Optimized Polynomial Neural Network Classifier Designed with the Aid of Space Search Simultaneous Tuning Strategy and Data Preprocessing Techniques

Wei Huang*** and Sung-Kwun Oh†

Abstract – There are generally three folds when developing neural network classifiers. They are as follows: 1) discriminant function; 2) lots of parameters in the design of classifier; and 3) high dimensional training data. Along with this viewpoint, we propose space search optimized polynomial neural network classifier (PNNC) with the aid of data preprocessing technique and simultaneous tuning strategy, which is a balance optimization strategy used in the design of PNNC when running space search optimization. Unlike the conventional probabilistic neural network classifier, the proposed neural network classifier adopts two type of polynomials for developing discriminant functions. The overall optimization of PNNC is realized with the aid of so-called structure optimization and parameter optimization with the use of simultaneous tuning strategy. Space search optimization algorithm is considered as a optimize vehicle to help the implement both structure and parameter optimization in the construction of PNNC. Furthermore, principal component analysis and linear discriminate analysis are selected as the data preprocessing techniques for PNNC. Experimental results show that the proposed neural network classifier obtains better performance in comparison with some other well-known classifiers in terms of accuracy classification rate.

Keywords: Polynomial Neural Network Classifier (PNNC), Principal component analysis, Linear discriminate analysis, Space search optimization, Simultaneous tuning strategy, Data preprocessing technique

1. Introduction

In the past decades, lots of neural network classifier have been developed in many research fields. Among these neural network classifiers, lot of architectures, learning abilities, and applications have been well documented, c.f. [4, 5], and [6-8]. Generally, there are three aspects in the design of neural network classifiers. They are as follows:

* Discriminate function;
* Lots of parameters in the design of classifier;
* High dimensional training data.

Linear discriminate function is generally used when constructing the neural network classifiers. However, such linear discriminate functions have a relatively simple geometry. For example, in the design radial basis function neural network classifier, such simple geometry is implied by the limited geometric variability of RBFs forming the hidden layer of the radial basis function neural network. Some enhancements have been shown by using two order polynomial discriminate functions [4], the problem of using high order discriminate function remains open.

In the design of neural network classifiers, there are generally several parameters needed to be estimated. A variety of evolutionary optimization algorithms such as genetic algorithm [9], particle swarm optimizer [10], differential evolution [11-12] have been successfully applied to adjust parameters when developing the neural network classifiers. For example, in the design of RBF neural network classifier, the parameters including the number of RBFs, the coordinates of RBF centers, the widths of RBFs are optimized by using evolutionary algorithms. In spite of advantages, these evolutionary algorithms cannot free from limitations. The first apparent limitation is that these evolutionary algorithms search solution space based on pure random mechanism. The second apparent limitation is that a problem of how to balance the structure optimization and parameter optimization, which are commonly encountered in the optimization of classifiers. Such problem plays an important role in the performance of neural network classifiers. Nevertheless, most of existing classifiers have not consider the above two limitations.

In this study, we propose a novel polynomial neural network classifier (PNNC) based on space search optimization algorithm using simultaneous tuning strategy to overcome all aforementioned limitations. The proposed
Optimized Polynomial Neural Network Classifier Designed with the Aid of Space Search Simultaneous Tuning Strategy and Data ~

classifiers employ higher order polynomials as discriminate functions. Space search optimization algorithm initialized by our previous study [13] is regarded as a vehicle to optimize the parameters when constructing the PNNC. Simultaneous tuning strategy is adopted in the process of optimization. Moreover, Principal component analysis (PCA) or linear discriminate analysis (LDA) is used to preprocess the original data set for solving the problem of high dimensional data. The rest of paper is organized as follows. Section II proposes the architecture of polynomial neural network classifiers. Section III presents the space search optimized polynomial neural network classifier with the use of simultaneous tuning. Section IV reports the experimental result of the proposed classifiers and Section V concludes this study.

2. Architecture of Polynomial Neural Network Classifier

Polynomial neural network classifier consists of four parts, namely preprocessing part, premise part, conclusion part, and aggregation part. The preprocessing part realized with the aid of principal component analysis or linear discriminate analysis aims at reducing dimension of input data. The premise part is formed by using fuzzy clustering method, while the conclusion part is determined via least square estimation method. A general topology of polynomial neural network classifier is described as shown in Fig. 1.

The overall polynomial neural network classifier can also be regarded as fuzzy models. In this case, the fuzzy rules can be outlined as follows:

\[ R_i^j : \text{IF } x \text{ is inclued in cluster } A_i, \text{THEN } y = \sum_{i=1}^{Q} w_i f_i(x_1, \ldots, x_q) \]  

(1)

Where \( R_i^j \) is the \( i \) th fuzzy rule ( \( i = 1, \ldots, n \) ), \( n \) represents the number of fuzzy rules, \( f_i(x_1, \ldots, x_q) \) is the consequent polynomial of the \( i \) th fuzzy rule, \( q \) is the dimension of the input after PCA or LDA transformation, \( w_i \) stands for membership of the \( i \) th cluster, and \( g_i(X) \) denotes the \( i \) th discriminant function.

2.1 Preprocessing part realized via PCA and LDA

The preprocessing part is realized based on principal component analysis and linear discriminate analysis. For convenience, the input data is denoted an \( N \times d \) matrix \( X \), and the output is expressed an \( N \times m \) matrix \( Y \). The steps of PCA are as follows:

[Step 1] Centralized the input data.
[Step 2] Calculate the \( d \times d \) covariance matrix \( C \), which can be expressed as follows:

\[ C = \frac{1}{N-1} X^T X \]  

(2)

Where \( N \) denotes the number of total data.
[Step 3] Calculate the eigenvectors of the covariance matrix.
[Step 4] Select \( m \) eigenvectors that correspond to the largest m eigenvalues to be the new basis.
[Step 5] Use the \( k \times d \) eigenvector matrix to transform the samples onto the new subspace.

2.2 Premise part realized via fuzzy C-Means

Fuzzy clustering method is exploited here to construct the premise part. The steps of FCM are as follows:

[Step 1] Compute the membership matrix \( U^{(r)} \)

\[ U^{(r)} = \{ u_{ik} | u_{ik} \in [0,1] \} \]  

(3)

Where two basic conditions should be met. They are as follows:

(1) For \( \forall k \), \( \sum_{i=1}^{c} u_{ik} = 1 \); and

(2) For \( \forall i \), \( 0 < \sum_{k=1}^{N} u_{ik} < N \).

[Step 2] Update the membership matrix \( U^{(r)} \) with the use of (5) and (6). Their expressions are as follows:
\[ u_{ik} = \frac{1}{\sum_{j=1}^{c} \left( \frac{x_k - s_{ij}^{(r)}}{x_k - s_{ij}^{(r)}} \right)^{2/(m-1)}} \]  \hspace{1cm} (4)

where

\[ s_{ij}^{(r)} = \left\{ s_{ij}^{(r)}, \ldots, s_{ic}^{(r)} \right\} \]  \hspace{1cm} (5)

\[ s_{ij}^{(r)} = \frac{\sum_{k=1}^{N} u_{ik} x_k}{\sum_{k=1}^{N} u_{ik}} \]  \hspace{1cm} (6)

[Step 3] Update the value of \( r \) using \( r = r + 1 \).

[Step 4] Repeat steps 2-4 and output the final results.

2.3 Conclusion part realized via linear and nonlinear polynomials

Polynomial discriminant functions are used to describe pattern classifiers. For any input \( X \), its output (class) is determined by the discriminant function \( g_i(X) \)

\[ g_i(X) > g_j(X), \text{ for all } j \neq i \]  \hspace{1cm} (7)

where

\[ g_i(X) = \sum_{l=1}^{c} u_{il} f_l(X) \]  \hspace{1cm} (8)

Here \( u_{il}, (l = 1, \ldots, c) \) are the values of membership matrix, \( f_l(X) \) are the outputs of polynomial functions. We use two types of functions that are as follows:

Type 1 (Linear): \( f(x) = a_0 + \sum_{i=1}^{n} a_i x_i \)  \hspace{1cm} (9)

Type 2 (Nonlinear): \( f(x) = a_0 + \prod_{i=1}^{n} g(x_i) \)  \hspace{1cm} (10)

Where \( g(x) = \exp\left( \frac{x_i - b_{ji}}{d_{ji}} \right)^2 \), \( b_{ji} \) denotes the centers (mean values), and \( d_{ji} \) represents the width (covariance values).

2.4 Aggregation part realized via least square error

Least square error approach is utilized to compute the parameter values of polynomials. The objective function can be expressed as follows:

\[ J_G = \sum_{k=1}^{m} \left( y_k - \sum_{i=1}^{n} w_{ik} f_i(x_k) \right)^2 \]  \hspace{1cm} (11)

Where \( w_{ik} \) are calculated based on membership matrix \( U^{(r)} \), and \( J_G \) can also be expressed as follows:

\[ J_G = (Y - Xa)^T (Y - Xa) \]  \hspace{1cm} (12)

Where \( a \) denotes the vector of polynomial coefficients, \( Y \) stands for the output vector, \( X \) is matrix. In case all consequent polynomials are linear polynomial, \( X \) and \( a \) can be expressed as follows

\[
\begin{bmatrix}
1 & w_0(a_0 - y_1) & \cdots & w_{n_0}(a_0 - y_1) \\
1 & w_1(a_1 - y_1) & \cdots & w_{n_1}(a_1 - y_1) \\
1 & w_2(a_2 - y_1) & \cdots & w_{n_2}(a_2 - y_1) \\
\vdots & \vdots & \ddots & \vdots \\
1 & w_{n_0}(a_{n_0} - y_n) & \cdots & w_{n_{n_0}}(a_{n_0} - y_n)
\end{bmatrix}
\]

\[ a = [a_0, a_1, \ldots, a_0, a_1, \ldots, a_m]^T \]

\[ Y = [y_1, y_2, \ldots, y_m] \]

The optimized coefficients \( a \) can be computed with the following formula:

\[ a = (X^T X)^{-1} X^T Y \]  \hspace{1cm} (13)

3. Space Search Optimization of Polynomial Neural Network Classifier

Unlike the conventional evolutionary algorithms, space search optimization algorithm is realized based on three basic space search operators: global search operator, local search operator, and opposition-based search operator [13]. The main steps of space search optimized PNNC are summarized as follows.

Step 1. Divide and preprocessing data set using PCA or LDA.

For simplicity, we divide a data set into three parts, namely training data, validation data, and testing data. The training and validation data are used for constructing classifiers, while the testing data is utilized for evaluating classifiers. Furthermore, principal component analysis and linear discriminant analysis are employed to preprocess the data sets.

Step 2. Initialize the structure parameters when constructing PNNCs.

Table 1 summarizes the basic parameters in the design of PNNCs. According to Table 1, one can obtain some...
different PNNCs. With the different selection of parameters.

**Step 3.** Optimize the PNNC using space search optimization algorithm with simultaneous tuning approach.

Fig. 2 depicts the solution compositions optimized PNNC and its interpretation using space search optimization algorithm with simultaneous tuning approach. As shown in Fig. 2(a), the first two parameters are used in the preprocessing part, the middle two parameters come from premise part, and the last parameter is utilized in the consequence part.

**Step 4.** Output the optimized PNNC. The PNNC with the best predictive capability will be selected as the optimized PNNC. Here an objective function based on training data and validation data is adopted here as the performance index. For convenience, TR denotes the performance index of training data, and TE stands for the performance index of testing data.

The false classification rate is used as performance index, whose expression is as follows:

\[
TR \text{ or } TE = \frac{(1 - F/N) \times 100\%}{1}
\]

Where \(F\) is the number of patterns in case of false classification, and \(N\) is the total number of patterns in the data set. TR is considered as the objective function.

**4. Experimental Studies**

This section reports the experiments of the proposed neural network classifier based on some well-known machine learning data [16-19]. Each data is divided into two parts, 80% is used for training, and the rest 20% is regarded as testing. Experiments in all cases were repeated 10 times leading to a random sub-sampling validation. The parameters of SSOA are summarized as follows: the number of generations is set as 150, and the population size is set as 30. The operator parameters of local space search and global space search are the same as the reference [13].

**4.1 WDBC data**

We first do experiments based on Wisconsin Diagnostic Breast Cancer data set, which is a medical classification
Table 2. Experimental results of PNNC on Benchmark data (WDBC)

<table>
<thead>
<tr>
<th>Classifier</th>
<th>TR</th>
<th>TE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP[31]</td>
<td>-</td>
<td>85.9</td>
</tr>
<tr>
<td>RBF2[32]</td>
<td>-</td>
<td>97.1</td>
</tr>
<tr>
<td>FFC</td>
<td>-</td>
<td>96.1</td>
</tr>
<tr>
<td>SVM[34]</td>
<td>-</td>
<td>96.2</td>
</tr>
<tr>
<td>RVM</td>
<td>-</td>
<td>97.2</td>
</tr>
<tr>
<td>Our classifier (with PCA+LDA)</td>
<td>99.1 ± 0.001</td>
<td>98.15 ± 0.016</td>
</tr>
<tr>
<td>Our classifier (without PCA+LDA)</td>
<td>94.7 ± 0.003</td>
<td>94.86 ± 0.02</td>
</tr>
</tbody>
</table>

Table 3. Experimental results of PNNC on Benchmark data (pima)

<table>
<thead>
<tr>
<th>Classifier</th>
<th>TR</th>
<th>TE</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNN[12]</td>
<td>-</td>
<td>70.3</td>
</tr>
<tr>
<td>TS-KNN[12]</td>
<td>-</td>
<td>77.7</td>
</tr>
<tr>
<td>RBF2[13]</td>
<td>-</td>
<td>77.0</td>
</tr>
<tr>
<td>SVM[14]</td>
<td>-</td>
<td>74.7</td>
</tr>
<tr>
<td>Our classifier (with PCA+LDA)</td>
<td>86.95±0.013</td>
<td>79.03±0.015</td>
</tr>
<tr>
<td>Our classifier (without PCA+LDA)</td>
<td>78.14±0.015</td>
<td>76.18±0.02</td>
</tr>
</tbody>
</table>

Fig. 3. Optimization process of SSOA-based neural network classifier on benchmark dataset (WDBC)

Fig. 4. Optimization process of SSOA-based Neural network classifier on benchmark dataset (PIMA)

4.2 PIMA data

The second well-known machine learning data is PIMA data set, which consists of 569 input-output pairs, and 32 input variables.

Fig 3 shows the optimization process of space search optimization algorithm based neural network classifier on WDBC data. As can be seen from Fig. 3, the proposed PNNCs optimized very fast before 20 generations and then obtains the stable model after 20 generations.

Table 2 further summarizes the experimental results between the proposed neural network classifier and previous models reported in the literatures.

4.3 Some other data

The proposed neural network classifier is further compared with some selected models on several sizes of well-known data sets, which has different input variables and patterns. The comparison results between the proposed classifier and some other models are summarized as shown in Table 4, which illustrates the effectiveness of the proposed neural network classifier.

5. Conclusions

This work is the extended paper of conference paper (Joint 8th international conference on soft computing and intelligent systems and 17th international symposium on advance intelligent systems). It contributes to the research on classification can be summarized as follows:

First, we have proposed a novel neural network classifier based on concept of polynomial discriminate functions and space search optimization algorithm.

Second, we have applied the PCA and LDA as the preprocessing techniques for the proposed neural network classifier.

For future study, the proposed neural network classifier may be improved by using new optimization algorithms. Furthermore, novel neural network classifier can be
constructed by combing the proposed neural network classifier and other classical classifiers.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 61301140, 61673295, 61562024, and 61272450), supported by the Open Foundation of State Key Laboratory of Digital Manufacturing Equipment and Technology (Grant No. DMETKF2015012), and also supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning [grant number NRF-2015R1A2A1A15055365] and by the GRRC program of Gyeonggi province [grant number GRRC Suwon 2016-B2, Centre for U-city Security & Surveillance Technology].

References


Wei Huang
He received the MS degree at school of Information Engineering from East China Institute of Technology, China, in 2006, and Ph.D. degree at State Key Laboratory of Software Engineering, Wuhan University, China, in 2011. He is currently an Associate Professor with the School of Computer and Communication Engineering, Tianjin University of Technology, Tianjin, China. His research interests include evolutionary computation, operations research, fuzzy system, fuzzy-neural networks, and software reliability.

Sung-Kwun Oh
He received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from Yonsei University, Seoul, Korea, in 1981, 1983, and 1993, respectively. During 1983-1989, he was a Senior Researcher of R&D Lab of Lucky-Goldstar Industrial Systems Co., Ltd. From 1996 to 1997, he was a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB, Canada. He is currently a Professor with the Department of Electrical Engineering, The University of Suwon, Suwon, South Korea. His research interests include fuzzy system, fuzzy-neural networks, automation systems, advanced computational intelligence, and intelligent control. He currently serves as an Associate Editor of the KIEE Transactions on Systems and Control, International Journal of Fuzzy Logic and Intelligent Systems of the KFIS, and Information Sciences.