Brill Academic Publishers
P.O. Box 9000, 2300 PA Leiden,
The Netherlands

Lecture Series on Computer and Computational Sciences
Volume 1, 2005, pp. 1-3

Computational Intelligence Methods for Financial Forecasting

N.G. Pavlidis[1], D.K. Tasoulis[1], V.P. Plagianakos[1], C. Siriopoulos[2] and M.N. Vrahatis[1]

[1]Department of Mathematics and Computational Intelligence Laboratory, University of Patras, GR-26110 Patras, Greece.

[2]Department of Business Administration, University of Patras, GR-26110 Patras, Greece.

Abstract: Forecasting the short run behavior of foreign exchange rates is a challenging problem that has attracted considerable attention. High frequency financial data are typically characterized by noise and non–stationarity. In this work we investigate the profitability of a forecasting methodology based on unsupervised clustering and feedforward neural networks and compare its performance with that of a single feedforward neural network and nearest neighbor regression. The experimental results indicate that the proposed combination of the two methodologies achieves a higher profit.

Keywords: Financial Forecasting, Unsupervised Clustering, Neural Networks, Nearest Neighbors

Mathematics Subject Classification: 62M10, 62M45, 91B84

1 Introduction

After the inception of floating exchange rates in the early 1970s, economists have attempted to explain and forecast the movements of exchange rates using macroeconomic fundamentals. Empirical evidence suggests that although these models appear to be capable of explaining the movements of major exchange rates in the long run and in economies experiencing hyperinflation, their performance is poor when it comes to the short run and out-of-sample forecasting [2]. Conventional time series models forecasting on global approximation models, employing techniques such as linear and non-linear regression, polynomial fitting and artificial neural networks. Such models are better suited to problems with stationary dynamics [13]. In the analysis of real–world systems two of the key problems are non–stationarity (often in the form of switching between regimes) and overfitting (which is particularly important for noisy processes) [13]. To improve forecasting performance numerous researchers have proposed methodologies that attempt to identify regions of the input space exhibiting similar dynamics and subsequently employ a local model for each region [5–8, 10, 13].

In [6, 7] the application of unsupervised clustering for the segmentation of the input space, and feedforward neural networks (FNNs) acting as local predictors for each identified cluster, was proposed. In this work we investigate the profitability of this methodology combined with a simple trading rule, and compare its performance with two other widely known forecasting methodologies, namely FNNs and the nearest neighbor regression. The profitability of these methodologies is evaluated on the time series of the daily spot exchange rate of the Euro against the Japanese Yen.

2 Algorithms

In this section we outline the unsupervised k-windows (UKW) algorithm, the nearest neighbors algorithm, and FNNs.

2.1 Unsupervised k-windows Algorithm

UKW generalizes the original k-windows algorithm [12] by approximating the number of clusters. The UKW algorithm employs a windowing technique to identify the clusters present in a dataset. Specifically, assuming that the dataset lies in d dimensions, UKW initializes a number of d–dimensional windows over the dataset. Next it moves and enlarges these windows so as to enclose all the patterns that belong to one cluster within a window. The movement and enlargement procedures are guided by the points that lie within each window at each iteration. As soon as the movement and enlargement procedures do not increase significantly the number of points within each window they terminate. The final set of windows defines the clustering result of the algorithm. UKW approximates the number of clusters by employing a sufficiently large number of initial windows, and at the final stage of its execution assigning windows that enclose a high proportion of common points to the same cluster [11].

2.2 Nearest Neighbors

Assume that the time series $\{z_t\}_{t=1}^{T}$ has been embedded in an n-dimensional space, yielding pattern vectors of the form $z_t^n = (z_t, z_{t-1}, \ldots, z_{t-n+1})$. To generate a prediction for z_t from the information available up to time $(t-1)$, first the k nearest neighbors of the pattern z_t^n are identified, and subsequently, an estimator of $E(x_t|x_{t-1}, \ldots, x_{t-n})$ by $\sum_{i=1}^{k} \omega_{ti} x_i$ is computed where ω_{ti} represents the weight assigned to the ith nearest neighbor. Alternative configurations of the weights are possible but we employed uniform weights as they are the most frequently encountered configuration in the literature.

2.3 Feedforward Neural Networks

In the context of time series modeling the inputs to the FNN typically consist of a number of delayed observations, while the target is the next value of the series. The *universal myopic mapping theorem* [9] states that any shift–invariant map can be approximated arbitrarily well by a structure consisting of a bank of linear filters feeding an FNN. An immediate implication of this theorem is that, FNNs alone can be insufficient to capture the dynamics of a non–stationary system. The supervised training process is a data driven adaptation of the weights that propagate information between the neurons. Learning in FNNs is achieved by finding a minimizer $w^* = (w_1^*, w_2^*, \ldots, w_n^*) \in \mathbb{R}^n$, such that $w^* = \min_{w \in \mathbb{R}^n} E(w)$, where E is the batch error measure of the FNN. The error function is based on the squared difference between the actual output value and the target value.

3 Experimental Results

The time series that was studied is the daily spot exchange rate of the Euro against the Japanese Yen. The 1682 available observations cover the period from 12/6/1999 to 29/6/2005. The first 1482 observations were used as a training set, while the last 200 observations were used to evaluate the profit generating capability of the different forecasting methodologies. Using the *false nearest neighbors* [4] method on the observations forming the trainig set we selected an embedding dimension of 5. Firstly we employed a global FNN with architecture 5-5-4-1 to model the time series. The FNN was trained for 200 epochs using the Improved Resilient Propagation algorithm [3] and subsequently its profit generating capability on the test set was evaluated. In particular, we assume that on the first day we have 1000 Euros available. The trading rule that we considered is the following: if the system at date t, holds Euros and $\widehat{x_{t+1}} > x_t$ (where $\widehat{x_{t+1}}$ is the predicted price for date $t+1$ and x_t is the actual price at date t) then the entire amount available is converted to Japanese Yen. On the other hand, if the system holds Japanese Yen and $\widehat{x_{t+1}} < x_t$, then the entire amount is converted to Euros. In all other cases, the holdings do not change currency at date t. The last observation of the series is employed to convert the final holdings to Euros. In all transactions we assume a cost of 0.25% [1]. The profitability from trading based on the predictions of the global FNN model is depicted with the red line (Global FNN) in Fig 1. A perfect predictor, i.e. a predictor that correctly predicts the direction of change of the spot exchange rate at all dates, achieves a total profit of approximately 9716 Euros excluding transactions costs, while including transactions costs reduces total profit to approximately 7301 Euros.

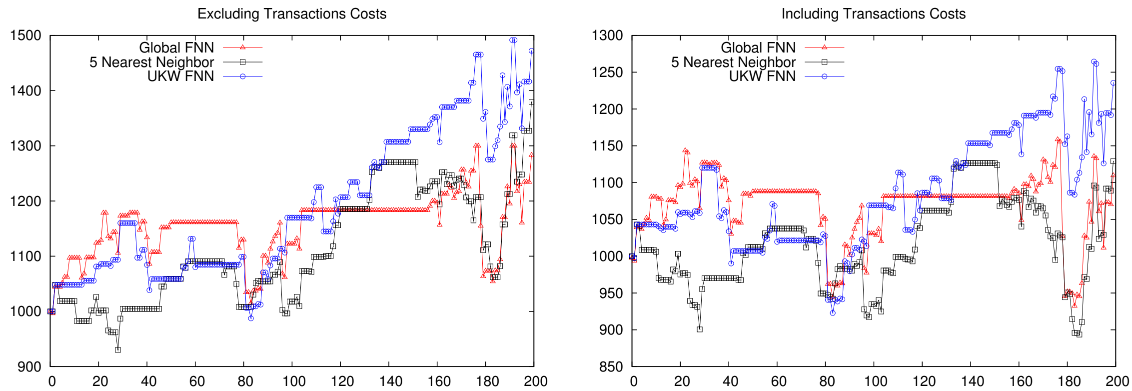

Figure 1: Trading performance of the different forecasting methodologies.

As can be seen from Fig. 1, excluding transactions costs the FNN is capable of achieving a profit of 282.94 Euros over the 200 days of the test set, while including transactions costs reduces total profit to 109.57 Euros. Next the results obtained by the k nearest neighbor regression method are presented. We experimented with all the integer values of k in the range $[1, 20]$. The best results were obtained for $k = 5$, and are illustrated with the black line (5 Nearest Neighbor) in Fig 1. The 5 nearest neighbor method achieved a profit of 379.51 excluding transctions costs and 129.16 including transactions costs. Finally, the performance of the forecasting system based on the segmentation of the input using the UKW algorithm and utilizing an FNN to act as a local predictor for each cluster, is illustrated with the blue line (UKW FNN) in Fig. 1. This approach achieved the highest profit: in the absence of transactions costs, 472.16 Euros, and 235.54 including transactions costs.

4 Concluding Remarks

In this paper, we report results concerning the profitability of a forecasting methodology that employs the UKW clustering algorithm and FNNs, acting as local predictors for each cluster, to predict the direction of movement of the daily spot exchange rate of the Euro against the Japanese Yen. The advantages of this methodology are that UKW automatically approximates the number of clusters present in a dataset, and that FNNs are capable of approximating an unknown non-linear relationship. The profit generating performance of this methodology compares favorably with that of a single FNN and nearest neighbor regression, which implies that it is capable of capturing more accurately the short run dynamics of the time series.

Acknowledgment

The authors thank the European Social Fund (ESF), Operational Program for Educational and Vocational Training II (EPEAEK II) and particularly the Program PYTHAGORAS, for funding this work. This work was supported in part by the "Karatheodoris" research grant awarded by the Research Committee of the University of Patras.

References

[1] F. Allen and R. Karjalainen, *Using genetic algorithms to find technical trading rules*, Journal of Financial Economics **51** (1999), 245–271.

[2] J. A. Frankel and A. K. Rose, *Empirical research on nominal exchange rates*, Handbook of International Economics (G. Grossman and K. Rogoff, eds.), vol. 3, Amsterdam, North-Holland, 1995, pp. 1689–1729.

[3] C. Igel and M. Hüsken, *Improving the Rprop learning algorithm*, Proceedings of the Second International ICSC Symposium on Neural Computation (NC 2000) (H. Bothe and R. Rojas, eds.), ICSC Academic Press, 2000, pp. 115–121.

[4] M. B. Kennel, R. Brown, and H. D. Abarbanel, *Determining embedding dimension for phase–space reconstruction using a geometrical construction*, Physical Review A **45** (1992), no. 6, 3403–3411.

[5] R. L. Milidiu, R. J. Machado, and R. P. Renteria, *Time-series forecasting through wavelets transformation and a mixture of expert models*, Neurocomputing **28** (1999), 145–156.

[6] N. G. Pavlidis, D. K. Tasoulis, and M. N. Vrahatis, *Financial forecasting through unsupervised clustering and evolutionary trained neural networks*, Proceedings of the Congress on Evolutionary Computation (CEC 2003), 2003, pp. 2314–2321.

[7] N.G. Pavlidis, D.K. Tasoulis, V.P. Plagianakos, and M.N. Vrahatis, *Computational intelligence methods for financial time series modeling*, International Journal of Biffurcation and Chaos (accepted for publication) (2005).

[8] J. C. Principe, L. Wang, and M. A. Motter, *Local dynamic modeling with self–organizing maps and applications to nonlinear system identification and control*, Proceedings of the IEEE **86** (1998), no. 11, 2240–2258.

[9] I. W. Sandberg and L. Xu, *Uniform approximation of multidimensional myopic maps*, IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications **44** (1997), no. 6, 477–485.

[10] A. Sfetsos and C. Siriopoulos, *Time series forecasting with a hybrid clustering scheme and pattern recognition*, IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans **34** (2004), no. 3, 399–405.

[11] D. K. Tasoulis and M. N. Vrahatis, *Unsupervised distributed clustering*, IASTED International Conference on Parallel and Distributed Computing and Networks, 2004, pp. 347–351.

[12] M.N. Vrahatis, B. Boutsinas, P. Alevizos, and G. Pavlides, *The new k-windows algorithm for improving the k-means clustering algorithm*, Journal of Complexity **18** (2002), 375–391.

[13] A. S. Weigend, M. Mangeas, and A. N. Srivastava, *Nonlinear gated experts for time series: Discovering regimes and avoiding overfitting*, International Journal of Neural Systems **6** (1995), 373–399.