabetes. The proposed system collected user's daily life data through the sensors on the smart home appliances. Then the periodicity and regularity of the data are analyzed. People at high risk of diabetes will be warned to see a doctor.

#### 3.1 Identifying the Characteristics of Diabetic Patients

Many diabetics prediction studies were based on a common dataset "The Pima Indians Diabetes Dataset" [2,13]. There are 768 instances in this dataset, and all instances have 8 input attributes (from  $X_1$  to  $X_8$ ) and 1 output attribute (Y), which are listed in Tab. 1.

**Table 1:** The attributes of "the pima indians diabetes dataset"

Attribute No.	Attribute	Description	Type
$X_1$	PREGNANT	Numbers of time pregnant	Numeric
$X_2$	GTT	Plasma glucose concentration in an oral glucose tolerance test	Numeric
$X_3$	BP	Diastolic blood pressure (mmHg)	Numeric
$X_4$	SKIN	Triceps skin fold thickness (mm)	Numeric
$X_5$	INSULIN	Serum insulin ( $\mu$ U/ml)	Numeric
$X_6$	BMI	Body mass Index (kg/m)	Numeric
$X_7$	DPF	Diabetes pedigree function	Numeric
$X_8$	AGE	Age of patient(years)	Numeric
Y	DIABETES	Diabetes diagnose results ("tested positive," "tested negative")	Numeric

The number of samples in this dataset is not enough for machine learning. Due to the difference of physical characteristics between Asians and Europeans, the dataset cannot accurately predict the diabetes of Asians (especially Chinese). In order to determine the correlation between diabetic patients' pathological characteristics and daily life data, we had carried out in-depth cooperation with a hospital. Patients with diabetes and impaired glucose tolerance usually monitor their blood glucose through lifestyle and physical changes. These changes include Body Mass Index (BMI), Waist to Hip Ratio (WHR), blood pressure, blood lipid, fasting blood glucose, OGTT, 2-hour Postprandial Blood Glucose (2hPBG), etc. Finally, we created a Chinese Diabetes Dataset (CDD) containing 575 patients. Some samples are shown in Tab. 2. An application was developed to collect data from intelligent device.

**Table 2:** Some samples in CDD

Sex	Age	Height	Weight	Pres_h	Pres_l	FPG	2hPBG	Diabetes
1	64	157	64	105/70	5.7	7.2	64	0
1	65	154	60	160/80	5.4	8	65	0
0	46	170	81	130/90	5.2	6.5	46	0
1	55	163	52	120/60	5.3	6.8	55	0
0	20	182	63	100/70	5.1	5	20	0
1	43	162	69	140/80	7.1	12.1	43	1
1	45	158	61	120/80	5	7.5	45	0
1	48	148	40	110/85	5.1	7.5	48	0
1	21	161	53	110/60	5.3	5.7	21	0
0	32	168	56	110/70	5.1	6	32	0
0	55	164	52	105/80	5.4	6.4	55	0
1	36	154	59	110/80	6.2	6.5	36	0
1	58	171	90	180/110	5	8.3	58	0
1	48	161	60	130/80	12.7	20.6	48	1
0	29	166	57	110/60	4.9	7	29	0

### 3.2 The Proposed Incremental Learning Network Model

This paper presents a method to adjust network parameters online by learning user feedback data. The accuracy and personalization of the network model can be improved. The goal of proposed deep learning model is to identify user feedback data which contains a lot of information iteratively. Using a large number of unlabeled samples to expand the training data set with large amount of information can improve the prediction accuracy. Thus, unlabeled samples are selected according to their potential contribution to training.

Both between-class and within-class insertions are important to the second layer of SOINN [31]. One drawback is that if the results of the first level change, all the learning results of the second level will be destroyed. The second layer must be retrained, which means that the second layer of SOINN is not suitable for online incremental learning. Since many nodes are generated in high-density regions, the distance between adjacent nodes in these regions will be shorter. Here, we define the range thresholds  $T_{s1}$  and  $T_{s2}$  as the Euclidean distance between two neurons.  $S1$  and  $S2$  are two neurons in neuron set which is close to new input data. The winner and second winner of the input data are located. Based on the range thresholds  $T_{s1}$  and  $T_{s2}$ , these

two winners will be added or removed the connection between them. Then we update the winner's density and weight after each learning. Nodes that caused by noise should also be deleted in the process. After learning, all nodes will be classified into different classes. We summarize the process of learning a new unlabeled data which is shown in Algorithm 1.

---

**Algorithm 1:** Proposed incremental learning algorithm
 

---

Input:

$\xi$ : new input data

A: neuron set

C: connection matrix

Begin

1: Initialize to neuron set  $A = \{C1, C2\}$ , where the weights of neurons  $C1$  and  $C2$  are  $W_1, W_2 \in R^n$ .

2: Initialize edge set  $C \subseteq A \times A$ , where C is an empty set, that is, there is no initial connection between neurons.

3: Establish a two-dimensional matrix  $age(m, n) = -1$ , where  $m, n$  are maximum integers.  $m, n$  can be increased automatically according to the actual situation. '-1' means that the no connection between neurons  $m$  and  $n$ .

4: Define the range thresholds  $T_{s1}$  and  $T_{s2}$  as the Euclidean distance between two neurons, that is,  $T_{S1} = T_{S2} = \|W_1 - W_2\|$ .

5: Find out two neurons  $S1$  and  $S2$  in  $A$  which is closest to  $\xi$ .

$$S1 = \arg \min_{C \in A} \|\xi - W_C\|, \quad S2 = \arg \min_{C \in A \setminus \{S1\}} \|\xi - W_C\|$$

6: If  $\|\xi - W_{S1}\| > T_{S1}$  or  $\|\xi - W_{S2}\| > T_{S2}$ , then create a new node  $R$  for  $\xi$ . Add  $R$  into  $A$ , and set  $W_r = X$ . Go to step 5.

7: If there is no connection between  $S1$  and  $S2$ , create a connection for them.

$$C = C \cup \{(S1, S2)\}, \quad age_{(S1, S2)} = 0$$

8: All the edge's age connected to the winner node is added '1'.

$$age_{(S1, i)} = age_{(S1, i)} + 1, \quad \text{where } i \text{ is the point connected with } S1.$$

9: Update the weights of two winner nodes.

$$W_{S1} = W_{S1} + \varepsilon(t) (\xi - W_{S1})$$

$$W_{S2} = W_{S2} + \varepsilon'(t) (\xi - W_{S2})$$

$$\varepsilon(t) = \frac{1}{t}, \quad \varepsilon'(t) = \frac{1}{100t}$$

where the parameter  $\varepsilon(t)$  is the different learning rate of each neuron, and  $t$  is the number of times that the neuron becomes the winner.

10: For connection  $e(i, j)$  in edge set C do

If  $age(i, j) > agemax$  then Remove  $e(i, j)$  from C

End For

where the value of  $agemax$  is set to 1/3 of the dimension of input data vector according to the experimental results.

---

(Continued)

---

11: Update the range thresholds  $T_{s1}$  and  $T_{s2}$  to the maximum distance from the adjacent datum points of S1 and S2.

$$T_{S1} = \operatorname{argmin}_{\xi, S1} \|\xi - S1\|,$$

$$T_{S2} = \operatorname{argmin}_{\xi, S2} \|\xi - S2\|$$

12: If the total number of input data samples is an integral multiple of  $\lambda$ , then

Check the neuron set A from whole neural network.

If  $\operatorname{Count}(\{si, sj\}) = 1, \{si, sj\} \in C$  then

Set  $j$  as the neuron connected with  $Si$ ,

Delete  $Si$  from neuron set A.

Endif

Endif

End

Output:

The neuron set A and connection matrix C.

---

#### 4 Experiments Results and Analysis

Experiments were carried out on a PC server with a GPU (Nvidia GeForce GTX 1080Ti with 11GB RAM) to evaluate the performance of the proposed method. Previous diabetics prediction studies were using a common dataset ‘‘The Pima Indians Diabetes Dataset’’ (PIDD) to test their performance [2,13]. PIDD was obtained from the UCI machine learning repository [33]. It is a subset of a bigger dataset held by the National Institute of Diabetes and Digestive and Kidney Diseases. The patient data provided in this dataset are Pima Indian genetic women over the age of 20. The output variable is 0 or 1, where 0 indicates negative detection and 1 indicates positive detection. All diabetes data have been normalized. 268 (34.9%) samples are positive and 500 (65.1%) samples are negative. However, as described in Section 3.1, PIDD is too small, and it is all non-Asian samples. We had created a Chinese Diabetes Dataset (CDD) for experiment with 5222 patients. 70% of the samples were used for training and the rest for testing.

**Table 3:** Results of different methods on PIDD

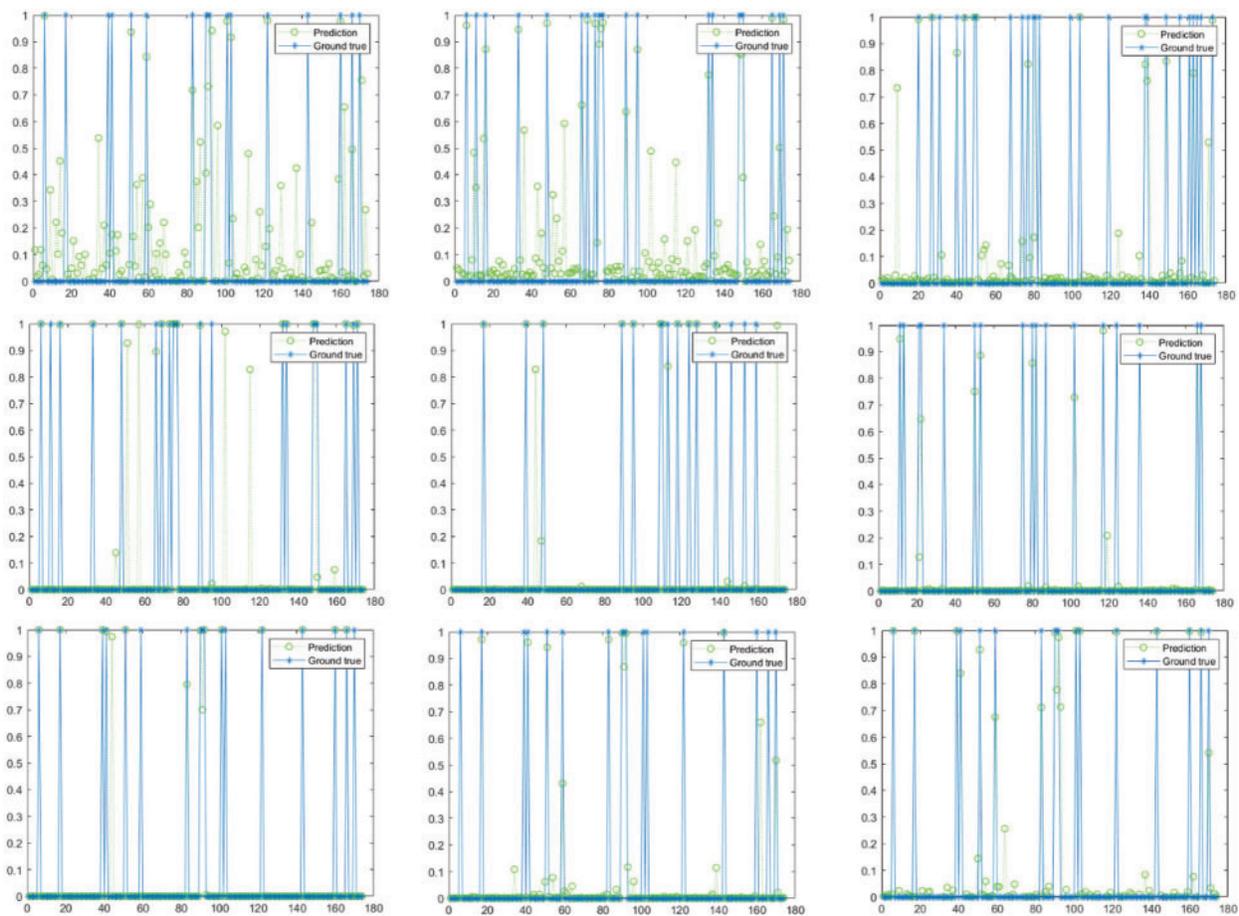
No	Method name	Remarks	Accuracy (%)
1	Yang et al. [2]	Bayes network	72.3
2	Jahangir et al. [34]	Automatic multilayer perceptron	88.7
3	Kim et al. [35]	DNN	86.7
4	Ashiquzzaman et al. [13]	DNN, with Dropout	88.4
5	Kayaer et al. [36]	General regression neural network	83.0
6	Polat et al. [37]	Generalized discriminant analysis and LS-SVM	82.1
7	Purushottam et al. [38]	C4.5 rules and partial tree	81.3
8	Nnamoko et al. [39]	Meta model of 5 classifier	77.0
9	Kalaiselvi and Nasira [40]	Neuro-Fuzzy inference system	80.0
10	Daho et al. [3]	Neuro-fuzzy Classifier	82.3
11	Proposed method	Incremental learning network	89.1

In order to evaluate the effectiveness of the proposed diabetes prediction method, four types of methods (Bayes network [2], automatic multilayer perceptron [34], DNN [35] and DNN with

Dropout [13]) were adopted. We used accuracy as the percentage of patients that are correctly diagnosed by prediction methods. The first experiment tested the accuracy diabetes prediction based on PIDD dataset. The comparison results of the previous methods discussed in the paper as well as result of the proposed method are depicted in Tab. 3. In this experiment, we only used 8 attributes provided in PIDD dataset to trained the proposed neural network.

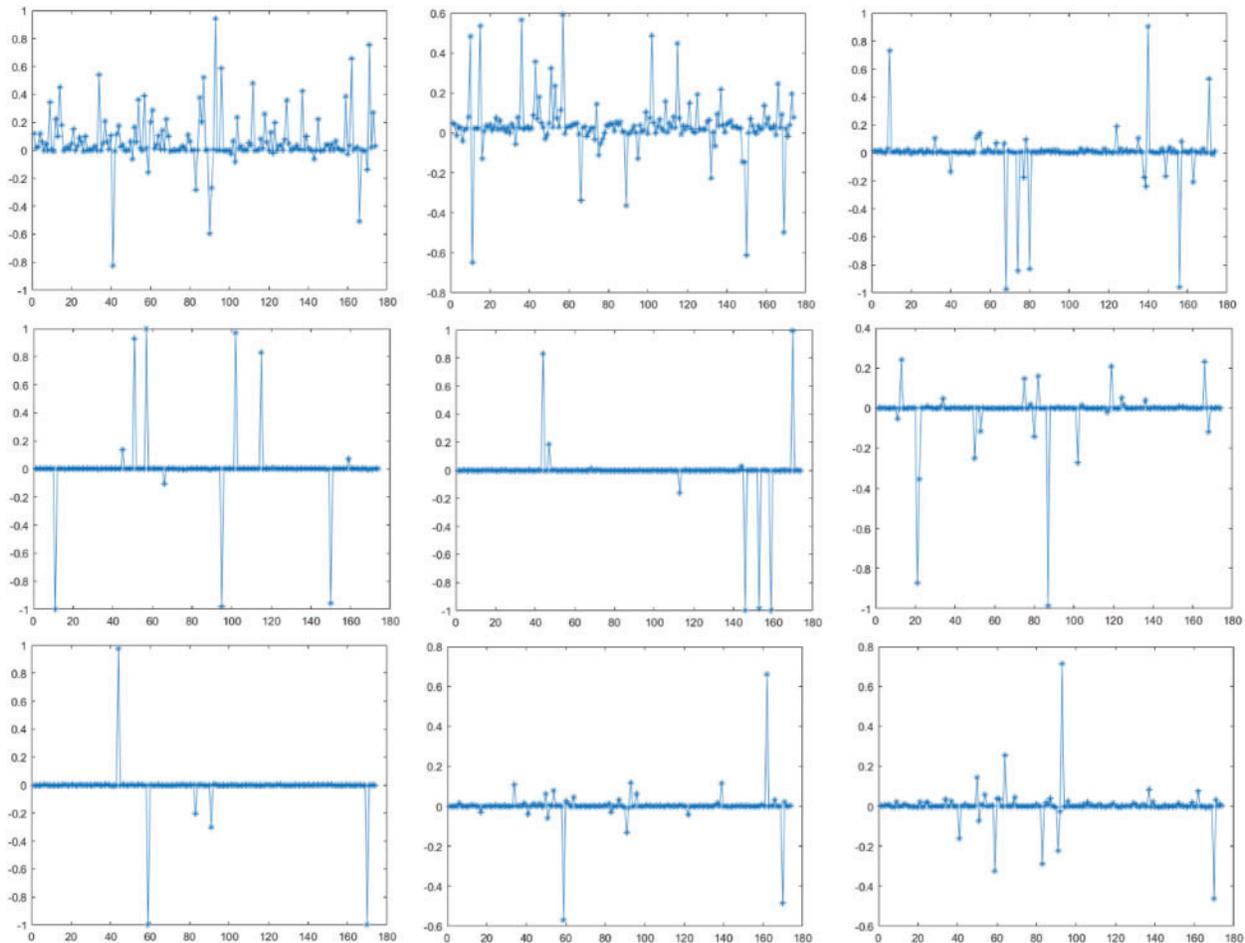
**Table 4:** Results of different methods on CDD

ID	Method name	Remarks	Accuracy (%)
1	Yang et al. [2]	Bayes network	78.2
2	Jahangir et al. [34]	Automatic multilayer perceptron	90.1
3	Kim et al. [35]	DNN	69.4
4	Ashiquzzaman et al. [13]	DNN, with dropout	84.7
5	Proposed method	Incremental learning network	92.4



**Figure 4:** The prediction results after 1 to 9 incremental training. In the first row of figures, from left to right are the 1st, 2nd and 3rd training respectively. In the second row of figures, from left to right are the 4th, 5th and 6th training respectively, and so on

It can be seen that in the first test on PIDD, our proposed method outperformed existing state-of-the-art methods. However, the experimental results of proposed method were not much better than those of others. This is because the depth of the neural network model used for prediction was not enough, resulting in the bottleneck of prediction accuracy of proposed method. Moreover, the lack of training attributes provided by PIDD data sets also limited the performance of artificial neural networks. Therefore, in the second experiment, we evaluated four well-known methods [2,13,34,35] to test the performance on CDD dataset as depicted in Tab. 4.



**Figure 5:** The prediction errors according to Fig. 4. In the first row of figures, from left to right are the 1st, 2nd and 3rd training respectively, and so on

Experimental results demonstrated that the more abundant attributes provided by CDD dataset can significantly enhance the accuracy of diabetes prediction. Meanwhile, a larger amount of data makes the deep neural network model easier to learn the characteristics of diabetes. Note that, the accuracy of [35], which using DNN for prediction, was quite low. This is because the depth neural network they used had excessive hidden layers and lacked of weight reduction means. In its training process, the neural network fitted the noise in the training data and the unrepresentative features in the training samples, so that the network was overfitted. The accuracy

of their method can be improved by using two-dimensional data projection technology [41]. In [13], the prediction was also based on DNN, and the performance was improved by using Dropout technique. In CDD dataset, the proposed method obtained a significant score. In order to further demonstrate the incremental learning characteristics of our proposed method, we used another diabetes daily dataset (DDD) containing daily life records (i.e., walking distance, sleeping time) of patients. The third experiment was conducted on DDD. In this experiment, samples were divided into 9 groups to test the efficiency of incremental learning. Prediction results after 1 to 9 incremental training processes were presented in Fig. 4.

Prediction errors according to Fig. 4 were also calculated and shown in Fig. 5. Experimental results demonstrate that the proposed incremental learning model is able to update its parameters continuously according to the feedback information.

## 5 Conclusions

Artificial intelligence and big data analysis are getting more and more attention in medical application research. In this paper, a novel incremental learning network model was proposed. Using this network model, a chronic disease prediction system can evaluate the possibility of chronic disease for users in real time. We have also developed a data collection process to continuously collect physiological information from users. The error will be fed back to the server system, and the neural network model will be adjusted by processing the misjudgment information. When users have similar feature information again, the prediction system will give a warning. In this case, the parameters of neural network would not be updated. After a period of incremental learning and feedback adjustment, neural network will be more and more adaptive to each user. A series of experiments verify the effectiveness of our proposed system.

**Acknowledgement:** We would like to thank Zhangjie Fu and Zhili Zhou for their contributions to this manuscript.

**Funding Statement:** This work has received funding from the Humanities and Social Sciences Projects of the Ministry of Education (Grant No. 18YJC760112, Bin Yang); the Social Science Fund of Jiangsu Province (Grant No. 18YSD002, Bin Yang) and Open Fund of Hunan Key Laboratory of Smart Roadway and Cooperative Vehicle Infrastructure Systems (Changsha University of Science and Technology) (Grant No. kfj180402, Lingyun Xiang).

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- [1] B. Wang, W. Kong, H. Guan and N. Xiong, "Air quality forecasting based on gated recurrent long short term memory model in Internet of Things," *IEEE Access*, vol. 7, no. 2, pp. 69524–69534, 2019.
- [2] Y. Guo, G. H. Bai and Y. Hu, "Using bayes network for prediction of type-2 diabetes," in *Int. Conf. for Internet Technology and Secured Transactions*, London, UK, pp. 471–472, 2012.
- [3] M. E. H. Daho, N. Settouti, M. E. A. Lazouni and M. A. Chikh, "Recognition of diabetes disease using a new hybrid learning algorithm for NEFCLASS," in *IEEE Proc. of the 8th Int. Workshop on Systems, Signal Processing and Their Applications*. Algiers, Algeria, pp. 239–243, 2013.
- [4] B. Yang, Z. Li and T. Zhang, "A real-time image forensics scheme based on multi-domain learning," *Journal of Real-Time Image Processing*, vol. 17, no. 1, pp. 29–40, 2020.

- [5] B. Wang, W. Kong, W. Li and N. N. Xiong, "A dual-chaining watermark scheme for data integrity protection in Internet of Things," *Computers, Materials & Continua*, vol. 58, no. 3, pp. 679–695, 2019.
- [6] F. Yu, L. Liu, L. Xiao, K. Li and S. Cai, "A robust and fixed-time zeroing neural dynamics for computing time-variant nonlinear equation using a novel nonlinear activation function," *Neurocomputing*, vol. 350, no. 20, pp. 108–116, 2019.
- [7] C. Sandeep, P. Chakrabarti and N. Chourasia, "Prediction of breast cancer biopsy outcomes—An approach using machine learning perspectives," *International Journal of Computer Applications*, vol. 100, no. 9, pp. 29–32, 2014.
- [8] L. Xiang, G. Guo, J. Yu, V. Sheng and P. Yang, "A convolutional neural network-based linguistic steganalysis for synonym substitution steganography," *Mathematical Biosciences and Engineering*, vol. 17, no. 2, pp. 1041–1058, 2020.
- [9] C. L. Yin, Y. F. Zhu, J. L. Fei and X. Z. He, "A deep learning approach for intrusion detection using recurrent neural networks," *IEEE Access*, vol. 5, pp. 21954–21961, 2017.
- [10] X. Chen, H. Zhong and Z. Bao, "A GLCM feature based approach for reversible image transformation," *Computers, Materials & Continua*, vol. 59, no. 1, pp. 239–255, 2019.
- [11] Z. Zhou, J. Q. M. Wu and X. Sun, "Multiple distance-based coding: Toward scalable feature matching for large-scale web image search," *IEEE Transactions on Big Data*, 2019.
- [12] H. T. H. Phan, A. Kumar, D. Feng, M. Fulham and J. Kim, "Unsupervised two-path neural network for cell event detection and classification using spatiotemporal patterns," *IEEE Transactions on Medical Imaging*, vol. 38, no. 6, pp. 1477–1487, 2019.
- [13] A. Ashiquzzaman, A. K. Tushar, M. R. Islam, D. Shon, K. Im *et al.*, "Reduction of overfitting in diabetes prediction using deep learning neural network," in *Proc. IT Convergence and Security*, Singapore, pp. 35–43, 2018.
- [14] Y. Cao, Z. Zhou, X. Sun and C. Gao, "Coverless information hiding based on the molecular structure images of material," *Computers, Materials & Continua*, vol. 54, no. 2, pp. 197–207, 2018.
- [15] L. Xiang, G. Zhao, Q. Li, W. Hao and F. Li, "TUMK-ELM: A fast unsupervised heterogeneous data learning approach," *IEEE Access*, vol. 6, pp. 35305–35315, 2018.
- [16] Y. Tu, Y. Lin, J. Wang and J. U. Kim, "Semi-supervised learning with generative adversarial networks on digital signal modulation classification," *Computers, Materials & Continua*, vol. 55, pp. 243–254, 2018.
- [17] J. Qin, W. Fu, H. Gao and W. X. Zheng, "Distributed k-means algorithm and fuzzy c-means algorithm for sensor networks based on multiagent consensus theory," *IEEE Transactions on Cybernetics*, vol. 47, no. 3, pp. 772–783, 2017.
- [18] H. Li, W. Luo, X. Qiu and J. Huang, "Identification of various image operations using residual-based features," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 1, pp. 31–45, 2018.
- [19] B. Yang, X. Sun, E. Cao, W. Hu and X. Chen, "Convolutional neural network for smooth filtering detection," *Iet Image Processing*, vol. 12, no. 8, pp. 1432–1438, 2018.
- [20] L. Mici, G. I. Parisi and S. Wermter, "A self-organizing neural network architecture for learning human-object interactions," *Neurocomputing*, vol. 307, no. 13, pp. 14–24, 2018.
- [21] Z. Zhou, Y. Cao and M. Wang, "Faster-RCNN based robust coverless information hiding system in cloud environment," *IEEE Access*, vol. 7, pp. 179891–179897, 2019.
- [22] L. Xiang, Y. Li, W. Hao, P. Yang and X. Shen, "Reversible natural language watermarking using synonym substitution and arithmetic coding," *Computers, Materials and Continua*, vol. 55, no. 3, pp. 541–559, 2018.
- [23] Z. Zhou, J. Qin, X. Xiang, Y. Tan, Q. Liu *et al.*, "News text topic clustering optimized method based on TF-IDF algorithm on spark," *Computers, Materials & Continua*, vol. 62, no. 1, pp. 217–231, 2020.
- [24] B. Yang, Z. Li, Y. Sun and E. Cao, "EM-FEE: An efficient multitask scheme for facial expression estimation," *Interacting with Computers*, vol. 32, no. 2, pp. 142–152, 2020.
- [25] Y. Xian and H. Hu, "Enhanced multi-dataset transfer learning method for unsupervised person re-identification using co-training strategy," *IET Computer Vision*, vol. 12, no. 8, pp. 1219–1227, 2018.

- [26] Y. Xing, F. Shen and J. Zhao, "Perception evolution network based on cognition deepening model—adapting to the emergence of new sensory receptor," *IEEE Transactions on Neural Networks & Learning Systems*, vol. 27, no. 3, pp. 607–620, 2017.
- [27] Y. Chen, X. Jie, W. Xu and J. Zuo, "A novel online incremental and decremental learning algorithm based on variable support vector machine," *Cluster Computing*, vol. 22, no. 8, pp. 7435–7445, 2019.
- [28] T. Furukawa, "SOM of SOMs: Self-organizing map which maps a group of self-organizing maps," in *Proc. Int. Conf. on Artificial Neural Networks*, Warsaw, Poland, 2005.
- [29] B. Yang, X. Sun, H. Guo, Z. Xia and X. Chen, "A copy-move forgery detection method based on CMFD-SIFT," *Multimedia Tools and Applications*, vol. 77, no. 1, pp. 837–855, 2018.
- [30] T. Yang, S. Yao and K. Xue, "Slicing point cloud incrementally for additive manufacturing via online learning," *Neural Computing and Applications*, vol. 21, no. 2, pp. 1–21, 2020.
- [31] F. Shen and O. Hasegawa, "An incremental network for on-line unsupervised classification and topology learning," *Neural Networks*, vol. 19, no. 1, pp. 90–106, 2006.
- [32] S. Furao, T. Ogura and O. Hasegawa, "An enhanced self-organizing incremental neural network for online unsupervised learning," *Neural Networks*, vol. 20, no. 8, pp. 893–903, 2007.
- [33] D. Dua and E. K. Taniskidou, "UCI machine learning repository," 2020. [Online]. Available: <http://archive.ics.uci.edu/ml/datasets.php>.
- [34] M. Jahangir, H. Afzal, M. Ahmed, K. Khurshid and R. Nawaz, "ECO-AMLP: A decision support system using an enhanced class outlier with automatic multilayer perceptron for diabetes prediction," 2017. [Online]. Available: <https://arxiv.org/abs/1706.07679>.
- [35] J. Kim, J. Kim, M. J. Kwak and M. Bajaj, "Genetic prediction of type 2 diabetes using deep neural network," *Clinical Genetics*, vol. 93, no. 4, pp. 822–829, 2018.
- [36] K. Kayaer and T. Yildirim, "Medical diagnosis on pima indian diabetes using general regression neural networks," in *Proc. Int. Conf. on Artificial Neural Networks and Neural Information Processing*, New York, NY, USA, pp. 181–184, 2003.
- [37] K. Polat, S. Güneş and A. Arslan, "A cascade learning system for classification of diabetes disease: Generalized discriminant analysis and least square support vector machine," *Expert Systems with Applications*, vol. 34, no. 1, pp. 482–487, 2008.
- [38] Purushottam, K. Saxena and R. Sharma, "Diabetes mellitus prediction system evaluation using C4.5 rules and partial tree," in *Proc. 4th Int. Conf. on Reliability, Infocom Technologies and Optimization (ICRITO)*, Noida, India, pp. 1–6, 2015.
- [39] N. A. Nnamoko, F. N. Arshad, D. England and J. Vora, "Meta-classification model for diabetes onset forecast: A proof of concept," in *Proc. IEEE Int. Conf. on Bioinformatics and Biomedicine (BIBM)*, Belfast, UK, pp. 50–56, 2014.
- [40] C. Kalaiselvi and G. M. Nasira, "A new approach for diagnosis of diabetes and prediction of cancer using ANFIS," in *Proc. World Congress on Computing and Communication Technologies*, Trichirappalli, India, pp. 188–190, 2013.
- [41] B. Yang, L. Wei and Z. Pu, "Measuring and improving user experience through artificial intelligence aided design," *Frontiers in Psychology*, vol. 11, 12381, 2020.