IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 10, 2019, accepted October 9, 2019, date of publication October 14, 2019, date of current version October 30, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2947307

Prediction of Negative Conversion Days of Childhood Nephrotic Syndrome Based on PCA and BP-AdaBoost Neural Network

YE LUO[1] AND BEIHONG WANG[2]

[1]Key Laboratory of Cloud Computing and Intelligent Information Processing of Changzhou City, Jiangsu University of Technology, Changzhou 213001, China
[2]Department of Pediatrics, Changzhou No. 2 People's Hospital, Changzhou 213003, China

Corresponding author: Beihong Wang (wang_beihong@163.com)

This work was supported in part by the Natural Science Foundation of Jiangsu Province in China under Grant BK20161199.

ABSTRACT The prognosis of childhood nephrotic syndrome directly hinges on the accurate prediction of negative conversion days (NCDs). Therefore, this paper designs a hybrid approach of principal component analysis (PCA) and backpropagation (BP)-adaptive boosting (AdaBoost) neural network (NN), and applies the method to predict the NCDs for children with nephrotic syndrome. Specifically, PCA method was employed for dimension reduction. Six principal components were extracted from multiple physiological features, and taken as input variables of three-tiered neural networks. The boosted predictor of BP-AdaBoost model, together with three predictors of BP NN, support vector machine (SVM) and radial basis function (RBF) NN, was trained for NCDs prediction. The experimental results show that the predictor of BP-AdaBoost NN achieves the mean absolute error of 0.2334, the mean relative error of 3.2789%, the SD of 4.6804 and the RSD of 50.8053% in NCDs prediction, and it outperforms other predictors of BP NN, SVM and RBF NN on both accuracy and precision. Furthermore, comparison experiments are conducted on PCA processed testing data and raw testing data for BP-AdaBoost NN and demonstrate the excellent effect of PCA. The hybrid approach of PCA and BP-AdaBoost NN is simple and reliable for NCDs prediction of childhood nephrotic syndrome, and it can help pediatricians prognose childhood nephrotic syndrome accurately and further provide patients with better care and treatment.

INDEX TERMS Childhood nephrotic syndrome, negative conversion days (NCDs), principal component analysis (PCA), backpropagation (BP)-adaptive boosting (AdaBoost) neural network (NN).

I. INTRODUCTION

Nephrotic syndrome, a collection of symptoms due to glomeruli damage, is a common childhood disease, which occurs in 1.5 per 100,000 children each year [1], [2]. The incidence of the disease varies in different populations [3]. The disease of nephrotic syndrome in children can run a frequently relapsing course, and consequently brings great harm to healthiness and growth of children, e.g. risk of acute kidney injury and growth retardation [4], [5]. Riar *et al.* [6] studied the lifetime prevalence of allergies in childhood nephrotic syndrome and discovered that two-thirds of children with nephrotic syndrome at presentation have allergic symptoms and asthma. Starc *et al.* [7] suggested that mesenchymal stromal cells may be used in autologous cellular therapy approaches for idiopathic nephrotic syndrome treatment in the pediatric population. Wang *et al.* [8] developed a novel text messaging system for disease monitoring in childhood nephrotic syndrome and it achieved better results than scheduled follow-up visits. Gipson *et al.* [9] pointed out that data management and analysis can support new therapies of nephrotic syndrome through clinical trials and improve patient health outcomes. Considering the poor self-care ability of children, it is important to prognose nephrotic syndrome of pediatric patients. The accurate prediction of negative conversion days (NCDs) helps to improve treatment and curing effect. The key to NCDs prediction is to find the complex nonlinear function that maps physiologic features of children with nephrotic syndrome to NCDs. Such features include gender, age, albumin, immunoglobulin, etc. [10]

In recent years, artificial neural networks have been widely applied in medical and clinical areas, including

The associate editor coordinating the review of this manuscript and approving it for publication was Sabah Mohammed.

the prediction of NCDs of childhood nephrotic syndrome. Liu *et al.* [11] applied the convolutional neural network (CNN) to automated segmentation of multi-spectral magnetic resonance (MR) image on stroke lesions, and achieved early definitive diagnosis of the patients with ischemic stroke. Gómez *et al.* [12] relied on the recurrent neural network (RNN) to estimate laryngeal pressure in the diagnosis of vocal cord dysfunction. Kawahara and Hamarneh [13] took the classification of dermoscopic features within super-pixels as a segmentation problem, and proposed a fully CNN to detect such features from dermoscopic skin lesion images. Choi *et al.* [14] proposed a new approach for real-time apnea-hypopnea event detection using CNNs, and the approach can both help reduce event detection time and be applied to screen sleep apnea hypopnea syndrome (SAHS) severity. With the aid of backpropagation (BP) neural network (NN), Fan *et al.* [15] effectively measured respiratory flow with portable pressure data, realized continuous respiratory monitoring, and thus provided an important tool for clinical monitoring. In particular, artificial neural networks have been successfully employed to diagnose and prognose various cancers [16]–[23]. For example, Alzubi *et al.* [16] suggested an ensemble of weight optimized neural network with maximum likelihood boosting (WONN-MLB) for lung cancer disease in big data and achieved accurate diagnosis of lung cancer. Horie *et al.* [17] developed deep learning method based on CNNs for detecting esophageal cancer and demonstrated the ability of neural networks in diagnosis of esophageal cancer. Li *et al.* [18] applied deep CNNs models on sonographic imaging data from clinical ultrasounds to diagnosis of thyroid cancer and successfully improved the diagnostic accuracy.

In study of Liu *et al.* [10], the NCDs of childhood nephrotic syndrome are predicted with a model based on BP NN. Specifically, the relationship between the physiologic features of patients and NCDs was established through BP NN, and then used to predict the NCDs of the disease. However, the method of Liu *et al.* [10] has two major defects:

(1) There are strong correlations between the multiple physiologic features. The strong correlations make the problem more complex and the modelling less efficient, damping the accuracy of NCDs prediction. Thus, it is inappropriate to use all physiologic variables to create the neural network prediction model.

(2) Traditional BP algorithm, as a local search optimization method, is prone to the local minimum trap. In the traditional algorithm, the minimum error function is taken as the approximation target, and the self-learning is completed by gradient descent. In this case, it is difficult for traditional BP algorithm to find the global optimum of complex nonlinear functions. Then, the neural network trained by BP algorithm may lead to low prediction accuracy and precision.

To overcome the two defects, this paper designs a hybrid approach of principal component analysis (PCA) and BP-adaptive boosting (AdaBoost) NN, and applies the method to predict the NCDs for children with nephrotic syndrome.

The remainder of the paper is organized as follows: Section II describes the dataset used in this research; Section III details the proposed approach and its application in NCDs prediction; Section IV presents the experimental results, main findings and limitations; Section V wraps up this paper with conclusions.

II. DATASET

In this research, the original dataset covers 90 samples of childhood nephrotic syndrome. The dataset was provided by Department of Pediatrics at Changzhou No.2 People's Hospital in China. First, blood tests were conducted and relevant biomarkers were obtained for diagnosis. Then medical treatments were offered to those children whose illnesses were diagnosed as childhood nephrotic syndrome. During treatment regular follow-up examinations were performed until patients had recovered and at the time NCDs were obtained. The dataset contains 9 biomarkers of children with nephrotic syndrome, including gender, age (years), immunoglobulin G (IgG), immunoglobulin A (IgA), immunoglobulin M (IgM), immunoglobulin E (IgE), complement component 3 (C3), albumin (ALB) and blood fat (BF). Except two biomarkers of gender and age, the remaining seven biomarkers were obtained through blood routine examinations. Figures 1 and 2 are respectively the multi-dimensional visualization and the box plot of the original samples. Note that 0 represents male and 1, female.

Because the size of the original samples which were clinically obtained is relatively small, the method known as data augmentation is used to produce virtual samples to enlarge the dataset. Data augmentation is formalized by Vicinal Risk Minimization (VRM), in which expert human is required to describe the vicinity of each example and the examples in the vicinity share the same calss. Zhang *et al.* [24] proposed a novel data-agnostic data augmentation approach termed *mixup* which incorporates the prior knowledge that linear interpolations of feature vectors lead to linear interpolations of the associated targets. The *mixup* constructs the virtual example of $(\tilde{x}, \tilde{y})$ by the following equation:

$$\begin{cases} \tilde{\boldsymbol{x}} = \lambda\, \boldsymbol{x_i} + (1-\lambda)\, \boldsymbol{x_j} \\ \tilde{\boldsymbol{y}} = \lambda\, \boldsymbol{y_i} + (1-\lambda)\, \boldsymbol{y_j} \end{cases} \tag{1}$$

where $(\boldsymbol{x}_\mathrm{i}, \boldsymbol{y}_\mathrm{i})$ and $(\boldsymbol{x}_\mathrm{i}, \boldsymbol{y}_\mathrm{i})$ are two examples randomly chosen from the original dataset, $\boldsymbol{x}_\mathrm{i}$ and $\boldsymbol{x}_\mathrm{j}$ are raw input vectors, $\boldsymbol{y}_\mathrm{i}$ and $\boldsymbol{y}_\mathrm{j}$ are one-hot label encodings, and $\lambda \in [0,1]$. In the *mixup*, $\lambda \sim \mathrm{Beta}(\alpha, \alpha)$, for $\alpha \in (0, \infty)$, and the hyper-parameter α controls the strength of interpolation. A detailed description of *mixup* can be found in literature [24].

In this research, the 300 virtual samples were constructed by using the data augmentation approach of *mixup*. Figures 3 and 4 are respectively the multi-dimensional visualization and the box plot of the virtual samples.

First 60 out of the 90 original samples were allocated to the testing dataset, and the 330 samples consisting of the remaining 30 original samples and the 300 virtual samples were allocated to the training dataset.

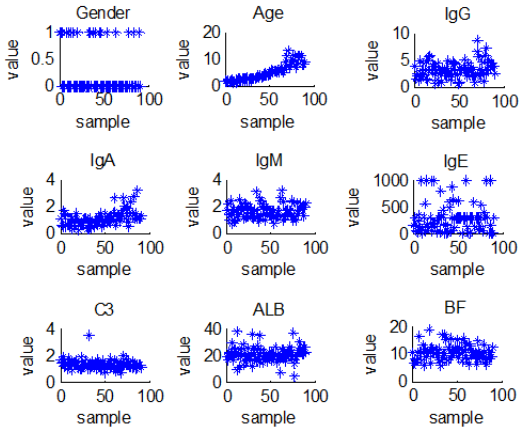

FIGURE 1. Multi-dimensional visualization of the original samples.

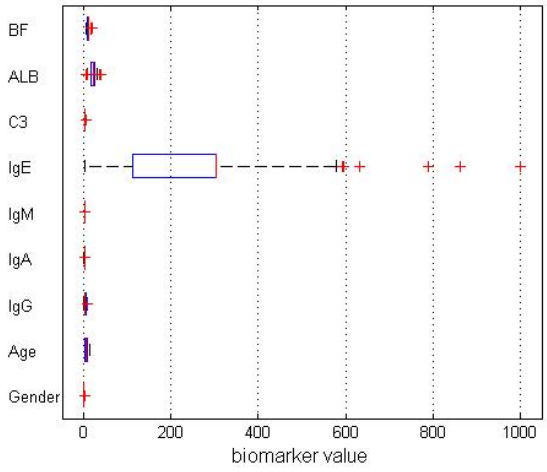

FIGURE 2. Bo plot of the original samples.

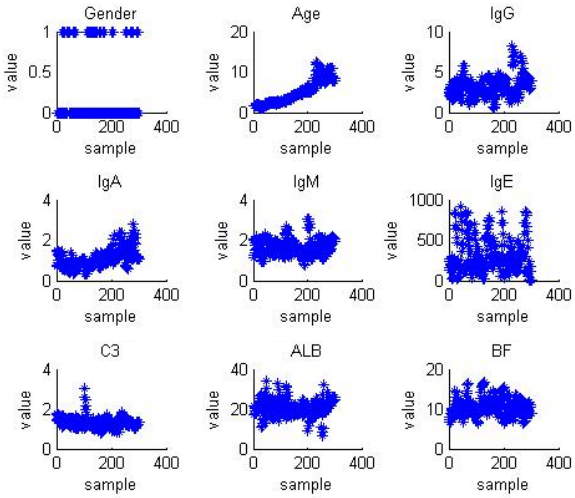

FIGURE 3. Multi-dimensional visualization of the virtual samples.

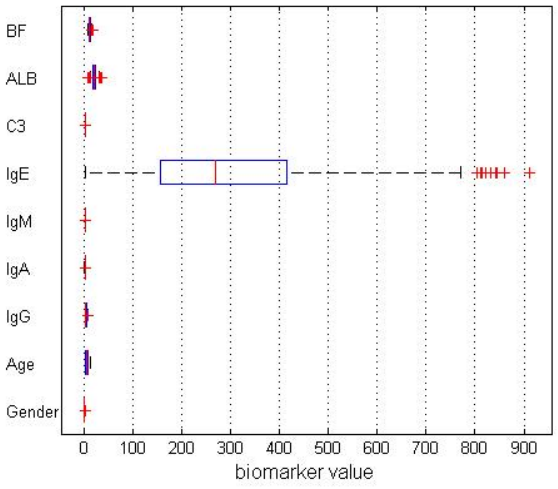

FIGURE 4. Bo plot of the virtual samples.

III. HYBRID APPROACH

A. PCA

PCA is a statistical method for data compression, dimension reduction and denoising based on the covariance matrix of variables. The basic principle is to find a set of orthogonal basis vectors and to convert noisy high-dimensional samples into low-dimensional ones [25]. By PCA, multiple interrelated variables can be transformed into a smaller number of uncorrelated variables, namely, principal components [26], [27]. This method has been widely adopted thanks to its excellent performance. For the children with nephritic syndrome, PCA can extract the principal components of physiological features in three steps:

Step 1: Calculate the covariance matrix of the training samples containing the physiological features of the children, compute the eigenvalues and their corresponding eigenvectors of the covariance matrix, and rank the eigenvalues in descending order.

Step 2: Calculate variance contribution rate and cumulative variance contribution rate.

Step 3: Determine the principal components of the physiological features according to discriminant criteria of cumulative variance contribution rate, difference of eigenvalues, etc.

B. BP-ADABOOST NEURAL NETWORK

BP NN is a multilayer feedforward neural network train by error BP algorithm. As an adaptive pattern recognition technique, BP NN features a simple structure, high maturity and wide popularity [28]. The performance of classical BP NN can be improved by simple ensemble realization, using ensemble methods like AdaBoost algorithm [29], [30].

AdaBoost algorithm combines multiple weak learners into a powerful regression machine. In general, a boosting ensemble outperforms a single neural network, and pushes up the accuracy [31]. AdaBoost NNs have delivered impressive performance in practical applications [32]–[34].

To overcome its defects in NCDs prediction (Section I), BP NN is replaced with BP-AdaBoost NN in this research.

TABLE 1. Total variance explained.

Component	Initial eigenvalues			Extraction sums of squared loadings		
	Total	Percent variance	Cumulative percentage	Total	Percent variance	Cumulative percentage
1	2.149	23.883	23.883	2.149	23.883	23.883
2	1.576	17.511	41.394	1.576	17.511	41.394
3	1.247	13.860	55.254	1.247	13.860	55.254
4	1.023	11.362	66.616	1.023	11.362	66.616
5	0.892	9.914	76.530	0.892	9.914	76.530
6	0.756	8.398	84.928	0.756	8.398	84.928
7	0.614	6.823	91.752			
8	0.450	4.996	96.747			
9	0.293	3.253	100.000			

The boosted predictor of BP-AdaBoost model was obtained by enhancing the weak predictor of BP models with AdaBoost algorithm.

Let m be the number of training samples, P be the number of weak predictors, $\boldsymbol{x}_\mathrm{i}$ be the input instance, and $\boldsymbol{y}_\mathrm{i}$ be the expected outputs of $\boldsymbol{x}_\mathrm{i}$. Then, the training set can be expressed as $S = \{(\boldsymbol{x}_\mathrm{i}, \boldsymbol{y}_\mathrm{i})\}$, $i = 1, 2, \ldots, m$. For the p^{th} weak predictor, $D_\mathrm{p}(i)$ denotes the weight of the i^{th} sample and $D_\mathrm{p}(i)$ is initialized to be $1/m$. In addition, let α denote the threshold of the prediction error. Then, the neural network structure can be determined by the dimensions of input and output. Next, BP-AdaBoost algorithm can be summarized as follows [35], [36].

Step 1: Initialize the weights and thresholds of the neural network, including the distribution weight $D_\mathrm{p}(i) = 1/m$ ($i = 1, \ldots, m$), α, P and $p = 1$.

Step 2: For $p = 1, 2, \ldots, P$, execute the following sub-steps:

Step 2.1: Train the p^{th} weak predictor, i.e. train the p^{th} BP NN by the training set S. Let $h_\mathrm{p}(\boldsymbol{x})$ be the function expressed by the newly-trained BP NN and $e_\mathrm{p}(i)$ be the absolute error of predictive output for the i^{th} sample. Calculate $e_\mathrm{p}(i)$ by the following equation:

$$e_\mathrm{p}(i) = \left|h_\mathrm{p}(x_\mathrm{i}) - y_\mathrm{i}\right| \quad i = 1, 2, \ldots, m \tag{2}$$

Step 2.2: Calculate ε_p by the following equation.

$$\varepsilon_\mathrm{p} = \sum_{i:e_\mathrm{p}(i)>0} D_\mathrm{p}(i) \quad i = 1, 2, \ldots, m \tag{3}$$

Step 2.3: Calculate the weight coefficient w_p of weak predictor according to the ε_p.

$$w_\mathrm{p} = 0.5 \times \ln\left[\frac{1-\varepsilon_\mathrm{p}}{\varepsilon_\mathrm{p}}\right] \tag{4}$$

Step 2.4: Adjust the weights of training samples in the next round of training according to the weight coefficient w_p:

$$D_{\mathrm{p}+1}(i) = \begin{cases} \dfrac{\exp\left(w_\mathrm{p}\right) \times D_\mathrm{p}(i)}{B_\mathrm{p}}, & e(i) > \alpha \\ \dfrac{D_\mathrm{p}(i)}{B_\mathrm{p}}, & e(i) \le \alpha \end{cases} \quad i = 1, 2, \ldots, m \tag{5}$$

where α is the threshold of prediction error; B_p is the normalization factor making the total distribution weight 1 under a constant weight ratio. The normalization factor can be expressed as:

$$B_\mathrm{p} = \frac{1}{\sum_{i=1}^{m} D_{\mathrm{p}+1}(i)} \tag{6}$$

Step 3: After P rounds of training, obtain the set of P weak predictive functions $\{o(h_\mathrm{p}(\boldsymbol{x}), w_\mathrm{p})\}$, $p = 1, 2, \ldots, P$. Then, derive the output function $O(\boldsymbol{x})$ of the boosted predictor obtained by combining weak predictive functions:

$$O(x) = \sum_{p=1}^{P} w_\mathrm{p} \times o\left(h_\mathrm{p}(x), w_\mathrm{p}\right) \tag{7}$$

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. EXPERIMENTAL RESULTS

PCA method was applied to extract the components of physiologic features of children with nephrotic syndrome. Total variance explained is presented in Table 1. It can be seen that six principle components were extracted, which explain 84.928% of the total variance. Tables 2 and 3 show the component matrix and the component score coefficient matrix, respectively.

BP-AdaBoost algorithm was run on MATLAB 2014a. The neural networks adopted a simple three-tier structure of 6-13-1, and were trained by the *traingd* function, which implements gradient descent learning. The maximum number of training rounds was set to 5,000 and the learning rate to 0.1. The number of weak predictors (BP models) P was set to 10. BP-AdaBoost model was the boosted predictor combined by 10 weak predictors of BP models. Besides BP-AdaBoost NN, three prediction models of BP NN, support vector machine (SVM) and radial basis function (RBF) NN were created for comparison.

As mentioned in Section II, the 330 samples were used to train the prediction models and the 60 samples were adopted to test the models. The method of 10-fold cross validation is performed to train the prediction models. The indices of Mean Squared Error (MSE) and Root Mean Square Error (RMSE) were used to estimate performance of the prediction models. The performance of cross validation for BP-AdaBoost NN, BP NN, SVM and RBF NN are tabulated in Table 4.

TABLE 2. Component matrix.

		Initial variables								
		Gender	Age	IgG	IgA	IgM	IgE	C3	ALB	BF
Components	1	-0.139	0.763	0.746	0.770	0.063	-0.023	0.110	0.394	-0.477
	2	0.244	0.386	-0.171	0.397	0.557	0.141	-0.632	-0.550	0.386
	3	0.713	-0.071	0.116	-0.234	0.553	-0.501	0.127	0.171	-0.250
	4	0.385	0.001	-0.074	-0.164	-0.122	0.666	-0.387	0.306	-0.374
	5	0.215	-0.083	0.272	0.019	0.108	0.475	0.570	-0.449	0.012
	6	0.389	0.162	0.303	-0.076	-0.426	-0.071	-0.048	0.113	0.528

TABLE 3. Component score coefficient matrix.

		Initial variables								
		Gender	Age	IgG	IgA	IgM	IgE	C3	ALB	BF
Components	1	-0.064	0.355	0.347	0.358	0.029	-0.011	0.051	0.183	-0.222
	2	0.155	0.245	-0.108	0.252	0.353	0.089	-0.401	-0.349	0.245
	3	0.572	-0.057	0.093	-0.188	0.444	-0.401	0.102	0.137	-0.201
	4	0.376	0.001	-0.072	-0.160	-0.119	0.652	-0.379	0.299	-0.366
	5	0.241	-0.093	0.305	0.022	0.121	0.533	0.639	-0.504	0.014
	6	0.515	0.215	0.401	-0.100	-0.564	-0.094	-0.063	0.149	0.698

TABLE 4. Performance of cross validation for BP-AdaBoost NN, BP NN, SVM, and RBF NN.

Performance indices	Predictors			
	BP-AdaBoost NN	BP NN	SVM	RBF NN
MSE	0.0012	0.0076	0.0022	0.0045
RMSE	0.0346	0.0872	0.0469	0.0671

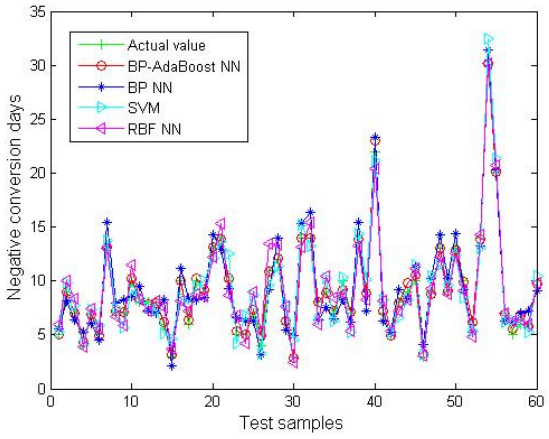

FIGURE 5. Prediction results of BP-AdaBoost NN, BP NN, SVM and RBF NN on PCA processed testing data.

Figure 5 displays the NCDs predicted by BP-AdaBoost NN, BP NN, SVM and RBF NN on the testing dataset which was processed by PCA, as well as the actual values of NCDs. Note that the output of BP NN is the mean of all outputs of 10 BP models. Figure 6 displays the NCDs predicted by BP-AdaBoost NN on PCA processed testing data and raw testing data respectively, as well as the actual values of NCDs.

B. RESULTS ANALYSIS

The absolute errors and relative errors of BP-AdaBoost NN, BP NN, SVM and RBF NN on PCA processed testing data

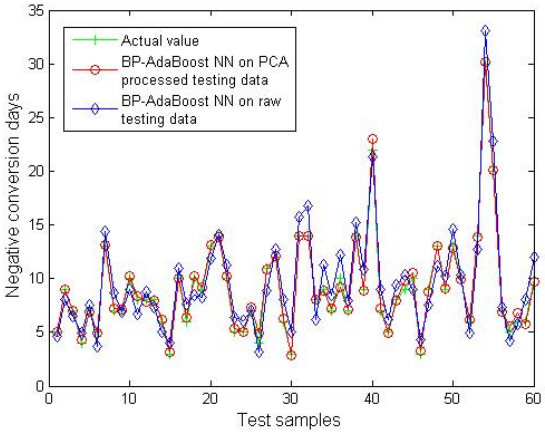

FIGURE 6. Prediction results of BP-AdaBoost NN on PCA processed testing data and raw testing data.

were calculated from the prediction results and the actual NDCs. The calculated absolute errors and relative errors are given in Figure 7 and Figure 8 respectively.

To demonstrate the effect of PCA, BP-AdaBoost NN was also trained and tested on raw data. Figure 9 displays the absolute errors of BP-AdaBoost NN on PCA processed testing data and raw testing data respectively. Figure 10 displays the relative errors.

The performance indices of accuracy and precision were calculated to estimate prediction models. Accuracy indices include mean/minimum/maximum absolute errors and relative errors. Precision indices include Standard Deviation (SD) and Relative Standard Deviation (RSD). Table 5 compares the prediction results of BP-AdaBoost NN, BP NN, SVM and RBF NN on PCA processed testing data. Table 6 compares the prediction results of BP-AdaBoost NN on PCA processed testing data and raw testing data.

As shown in Table 5, the absolute errors of BP-AdaBoost NN on PCA processed testing data averaged at 0.2334,

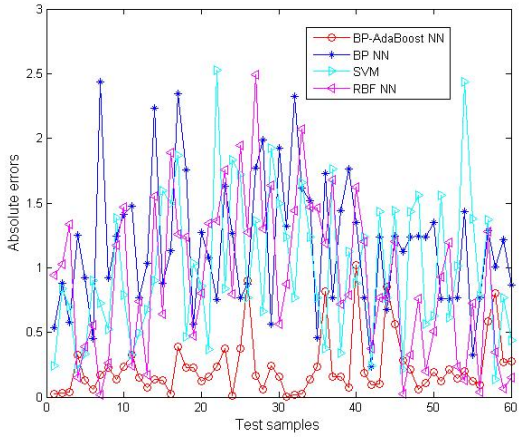

FIGURE 7. Absolute errors of BP-AdaBoost NN, BP NN, SVM and RBF NN on PCA processed testing data.

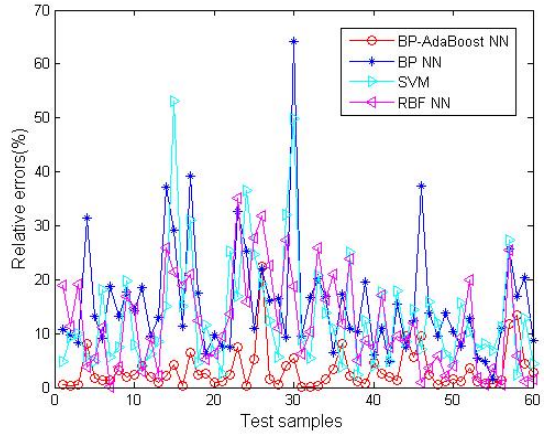

FIGURE 8. Relative errors of BP-AdaBoost NN, BP NN, SVM and RBF NN on PCA processed testing data.

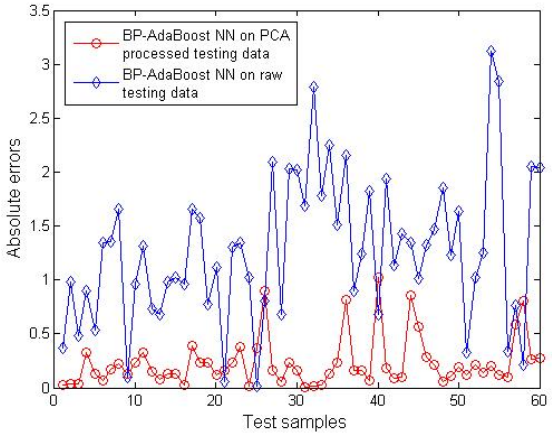

FIGURE 9. Absolute errors of BP-AdaBoost NN on PCA processed testing data and raw testing data.

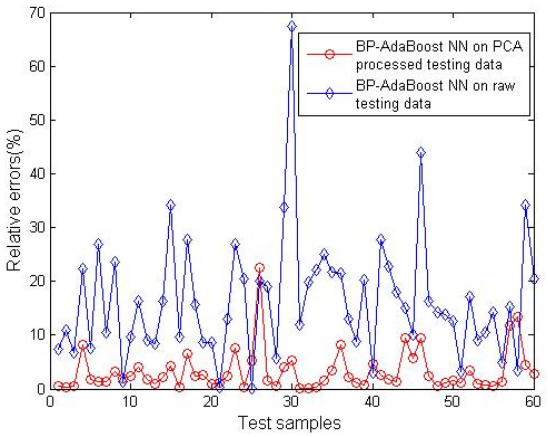

FIGURE 10. Relative errors of BP-AdaBoost NN on PCA processed testing data and raw testing data.

TABLE 5. Comparisons of BP-AdaBoost NN, BP NN, SVM and RBF NN on PCA processed testing data.

Performance indices		Predictors			
		BP-AdaBoost NN	BP NN	SVM	RBF NN
Absolute error	Mean	0.2334	1.1752	1.0003	0.9226
	Minimum	0.0008	0.2378	0.1383	0.0145
	Maximum	1.0218	2.4359	2.5234	2.4927
Relative error (%)	Mean	3.2789	15.5658	13.2794	12.0275
	Minimum	0.0057	1.6235	2.3050	0.1115
	Maximum	22.4350	64.1233	53.1133	35.0900
SD		4.6804	5.0020	4.8309	4.7436
RSD (%)		50.8053	53.5912	52.3232	50.9819

minimized at 0.0008 and peaked at 1.0218, the relative errors averaged at 3.2789%, minimized at 0.0057% and peaked at 22.4350%, the SD is 4.6804 and the RSD is 50.8053%, indicating that BP-AdaBoost NN outperforms other predictors and has a good performance in the prediction of NCDs of childhood nephrotic syndrome. Compared with BP NN, SVM and RBF NN, BP-AdaBoost NN boasts strong generalization ability and predicts the NCDs in an accurate and precise manner. This is because the weak predictors of BP models are improved by AdaBoost algorithm into the boosted predictor of BP-AdaBoost model. Specifically, the weights of training samples and those of weak predictors of BP models are adjusted according to prediction errors for training samples; the weight of the sample with larger prediction error is increased, and the weak predictor with good performance is assigned a relatively large weight in the final weighted combination. That is why the boosted model of BP-AdaBoost NN achieves better learning and prediction results than the BP predictor.

As shown in Table 6, BP-AdaBoost NN on PCA processed testing data outperforms BP-AdaBoost NN on raw testing data in all indices. The excellent performance of the hybrid approach of PCA and BP-AdaBoost NN can be attributed to the application of PCA. PCA method reduces the number of input variables of the neural network and simplifies the

TABLE 6. Comparisons of BP-AdaBoost NN on PCA processed testing data and raw testing data.

Performance indices		BP-AdaBoost NN	
		On PCA processed testing data	On raw testing data
Absolute error	Mean	0.2334	1.2662
	Minimum	0.0008	0.0116
	Maximum	1.0218	3.1258
Relative error (%)	Mean	3.2789	16.3667
	Minimum	0.0057	0.1657
	Maximum	22.4350	67.3767
SD		4.6804	5.0629
RSD (%)		50.8053	52.8515

network structure, leading to high prediction accuracy and precision.

Of course, the research results may not apply to all children with nephrotic syndrome, owing to two reasons. First, the sample size is not big although the data augmentation method is used. Second, the multiple physiologic features of patients are considered as a whole in the PCA-based dimension reduction, calling for differentiation between the effect of each physiologic feature.

V. CONCLUSION

The NCDs prediction is crucial for prognosis of childhood nephrotic syndrome. In view of this, PCA and BP-AdaBoost neural network were combined into a hybrid approach for automated prediction of the NCDs. Through experiments, the proposed BP-AdaBoost model was proved suitable for accurate prediction of the NCDs. Thus, this research provides a potential adjunct tool for pediatricians, enabling them to prognose childhood nephrotic syndrome accurately and provide patients with better care and treatment.

To further enhance the generalization ability of the prediction model, the future research will focus on the following two issues. First, more clinical cases will be collected to train the model, aiming to improve the reliability of NCDs prediction. Second, the physiological features closely related to NCDs will be determined through multi-factor analysis.

ACKNOWLEDGMENT

The authors are indebted to the supports and encouragements received from the staff and colleagues.

REFERENCES

[1] A. Al Talhi, K. Al Saran, E. T. Osman, A. Al Shatri, M. Osman, and K. Mirza, "A randomized study on a 3-month versus a 7-month prednisolone regimen for the initial episode of childhood idiopathic nephrotic syndrome at a large Saudi center," *Int. J. Pediatrics Adolescent Med.*, vol. 5, no. 1, pp. 18–23, 2018.

[2] L. El Bakkali, R. R. Pereira, D. J. Kuik, J. C. F. Ket, and J. A. E. van Wijk, "Nephrotic syndrome in The Netherlands: A population-based cohort study and a review of the literature," *Pediatric Nephrol.*, vol. 26, no. 8, pp. 1241–1246, 2011.

[3] J. D. Varner, A. Matory, and R. A. Gbadegesin, "Genetic basis of health disparity in childhood nephrotic syndrome," *Amer. J. Kidney Diseases*, vol. 72, no. 5, pp. S22–S25, 2018.

[4] D. G. Noone, K. Iijima, and R. Parekh, "Idiopathic nephrotic syndrome in children," *Lancet*, vol. 392, no. 10141, pp. 61–74, 2018.

[5] G. K. Aksoy, M. Koyun, E. Çomak, and S. Akman, "Early or late transplantation in congenital nephrotic syndrome: Which is effective for optimal growth?" *Transplantation Proc.*, vol. 51, no. 7, pp. 2283–2288, 2019.

[6] S. S. Riar, T. H. M. Banh, K. Borges, P. Subbarao, V. Patel, J. V. Ristovska, R. Chanchlani, N. H. Shamsy, D. Noone, D. Hebert, C. P. B. Licht, V. Langlois, R. J. Pearl, and R. S. Parekh, "Prevalence of asthma and allergies and risk of relapse in childhood nephrotic syndrome: Insight into nephrotic syndrome cohort," *J. Pediatrics*, vol. 208, pp. 251–257, May 2019.

[7] N. Starc, M. Li, M. Algeri, A. Conforti, L. Tomao, A. Pitisci, F. Emma, G. Montini, P. Messa, F. Locatelli, M. E. Bernardo, and M. Vivarelli, "Phenotypic and functional characterization of mesenchymal stromal cells isolated from pediatric patients with severe idiopathic nephrotic syndrome," *Cytotherapy*, vol. 20, no. 3, pp. 322–334, 2018.

[8] C.-S. Wang *et al.*, "Text messaging for disease monitoring in childhood nephrotic syndrome," *Kidney Int. Rep.*, vol. 4, no. 8, pp. 1066–1074, 2019.

[9] D. S. Gipson *et al.*, "Nephcure accelerating cures institute: A multidisciplinary consortium to improve care for nephrotic syndrome," *Kidney Int. Rep.*, vol. 3, no. 2, pp. 439–446, 2018.

[10] Y. Liu, B. Wang, J. Tang, M. Zhu, D. Chen, H. Jiang, and X. Chen, "Prediction of negative conversion days of childhood nephrotic syndrome based on the improved backpropagation neural network with momentum," *Int. J. Bioautomat.*, vol. 19, no. 4, pp. 543–554, 2015.

[11] Z. Liu, C. Cao, S. Ding, Z. Liu, T. Han, and S. Liu, "Towards clinical diagnosis: Automated stroke lesion segmentation on multi-spectral MR image using convolutional neural network," *IEEE Access*, vol. 6, pp. 57006–57016, 2018.

[12] P. Gómez, A. Schützenberger, M. Semmler, and M. Döllinger, "Laryngeal pressure estimation with a recurrent neural network," *IEEE J. Transl. Eng. Health Med.*, vol. 7, 2019, Art. no. 2000111.

[13] J. Kawahara and G. Hamarneh, "Fully convolutional neural networks to detect clinical dermoscopic features," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 2, pp. 578–585, Mar. 2019.

[14] S. H. Choi, H. Yoon, H. S. Kim, H. B. Kim, H. B. Kwon, S. M. Oh, Y. J. Lee, and K. S. Park, "Real-time apnea-hypopnea event detection during sleep by convolutional neural networks," *Comput. Biol. Med.*, vol. 100, pp. 123–131, Sep. 2018.

[15] D. Fan, J. Yang, J. Zhang, Z. Lv, H. Huang, J. Qi, and P. Yang, "Effectively measuring respiratory flow with portable pressure data using back propagation neural network," *IEEE J. Transl. Eng. Health Med.*, vol. 6, Jan. 2018, Art. no. 1600112.

[16] J. A. Alzubi, B. Bharathikannan, S. Tanwar, R. Manikandan, A. Khanna, and C. Thaventhiran, "Boosted neural network ensemble classification for lung cancer disease diagnosis," *Appl. Soft Comput.*, vol. 80, pp. 579–591, Jul. 2019.

[17] Y. Horie, T. Yoshio, K. Aoyama, S. Yoshimizu, Y. Horiuchi, A. Ishiyama, T. Hirasawa, T. Tsuchida, T. Ozawa, S. Ishihara, Y. Kumagai, M. Fujishiro, I. Maetani, J. Fujisaki, and T. Tada, "Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks," *Gastrointestinal Endoscopy*, vol. 89, no. 1, pp. 25–32, 2019.

[18] X. Li *et al.*, "Diagnosis of thyroid cancer using deep convolutional neural network models applied to sonographic images: A retrospective, multicohort, diagnostic study," *Lancet Oncol.*, vol. 20, no. 2, pp. 193–201, Feb. 2019.

[19] R. Zemouri, N. Omri, B. Morello, C. Devalland, L. Arnould, N. Zerhouni, and F. Fnaiech, "Constructive deep neural network for breast cancer diagnosis," *IFAC-PapersOnLine*, vol. 51, no. 27, pp. 98–103, 2018.

[20] H. Chougrad, H. Zouaki, and O. Alheyane, "Deep convolutional neural networks for breast cancer screening," *Comput. Methods Programs Biomed.*, vol. 157, pp. 19–30, Apr. 2018.

[21] M. A. Mohammed, B. Al-Khateeb, A. N. Rashid, D. A. Ibrahim, M. K. A. Ghani, and S. A. Mostafa, "Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images," *Comput. Elect. Eng.*, vol. 70, pp. 871–882, Aug. 2018.

[22] S. Walczak and V. Velanovich, "Improving prognosis and reducing decision regret for pancreatic cancer treatment using artificial neural networks," *Decis. Support Syst.*, vol. 106, pp. 110–118, Feb. 2018.

[23] A. Lopez-Rincon, A. Tonda, M. Elati, O. Schwander, B. Piwowarski, and P. Gallinari, "Evolutionary optimization of convolutional neural networks for cancer miRNA biomarkers classification," *Appl. Soft Comput.*, vol. 65, pp. 91–100, Apr. 2018.
[24] H. Zhang, M. Cisse, Y. N. Dauphin, and D. L. Paz, "mixup: Beyond empirical risk minimization," in *Proc. 6th Int. Conf. Learn. Represent. (ICLR)*, Vancouver, BC, Canada, Apr./May 2018, pp. 1–13.
[25] S. Duan, K. Chen, X. Yu, and M. Qian, "Automatic multicarrier waveform classification via PCA and convolutional neural networks," *IEEE Access*, vol. 6, pp. 51365–51373, 2018.
[26] X. Y. Shang, X. B. Li, A. Morales-Esteban, and G. H. Chen, "Improving microseismic event and quarry blast classification using artificial neural networks based on principal component analysis," *Soil Dyn. Earthquake Eng.*, vol. 99, pp. 142–149, Aug. 2017.
[27] W. N. L. Santos, M. C. S. Sauthier, A. M. P. Santos, D. A. Santana, R. S. A. Azevedo, and J. de Cruz Caldas, "Simultaneous determination of 13 phenolic bioactive compounds in guava (Psidium guajava L.) by HPLC-PAD with evaluation using PCA and neural network analysis (NNA)," *Microchem. J.*, vol. 133, pp. 583–592, Jul. 2017.
[28] J. Zhang, Y. Yang, and J. Zhang, "A MEC-BP-AdaBoost neural network-based color correction algorithm for color image acquisition equipments," *Optik*, vol. 127, no. 2, pp. 776–780, 2016.
[29] L. K. Hansen and P. Salamon, "Neural network ensembles," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 10, pp. 993–1001, Oct. 1990.
[30] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, Aug. 1997.
[31] D. W. Opitz and R. F. Maclin, "An empirical evaluation of bagging and boosting for artificial neural networks," in *Proc. Int. Conf. Neural Netw. (ICNN)*, Houston, TX, USA, Jun. 1997, pp. 1401–1405.
[32] L. X. Liu, Y. Hua, Q. J. Zhao, H. Huang, and A. C. Bovik, "Blind image quality assessment by relative gradient statistics and AdaBoosting neural network," *Signal Process., Image Commun.*, vol. 40, pp. 1–15, Jan. 2016.
[33] Y. S. Zhao, L. Gong, B. Zhou, Y. X. Huang, and C. L. Liu, "Detecting tomatoes in greenhouse scenes by combining AdaBoost classifier and colour analysis," *Biosyst. Eng.*, vol. 148, pp. 127–137, Aug. 2016.
[34] G. Saad, A. Khadour, and Q. Kanafani, "ANN and AdaBoost application for automatic detection of microcalcifications in breast cancer," *Egyptian J. Radiol. Nucl. Med.*, vol. 47, no. 4, pp. 1803–1814, 2016.
[35] L. Li, Y. Huang, J. Tao, C. Liu, and K. Li, "Featured temporal segmentation method and AdaBoost-BP detector for internal leakage evaluation of a hydraulic cylinder," *Measurement*, vol. 130, pp. 279–289, Dec. 2018.
[36] M. M. Baig, M. M. Awais, and E. S. M. El-Alfy, "AdaBoost-based artificial neural network learning," *Neurocomputing*, vol. 248, pp. 120–126, Jul. 2017.

YE LUO received the B.Sc. degree in computer education from Nanjing Normal University, in 1997, and the M.Sc. degree in computer application technology from Soochow University, in 2008. She is currently a Lecturer with the School of Computer Engineering, Jiangsu University of Technology, China. Her current research interests include data mining and intelligent information systems.

BEIHONG WANG was born in Changzhou, China, in 1978. She received the B.Sc. and M.Sc. degrees in pediatrics from Nanjing Medical University, in 2001 and 2011, respectively. She is currently an Associate Senior Doctor with the Department of Pediatrics, Changzhou No. 2 People's Hospital, China. Her research interests include childhood nephrotic syndrome, Henoch-Schönlein purpura in childhood, and biomedical signal processing.