ISSN 2395-1621

Predict PC Data Using Artificial Neural Network and Support Vector Machine

#1Harshali S. Lidhure.

1harshali_3691@rediffmail.com

#1Department of Electronics and Communication.

S.S.G.B.C.O.E.T, BHUSAWAL.
Dist-Jalgoan, Maharashtra, India.

ABSTRACT

Once a network has been planned for a particular application, that network is ready to be trained. Then, the training, or learning, begins. In this paper we offer methods that select a separation of data for SVM training. Several methods have been proposed, one of which is to decrease the training set size called training data selection. Recently, there has been great study on data selection for SVM training.

However, the training of neural networks can take many hours on a PC. In this paper we suggest a technique for training difficult neural networks for a general purpose computer. Our system allows the training to be done with minimal CPU utilization time. This allows the user to carry out other tasks while the training is in growth. We balance several avenues of neural network training on a general purpose computer. In this paper we propose methods that select a separation of data for SVM training. To speed up training thereby reduction the time for model selection, several methods have been proposed, one of which is to reduce the training set size called training data selection. Recently, there has been great research on data selection for SVM training.

In particular, prediction of time series using multi-layer feed-forward neural networks will be described. First, the topic of prediction will be described together with classification of prediction into types. After that, the prediction using neural networks (NNs) will be described. The focus will be on the creation of a training set from a time series. The paper is concluded with a summary of an experiment consisting of running an implementation of an automated ANN based prediction system on examples of real life time series data, and basic conclusions drawn from some of the results of this experiment.

Support Vector Machine is used for Time Series Prediction and Compared to neural network. The performances of these techniques are compared and it is concluded that SVM provides better performance as compared to NN Networks. The implementation is done by using MATLAB.

ARTICLE INFO

Article History

Received: 19th December 2016

Received in revised form :

19th December 2016

Accepted: 21st December 2016

Published online :

22nd December 2016

***Keywords:*-Neural network, artificial neural network, support vector machine.**

I. INTRODUCTION

Neural Networks have been widely used as time series forecasters: most often these are feed-forward networks which employ a sliding window over the input sequence.

ARTIFICIAL NEURAL NETWORKS AND TIME SERIES PREDICTION

Artificial neural networks are relatively simple electronic networks of neurons based on the neural structure of the brain. They process report one at a time, and learn by comparing their prediction of the record (largely arbitrary) with the known actual record. The errors from the initial prediction of the first record is fed back to the network and used to adjust the network's algorithm for the second interval.

2.1. Neural network for forecasting: - Let an ANN have k inputs and 1 output. An ANN consists of multiple artificial neurons, each having multiple weighted inputs, and an output. The output of the entire network, as a response to an input vector, is generated by applying certain arithmetic operations (as shown in fig 1).

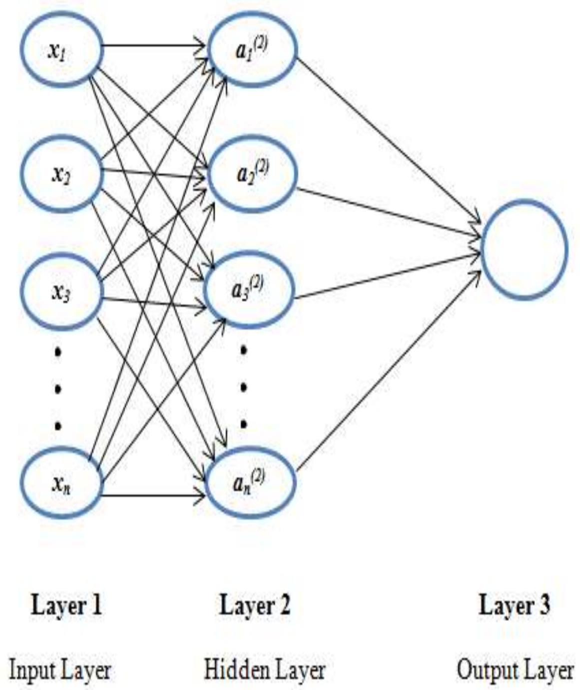

Fig 1 Structure of Artificial Neural Network

2.2 Issues with finding proper network:- One issue is that machine learning, in case of ANNs, requires important computational assets.
Another problem is finding the right network structure. Neurons need nonlinear continuous activation functions

2.3. Structuring the Network:-The number of layers and the number of handing out elements per layer are important decisions. To a feed forward, these parameters back-propagation topology, are also the most delicate - they are the art of the network designer. There is no experimental, best answer to the layout of the network for any particular application. There are only three general rules picked up over time and followed by most researchers and engineers applying this architecture to their problems.
Rule One: As the difficulty in the relationship between the input data and the desired output increases, the number of the giving out elements in the hidden layer should also increase.
Rule Two: If the process being modeled is discrete into multiple stages, then additional hidden layer(s) may be required. If the process is not separable into stages, then other layers may simply permit memorization of the Training Set, and not a true general solution.
Rule Three: The amount of training data available sets an upper leap for the number of giving out elements in the hidden layer(s). To calculate this upper leap, use the number of cases in the Training Set and divide that number by the addition of the number of nodes in the input and output layers in the network. If too many artificial neurons are used, the Training Set will be memorized, not global, and the network will be useless on new data sets.

2.4. Support Vector Machine:-Time series prediction techniques have been used in many real-world applications such as financial market prediction, electric utility load forecasting, weather and environmental state prediction, and reliability forecasting. The underlying system models and time series data generating processes are generally complex for these applications and the models for these systems are usually not known a priori. Exact and balanced estimation of the time series data produced by these systems cannot always be achieved using well known linear techniques, and thus the estimation process requires more advanced time series prediction algorithms. This paper provides a survey of time series prediction applications using a novel machine learning approach: support vector machines (SVM). The primary motivation for using SVMs is the ability of this methodology to exactly forecast time series data when the primary system processes are typically nonlinear, non-stationary and not defined a-priori. SVMs have also been proven to better other non-linear techniques including neural-network based non-linear prediction techniques such as multi-layer perceptrons.The ultimate aim is to provide the reader with approaching into the applications using SVM for time series prediction, to give a brief discussion on SVMs for time series prediction, to outline some of the advantages and challenges in using SVM.

III. RELATED WORK

3.1. Data preparation: -
Training set:-Training set can be made easily openly from the time series. Certain number of measured values is used as inputs and the value to be predicted (i.e., the value in the future, in some chosen distance after these input measured values) is used as required output. Input part of the time series is called window, the output part is the predicted value. By shifting the window over time series the items of training set are made (see figure_2). It is advised to left part of time series for testing, i.e., to not use this part during learning, but to use it to test how effectively the network learned to predict our data.

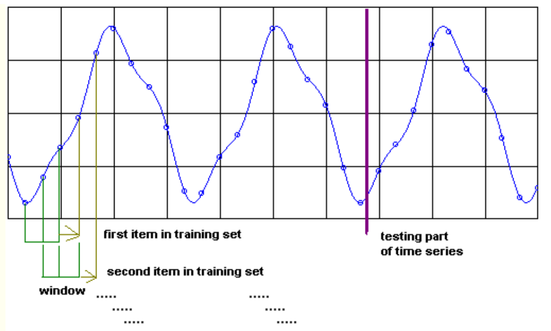

Fig2 Training set

The training set obtained in this way can be then adjusted for the needs of a particular neural network. For example, it may be necessary to adjust the values to a certain interval, such as (0,1).

Available data are often divided into three set: learning set, validating set and testing set. These sets can overlap (see the figure 3) and do not have to be constant. The learning set is a sequence that is shown to the neural network during the learning phase. The network is adapted to it to achieve required outputs (in other words, weights in the network are changed based on this set). The difference to required output is measured using the validating set and this difference is used to validate whether the learning of the network can be finished. The last set, testing set, is then used to test whether the network is able to work also on the data that were not used in the previous process.

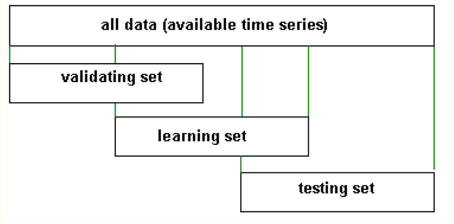

Fig 3 datasets

To summarize, the learning set is used for creating a model, validation set is used for verifying the model, and the testing set is used for testing of the usability of the model.

Especially for neural networks that can have outputs only in a certain interval it is important to realize that it is not possible to predict values outside of this interval. Data normalization is then required for the network to be able to get meaningful outputs.

3.2 Artificial Neural Networks:-Time series are sequences of data points in time, usually created by measuring output of some process in discrete time intervals. The goal of prediction is to successfully estimate output of the process in next time step or several steps. It is assumed that the process is at least partially observable and to some extent, future values can be determined by observing past values. Prediction then reduces to problem of process approximation. Since value of stored data decreases over time, effort is put into gaining insight from data as they come. Many applications don't require storing data at all, or storing it would be impractical, therefore it is advantageous to create models and update them with every data point. Continuous online learning is well suited for such tasks and can adapt to changing environments without human intervention.

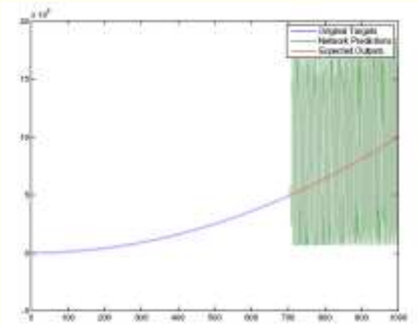

Fig 4 output of artificial neural network.

3.3. Support Vector Machine: - Even though various techniques have been used in the literature survey, still there is a need of best techniques to solve the following research issues. The Learning phase of SVM scales with the number of training data points. If the data set size increases, the learning phase can lead to a slower process. Some activities like Data cleaning, Data Transformation and Outlier detection are important problem for any type of data sets since some of the attributes values cannot be obtained usually. So handling of missing values for classification or forecasting problem is a challenging task. Complexity of handling the large dataset for any application is a recent issue since most of classification algorithms are not suitable to handle it. Selection of most apposite sample of data for classification instead of the intact data is another risk for getting better result. Selecting the suitable classification techniques without much computation complexity is another positive direction but the effectiveness should not be affected.

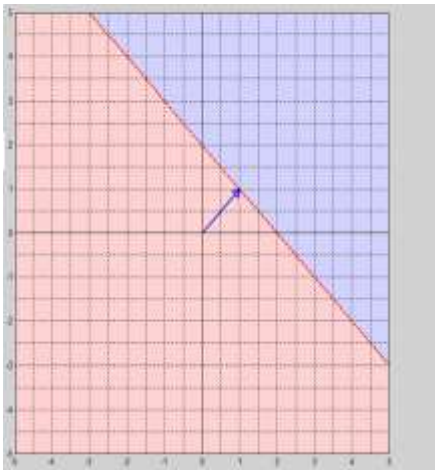

Fig 5 output of support vector machine.

IV. CONCLUSION

The main purpose of this study was to investigate the applicability and capability of the SVM methods for modeling of time-series forecasting. To verify the application of this approach, the benchmarked data are used in this study. The well known data are five well-known data sets that always handled in real life time series application. There are the chemical process concentration, the IBM common stock closing prices, the chemical process temperature, the Wolf's sunspot data and the international airline passengers. The prediction result by SVM method is compared with those by ANN.

In this research work, we have examined the feasibility of applying two machine learning models, Support Vector Machines (SVM) and Back Propagation Neural Network (BPN), to financial time-series forecasting for the futures trading in Indian derivative markets. Our experiments demonstrated that: SVMs provide a promising alternative tool to the neural network for financial time series forecasting. This is because SVMs adopt the Structural Risk Minimization Principle, eventually leading to better generalization than that conventional technique.

Finally, further research is needed to determine how long a trained ANN system remains valid and effective in prediction before it is found to be in need of retraining.

So by this we can conclude this system works with better efficiency considering SVM as a classifier either then ANN.

V. ACKNOWLEDGEMENT

I wish to express my profound thanks to all who helped us directly or indirectly in making this paper.

Finally I wish to thank to all our friends and well wishers who supported us in completing this paper successfully

I am especially grateful to our guide **Prof. G. A. Kulkarni** Sir for him time to time, very much needed, valuable guidance. Without the full support and cheerful encouragement of my guide, the paper would not have been completed on time.

REFERENCES

[1] Brillinger D.R., Time series. Data analysis and theory, Holt, Rinehart & Winston, 1974.

[2] Russel S., Norvig P., Artificial Intelligence: A Modern Approach(3rd Editioon),Prentice Hall,2009.

[3] Haykin,S.,1999,Neural Networks: A Comprensive Foundation.2nd Ed.(Englewood Cliffs, NJ: Prentice-Hall).

[4] V. Vapnik., 1995, The Nature of Statistical Learning Theory. Springer,N.Y.ISBN 0-387-94559-8.

[5] Elena Sanchez-Nielsen, - Hand Gesture Recognition for Human – Machine Interaction,Department of Statistic O.R. and Computer science, University of La Laguna Edificio de Fisica y Mathematics 38271, La Laguna, Spain Vladimir I. Pavlovic, Rajeev Sharma, IEEE transactions on pattern Analysis and Machine Intelligence, vol. 19, no. 7,July 1997

[6] S. M. Hosseini-Nezhad, T. S. Yamashita, R. A. Bielefeld, S. E. Krug and Y. H. Pao, “A neural network approach for the determination of inter hospital transport mode”, Comput Biomed Res., vol. 28, no. 4, (1995), pp. 319-334.

[7] C. J. Burges, “A tutorial on support vector machines for pattern recognition”, Data Mining and Knowledge Discovery, vol. 2, (1998), pp. 1-43. [24] V. Vapnik, “Statistical Learning Theory”, John Wiley and Sons, (1998).

