A Novel Feature Extraction-based Selective & Nonlinear Neural Network Ensemble Model for Economic Forecasting

Zhu Bangzhu[†, ††] and Lin Jian[††]

[†] School of Economics and Management, Beijing University of Aeronautics and Astronautics, Beijing, China
[††]Institute of System Science and Technology, Wuyi University, Jiangmen, China

Summary
In this study, a novel selective & nonlinear neural network ensemble model, i.e. NSNNEIPCABag, is proposed for economic forecasting. In this model, some different training subsets are first generated by bagging algorithm. Then the feature extraction technique, improved principal component analysis (IPCA), and then the IPCA approach is also used to extract their data features to train individual networks, and to select the appropriate number of ensemble members from the available networks. Finally, the selected members are aggregated into a nonlinear ensemble model with support vector regression (SVR). For illustration and testing purposes, the proposed ensemble model is applied for economic forecasting.

Key words:
Ensemble learning, Neural network ensemble, Nonlinear ensemble, Selective ensemble, Feature extraction.

1. Introduction

Economic forecasting has been a common research problem in the past few years. Much attention has been paid to it by many scholars. Various approaches have been adopted to solve the problem. However, it is not easy to predict economic values due to its inherent high complexity and noise, which has encouraged us to developed more predictable models. As a result, artificial neural network (ANN), especially back-propagation neural network (BPNN), is regarded as a good method. ANN has a good learning ability and generalization with its simple structure. Although a large number of successful applications have showed ANN can be a very useful tool for economic forecasting, some studies, however, showed that ANN also had some limitations such as the difficulty in determining the number of neural cells of the hidden layer, its initialization and easily drop local minima and so on, which could influence its generalization and make the forecasting results not very good [1].

In order to overcome the main limitations of ANN, recently a novel ensemble forecasting model, i.e. neural network ensemble (NNE) [2], has been developed. Because of combining multiple neural networks learning from the same training samples, NNE can remarkably enhance the forecasting ability and outperform any individual neural network. However, in the previous work on ensemble forecasting models, three main problems can be found. First, most of neural network ensembles took use of all the input attributes. However, in essential, economic forecasting is a very complicated system with many influential factors, and every factor isn't equaled important for forecasting [3], i.e. not all the factors are useful for the economic forecasting (Problem Ⅰ). In view of feature extraction, forecasting is mostly determined by several key factors. Therefore, the key attributes for forecasting can be picked out by feature extraction as the input ones of individual neural networks. Second, almost every research combined all the available individual neural networks to constitute an ensemble model. However, not all the circumstances are satisfied with the rule of "the more, the best" [4] [5] (Problem Ⅱ). Thus, how to determine the number of individual forecasting models is not easy, which urges us to develop an appropriate method to resolve the problem. Third, in most existing literatures, the ensemble models were limited to linear combination form. However, linear combination is not necessarily appropriate for all the circumstances (Problem Ⅲ).

Considering the above-mentioned three main problems, this paper proposes a novel feature extraction-based selective & nonlinear neural network ensemble model for economic forecasting. This ensemble approach utilizes the improved principal component analysis (IPCA) [1] for feature extraction, i.e. discovering and remaining the main information with fewer principal components on the results of the sampling process. Then, the principal components are used as inputs to train individual neural networks. Finally, some selected neural networks are used to constitute a nonlinear ensemble model with support vector regression (SVR). Thus, the mechanism of disturbing the training data and the input principal components are combined to generate accurate and diverse individual neural networks, which can help to constitute a novel selective & nonlinear neural networks ensemble forecasting model.

The main aim of this study is to show how to carry out economic forecasting using the proposed novel feature extraction-based selective & nonlinear neural network ensemble model. Therefore, this paper mainly describes the building process of the proposed ensemble model and

Manuscript received February 5, 2007.
Manuscript revised February 25, 2007.

the application of the ensemble approach in forecasting the Gross Domestic Product (GDP) of Jiangmen, Guangdong, China.

The rest of the paper is organized as follows. The next section describes the building process of the proposed ensemble model in detail. In order to verify the effectiveness of the proposed model, empirical analysis of the GDP of Jiangmen, Guangdong, China is reported in section 3. Finally, a few conclusions are contained in section 4.

2. The Building Processes of the Proposed Selective & Nonlinear Ensemble Model

In this section, a three-phase feature extraction-based selective & nonlinear neural network ensemble model, i.e. NSNNEIPCABag, is described step by step and in detail. First multiple individual BPNNs are generated. Then an appropriate number of BPNNs are selected from the candidate predictors generated by the former phase. Finally, the selected BPNNs are aggregated into an ensemble predictor in a nonlinear way.

2.1 Generating individual BPNN predictors

According to the bias-variance trade-off principle [6], in order to build a strong ensemble model, the individual models should be with high accuracy as well as high diversity. Same or similar individual models may be of no help for enhancing the generalization. Therefore, how to generate the diverse models is a crucial factor. As for neural network ensemble model, some methods have been put brought for generation of ensemble members making different errors, most of which depend on varying the parameters of neural networks.

In this study, we adopt bagging algorithm to generate different training subsets, and the IPCA approach [1] to extract their data features used to train the individual networks, which can resolve the above-mentioned first problem (Problem Ⅰ).

2.2 Selecting an appropriate number of ensemble members

After training, each individual network has generated its own forecasting result. According to the selective ensemble learning theory [5], assembling many of the available neural networks may be better than assembling all of those networks. If there are a great number of available individual neural networks, it is very necessary to select a subset of representatives in order to improve ensemble efficiency and have a good performance. In order to show that those “many” neural networks can be efficiently selected from all of the available neural networks. In view of the previous work [4], in this paper the IPCA approach [1] is presented to select appropriate ensemble members. This approach selects some neural networks without linear correlation to constitute a selective ensemble model, which may resolve the above-mentioned second problem (Problem Ⅱ).

Supposed that there are n neural predictors with p forecasting results. Then the forecasting matrix (Y) can be described as

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & y_{22} & \cdots & y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{np} \end{bmatrix}$$

where y_{ij} is the jth forecasting value with ith neural predictor.

Next, we deal with the forecasting matrix using the IPCA method. First, eigenvalues ($\lambda_1, \lambda_2, \cdots, \lambda_n$) and their corresponding eigenvectors A= ($a_1, a_2, \cdots, a_n$) can be solved from the forecasting matrix. Then the new principal components are calculated as

$$F_i = a_i^T Y \quad (i = 1, 2, \cdots, n) \ .$$

Subsequently, we choose $m (m \le n)$ principal components from the existing n components. Therefore, the remained information continent is judged by

$$\theta = \sum_{i=1}^{m} \lambda_i / \sum_{i=1}^{n} \lambda_i$$

If θ is a larger enough (e.g. $\theta \ge 0.90$), we can believe that sufficient information has been saved after feature extraction. Therefore, re-combining the new information can further enhance the forecasting ability of the proposed ensemble model.

2.3 Combining the selected members

Based on the previous two-phase work, some appropriate ensemble members can be picked out. The next task is to combine these members into an aggregated predictor in an appropriate ensemble strategy. Generally, there are two ensemble strategies: linear ensemble and nonlinear ensemble. Linear ensemble includes two approaches: simple average or weighted average [1] [7] [8], which have been widely applied in the existing literatures. Different from the previous work, in this study we adopt the SVR method [9] to combine the selected members into an nonlinear ensemble model, which may resolve the above-mentioned third problem (Problem Ⅲ).

By solving the earlier three problems, an IPCA-based selective & nonlinear neural network ensemble model has been built for economic forecasting.

To summarize, the proposed selective & nonlinear neural network ensemble model can be described in detail as follows. At the beginning, for a given training set S, a series of training subsets S_1, S_2, …, S_n can be determined by bagging method. Then, the key attributes F_1, F_2, …, F_m, extracted by the IPCA method on the respective training subsets, should be taken as the inputs of neural networks to create n individual neural networks. Next, the IPCA approach is also applied to select k ensemble members. Finally, SVR is used to combine the selected individual neural networks into a nonlinear ensemble model. The basic flow diagram can be shown in Fig.1.

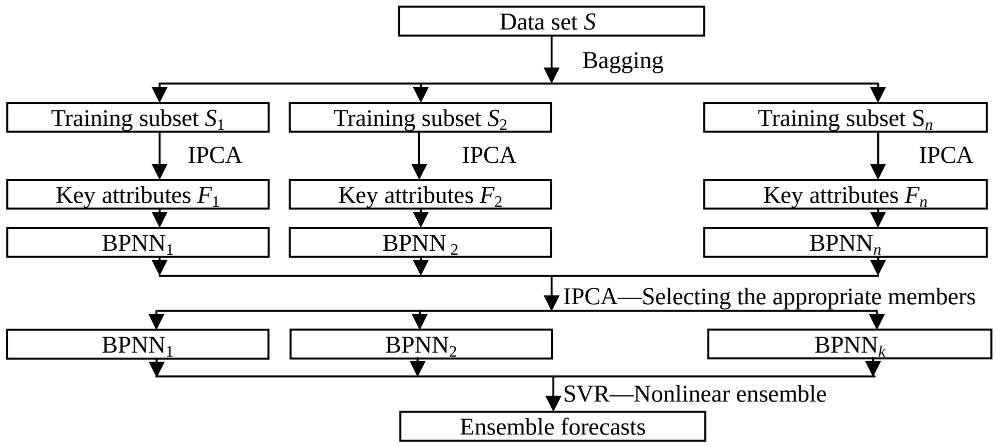

Fig.1 A flow diagram of the proposed ensemble forecasting model

3. Empirical analysis

The proposed selective & nonlinear ensemble model is applied to forecast the increasing rates of Gross Domestic Product (GDP) of Jiangmen, China to illustrate its validity and feasibility. In this study, multiple BPNNs are used to build the NSNNEIPCABag.

3.1 Data description

The economic forecasting data used in this paper are yearly and are obtained from Jiangmen Statistical Yearbooks from 1982 to 2004. This paper selects 8 forecasting indexes as follows [10]: x_1: GDP of Jiangmen; x_2: GDP of China; x_3: GDP of Guangdong; x_4: Government expenditures; x_5: Total exports; x_6: Total investment in fixed assets; x_7: Total sales of consumer goods; x_8: Foreign capital actually utilized.

While carrying on the importation of the input variables, this paper uses the increasing rate of the index in order to get rid of the long-term growth trend influence, price fluctuation influence and the difference of index scale. That means:

$$x_i(t)' = \frac{x_i(t)/w(t)}{x_i(t-1)/w(t-1)} - 1, i = 1,2,\cdots,8; t = 1982,1983,\cdots,2004$$

where $x_i(t)'$: the increasing rate of indicator i in t year; $x_i(t)$: the real value of indicator i in t year; $w(t)$: the price index in t year; $w(t\text{-}1)$: the price index in t -1 year. In addition, for convenience, let $x_i(t)'$ still be denoted as $x_i(t)$. In order to save space, the original data are not listed here, and detailed data can be obtained from the yearbooks or from the authors.

We take the data from1983 to 1999 corresponding to the increasing rates of GDP from 1984-2000 as the training set, that is, the number of the training samples is 17, i.e. $\|S\| = 17$. We also take the data from 2000 to 2003 corresponding to the increasing rates of GDP from 2001 to 2004 as the test set. In addition, we use the data in 2004 to forecast the increasing rate of GDP in 2005. On the previous work, five training subsets, namely S_1, S_2, ---, S_5 are bootstrap sampling from all the available data, and $\|S_1\| = 17$, $\|S_2\| = 16$, $\|S_3\| = 15$, $\|S_4\| = 14$ and $\|S_5\| = 13$.

The IPCA method has been applied for dimension reduction and feature extraction of each training subsets. Three principal factors denoted as F_1, F_2 and F_3 of the above five training subsets can be respectively extracted. The accumulative total variances respectively achieve 92.34%, 91.65%, 92.01%, 92.31% and 92.07%, so to a certain extent, these three factors can be used to describe the main development information of Jiangmen economic in past years. F_i (i=1, 2, 3) can be extracted, which can be taken as the 3 input neural cells of individual neural networks. Likewise, we also adopt the IPCA method to extract the date features from 2000 to 2004.

3.2 Empirical results

Matlab 7.01 neural network toolbox has been used to generate and train the individual neural networks. In this thesis, we use tansig function as the hidden layer’s threshold function, purelin as the output layer’s threshold function and trainlm as the training function. Five neural networks, i.e. BPNN$_i$ (i=1, 2,…, 5) are not trained until they all satisfy the request of error. F_i trains BPNN$_i$. These five neural networks are all 3-5-1 BPNNs with training epochs of 5000, learning rate of 0.001, target error of 0.000001 and initial weights of the random values among [-1,1].When all the training results satisfy the request of error, all the component neural networks have been trained well.

After training, the IPCA approach is adopted to select some neural networks without linear correlation from the available neural networks to constitute an ensemble, and SVR is used to combine their forecasting results.

For the purpose of comparison, we have also built four other ensemble forecasting models: (1) directly using the subsets to train the individual neural networks, and linearly combining all the available neural networks into an ensemble model only with bagging, i.e. NNEBag; (2) linearly combining all the available neural networks into an ensemble model with feature extraction by IPCA and bagging, i.e. NNEIPCABag; (3) selecting some available neural networks to linearly combine an ensemble model only with bagging, i.e. SNNEBag; and (4) selecting some available neural networks to linearly combine an ensemble model with feature extraction by IPCA and bagging, i.e. SNNEIPCABag. In addition, the root mean squared error (RMSE) is used as the evaluation criteria over each of the two different ensemble models and corresponding results are reported in Table 1.

As are shown in the Table 1, we can conclude that (1) all the forecasting results are close to their real values, even though the number of the training samples is only 17; (2) the forecasting ability of NNE is generally much higher than that of each individual neural network; (3) the forecasting ability of NSNNEIPCABag is superior to that of the other NNEs on the whole; (4) Jiangmen increasing rate of GDP in 2005 is 11% or so and (6) NSNNEIPCABag is more feasible and effective for economic forecasting.

Table 1: A comparison of results of ensemble forecasting models

Ensemble models	RMSE	Rank	2005
NNEBag	0.0087	5	0.11180
SNNEBag	0.0071	3	0.10947
NNEIPCABag	0.0081	4	0.11235
SNNEIPCABag	0.0053	2	0.11249
NSNNEIPCABag	0.0047	1	0.11238

4. Conclusion

All the above is an attempt to presenting a novel feature extraction-based selective & nonlinear neural network ensemble forecasting model, and applying it for economic forecasting. The experiment results reported in this paper demonstrate the effectiveness of the proposed ensemble model, implying the proposed feature extraction-based selective & nonlinear ensemble forecasting model can provide a new way for economic forecasting under small samples.

Acknowledgment

This work is partially supported by the National Natural Science Foundation of China under Grant 70471074 and Guangdong Provincial Department of Science and Technology under Grant 2004B36001051 to Lin Jian.

References

[1] J. Lin, B.Z. Zhu, “Improved principal component analysis and neural network ensemble based economic forecasting,” Lecture Notes in Computer Science 4113 (2006) 135–145.

[2] L.K., Hansen, P. Salamon, “Neural network ensembles,” IEEE Trans Pattern Analysis and Machine Intelligence, vol.12, no.10, pp.993-1001, 1990.

[3] J.J., Ling, Z.Q., Chen, Z.H. Zhou, “Feature selection based neural network ensemble method,” Journal of Fudan University (Natural Science), vol.43, no.5, pp.685-688, 2004.

[4] L.A., Yu, S.Y., Wang, K.K., Lai, “A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates,” Computer &Operation Research, no.32, pp. 2523-2541, 2005.

[5] Z.H., Zhou, J.X., Wu, W. Tang, “Ensemble neural networks: many could be batter than all,” Artificial Intelligence, vol.137, no.1-2, pp. 239-263, 2002.

[6]L.A., Yu, S.Y., Wang, K.K., Lai, W., Huang, “A bias-variance-complexity trade-off framework for complex system modeling,” Lecture Notes in Computer Science 3980(2006) 518-527.

[7] J. A. Benediktsson, R. J. Sveinsson, O.K. Ersoy, P.H. Swain, “Parallel consensual neural networks,” IEEE Transactions on Neural Networks, vol. 8, pp. 54-64, 1997.

[8] M.P. Perrone, L.N. Cooper, “When networks disagree: ensemble methods for hybrid neural networks,” In Mammon, R. J. (ed.): Neural networks for speech and image processing, Chapman-Hall (1993)126-142.

[9] V.Vapnik, The nature of statistics learning theory, Spring-Verlag, New York, 1995.

[10] J.H. Xiao, “Intelligent forecasting for the regional economic,” Mathematics in Economics, vol.22, no.1, pp. 57-63, 2005.

Zhu Bangzhu received the B.S. and M. S. degrees in Industrial Engineering from Zhengzhou Institute of Aeronautical Industry Management and Guilin University of Electronic Technology in 1999 and 2004, respectively. He is a Ph.D Candidate of Beijing University of Aeronautics and Astronautics. His research interests include complex system modeling and simulation, intelligent information process and industrial engineering.

Lin Jian received the B.S., M. S., and Dr. degrees in Management Science and Engineering from Fuzhou University, Beijing University of Aeronautics and Astronautics and Lancaster University in 1982, 1987, and 1993, respectively. He has been a professor at Wuyi University since 1998. His research interests include complex system modeling and simulation, management information system and industrial engineering.