

Chapter 22

Escalation of Prediction Accuracy With Virtual Data: A Case Study on Financial Time Series

Sarat Chandra Nayak

Kommuri Pratap Reddy Institute of Technology, India

Bijan Bihari Misra

Silicon Institute of Technology, India

Himansu Sekhar Behera

Veer Surendra Sai University of Technology, India

ABSTRACT

Random fluctuations occur in the trend of financial time series due to many macroeconomic factors. Such fluctuations lead to sudden fall after a constant raise or a sudden rise after a constant fall, which are difficult to predict from previous data points. At the fluctuation point, previous data points that are not too close to the target price adversely influence the prediction trend. Far away points may be ignored and close enough virtual data points are explored and incorporated in order to diminish the adverse prediction trend at fluctuations. From the given data points in the training set, virtual data positions (VDP) can be explored and used to enhance the prediction accuracy. This chapter presents some deterministic and stochastic approaches to explore such VDPs. From the given data points in the training set, VDPs are explored and incorporated to the original financial time series to enhance the prediction accuracy of the model. To train and validate the models, ten real stock indices are used and the models developed with the VDPs yields much better prediction accuracy.

1. INTRODUCTION

Artificial Neural Networks (ANNs) are data driven and requires sufficient number of examples for training. Insufficient number of training examples reduces the generalization and approximation capability of the model which may leads to suboptimal solutions. In many real life situations sufficient amount of training data may not be available or if so, the correlation between the data points may not be strong.

DOI: 10.4018/978-1-5225-2857-9.ch022

Particularly in case of financial time series, random variations occur in the movement of stock market due to several socio-economical factors. Such random fluctuations lead to sudden fall after a steady increase or a sudden rise after a gradual fall, which are difficult to predict from previous data points. At the fluctuation point, previous data points that are not too close adversely influence the prediction trend. Some researchers attempted to enrich the training volume by adopting the deterministic and stochastic schemes for generating artificial training examples. These schemes are developed for training back propagation neural network for approximating ordinary time series. The details of such schemes will be discussed in the first part of the chapter. In the schemes attempted by the previous researches, artificial training examples are generated by manipulating existing training data and these schemes were validated in function approximation, solving toy problems as well as some benchmark functions. In these methods each training example consists of only artificial or natural data points. Since there is the chance of existence of noise in the artificial training examples, the overall performance of the system may hamper. There are some methods generating artificial training samples by local interpolation of consecutive data points. The coexistence of these artificial sample points with the original data may be able to retain the trend or changes a little. These schemes are claimed to be effective for time series analysis and suggested to adopt where computation cost is not concerned due to increase in volume of training samples. However, the stock movements, which can be visualized as financial time series do not behave like ordinary time series. The existing artificial training sample generation schemes which follow local interpolation may not be able to handle the random fluctuations which occur frequently in the stock movement.

The back propagation neural networks lead to poor performance where there is insufficiency of training examples. One way of achieving improved performance is by generating derived training patterns from the original training data and incorporating them to the original training pattern. Some previous attempts (Abu-Mustafa, 1995; An, 1996; Grandvalet et al., 1997) made in order to address this lacking for time series forecasting. They adopted a method of adding artificial training example points to the training set. But this may reduce the performance of forecasting if they have much noise.

Cho et al. (1996) proposed a scheme of generating artificial training examples randomly within a space of original training examples. Their scheme has been validated empirically that the proposed scheme improves the generalization performance of back propagation in nonlinear regression problems. The scheme had some disadvantage in the sense that it is fragile to very few original training examples, since a committee of neural networks trained by very few training examples labels the generated artificial data. The research work conducted by Taeho Jo (2013) proposed three virtual term generation schemes and validated their effect in using the back propagation for the tasks of multivariate time series prediction. The work considered three artificial and one real data set and it is observed that the prediction errors were reduced at least by 30% as the effect of the virtual term generation schemes. As stock market prediction has been an important challenge due to its uncertainty and nonlinear behavior, in order to forecast the future trend accurately, dependability of such models should be improved. However, hardly any literature is found in support of research attempts in the domain of financial forecasting to enhance the performance of neural based models by exploring and incorporating virtual data points. Nayak et al. (2014) explored the scope of improving prediction accuracy in financial time series by incorporating Virtual Data Position (VDP) in the actual datasets. They observed that incorporation of VDPs helps in reducing prediction error to a substantial extent. The same authors applied VDPs with adaptive neuro-fuzzy inference system for the task of forecasting next day's closing prices of some fast growing stock markets and established the usability of VDPs (Nayak et al., 2014).

27 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/escalation-of-prediction-accuracy-with-virtual-data/187699

Related Content

Structural Alignment of RNAs with Pseudoknots

Thomas K. F. Wong and S. M. Yiu (2011). *Handbook of Research on Computational and Systems Biology: Interdisciplinary Applications* (pp. 550-571).

www.irma-international.org/chapter/structural-alignment-rnas-pseudoknots/52332

Materialized View Selection using Marriage in Honey Bees Optimization

Biri Arunand T.V. Vijay Kumar (2015). *International Journal of Natural Computing Research* (pp. 1-25).

www.irma-international.org/article/materialized-view-selection-using-marriage-in-honey-bees-optimization/146882

Improved Evolutionary Extreme Learning Machines Based on Particle Swarm Optimization and Clustering Approaches

Luciano D. S. Pacifico and Teresa B. Ludermir (2012). *International Journal of Natural Computing Research* (pp. 1-20).

www.irma-international.org/article/improved-evolutionary-extreme-learning-machines/76374

Dealing with Interaction for Complex Systems Modelling and Prediction

Walter Quattrociocchi, Daniela Latorre, Elena Lodi and Mirco Nanni (2010). *International Journal of Artificial Life Research* (pp. 1-11).

www.irma-international.org/article/dealing-interaction-complex-systems-modelling/38930

Multi-Objective Higher Order Polynomial Networks to Model Insertion Force of Bevel-Tip Needles

Hashem Yousefi and Mehdi Fallahnezhad (2015). *International Journal of Natural Computing Research* (pp. 54-70).

www.irma-international.org/article/multi-objective-higher-order-polynomial-networks-to-model-insertion-force-of-bevel-tip-needles/146884