From: Proceedings of the Third International Conference on Multistrategy Learning. Copyright © 1996, AAAI (www.aaai.org). All rights reserved.

# Application of Multistrategy Learning in Finance

**M. Westphal and G. Nakhaeizadeh**

SGZ-Bank, Karlsruhe, Karl-Friedrich Str. 23, D-76133 Karlsruhe, Germany
martin@pnn.sgz-bank.com
and
Daimler-Benz, Research and Technology, Postfach 2369, D-89013 Ulm, Germany
nakhaeizadeh@dbag.ulm.daimlerbenz.com

**Abstract**

Although one can find in literature some contributions reporting on application of Multistrategy Learning (MSL) in different domains, there are only few studies dealing with application of MSL in financial fields. This paper gives an overview about the possibilities of the application of MSL in finance. Presenting some recent empirical results achieved by the authors, we discuss some advantages of application of MSL to financial domains and suggest further research topics.

## Introduction

The theoretical aspects of MSL are discussed in many publications, among them in Michalski and Tecuci (1991, 1993). There are also some works in the statistic community dealing with MSL (see for example Dasarathy, Sheela, 1979). The works of Wolpert (1992 a, 1992 b) and Henery (1996) are related to the context of MSL as well, although they use a horizontal combination of different approaches.

Furthermore, one can also find in literature some empirical evidence showing that the application of MSL can improve the results obtained by single approaches. In this connection, Esposito et al. (1993) discuss the application of MSL to document understanding, Sheppard (1993) applies MSL to classifying public health data and Hunter (1991) uses MSL to predicting protein structure.

Although in the literature there are several works applying the single-strategy learning in finance, only a few of them try to use the advantages of MSL. In this paper we present some of the works dealing with application of MSL in finance - which are certainly unknown to the MSL community - and based on interpretation of their results, we will make for some suggestions further research.

## MSL approaches in finance

### Prediction of exchange rates

The number of the works dealing with the prediction of exchange rates using a single forecasting approach is too high and we can not discuss all of them in this study. Interested readers can however find some of the recent works in Pesaran and Potter (1993), Refenes (1995) and Bol et al. (1996).

On the contrary, the number of the contributions that use MSL to forecasting of exchange rates is too small. In following we will discuss some of the recent works.

The central question addressed by Steurer (1996) is whether nonlinear methodologies can outperform econometric methods and the naive prediction, respectively. Therefore he conducts a performance comparison concerning the prediction of the monthly DM/US-Dollar exchange rate. Furthermore he examined the question whether a combination of the single prediction approaches can improve the performance. To develop his combined approaches he applies in a first step a cointegration study to find an adequate model for exchange rate prediction. This leads to a linear regression model that determines a long-run relationship between a set of nonstationary variables. Then the residues of this model, the error-correction term, together with other stationary and statistical important variables are used in a second step as explanatory variables for a linear regression forecasting models. Instead of the regression model of the second step, he uses also some machine learning approaches as alternatives, among them Neural Nets and M5-algorithm (Quinlan, 1992).

Steurer compares the results achieved by his MSL approach with the results of the random walk model and concludes that in several used versions the performance of the combined approach is superior to those of the random walk. The above MSL approach is used also in the work of Hann and Steurer (1996). In this study they emphasis on the short term prediction using weekly data. The authors conclude that their MSL approach should go a long way towards a more accurate prediction.

### Prediction of stock prices

There are many single-strategy approaches to predict stock prices. Graf and Nakhaeizadeh (1994) and Cao and Tsay (1993) are examples of such studies. The number of used MSL approaches in this domain is, however, too low as

From: Proceedings of the Third International Conference on Multistrategy Learning. Copyright © 1996, AAAI (www.aaai.org). All rights reserved.

well. A comprehensive study using MSL approach for forecasting stock prices can be found in Westphal and Nakhaeizadeh (WN) (1996). This work is motivated by Quinlan (1993) who suggested the combination of model-based and instance-based approaches. Furthermore they use various single approaches.

## Single approaches

Concerning the k-NN method WN discuss the different theoretical aspects, among them the various distance measures and the selection of the best value for *k*. Furthermore they discuss the runtime behavior of k-NN that is of interest in practical applications. In their empirical study WN implement a variation of Yuncks k-NN algorithm (Yunck, 1976). This new version avoids some disadvantages of the Yunck approach, especially concerning the selection of the start value and the choice of an appropriate step-size for the iteration.

In dealing with single approaches, the performance of classical k-NN methods is compared with those of IBL's. Further WN discuss some theoretical shortcomings of IBL approaches. Using an empirical example they show that the weighting strategy used by Aha, et al. (1991) might lead to wrong results. They discuss also that there is no significant difference between IB1 and a k-NN for $k$=1, because of this reason one can not claim that IB1 is a new algorithm.

Furthermore, IB2-algorithm belongs to the class of edited k-NN methods, discussed in detail in Dasarathy (1991). WN argue also, that in contrast to IB2 the algorithm of Chang (1974) leads to a lower number of prototypes and as a result to a better performance. It is also independent of the order of the used training data. It means that IB2 can lead to a case-base that is totally different although the same population is trained.

WN criticizes that in contrast to Wilson (1972) the power of IB3 in noise-reduction is unknown. Furthermore IB3 does not offer the possibility to classify correctly the cases located near the concepts-boundaries. This because IB3 eliminates the corresponding prototypes due to amount of points classified with these prototypes.

According to WN the different versions of IBL-algorithms only IB4 and IB5 can be regarded as new algorithms.

Another single approach used by WN is M5 algorithm. The most advantages of M5 is the usage of linear regression models in the nodes of the tree rather than the average values as for example in CART and NEWID.

The applied Neural Nets used by WN is a multi-layer perceptron net with a topology of 43 inputs, 7 hidden and 1 output neuron. The used learning algorithm is standard Backpropagation with a learning rate of 0.05 updated after each learning epoch. To these 43 explanatory variables belong some fundamental indicators, like the Dow Jones Index, the Nikkei Index or US$-DM exchange rate, which come from an econometric model and some technical indicators, like relative strength index, momentum or Williams R%. For each of the input series a set of 4 lags was used.

## Combined methods

The work of Quinlan (1993) is extended by WN in different directions. On one hand they use a classical k-Nearest Neighbor (k-NN), which has a long tradition in Statistics, as additional instance-based method. On the other hand they use the average of the predictions of two or more different methods and apply the prediction made by one technique as additional input to an alternative prediction. For example, the prediction of the k-Nearest Neighbors can be considered as an additional attribute to train a Neural net. Another extension made by WN is the application of other versions of IBL-algorithms due to Aha, et al. (1991) and Aha (1992).
The IBL-Algorithms used in the empirical study are made available by David Aha and M5 by Ross Quinlan.

# Empirical results

WN apply learning algorithms to 4 different forecasting tasks, the prediction of the change of the DAX (German Stock Index) one day ahead, the change in 3 and 5 days and the change during 5 days. The results were compared to a benchmark, the naive prediction that predicts that the value will be the same as last period.
To examine the performance of the single and combined approaches the future development of DAX is predicted. Besides the technical indicators the following explanatory variables are used as input:

- the Dow Jones index
- the Nikkei index
- the exchange rate between US $ and DM,
- the German floating rate

The target variable to be predicted is the daily changes of the DAX. The period from 01.01.1990 till 01.01.1993 was taken as the training set. The remaining data of the year 1993 formed the test set. All the time series are converted to logarithmic differences of two following days.
The correct forecast of the direction of a change in price is more important. Technical indicators were used as additional input. This has two advantages: First, not so many real time series are necessary for the test, because the technical indicators are estimated from the original series. Second, the forecasting method gains access to the tools common in the praxis of financial markets, even if these tools are not scientifically established. The "technical analysis", in contrast to the "fundamental analysis", does not attempt to explain the level or the development of a series. It tries instead to identify early symptoms for changes in the market. Table 1 gives an overview about the economic variables used in the study

From: Proceedings of the Third International Conference on Multistrategy Learning. Copyright © 1996, AAAI (www.aaai.org). All rights reserved.

of Westphal and Nakhaeizadeh (1996). The variable "Umlaufrendite" means long term average interest rate, RSI 5 and RSI 14 mean Relative Strength Index estimated for 5 and 14 days, respectively. OBOS stands for Over Bought Over Sold or "Williams R%".

Table 1: Time series used for prediction

| fundamental | technical indicators |
|---|---|
| DOW | RSI 5 |
| US $ / DM | RSI 14 |
| Nikkei | Moving Average 12-26 days |
| Umlaufrendite | Momentum 1 |
| DAX | Momentum 5 |
| | Momentum 10 |
| | OBOS 5 |
| | OBOS 10 |
| | OBOS 15 |
| | OBOS 20 |
| | Deviation for 2 days |
| | Deviation for 3 days |
| | Deviation for 4 days |

Besides Mean Square Error (MSE), WN use accuracy rate as evaluation measure where the prediction was reduced to the statements the DAX will "rise" or "fall".

## Some of the prediction results of single methods

WN observe that the k-NN algorithm yield a nearly constant and the achieved accuracy-rate remains independent from the forecasting horizon. Using the single approaches the best performance was obtained by the more simple versions of the IBL-algorithms, IB1 and IB2. These results were unexpected, because at least from the theoretical point of view IB4 and IB5 should lead to better results. An explanation for these unexpected results could be that the reduction of cases to prototypes by IB3-IB5 was connected with lost of information existing in the eliminated cases. Table 2 to 4 show the results for the DAX-prediction of the next day, 3rd day and 5th day, respectively. The best performance of the k-Nearest Neighbor method was achieved with k=5. The technical indicators had mostly a high significance. Especially including of the deviations was advantageous.

Table 2: Results for the DAX-prediction of the next day:

| | Neural Net | M5 | IBL1 | k-NN | Naive Prediction |
|---|---|---|---|---|---|
| MSE | 0.0106 | 0.0074 | * | 0.0090 | 0.0111 |
| Accuracy-Rate | 66.80 % | 69.53 % | 54.50 % | 53.13 % | 51.95 % |

Table 3: The results achieved by the different methods for the DAX-prediction for the 3rd day:

| | Neural Net | M5 | IBL1 | k-NN | Naive Prediction |
|---|---|---|---|---|---|
| MSE | 0.0101 | 0.0087 | * | 0.0092 | 0.0111 |
| Accuracy-Rate | 57.03 % | 47.27 % | 45.80 % | 54.30 % | 51.56 % |

Table 4: The results for the 5th day prediction:

| | Neural Net | M5 | IBL1 | k-NN | Naive Prediction |
|---|---|---|---|---|---|
| MSE | 0.0103 | 0.0087 | * | 0.0090 | 0.0113 |
| Accuracy-Rate | 55.08 % | 46.09 % | 46.00 % | 52.73 % | 51.56 % |

* The IBL-Software does not allow to estimate this value.

The results of single methods are partially very good. The employed methods seem to be able to identify hidden structures in the data. An important difference between the methods is, that in contrast to the k-NN and IBL-algorithms the Neural Net and M5 are able to estimate values also outside the range of the target variable during the training process.

Using M5, the predictions were, however, not so volatile as the ones of the Neural Net. Furthermore, using the default version of M5 showed a tendency to predict a constant value for whole time. Dealing with the used Neural Net, it becomes obvious that the choice of the net input is of greater importance than the optimal architecture and complexity of the net. The pre-choice of the input time series highly determines the performance of the prediction.

Concerning k-NN, the parameters of the applied algorithms had to be adjusted individually for each task to achieve the best performance. This was not the case for the IBL-algorithms. The versions IB1 and IB2 mostly outperformed their further developed versions namely IB3 and IB4. In theory, IB3 should be better because of the filtering of noisy data and IB4 should have been the best because it learns not only the best prototypes but also the most important attributes for the task, confirmed by Aha, (1991) and (1992). The results of the IBL-algorithms were generally inferior compared to the results of the other methods.

## Selected prediction results of MSL-approaches

Some of the results achieved by WN using the combined approaches are very interesting. Specially using the prediction of one method as additional input attribute for the Neural Net or M5. The results of the combination with the classical k-NN rule are presented in table 5.

From: Proceedings of the Third International Conference on Multistrategy Learning. Copyright © 1996, AAAI (www.aaai.org). All rights reserved.

Table 5: Neural Net and M5 get the prediction of the k-NN rule as additional input:

| Sort of Prediction | NN with input from k-NN Accuracy | MSE | T* | M5 with input from k-NN Accuracy | MSE | T |
|---|---|---|---|---|---|---|
| 1 Day | 65.23% | 0.0088 | 0.7965 | 64.06% | 0.0076 | 0.6829 |
| 3. Day | 55.47% | 0.0107 | 0.9686 | 46.09% | 0.0089 | 0.7879 |
| 5. Day | 59.38% | 0.0090 | 0.7987 | 45.70% | 0.0089 | 0.7918 |
| in 5 Days | 66.80% | 0.0205 | 1.7012 | 82.03% | 0.0121 | 0.9902 |

*With T is Theil's coefficient: $T = \sqrt{\frac{\sum_t (x_t - \bar{x}_t)^2}{\sum_t (x_t - x_{t-1})^2}}$ with the prediction $\bar{x}_t$

The accuracy-rates of the values at the 5th day and in 5 days are the best of all. The results of the MSE are also very good. The values of T are also very well. The prediction of the level of the DAX in 5 days by the combination with the Neural Net is the only surpassed by the "naive prediction". This case is also the only one where the combination with M5 performs better than the one with the Neural Net. Examining of the significance of the inputs of the Neural Nets shows that the prediction of the k-NN method is an important input attribute for the Neural Net model. Furthermore the past values of the DAX and the Nikkei were important as well.

Another alternative method of combination used by WN is using the average of two predictions from different methods. This average is then the new prediction. Although this is a very simple combination method its results are very interesting because nearly all the combinations have an MSE lower than that of the single methods (See Westphal and Nakhaeizadeh 1996 for more details).

## Conclusions and final remarks

Although application of the MSL in finance is a rather new applied research field, the results achieved up to now are very promising. Generally, prediction of the development of financial time series is not an easy task, because the structure of many of financial time series does not differ significantly from a random walk process. An additional problem is changing the structure of data over the time. The motivation behind the application of MSL to finance has been to use the advantages of the alternative approaches to improve the forecasting results. It means, the alternative approaches are considered not as competitors but they are used in a cooperatively manner.

The empirical results reported in these works encourage further research in this area. Specially the application of the methodology used in Westphal and Nakhaeizadeh (1996) could be interesting for forecasting of the other financial time series, for example, exchange and interest rates. The predictability of the exchange rates is still an open and controversy problem. Most of the applied approaches to solve this problem are based on the statistical methods. It would be quite interesting to examine weather combination of such statistical approaches with AI-based methods within a MSL concept can contribute significantly to this controversy debate.

Another open problem is the adaptation of forecasting approaches as the structure of data changes over the time. Although there are many single-methods which examine the so called „structural change" in time series, there is no comprehensive work examining the practicability of MSL-approach to this domain. Further research in this connection would be quite interesting as well.

**Acknowledgments**: The authors would like to thank David Aha and Ross Quinlan for providing IBL- and M5-Software.

## References

Aha, D. W., 1992. Tolerating noisy, irrelevant and novel attributes in instance-based learning algorithms. International Journal of Man-Machine Studies, 267-287, Vol 36.

Aha, D. W., Kibler, D., Albert, M. K. 1991. Instance-Based Learning Algorithms. Machine Learning, 37-666,. Kluwer Academic Publishers, Boston.

Bol, G. Nakhaeizadeh, G. and Vollmer, K. H. 1996. (Eds.). Finanzmarktanalyse und -Prognose mit innovativen quantitativen Verfahren. Physica-Verlag, Berlin.

Cao, C. Q. and Tsay, R. S. 1993. Nonlinear Time-Series Analysis of Stock Volatilities. In: Nonlinear Dynamics Chaos and Econometric, 157-178, edited by Pesaran, M. H. and Potter, S. M. Wiley.

Chang, C. L., 1974. Finding Prototypes for Nearest Neighbor Classifiers. IEEE Transactions on Computers, Volume C-23, No. 11, 1179-1184. Reprinted in Dasarathy, 1991.

Dasarathy, B. V., (Ed.), 1991. NN Pattern Classification Techniques. IEEE Computer Society Press, Los Alamitos.

Dasarathy, B. and Sheela, B. (1979). A. Composite Classifier System Design: Concept and Methodology. Proceedings of the IEEE, 708-713, Volume 67, Number 5.

Esposito, F., Malerba, D. Semeraro, G. and Pazzani, M. 1993. A machine Learning Approach to Document Understanding. Proceedings of the second International Workshop on Multistrategy Learning, 276-292. Center of Artificial Intelligence, George Mason University

Graf, J., and Nakhaeizadeh, G. 1994. Application of Learning Algorithms to Predicting Stock Prices. In: Frontier Decision Support Concepts, 241-260, edited by Plantamura, V. Soucek, B. and Visaggio, Wiley.

Hann, T. and Steurer, E. 1996. Much ado about nothing ? Exchange rate forecasting: Neural Networks vs. linear models using monthly and weekly data, 1-17. Neurocomputing.

Henery, R. 1996. Combination Forecasting Procedures. Forthcoming.

Hunter, L. 1993, Classifying for Prediction: A multistrategy Approach to Predicting Protein Structure, Proceedings of the second International Workshop on Multistrategy Learning, 394-402. Center of Artificial Intelligence, George Mason University.

Michalski, R. and Tecuci, G. Eds., 1993. Proceedings of the second International Workshop on Multistrategy Learning, Center of Artificial Intelligence, George Mason University.

Michalski, R. and Tecuci, G. Eds. 1991. Proceedings of the first International Workshop on Multistrategy Learning, Center of Artificial Intelligence, George Mason University.

Pesaran, M. H. and Potter, S. M. Eds. 1993. Nonlinear Dynamics Chaos and Econometrics, Wiley

Quinlan, J. R., 1993. Combining Instance-Based and Model-Based Learning. Proceedings of the International Conference of Machine Learning, 236-243, Morgan Kaufmann.

Quinlan, J. R., 1992. Learning with Continuos Classes. Proceedings of the 5th Australian Joint Conference on Artificial Intelligence, 343-348. Singapore: World Scientific.

Refenes, A. (Ed.), 1995. Neural Networks in the Capital Markets, Wiley, New York.

Steurer, E. 1996. Ökonometrische Methoden und maschinelle Lernverfahren zur Wechselkursprognose: Theoretische Analyse und empirischer Vergleich. Ph. D. diss. University of Karlsruhe, Germany.

Sheppard, J. 1993. Applying Multiple Learning Strategies for Classifying Public Health Data. In: Proceedings of the second International Workshop on Multistrategy Learning, 309-323. Center of Artificial Intelligence, George Mason University

Westphal, M. and Nakhaeizadeh, G. 1996. Combination of Statistical and other learning methods to predict financial time series. Discussion Paper, Daimler-Benz Research Center Ulm, Germany

Wilson, D. L., 1972. Asymptotic Properties of Nearest Neighbor Rules Using Edited Data. IEEE Transactions on Systems, Man and Cybernetics, 408-421, Volume SMC-2, No 3. Reprinted in Dasarathy, 1991.

Wolpert, D. 1992a. Stacked Generalization. Neural Network, 241-259, Vol. 5.

Wolpert, D. 1992b. Horizontal Generalization. Working Paper, Santa Fe Institute, 92-07-033.

Yunck, T. P. 1976. A Technique to Identify Nearest Neighbors. IEEE Transactions on Systems, Man and Cybernetics, 678-683, Volume SMC-6, No 10. Reprinted in Dasarathy, 1991.