Exchange Rate Forecasting with Ensemble K-PLS and Ensemble Neural Networks: A case study for the Indian Rupee/US Dollar

Mark J. Embrechts (embrem@rpi.edu) and Christopher J. Gatti
Department of Industrial and Systems Engineering
Rensselaer Polytechnic Institute, Troy, NY 12180, USA

Jonathan Linton
Telfer School of Management
University of Ottawa, Ottawa, ON K1N 6N5, Canada

The purpose of this presentation is to evaluate and benchmark ensemble methods for time series prediction for daily currency exchange rates using ensemble methods based on feedforward neural networks and kernel partial least squares (K-PLS). Ensemble methods reduce the variance on the forecasts and allow for the assessment of confidence metrics and risk for the forecasting model. The use of neural networks for time series forecasting has been well established and best practice methods are summarized in [1-3]. An obvious advantage of artificial neural networks is that the models are nonlinear and relatively easy to train. Shortcomings of the neural network literature for time series forecasting include (i) the lack of consensus for parameter settings; (ii) a lack of established standards for training neural networks; (iii) a lack of consistent evaluation metrics for time series forecasting; and (iv) a lack of clearly established benchmark problems. It has been shown that averaging neural network forecasts leads to more robust models that furthermore allow for an estimate of the confidence level [4]. Zimmermann reported that typically ensembles of 200 neural networks with the same neural network architecture, but with different seeds for the random weight initialization, are sufficient. Whereas Zimmermann applies ensembles of recurrent neural networks, ensembles of feedforward neural networks trained with the backpropagation algorithm will be applied in this work. In order to let training proceed in an automated fashion an extension of the Efficient BackProp strategy introduced by LeCun was applied [5, 6]. In addition, we propose two different types of ensemble methods: (i) an approach similar to that of Zimmermann, where the different neural networks have the same architecture, but are initialized with different random weights, and (ii) a novel ensemble strategy, where the training models use different weight initializations, but in addition, multiple cross-validation folds are used for training the neural networks. Two different types of time series forecasting methods will be investigated: (i) a one-step ahead prediction, and (ii) a roll-out prediction that will lead to long-term forecasts by feeding predictions back to the input space (just as if they were the actual values) and bootstrapping over successive steps to make multi-step ahead forecasts.

A novel ensemble method for time series prediction based on Kernel Partial Least Squares (K-PLS) is also introduced. Kernel partial least squares [7-8] is a ”kernelization” of the (linear) Partial Least Squares method. PLS is widely applied in chemometrics and was first introduced by Herman Wold for latent variable analysis of socio-economical models [9]. Svante Wold showed that PLS is a robust linear method, with few parameters to tune, except for the number of latent variables [10].

This study found that daily exchange rates such as the US Dollar per Euro and the Australian Dollar per Euro are generally not very predictable from single time series data. However, the Indian Rupee per US Dollar returns for late 2011, were surprisingly predictable, even on a relatively long time-scale (see Fig. 1). The ensemble method results for the two neural network approaches and K-PLS very similar to each other for a variety of metrics.

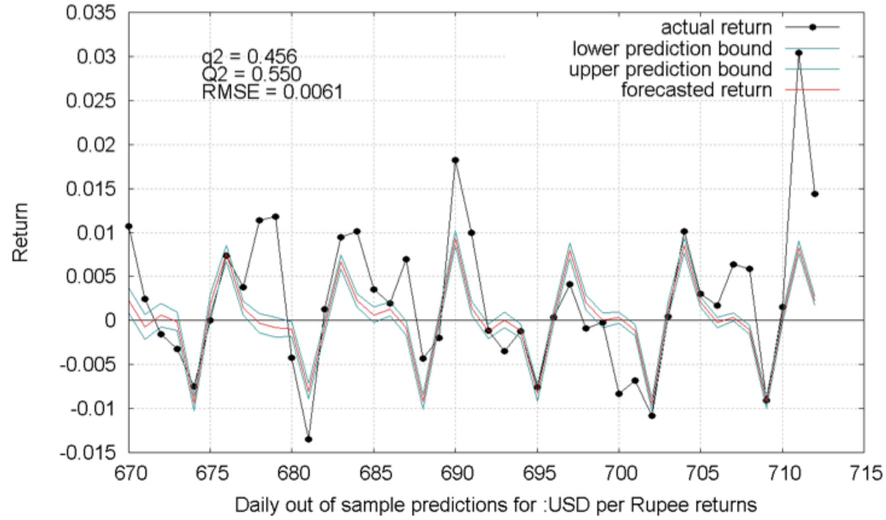

Fig. 4. Long-term roll-out forecasts with error bounds for the Indian Rupee per US Dollar Returns for the period November 1, 2011 to December 14, 2011 based on neural network ensemble averaging by weight initialization, error bounds for the forecasts are also indicated.

References

[1] J. Faraway and C. Chatfield. Time series forecasting with neural networks: A comparative study. *Applied Statistics*, Vol. 47(2), pp. 231 - 250, 1998.

[2] Y. Kajitani, K. H. Hipel, and A. I. McLeod. Forecasting nonlinear time series with feed-forward neural networks: A case study of Canadian lynx data. *Journal of Forecasting*, Vol. 24, pp. 105 - 117, 2005.

[3] H. G. Zimmermann, R. Grothmann, Historical Consistent Neural Networks for Forecasting and Risk Analysis in Economics. *Advances in Intelligent Signal Processing and Data Mining: Theory and Applications*, P. Georgieva, L. Mihaylova, L. C. Jain, Eds., Springer Verlag, Berlin, 2012.

[4] H. G. Zimmermann, C. Tietzand, and R. Grothmann. Yield curve forecasting by error correction neural networks and partial learning. *Proceedings of the 10th European Symposium on Artificial Neural Networks (ESANN 2012),* 2002.

[5] Y. LeCun, L.Bottou, G. Orr,and K. Müller. Efficient BackProp, in *Neural Networks: Tricks of the Trade*, G. Orr, and K. Müller, Eds., Springer, 1998.

[6] S. Haykin. *Neural Networks and Learning Machines* (3rd Ed), Prentice Hall, 2008.

[7] F. Lindgren, P. Geladi, and S. Wold. The kernel algorithm for PLS. *Journal of Chemometrics,* Vol. 7, pp. 45 - 59, 1993.

[8] R. Rosipal and L. J. Trejo. Kernel partial least squares regression in reproducing kernel Hilbert spaces. *Journal of Machine Learning Research*, Vol. 2, pp. 97 - 123, 2001.

[9] H. Wold. Soft Modeling by Latent Variables: The Nonlinear Iterative Partial Least Squares Approach. In *Perspectives in Probability and Statistics*, J. Gani, Ed., pp. 520–540, Academic Press, London, 1975.

[10] S. Wold, M. Sjöström, and L. Eriksson. PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, Vol. 58, pp. 109 - 130, 2001.

Topic: learning theory
Preference: poster