

Modeling and Trading the EUR/USD Exchange Rate Using Machine Learning Techniques

Konstantinos Theofilatos

Dpt of Computer Engineering and
Informatics
University of Patras
Patras, Greece
theofilk@ceid.upatras.gr

Spiros Likothanassis

Dpt of Computer Engineering and
Informatics
University of Patras
Patras, Greece
likothan@ceid.upatras.gr

Andreas Karathanasopoulos

London Metropolitan Business School
London,
United Kingdom
A.Karathanasopoulos@londonmet.ac.uk

Abstract— The present paper aims in investigating the performance of state-of-the-art machine learning techniques in trading with the EUR/USD exchange rate at the ECB fixing. For this purpose, five supervised learning classification techniques (K-Nearest Neighbors algorithm, Naïve Bayesian Classifier, Artificial Neural Networks, Support Vector Machines and Random Forests) were applied in the problem of the one day ahead movement prediction of the EUR/USD exchange rate with only autoregressive terms as inputs. For comparison reasons, the performance of all machine learning techniques was benchmarked by two traditional techniques (Naïve Strategy and moving average convergence/divergence model). Trading strategies produced by the machine learning techniques of Support Vector Machines and Random Forests clearly outperformed all other strategies in terms of annualized return and sharp ratio. To the best of our knowledge, this is the first application of Random Forests in the problem of trading with the EUR/USD exchange rate providing extremely satisfactory results.

Keywords- EUR/USD Exchange Rate; future direction prediction; Naïve strategy; MACD strategy; Naïve Bayesian Classifier; K-Nearest neighbors classifier; SVM; Random Forests; leverage; transaction costs

I. INTRODUCTION

The application of machine learning techniques for market predictions has been widely established in the scientific community. This paper deals with the application of a variety of state-of-the-art machine learning techniques in the problem of predicting the one day ahead movement direction of the EURO-USD exchange rates. Developing high accuracy techniques for predicting financial time series is a very crucial problem for economists, investigators and analysts. The traditional statistical methods, used by economists in the past years, seem to fail to capture the discontinuities, the nonlinearities and the high complexity of financial time series. Complex machine learning techniques like Artificial Neural Networks, Support Vector Machines (SVM) and Random Forests provide enough learning capacity and are more likely to capture the complex non-linear models which are dominant in the financial markets.

Some approaches examining the performance of machine learning techniques in trading with the EURO-USD exchange rate have already been developed. In [1], Dunis and Williams demonstrated the ability of Multi Layer Perceptron (MLP) Artificial Neural Networks in modeling and trading with the EUR/USD exchange rate. Their empirical results showed that the MLP outperformed all other benchmark models used. Next, in [2], Ullrich et al. used SVMs to trade with a variety of foreign exchange rates including EUR/USD. Their results indicated that SVMs outperformed Artificial Neural Networks and all other traditional techniques used in this paper for comparative reasons. Finally, in [3], Dunis et al. compared Higher Order Neural Networks, Psi Sigma Networks, Recurrent Networks and MLP in the task of trading with the EURO/USD exchange rate. MLP was proved to outperform all other neural networks variation.

The rest of the paper is organized as follows: In section II, the dataset of EURO-USD exchange rates is presented. In section III, all the traditional and machine learning techniques used in the present paper are briefly described. In section IV the comparative results are presented and in section V they are discussed and some future research directions are proposed.

II. THE EUR/USD EXCHANGE RATES AND RELATED FINANCIAL DATA

The European Central Bank (ECB) publishes a daily fixing for selected EUR exchange rates: these reference mid-rates are based on a daily concentration procedure between central banks within and outside the European System of Central Banks, which normally takes place at 2.15 p.m. ECB time. The reference exchange rates are published both by electronic market information providers and on the ECB's website shortly after the concentration procedure has been completed. Although only a reference rate, many financial institutions are ready to trade at the EUR fixing and it is therefore possible to leave orders with a bank for business to be transacted at this level.

The ECB daily fixing of the EUR/USD is therefore a tradable level which makes using of a more realistic alternative to, say, London closing prices and this is the series that we

investigate in the present paper. EUR/USD is quoted as the number of USD per Euro.

Specifically, the data which are used for our problem were downloaded from [4], and they were split as shown in Table 1 in order to train and evaluate our models.

TABLE I. EURO/USD DATASETS.

Name of period	Trading days	Beginning	End
Total dataset	2294	17 January 2002	30 December 2010
Training dataset	1363	17 January 2002	16 May 2007
Validation dataset	459	17 May 2007	2 March 2009
Testing dataset [out of sample dataset]	471	3 March 2009	30 December 2010

The graph in Figure 1 shows the total dataset for the EUR/USD. As inputs to our models we selected a set of autoregressive terms of the EUR/USD exchange rate returns presented in Table 2.

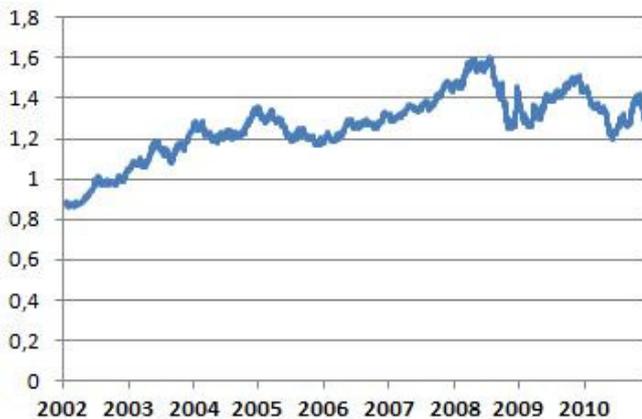


Fig. 1. Euro/USD daily fixing prices (total dataset)

TABLE II. EXPLANATORY VARIABLES.

Number	Input variable	Lag
1	EUR/USD exchange rate return	1
2	EUR/USD exchange rate return	2
3	EUR/USD exchange rate return	3
4	EUR/USD exchange rate return	4
5	EUR/USD exchange rate return	5
6	EUR/USD exchange rate return	6
7	EUR/USD exchange rate return	7
8	EUR/USD exchange rate return	8
9	EUR/USD exchange rate return	9
10	EUR/USD exchange rate return	10

III. FORECASTING MODELS

A. Benchmark Models

In the present paper, the machine learning methods presented in section 3.2 were benchmarked with 2 traditional strategies, namely Naïve strategy [3] and a moving average convergence/divergence technical model (MACD) [3].

The Naïve strategy takes the most recent period change as the best prediction of the future change. The MACD strategy used is quite simple. Two moving average series are created with different moving average lengths. The decision rule for taking positions in the market is straightforward. Positions are taken if the moving averages intersect. If the short-term moving average intersects the long term moving average from below a ‘long’ position is taken. Conversely, if the long-term moving average is intersected from above a ‘short’ position is taken.

B. Machine Learning Models

Some of the state-of-the-art machines learning classification techniques were applied in the problem of the one day ahead prediction of the direction movement of the EUR/USD exchange rate. These machine learning techniques are: K-nearest neighbor classifier (KNN), Naïve Bayesian classifier, Support Vector Machines (SVM) and Random Forests. In order to find the optimal parameters for each machine learning technique we used only the training and validation datasets leaving the test set for the final evaluation of the algorithms. Doing this we avoid specializing in our dataset and getting misleading results. The parameters were optimized using genetic algorithms [5]. All implementations were done using Matlab R2009a edition.

K-nearest neighbors [6] is a method for classifying objects based on closest training examples in the feature space. The KNN method is considered one of the simplest machine learning techniques. An object is classified by a majority vote of its neighbors. If K=1, then the object is assigned to the class of its nearest neighbor. In our case study the optimal K found was 8.

The Naïve Bayesian classifier [7] is a simple probabilistic classifier with strong assumptions of independence among input variable. It is the classifier derived from the use of Bayes’ theorem. Bayes’ Theorem expresses the conditional probability, or “posterior probability”, of a hypothesis H, (i.e. its probability after evidence E is observed) in terms of the “prior probability” of H, the prior probability of E and the conditional probability of E given H. It implies that evidence has a stronger confirming effect if it was more likely before being observed. Bayes’ theorem is valid in all common interpretations of probabilities, and it is commonly applied in science and engineering. The main disadvantages of Naïve Bayesian classifiers are their assumptions about the independency of the input variables, which is not usually the case in real problems.

Neural networks [8] exist in several forms in the literature. The most popular architecture is the Multi-Layer Perceptron (MLP). A MLP consists of at least three layers of nodes. The network processes information as follows: the input nodes

contain the value of the explanatory variables. Each node of the hidden layers passes incoming information through a nonlinear activation function and passes it to the output layer if the calculated value is above a threshold. Each node of one layer has weighted connections to all other nodes of the next layer. The training of the network (which is the adjustment of its weights in the way that the network maps the input value of the training data to the corresponding output value) starts with randomly chosen weights and proceeds by applying a supervised learning algorithm, called back propagation of errors. Since networks with sufficient hidden nodes are able to learn the training data (as well as their outliers and their noise) by heart, it is crucial to stop the training procedure at the right time to prevent overfitting (this is called 'early stopping'). In the present paper the best network's architecture found was the one using one hidden layer with 19 hidden neurons.

Support vector machines (SVM) are a group of supervised learning methods that can be applied in classification and regression problems. SVMs represent an extension to non linear models of the generalized algorithm developed by Vapnik [9]. They have already been applied in many scientific problems. Specifically, SVM have already been used in many prediction and classification problems in finance and economics although they are still far from mainstream. The few financial applications so far have only been published in statistical learning and artificial intelligence journals. SVM models were originally defined for the classification of linearly separable classes of objects. For any original separable set of two-class objects SVM are able to find the optimal hyperplanes that separates them providing the bigger margin area between the two hyperplanes. Furthermore they can also be used to separate classes that cannot be separated with a linear classifier. In such cases, the coordinates of the objects are mapped into a feature space using nonlinear functions. The feature space in which every object is projected is a high dimensional space in which the two classes can be separated with the linear classifier. In the present work we used the Radial Basis Function (RBF) as Kernel function for the SVM models because of its efficiency in providing very high performance classification results. The optimal RBF parameters C and gamma were found to be 64 and 2 respectively re-assuring that the model does not over fit.

Another sophisticated machine learning method, such as the support vector machines (SVM), is the random forest method. Random forests [10] are ensemble classifiers that "grow" many decision trees simultaneously where each node uses a random subset of the features considered. Specifically, Random forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. The idea of growing an ensemble of trees and letting them vote for the most popular class has led to significant improvements in classification accuracy. The generalization error for the forest converges to a limit as the number of trees in the forest becomes large. The generalization error for the forest of tree classifiers depends on the strength of the individual trees in the forest and the correlation between them. The optimal number of classification trees found for our case study was 51 which seem to be enough for achieving a very good generalization

performance. The applications of Random Forests in predicting the movement direction of financial time series remains nowadays quit limited despite of their high classification performance and their ability to generalize in data that have not been used to train their classifiers [11].

IV. EMPIRICAL TRADING SIMULATION RESULTS

The trading performance of all models considered in the validation subset is presented in Table 3. Due to the stochastic nature of the machine learning techniques, every method was executed 10 times and the results presented in Table 3 present the mean performance of these executions. The trading strategy derived for all the machine learning classification techniques used in the present paper is simple and identical for all of them: go or stay long if the classification model forecasts a positive movement and go or stay short if the classification model forecasts a negative movement. Since some of our models trade quite often, taking transaction costs into account might change the whole picture.

The transaction costs for a tradable amount, say USD 5-10 million, are about 1 pip (0.0001 EUR/USD) per trade (one way) between market makers. But since we consider the EUR/USD time series as a series of bid rates, we have to pay the costs only one and not two times per position taken. With an average rate of EUR/USD of 1.332 for the testing period, a cost of 1 pip is equivalent to an average cost of 0.008% per position.

TABLE III. TRADING PERFORMANCE RESULTS.

	Naïve Strategy	MACD	KNN	Naïve Bayes
Information Ratio (excluded costs)	-0.48	0.27	-0.12	-0.21
Annualized Volatility (excluded costs)	11.44%	11.46%	11.46%	11.46%
Annualized Return (excluding costs)	-5.51%	3.08%	-1.36%	-2.46%
Maximum Drawdown (excluding costs)	-26.54%	-16.96%	-13.98%	-15.14%
Correct Directional Prediction	49.24%		50.11%	48.83%
Transaction costs	1.04%	0.28%	0.90%	0.62
Annualized Return (including costs)	-6.59%	2.80%	-2.26%	-3.08

	Artificial Neural Networks (MLP)	SVM	Random Forests
Information Ratio (excluded costs)	0.22	0.43	0.72
Annualized Volatility (excluded costs)	11.46%	11.46%	11.45%
Annualized Return (excluding costs)	2.51%	4.90%	8.29%
Maximum Drawdown (excluding costs)	-18.89%	-14.74%	-9.94%
Correct Directional Prediction	50.12%	52.65%	53.50%
Transaction costs	0.92	0.92%	1.01%
Annualized Return (including costs)	1.59%	3.98%	7.28

As observed in Table 3, SVM and Random Forests outperform all models in terms of annualized return and information ratio even when transaction costs are considered. Random Forests are the dominant model presenting the higher minimum drawdown and thus reducing the trading risk. Furthermore, as a classifier Random Forests demonstrated the highest correct directional prediction in the out of sample period and thus being the most accurate one.

V. CONCLUSIONS

In the present work, we applied a variety of machine learning techniques in the problem of modeling and trading with the EURO/USD exchange rate. From all the applied machine learning techniques, Random Forests has not been applied in this problem again while being one of the most accurate classifiers. The machine learning techniques were benchmarked with two traditional trading strategies: Naïve strategy and MACD strategy.

Except from the simple methods of KNN and Naïve Bayesian classifiers all other machine learning techniques outperformed the traditional strategies that are even until now used by economists. Thus, our empirical results encourage future research in applying machine learning techniques in trading with financial time series. From all the machine learning techniques applied in the present paper, Random Forests indicated the best trading performance in terms of annualized return and information ratio even when the transaction costs were considered.

As a future direction we propose the application of random forests and other machine learning techniques in trading with other financial time series in order to test their performance and establish them as reliable quantitative trading tools.

REFERENCES

- [1] C. Dunis, M. Williams, "Modeling and trading the Euro/Us Dollar exchange rate: do neural networks perform better?", *Derivatives Use, Trading and Regulation*, Vol. 8, No. 3, pp. 211-240, 2002.
- [2] C. Ullrich, D. Seese, S. Chalup, "Foreign exchange trading with support vector machines", *Advances in Data Analysis: Studies in Classification, Data Analysis and Knowledge Organization, Part VII*, pp. 539-546, 2007.
- [3] C. Dunis, J. Laws, G. Sermpinis, "Modelling and trading the EUR/USD exchange rate at the ECB fixing", *The European Journal of Finance*, Vol. 16, No. 6, pp. 641-561, 2010.

- [4] Thomson Reuters Datastream: <http://online.thomsonreuters.com/datastream/>.
- [5] J. Holland, *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*, Cambridge: Mass, MIT Press, 1995.
- [6] T. Cover, P. Hart, "Nearest neighbor pattern classification", *IEEE Trans. Inform. Theory*, Vol. 13, No. 1, pp. 21-27, 1967.
- [7] C. Howson, P. Urbach, *Scientific Reasoning: The Bayesian Approach*, Third Edition, Open Course Publishing Company, 1993.
- [8] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, 1998.
- [9] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer 2000.
- [10] L. Breiman, "Random Forests", *Machine Learning*, Vol. 45, No. 1, pp. 5-32 2001.
- [11] K. Manish, M. Thenmozhi, "Forecasting stock index movement: A comparison of support vector machines and random forest". In *Proceedings of ninth Indian institute of capital markets conference*, Mumbai, India, 2005.

AUTHORS PROFILE

Konstantinos Theofilatos is a PhD candidate in the Department of Computer Engineering and Informatics of the University of Patras, Greece. In 2009, he received a Master's degree from the Department of Computer Engineering and Informatics of the University of Patras. His research interests include computational and artificial intelligence, evolutionary computation, time series modeling and forecasting, bioinformatics, data mining and web technologies.

Spiros Likothanassis is an Electrical Engineer with a Ph.D. in Computer Engineering and Informatics. He currently is a professor in the Department of Computer Engineering and Informatics of the University of Patras. He is a reviewer in many scientific journals and conferences including *IEEE Neural Networks*, *Journal of Knowledge and Information Systems*, *Information Sciences Journal* and *Journal of Computational and Applied Mathematics*. He has a very good knowledge on Adaptive Signal Processing, Adaptive Control, Computational and Artificial Intelligence, Intelligent Agents and Bioinformatics. He has also experience in designing Information Systems, E-commerce Applications, Distance Learning Systems and Teleworking Applications

Andreas Karathanasopoulos is a senior Lecturer in London Metropolitan Business School. In 2008 he received the Master of Science in International Banking and Finance from the Department of Banking and Finance of Liverpool John Moore's University. His research interests include financial forecasting, trading strategies, time series prediction, artificial and computational intelligence and neural networks.