AN INVESTIGATION OF FINANCIAL TIME SERIES PREDICTION USING BACK PROPAGATION NEURAL NETWORKS

K. Jayanthi, Dr. P. Suresh

1 Department of Computer Science, Thiruvalluvar Govt. Arts College, Rasipuram, Tamilnadu, India
2 Department of Computer Science, Salem Sowdeswari College, Salem, Tamilnadu, India

1jaysureshlaya@gmail.com, 2sur_bh0071@rediffmail.com

ABSTRACT:

The stock market forecasting is a significant challenging research area in the application of time series analysis since the stock market data are non-stationary and naturally noisy. There are two stock market indices used in this study such as CNX Nifty and S&P BSE Sensex. In this paper, proposed Discriminant Independent Component Analysis for eliminating noise from the stock market data in order to enhance the performance of the Back Propagation Neural Networks. The performance of the proposed model is formed superior generalization performance compared with the ICA-BPNN and standard Back Propagation Neural Networks.

Keywords: Artificial Neural Network, Back Propagation Neural Network, Discriminant Independent Component Analysis, Data Smoothing

[1] INTRODUCTIONS

The stock market forecasting is a major task and important topic in financial time series analysis, which is a wide-ranging concentration in research society. Conversely, the financial market data are inherently noisy and non-stationary[1]. The noise personality of stock market data which is referred to the unavailability of wide-ranging in a row data from past behavior of financial markets to entirely capture the dependency between future prices and past prices[2]. As
well, the nonlinear and non-stationary personality of the financial market data that makes it complicated to predict stock price and suffers from the hypothesis of linear variation of the stock prices at some stage.

Artificial Neural Networks is an alternating solution to capture the non-linearity among data in a better way than the primary statistical prediction methods. ANNs and SVM are two most used machine learning algorithms for predicting for stock future behaviour [1, 3-6]. Artificial Neural Networks is a broadly used approach to investigate the stochastic nonlinear system, show good performance. In Artificial Neural Networks, the most popular neural network supervised feed forward training algorithms, namely Back Propagation Neural Networks for financial time series analysis that is a simple architecture with powerful problem solving capability[1, 7]. This paper has concentrates the back propagation neural networks for predicting stock indices. Thus, the stock market data is inherently noisy and non-stationary since that may suffer the approximation function. This will be disappointment the generalization capacity and could direct to over-fitting/under fitting.

Real time has not easy task since the noise of the natural data which is used as well as in native format will suffer the computation cost of the machine learning and accuracy of the given problems. Consequently, need to pay attention for takeout the noise from the raw data in order to improve the generalization performance of the machine learning algorithms. However, the de-noise scheme is very significant task for improving generalization performance of the BPNN forecasting model[8]. For eliminating the noise from stock market data, there are many approaches were used such as Principle Component Analysis[9], Independent Component Analysis [5], Non-Linear ICA[3].

Although PCA, ICA and NLICA methods are finds the independent or orthogonalized hidden noise from stock market data these noise may not be best possible for predictions. To solve abovementioned weakness of the presented algorithms, Discriminant Independent Component Analysis is used to removing noise from the stock market data. Wherein, fisher linear discrimination and Negentropy are used to optimize and the latent independent factors of the noise are uninvolved from multivariate noise data. The Discriminant ICA method has provides higher prediction accuracy and little computational cost. The Discriminant function is combining with ICA method is called DICA for removing more noise. In this paper, the proposed Discriminant ICA[10] is integrated with BPNN for make a prediction model which is mentioned as a DICA-BPNN prediction model.

[2] DISCRIMINANT INDEPENDENT COMPONENT ANALYSIS

Discriminant ICA is a semi-supervised approach to discover a linear transformation to a low-dimensional feature space where the features are statically independent with maximally discriminant[10]. In Discriminant ICA, the fisher criterion and the sum of the marginal negentropy of the pull out independent features are maximized concurrently. Therefore, Discriminant ICA combines the properties of ICA with discriminant function in order to enhance the accuracy of the financial forecasting[11].
a) Negentropy Maximizations

Negentropy is superior statistical scaling technique for non-Gaussian variable and approximations of marginal negentropy can be obtained as follows,

\[ J(y_i) \approx k_1(E(G^1(y_i))^2 + k_2(E(G^1(y_i)) - E(G^1(v))^2) \]  

(1)

Where, \( v \) gives a univariate Gaussian distribution with the same mean and standard deviation as \( y_i \). The value of \( k_1 \) and \( k_2 \) are \( k_1 = 36/(8\sqrt{3-9}) \), \( k_2 = 24/(16\sqrt{3-27}) \). \( G^1 \) and \( G^2 \) are the non-quadratic functions. Some choices of \( G^1 \) and \( G^2 \) for random vector \( y_i \) which have been proved to be used are given as follows,

\[ G^2(y_i) = \frac{1}{\sqrt{k}} \]  

(2)

\[ G^2(y_i) = -\log\cosh a_i y_i, \quad 0 < a_i \leq 1 \]  

(3)

\[ G^2(y_i) = \frac{a_i \exp(y_i^2/2)}{k} \]  

(4)

Maximization of the sum of the marginal negentropy with element of covariance can be achieved through a Lagrange formula are given as follows,

\[ L(W) = \sum [E(G(w_i^T z)) - E(G(v))]^2 + \sum \beta_i (w_i^T w_i - 1) \]  

(5)

Features are obtained during maximizing the target function in equation (5). Optimization problem that maximizes the functional principle of prediction performance and negentropy of the take out the independent features can be defined concurrently through a Lagrange formula in the following form,

\[ L = L(W) + k\phi(W, Z, C) \]  

(6)

Where \( k \) is a constant value, \( \phi \) is the function calculates the efficiency of the classification of the features of \( Y \) with the particular \( C \), and \( L(W) \) is same as equation (5). As a final point, the learning rule is defined in the following form:

\[ \Delta w_i = \eta \gamma_i (E(Z_i w_i^T) + k(\partial\phi(W, Z, C)/\partial w) + 2\beta_i w_i) \]  

(7)

\[ \beta_i = -1/2\gamma_i E(y_i g(y_i)) \]  

(8)

\[ \gamma_i = 2(-\sum_{n=1}^{N} \exp(-y_i^2/2))^{-1/2} \]  

(9)

\[ W \] of symmetric orthogonalization is applied.

\[ W \leftarrow (WW^T)^{-.5}W \]  

(10)

[3] BACK PROPAGATION NEURAL NETWORKS

The back propagation neural network (BPN) is a commonly used supervised multi-layered feed forward neural network model that constructed with collection of interconnected layer such as input layer, hidden layer, and one output layer [12]. An input layer receive inputs signal from the external source and output layer provides output of the target signals. A layer between input
layer and output layer is called hidden layer. Though, single hidden layer of neural networks is sufficient for solving any complex problem with most wanted accuracy of the given problem[3].

The optimal number of hidden neurons of the hidden layer is generally determined through a trial-and-error method based on network errors found in the training and test data[12, 13]. Each layer of the BPNN is connected with set of weights. The adjustment of weight in the BPNN is done with help of gradient steepest descent training algorithm in order to minimize error between desired output and network output. The learning process of BPNN proceeds repeatedly until the network error associated with the training data sufficiently meet to a global minimum[14] [7].

[4] THE PROPOSED MODEL

In this paper, the BPNN model has been integrated with data smoothing scheme such as DICA[11] is proposed as a proposal. The Discriminant ICA is applied to eliminate inherently noise from the stock market indices in order to improving the generalization performances and provides higher prediction accuracy of traditional BPNN. The graphical representation of the proposed model has shown in Figure 2.

While using Discriminant ICA for eliminating noise, the basic DICA model is first utilized on the mixture matrix X size of $n \times m$ combined with m forecasting variable $x_i$ of size $1 \times n$ for estimating de-mixing matrix $W$ of size $m \times m$ and independent components $y_i$ of size $1 \times n$. One hidden layer is adequate to represent in any complex problems with desired accuracy when constructing the BPNN model[7]. In this study, the BPNN model is used one hidden layer. There are no rules for selecting appropriate topology when building BPNN. The selection of the topology is followed by trial-error method which is used in this paper for building topology of the BPNN model.

[5] EXPERIMENTAL RESULTS AND DISCUSSIONS

a) Data Collections

The efficiency of the proposed DICA-BPNN is evaluated on a set of benchmark time series analysis data such as S&P BSE Sensex, Nifty 50. Totally, 3673 data points are collected from 3rd January 2003 to 30th September 2017 for all datasets. In this study, 3000 data points are used as a training set and left over 673 data points are used as a testing set from the collected datasets. The datasets contain an open price of the day, high price of the day, average price of data, and closing price of the day.

b) Performance Measurements

In order to analyzing the performance of the proposed DICA-BPNN is applied to well standard time series analysis datasets and its performance measured by statistical indicators namely Root Mean Square Error (RMSE), Mean Absolute Error (MAE) as follows,
\[ RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (t(i) - y(i))^2} \]  
\[ MAE = \frac{1}{N} \sum_{i=1}^{N} |t(i) - y(i)| \]  
Where, \( t(i) \) – target value, \( y(i) \) – predict value, \( N \) - No. of trading days.

c) Results Analysis

To investigate the performance of the proposed neural network model, we focus on prediction accuracy is considered. In this study, the weights of the learning algorithm are initiated by small random values that range between [-1.0, 1.0] and its termination conditions is set 0.005. The original datasets are normalized into the range [-1.0, 1.0].

Table 1: Performance comparisons of the proposed model for Nifty 50

<table>
<thead>
<tr>
<th>Methods</th>
<th>RMSE</th>
<th>MAE</th>
<th>Iteration(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DICA_BPNN</td>
<td>0.0107</td>
<td>0.0079</td>
<td>1830.5</td>
</tr>
<tr>
<td>ICA_BPNN</td>
<td>0.0108</td>
<td>0.0081</td>
<td>2589.3</td>
</tr>
<tr>
<td>BPNN</td>
<td>0.0204</td>
<td>0.0166</td>
<td>3210.9</td>
</tr>
</tbody>
</table>

Table 2: Performance of the comparison of the proposed model

<table>
<thead>
<tr>
<th>Methods</th>
<th>RMSE</th>
<th>MAE</th>
<th>Iteration(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DICA_BPNN</td>
<td>0.0106</td>
<td>0.0076</td>
<td>1645.7</td>
</tr>
<tr>
<td>ICA_BPNN</td>
<td>0.0112</td>
<td>0.0073</td>
<td>2439.4</td>
</tr>
<tr>
<td>BPNN</td>
<td>0.0204</td>
<td>0.0166</td>
<td>2978.6</td>
</tr>
</tbody>
</table>

The principle of linear scaling is to independently normalize each feature component to the accurate range. It makes sure that the large values of an input feature do not overwhelm smaller value inputs, which is helps to reduce prediction errors of predictors. The sigmoid functions \( g(x) = (1 - \exp(-x))/(1 + \exp(-x)) \) is used as an activation function of the both input and output phase. The result of the proposed model has been obtained by its statistical parameters and which is calculated its average values of the different 10 runs on 10 different weights.

d) Discussions

The Nifty 50 closing index dataset is used to predicting closing index price and those results are showed in Table-2. In Table-2, the proposed model achieved 0.0107 and 0.0079 in terms of RMSE and MAE respectively and their learning process is completed over the 1930.5 iterations. In terms of S&P BSE Sensex datasets, the RMSE and MAE is produced 0.0106, 0.0076 respectively and their learning process is completed over the 1645.7 iterations. The proposed DICA_BPNN model is produced model provides a better accuracy then the ICA-BPNN and Standard BPNN models.
[6] CONCLUSIONS

This paper presented DICA-BPNN model is used to predict stock indices in future behaviors based on their past behavior. The Discriminant ICA method has used to removing noise from the stock market data in order to avoid the over fitting/under fitting of the neural networks learning process. The proposed method is compared with ICA-BPNN prediction model and Conventional BPNN prediction model. The experimental results show that the proposed model can produce lowest prediction error, higher prediction accuracy and extremely high-speed convergence rate.

REFERENCES

Figure 1: graphical representation of the Proposed DICA-BPNN Prediction Model

Figure 1

Predictions vs Target comparisons of Nifty 50
AN INVESTIGATION OF FINANCIAL TIME SERIES PREDICTION USING BACK PROPAGATION NEURAL NETWORKS

Figure 2: Predictions vs Target comparisons of S&P BSE Sensex