International Journal of Information and Computation Technology.
ISSN 0974-2239 Volume 3, Number 5 (2013), pp. 383-390

http://www. irphouse.com /ijict.htm

Time Series Forecasting Using Artificial Neural Networks under Dempster–Shafer Evidence Theory and Trimmed-winsorized Means

Hari Teja.A[1] and S.Vasundra[2]

[1]CSE *Department, JNTUACEA, AP, INDIA.*
[2]CSE *Department, JNTUACEA, AP, INDIA.*

Abstract

Artificial neural networks are adequate for multiple jobs in design recognition and machine-learning. Contrasting to the traditional approaches for time series evaluation, a synthetic neural system needs advice regarding the time series data and could be placed on a comprehensive variety of issues. For the use, we conducted experiments to obtain the appropriate guidelines for a forecasting system. The artificial neural networks which were discovered had provided a better forecasting efficiency under well known feature selection approach. An Accurate time series forecasting using which is significant to demonstrate the approach within the past continues in the direction to influence the future as well as to schedule our daily actions. At present, a huge literature is in progress on the usage of artificial neural networks (ANNs) in many forecasting applications. In this regard the proposed artificial neural networks based on Time Series Forecasting use Dempster–Shafer evidence theory and trimmed-Winsorized means for feature selection. A comparative study between these two methods, with a set of referenced time series will be shown.

Keywords: Time series Forecasting; Dempster–Shafer; Trimmed-Winsorized; Artificial Neural networks; Machine Learning.

1. Introduction

Normally, the guidelines of the typical approach [7] are based on frequency range and the auto connection of time series. The utilization of artificial neural networks for time sequence analysis relies just about the data which were discovered. A synthetic neural system is strong enough to signify any kind of time series, as multiple layer feed forward systems with an adequate amount of hidden models and one or more hidden layers are capable of approximating function [5, 6] to any considerable. The forward and backward paths of the completely connected feed forward network may be applied by internal and external products of matrices and vectors in several outlines of APL code. For the application, we choose to make use of a completely joined, layered, feed forward artificial neural system with a single hidden layer and the back-propagation learning algorithm[1,2].

The fundamentals of the method for understanding in neural nets were set by [2]. Artificial neural nets contain many simple processing devices (called Processing factors or neurons) grouped in levels. Each layer is recognized from the index l=O,......,M. The processing elements are interlocked as follows, Conversation between processing elements is permitted for processing elements of nearby levels [3]. Neurons in a level cannot convey, each neuron has a particular activation level 'a' as shown in Fig.1, Data is processed by the network from the exchange of service levels between neurons .

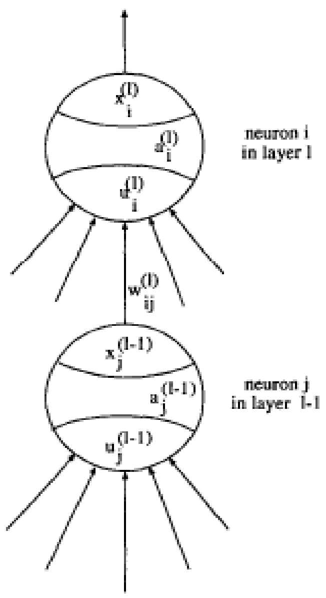

Figure 1: Interchanging of activation values among neurons.

The output value of the i-th neuron in sheet l is denoted by $\chi_i^{(l)}$, It is calculated with the formula $\chi_i^{(l)} = g(a_i^{(l)})$,

where $g(a_i^{(l)})$ is a monotone growing function. For instance, let us consider the function *g(y)* as the squashing function where $g(y) = \dfrac{1}{1+e^{-y}}$.

The commencement level $a_i^{(l)}$ of the neuron i in layer l is considered by $a_i^{(l)} = f(u_i^{(l)})$, Where $f(u_i^{(l)})$ is the activation function (in our case the characteristics function is used). The net input $u_i^{(l)}$ of neuron i in layer l is projected as follows,

$$u_i^{(l)} = (\sum_{j=1}^{n^{(l)}} \omega_{ij}^{(l)} \chi_j^{(l-1)}) - \theta_i^{(l)}$$

Where $\omega_{ij}^{(l)}$ is the weight of neuron j in layer $l-1$ connected to neuron i in layer l, $\chi_j^{(l-1)}$ the output of neuron j in layer $l-1$. $\theta_i^{(l)}$ is a prejudice allege that is removed from the total amount of the weighted activations.

2. Implementation

The time series modeling and forecasting classification was implemented using core 2 duo processor with 4 GB RAM of an auxiliary workstation. The system consists of two main components:

- A toolkit of APL functions that constrain the neural network and log parameters and results of the replication runs to APL component files.
- A graphical user interface that allows the user to navigate during the simulations and to evaluate the actual time series generated by the neural network.

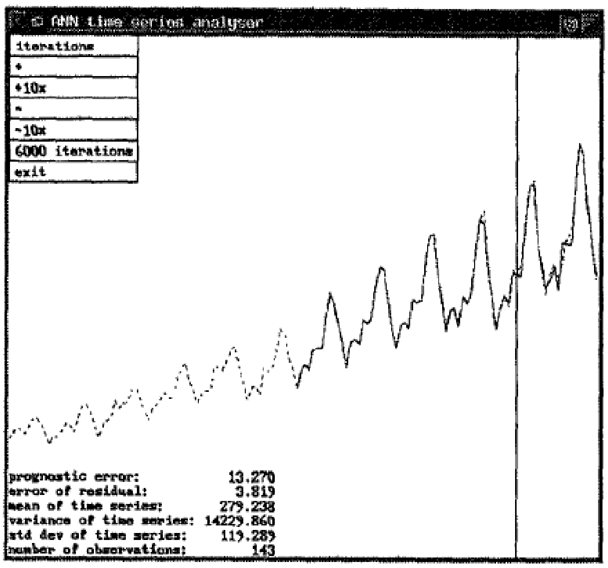

Figure 2: User interface of the forecasting system.

Fig. 2 portrays a display dump of the interface, the damaged line displays the real time sequence and the strong line signifies the network's output. The outlook data is divided from the historic data - the network's training set - by the vertical bar in the best quarter of the chart, the menu in the top left part of number 3 enables the person to choose a perspective of the network's forecasting capacity at different countries through the entire learning stage. By looking at the record files of the simulator runs, current and previous outcomes could be assessed and compared.

2.1 The Training Sets

As test bed for our forecasting program we utilized two well-known time sequence from [7], the monthly totals of international flight passengers (thousands of passengers) from 1949 to 1960, as well as the daily closing costs of share inventory.

Table 1: Properties of time series.

Time Series	σ	μ	n
Share Price	84.11	478.47	369
Grid power unit Price	119.55	280.30	144

Where σ is the standard deviation, μ is the mean and n is the number of observations. Following section will explain about how a neural network learn a time series.

2.2 Back-Propagation Algorithm

The neural network sees the time series X1,. . . . ,Xm in the form of many mappings of an input vector to an output value . This technique was presented in [5].

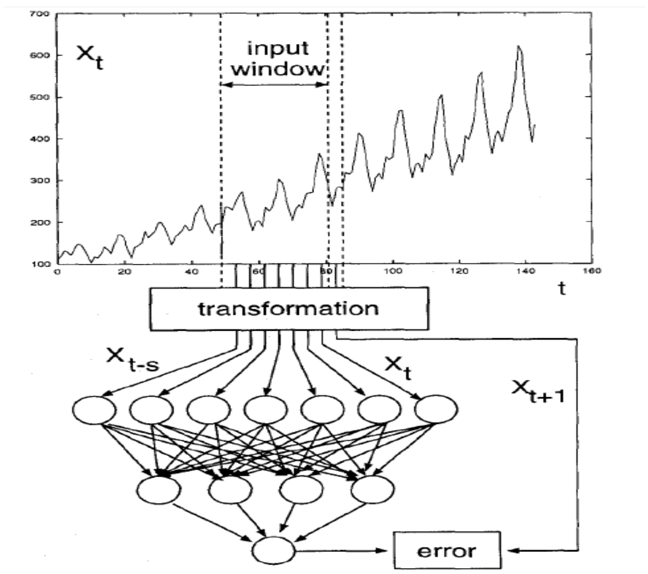

Figure 3: Learning a Time Series.

A number of adjoining data points of the time series (the input window X_t to X_{t-s}) as shown in Fig.3 are planned to the interval [0,1] and utilized as service levels for the models of the entire level. The problem useful for that back-propagation learning algorithm is currently computed by evaluating the worth of the result unit together with the value of the time sequence at time' t_1'. Training a sensory network with all the back-propagation algorithm generally demands that representations of the entire set (called one epoch) which is offered for several times.

To understand the period of series data, the representations were offered in a random fashion, As reported in [5], picking a random place for every portrayal's input window ensures better system functionality and prevents local minima.

3. Comparison between Divergent Feature Selection Modeling's

On comparing the results gained under feature selection methods called Dempster–Shafer evidence theory and trimmed-Winsorized means, the execution of the procedure uses the applications explained by Jenkins and Box in Component V of their classic [7]. The approach is known as an autoregressive integrated moving average method of order (p, d, and q). It is described by Eqs (1-2) as follows,

$$a(z)\nabla^d X_t = b(z)U_t \tag{1}$$

Where X_t stands for in time ordered values of a time series, $t = 1...N$ for n observations. U_t Is a sequence of random values called "white noise" process. The backward difference operator V is defined as

$$\nabla X_t = X_t - X_{t-1} = (1-z)X_t \tag{2}$$

We fixed these feature selection models for each time series and let it predict the next 20 observations of the time series. The last 20 observations were dropped from the time series and used to calculate the prediction error of the models.

The following feature selection models were calculated for the power unit price of the time series after a logarithmic transformation as,

$(1 - Z)(1 - z^{12})\ X_t = (1 - 0.24169,2 - 0.47962.z^{12})\ U_t$ and for the Share time series,

$(1 - z)\ X_t = (1 - 0.10538z) U_t$

In the forecasted errors of the man-made neural network (ANN), the artificial neural network utilizing the logarithmic and change (ANN record,) in the feature selection models opted are compared. The artificial neural network utilizing the logarithmic and changed time series out-performed for both time series, whereas the "simple" artificial neural network expected more precisely just for the Shares time series [4]. This behavior could be described as, the bigger the data selection in the time series, It results in a reduced precision for the untransformed place set. Logarithmic and differencing transformations helped to remove the tendency and planned the time series data in to a smaller variety.

Table 2: Forecasting errors for ANN under Dempster–Shafer evidence theory and Trimmed-Winsorized means.

Time series	Dempster–Shafer evidence theory	Trimmed-Winsorized means
power unit prices	20.97	18.72
Share price	10.97	11.35

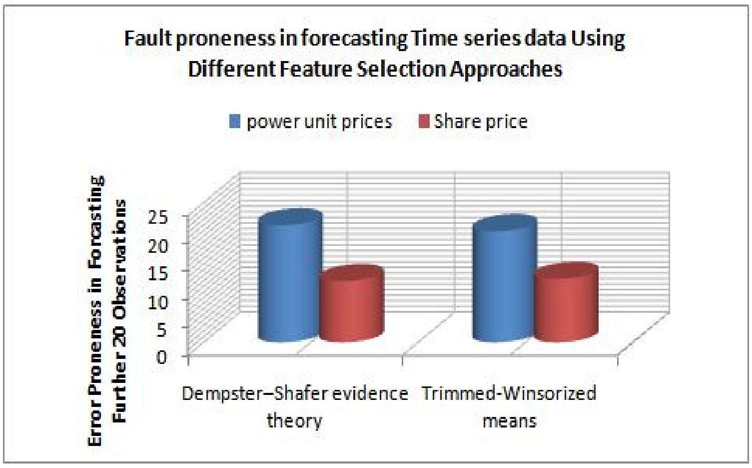

Figure 4: Fault proneness of ANN under Dempster–Shafer evidence theory and Trimmed-Winsorized means.

4. Conclusion and Future Work

We've offered a predicting program for uni-variant time sequence that utilizes artificial neural networks. These processing devices proved themselves to be feasible options to classic approaches. The machine may be utilized along with other approaches for time series evaluation or just as a standalone device.

Enhancement of the work may include the assessment with additional TSA (time series analysis) approaches, development of hybrid process that blend the power of standard methods with artificial neural networks and the use of our program to multivariate time series.

References

[1] Weizhong Yan and Senior Member "Toward Automatic Time-Series Forecasting Using Neural Networks " in the proceedings of IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, VOL. 23, NO. 7, JULY 2012

[2] G. P. Zhang and D. M. Kline, “Quarterly time series forecasting with neural networks,” IEEE Trans. Neural Newt., vol. 18, no. 6, pp. 1800 – 1814, Jun. 2007.

[3] S. Crone and P. Grafeille, “An evaluation framework for publications on artificial neural networks in sales forecasting,” in Proc. Int. Conf. Artificial intelligence ., Las Vegas, NV, Jun. 2004, pp. 221–227.

[4] S. D. Balkin and J. K. Ord, “Automatic neural network modeling for univariate time series,” Int. J. Forecast., vol. 16, no. 4, pp. 509–515,2000.

[5] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer Feed forward Networks are Universal Approximators. Neural Networks, 2:359-366, 1989.

[6] Halbert White. Economic prediction using neural networks: the case of ibm daily stock returns. In Proceedings of the IEEE International Conference on Neural Networks, pages II—451—11—459, 1988.

[7] George E. P. Box and Gwilym M. Jenkins. Time Series Analysis - forecasting and control. Series in Time Series Analysis. Holden-Day, 500 Sansome Street, San Francisco, California, 1976.