ORIGINAL RESEARCH PAPER | Mathematics | Volume : 6 | Issue : 9 | September 2016 | ISSN - 2249-555X | IF : 3.919 | IC Value : 74.50

Analysis of Wavelet Neural Networks for Time Series

KEYWORDS Stock market, Wavelet Transform, Neural Networks, Wavelet networks

Amandeep Kaur

Department of Mathematics, Guru Nanak Dev University, Amritsar-143005, India.

ABSTRACT *Wavelet neural networks (WNNs) combine the properties of wavelet decompositions along with the characteristics of neural networks. WNNs include wavelet functions in the neurons of hidden layer of neural network. It has good generalization ability, can approximate complex functions and can be easily trained than other networks. This paper describes the development of wavelet neural networks for time series.*

Introduction

Stock market prediction is hard because financial time series is highly irregular and nonlinear. Various techniques have been developed for the analysis of the nonlinearity of time series. These are linear regression (LR), Box-Jenkins method, autoregressive random variance (ARV) model, autoregressive conditional hetroskedasiticity (ARCH), general autoregressive conditional hetroskedasiticity (GARCH). While these techniques may be good for a particular situation, they do not give satisfactory results for the nonlinear data series. Therefore, these traditional linear models are not suited to model financial time series. So the idea of applying non-linear models, like soft computing technologies, such as neural networks, fuzzy systems, have become important for the analysis of non-linear and non-stationary data. Neural networks are effective in realizing the input–output mapping and can approximate any continuous function given an arbitrarily desired accuracy. In addition, there is no prior assumption of the model form required in the model building process. Because of the attractiveness of neural networks, a large number of applications have been proposed in recent decades for predicting stock markets using neural networks. The backpropagation (BP) is the most widely used algorithm to train the artificial neural networks

Wavelet neural network is an important tool for analyzing nonlinear and non-stationary data. It takes advantages of high resolution of wavelets and learning and feed forward nature of neural networks. Wavelet networks are as neural network for training and structural approach. But, training algorithms of wavelet networks is required a smaller number of iterations when the compared with neural networks. Gaussian based mother wavelet function is used as an activation function. Wavelet networks have three main parameters; dilation, translation, and connection parameters (weights). Initial values of these parameters are randomly selected. They are optimized during training (learning) phase.

Recently wavelet network has been used as an alternative of the neural networks because of interpretation of the model with neural networks is so hard. On the other hand training algorithms for wavelet networks require less number of iterations than neural networks. The wavelet network is an approach for system identification in which nonlinear functions are approximated as the superposition of dilated and translated versions of a single function. There is another approximation method except the neural network. This is wavelet decomposition. In the wavelet decomposition, only the weights are identified, while the dilation and translations will follow the regular grid structure. In contrast, in the wavelet network, weights dilations and translations are jointly fitted from data. Wavelet network use a wavelet like an activation function.

Wavelets

Wavelet transform is based on a finite support test function, called wavelet, which acts as a localised filter of the original signal. It means wavelet gives local information of the signal in terms of frequency and time. Each wavelet is derived from a zero mean mother function ψ through two linear transformations (i) translation by b (ii) dilation by scale parameter a. These parameters determine the width of the window and hence define the resolution of the transform

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\psi(\frac{x-b}{a})$$

The wavelet transform of a function $f(x) \in L^2(R)$ for a given resolution (a, b) is defined as

$$WT_{a,b}f(x) = \int_{-\infty}^{\infty} f(x)\psi_{a,b}(x)\ dx$$

which, in the Hilbert space L^2(R), corresponds to the inner product $\langle f, \psi_{a,b}\rangle$ i.e. the projection of the function in the direction of vector $\psi_{a,b}$.

As a matter of fact, the set $\{\psi_{a,b}\}_{a,b \in R}$ represents a basis of $L^2(R)$ over which any given signal with finite energy may be decomposed into its different frequency bands

$$\tilde{f}(x) = \sum_{a,b \in R} \langle f, \psi_{a,b}\rangle\, \psi_{a,b}(x) \qquad (1)$$

Here the function $\tilde{f}(x)$, approximation of the original function $f(x)$, is as close as desired to the original function for a finite number of components N i.e.

$$\forall \in, \exists\, N \ \ such\ that$$

$$\left\| f(x) - \sum_{n=1}^{N} \langle f, \psi_n\rangle \psi_n \right\| < \in$$

These combine parallel decompositions over different resized wavelets in an attempt to approximate non linear signals while ensuring that certain desired components are maintained. Wavelet Networks derive logically from this structure and go beyond, offering the possibility to be applied for learning or classification. They present interesting characteristics that will be detailed in the following section and which make them especially relevant for applications in computer vision such as face recognition or tracking biology etc.

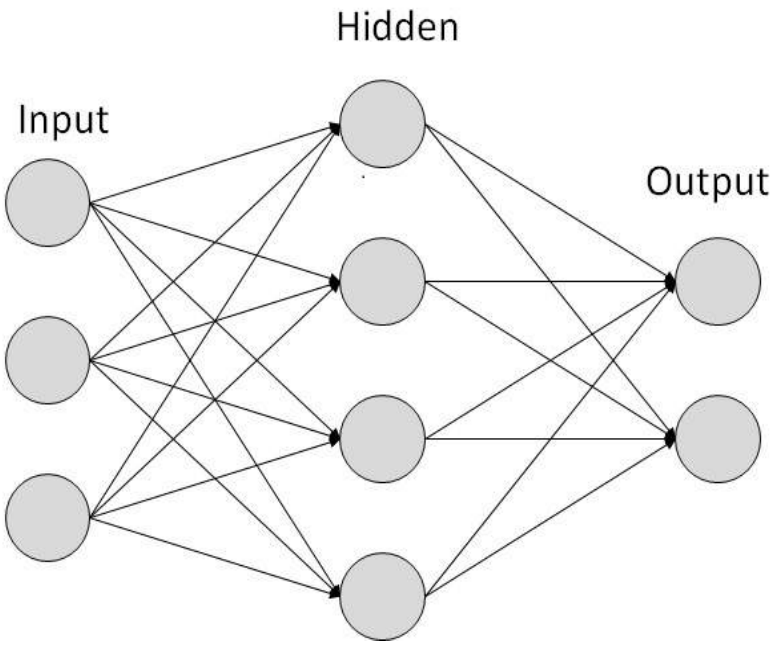
There are two types of wavelet transforms: continuous wavelet transform and discrete wavelet transform. The first is designed to work with functions defined over the whole real

Volume : 6 | Issue : 9 | September 2016 | ISSN - 2249-555X | IF : 3.919 | IC Value : 74.50

axis. The second works with functions that are defined over a range of integers. Wavelets show local characteristics which is the main property in both space and spatial frequency. It provides a unified framework for a number of techniques that had been developed independently for various signals processing application, e.g., multiresolution signal processing used in computer vision; subband coding, developed for speech and image compression; and wavelet series expansions, developed in applied mathematics.

Neural Network

Neural Networks are relatively crude electronic models based on the neural structure of the brain. The brain basically learns from experience. The human brain has capabilities in processing information and marking instantaneous decision. The many researchers have shown that the human brain makes computations in a radically different manner to that done by binary computers. The neurons is a massive network of parallel and distributed computing elements, many scientists are working last few decades to build computational system called neural network, which is also called as connectionist model. A neural network is composed of set of parallel and distributed processing units called nodes or neurons, these neurons are interconnected by means of unidirectional or bidirectional links by ordering them in layers.



Neural networks are flexible, nonparametric modeling tools. They can perform any complex function mapping with arbitrarily desired accuracy. Neural network is typically composed of several layers of many computing elements called nodes. Each node receives an input signal from other nodes or external inputs and then after processing the signals locally through a transfer function, it outputs a transformed signal to other nodes. They are characterized by the network architecture, that is, the number of layers, the number of nodes in each layer and how the nodes are connected. In a popular form of neural network called the multi-layer perceptron (MLP), all nodes and layers are arranged in a feedforward manner. The first layer is called the input layer where external information is received. The last layer is called the output layer where the network produces the model solution. In between, there are one or more hidden layers which are critical for neural network to identify the complex patterns in the data.

Like in any statistical model, the parameters or weights of a neural network model need to be estimated before the network can be used for prediction purposes. The process of determining these weights is called training. The training phase is a critical part in the use of neural networks. In using neural networks, the entire available data set is usually randomly divided into a training set and a test set. The training set is used for neural network model building and the test set is used to evaluate the predictive capability of the model.

There are many areas where neural network has been successfully applied which are described as follows:

- Neural network for process monitoring and optimal control.
- Neural network is a semiconductor manufacturing process.
- Neural network is a power system.
- Neural network in robotics.
- Network in communications
- Neural network in pattern recognition.

Structure of wavelet network

Wavelet Networks attempt to combine the properties of the Wavelet decomposition along with the characteristics of neural networks. Their structure relies on the aforementioned principles underlying non-linear function approximation and is given by the equation:

$$\check{f}(x) = \sum_{i} w_i \psi_{n_i}(x) \qquad (2)$$

Where the weights w_i represent the coefficients of the network. These are to be tuned as the network learns, in order to give preference to relevant components among the set of N wavelet functions $\Psi = (\psi_{n1}, \psi_{n2}, \dots \psi_{nN})$, whereas non-relevant ones are to be penalised. In this notation, the vector n_i for each wavelet gathers its corresponding parameters, i.e. $n = (s, u)$ in the case of the 1D decomposition of functions in $L^2(R)$ or $n = (s_x, s_y, u_x, u_y, \theta)^1$ for 2D images in $L^2(R^2)$.

From the equation (2), the wavelet network is completely defined by the tuple (ψ, w). Its optimised components may be obtained by calculating the weights w_i and wavelet parameters n that minimise the least-square error function for $f(x)$, i.e. the ones that make the model fit better to the original function f

$$min \left\| f(x) - \sum w_i \Psi_{n_i} \right\|^2$$

This calculation hence implies finding the most suitable N wavelets on which to project, along with the weight that each component ought to be given how much it ought to contribute to the overall description of $f(x)$ in order to maximise the approximation $\hat{f}(x)$. Such procedure is to be carried out during the learning phase so as to adapt the network to the set of training data points. Let us remind that minimising the least-square error function in $L^2(R^2)$ corresponds to finding the function that reduces the euclidean distance to each point in the training data, i.e. the one that minimises the lost energy that is due to the approximation.

We can see from equations (1) and (2) that there is a tight link between the weights of the network w_i and the wavelet decomposition $\langle f, \psi_{a,b} \rangle$. As a matter of fact, the values for w_i are automatically provided by the wavelet transform. As a result, the final distribution of weights (after optimisation) inherits the suitability of wavelets for 'feature detection'. The w_i are thus automatically tuned so as to prioritize projections that highlight transients, whereas wavelets encoding less relevant parts of images are penalised. The result is a network which is directly related to the underlying image structure and in which all parameters (w_i, n) are jointly fitted from data.

Applications of Wavelet Neural Networks.

The properties of Wavelet Neural Networks especially are interesting for a wide range of applications in engineering, computer science or biology. These may range from classification, to feature extraction or approximation of complex nonlinear functions. There are some relevant examples that will be further discussed in this section, in order to give a flavour of the possibilities offered by this technique:

Computer Vision

Many probabilistic approaches have been developed for computer vision, such as neural networks. These methods tend to learn the variance of grey-value pixels 5 over a set of training data, and then make use of that pixel-basis knowledge to classify new images. This is performed independently of the object itself.

The interest of Wavelet Networks relies on their ability to be directly related to the underlying structure of the image. The wavelets functions – on which the model is built – are 'natural' feature detectors and are independent, for instance, of illumination changes. Moreover, they provide a resolution that may be tuned in order to concentrate on given regions, making them especially suitable for surveillance or tracking applications.

Engineering

In Engineering, Robot motion is described by complex non-linear dynamics equations which include time-dependent parameters and system uncertainties, as observed in the differential description:

$$u = M(q)\ddot{q} + C(q, \ddot{q})\dot{q} + g(q)$$

In this framework, approximations through non-linear networks turn out to be very useful for learning control patterns, solving inverse kinematics problems and synthesising correct behaviour. This purpose has been previously studied and solved by means of neural networks based on radial basis functions (RBF), i.e. functions that depend only on the distance to a reference point (or centre) c_i. However, for a given function, the RBF network may not be unique, nor particularly efficient. The model developed by Katic, for example, replaces this activation function by a wavelet-based network which then plays the role of a robust controller,

Volume : 6 | Issue : 9 | September 2016 | ISSN - 2249-555X | IF : 3.919 | IC Value : 74.50

helps compensate uncertainties when the system is in contact with the environment, and yields much more computational efficient results.

Thomas extends the concept of wavelet networks to cope with the requirements of production lines. This purpose entails monitoring large number of non-stationary signals which are obtained from sensors, and performing feature extraction and classification so as to come up with a diagnosis system.

Conclusion

This paper presents an overview of Wavelet networks. Wavelet networks are an alternative to neural networks for nonlinear function learning. Wavelets show local characteristic so the initial values of translations and dilations requires more care than the initial value of the weights. If selecting initial values of the translation and the dilation properly, training time is shorter than neural networks. The architecture of a wavelet network is exactly specified by the number of wavelets required for a given classification or regression application. The optimal wavelet network structure is achieved the best approximation and prediction capability.

References

1. Mallat, S.G, A Wavelet Tour Signal Processing, Academic Press, 1998
2. Gilbert Strang and Truong Nguyen, Wavelet and Filter Banks, Wellesley-Cambridge Press,1996
3. Zhang, Q.; Benveniste, A. "Wavelet networks," IEEE Transactions on ," Neural Networks, Vol: 3 6, pp. 889 -898, Nov. 1992
4. Wen J, Zhao JL, Luo SW, Han Z (2000). The Improvements of BP Neural Network learning Algorithm. Proceedings of the 5th Int. Conf. On Signal Processing WCCC-ICSP. 3:1647-1649.
5. Zhang G, Patuwo BE, Hu MY (1998). Forecasting with Artificial Neural Networks: The state of The Art. International Journal of Forecasting. 14: 35-62.
6. Giordano F, La Rocca M, Perna C (2007). Forecasting Nonlinear Time Series with Neural Network Sieve Bootstrap. Computational Statistics and Data Analysis. 51: 3871-3884.
7. Hornik K, Stinnchcombe M, White H (1989). Multilayer Feed Forward Networks are Universal Approximators. Neural Networks. 2: 359-366.
8. Huang W, Lai. KK, Nakamori Y, Wang SY, Yu L (2007). Neural Networks in Finance and Economics Forecasting. International Journal of Information Technology and Decision Making. 6: 113-140.
9. Simon Haykin; Neural Networks: A Comprehensive Foundation, Macmillan College Publishing Company, Inc. 1994.
10. Oussar Y, Dreyfus G (2000) Initialization by selection for wavelet network training. Neurocomputing 34:131–143.
11. Pati YC, Krishnaprasad PS (1993) Analysis and synthesis of feedforward neural networks using discrete affine wavelet transformations. IEEE Trans Neural Netw 4(1):73–85.
12. Daubechies I (1990). The wavelet Transform, Time-Frequency Localization and Signal Analysis. IEEE Transactions in Information Theory. 36 (5): 961-1005.