

Different Stock Market Models Using Support Vector Machines

Rafael Rosillo, Javier Giner, Javier Puente, and Borja Ponte

Abstract—The goal of this research is to analyse the different results that can be achieved using Support Vector Machines to forecast the weekly change movement of the different simulated markets. The data cover 3000 daily close for each simulated market. The main characteristic of these markets are: high volatility, bearish movement, bullish movement and low volatility. The inputs of the SVM are the Relative Strength Index (RSI) and the Moving Average Convergence Divergence (MACD). SVM-KM is used by Matlab in order to design the algorithm. The outputs of the SVM are the degree of set membership and the market movement (bullish or bearish). The configuration for the SVM shows that results are better in high volatility markets or low volatility markets than trend markets.

Index Terms—Support vector machines, quantitative trading, stock market models, technical analysis.

I. INTRODUCTION

Different market situations such as high volatility, low volatility, bullish movements and bearish movements are shown in this paper. The SVM helps to investor in the quantitative decision making choosing a weekly forecast (bullish or bearish). We analyse in which market situation the SVM can achieve the best results.

The rest of the paper is structured as follows. Section II, the literature review of SVM is presented. Section III explains the design of the trading rule. The results are shown in section IV. Finally section V provides some concluding remarks.

II. LITERATURE REVIEW OF SVM

A resume of the state of the art that is presented in [1] is described below.

SVMs were originally developed by [2]. For a detailed introduction to the subject, [3] and [4] are recommended.

The biggest difference between SVMs and other traditional methods of learning is that SVMs do not focus on an optimisation protocol that makes few errors like other techniques. Traditionally, most learning algorithms have

focused on minimising errors generated by the model. They are based on what is called the principle of Empirical Risk Minimization (ERM). The goal of SVM is different. It does not seek to reduce the empirical risk of making just a few mistakes, but pretends to build reliable models. This principle is called Structural Risk Minimization. The SVM searches a structural model that has little risk of making mistakes with future data.

The main idea of SVMs is to construct a hyperplane as the decision surface so that the margin of separation between positive and negative examples is maximised [5]; it is called the Optimum Separation Hyperplane (OSH), as shown in Fig. 1.

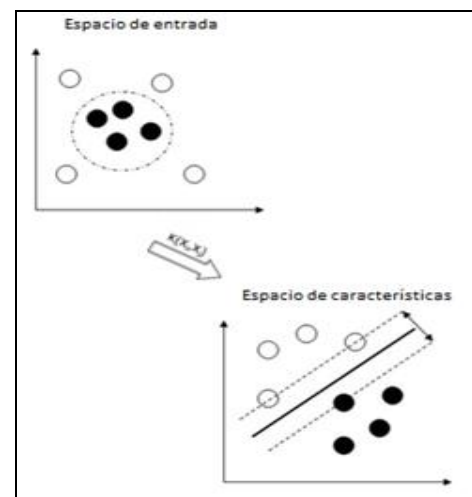


Fig. 1. An example of how a kernel function works.

The SVMs can be used in two different ways: classification or regression.

At the beginning of twentieth century, investors start to use SVMs in stock markets. The most relevant researches are shown below.

In 2001, [6] compare SVMs with other techniques such as back-propagation neural networks. Two applications for financial series prediction with SVMs were developed in 2003 [7], SVMs are applied to the problem of forecasting several futures contracts from the Chicago Mercantile Market showing the superiority of SVMs over back-propagation and regularised Radial Basis Function Neural Networks; in [8], SVMs are used to predict the direction of change in the daily Korean composite stock index and they are benchmarked against back-propagation neural networks and Case Base Reasoning. The experimental results show that SVMs outperform the other methods and that they should be considered as a promising methodology for financial time-series forecasting. In [9], a SVM Classifier is used to predict the directional movement of the Nikkei225 index

Manuscript received June 22, 2013; revised August 19, 2013. Financial support given by the Government of the Principality of Asturias is gratefully acknowledged.

Rafael Rosillo, Javier Puente, and Borja Ponte are with the Polytechnic School of Engineering, University of Oviedo, Campus de Viesques s/n, CP 33204, Gijón, Asturias, Spain (e-mail: rosillo@uniovi.es, jpuente@uniovi.es, uo183377@uniovi.es).

Javier Giner is with the Faculty of Economics and Business, University of La Laguna, Campus de Guajara s/n, CP 38071, La Laguna, Islas Canarias, Spain (e-mail: jginer@ull.es).

with extremely promising results. In this study, we apply SVM in the classification way.

III. TRADING RULE

The design of the trading rule is presented in this section.

A. Simulated Markets

The simulated markets are 4 in this experiment. An especially characteristic has been chosen in order to differentiate each series. We want to analyse how SVM help to investors in bullish trend (Fig. 2), bearish trend (Fig. 3), high volatility (Fig. 4) and low volatility (Fig. 5).

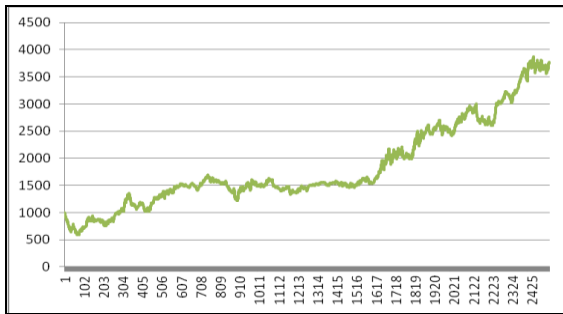


Fig. 2. Bullish trend.

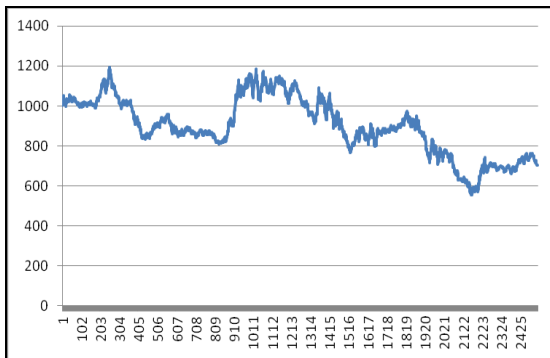


Fig. 3. Bearish trend.

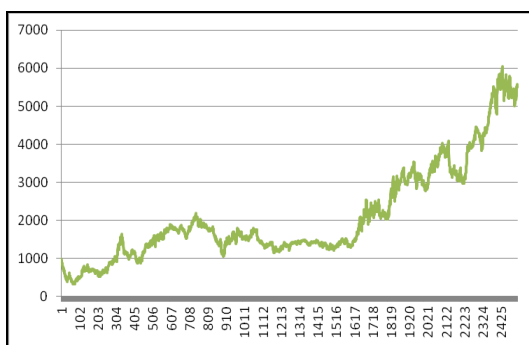


Fig. 4. High volatility.

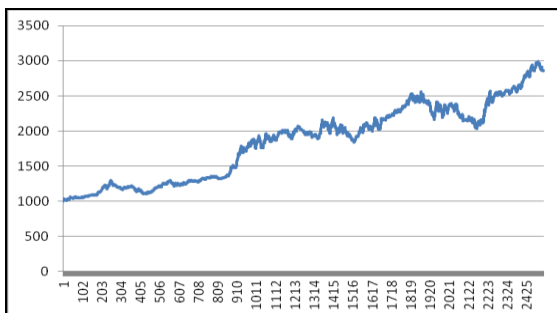


Fig. 5. Low volatility.

B. Inputs

The inputs of the SVM are the quantitative analysis indicators RSI and MACD. In [10] it is explained that RSI gets good profits in blue chips and Momentum indicator gets good profits in Small Caps, MACD and Stochastic indicators have been analysed over the Spanish Continuous Market too.

- Relative Strength Index (RSI)

It was designed by [11]. A brief explanation of this indicator is shown below in equation 1. If more details are needed it can be seen in [11].

The RSI is an oscillator that shows the strength or speed of the asset price by means of the comparison of the individual upward or downward movements of the consecutive closing prices.

For each day, an upward change (U) or downward change (D) is calculated. “Up days” are characterised by the daily close S_t being higher than the close of previous day S_{t-1} ($U_t = S_t - S_{t-1}$, $D_t = 0$).

“Down days” are characterised by the daily close being lower than the close of the previous day ($U_t = 0$, $D_t = S_{t-1} - S_t$).

The average U_t and D_t are calculated using an n -period exponential moving average (EMA_n).

Relative Strength Index at time t (RSI_t) is the following ratio between 0 and 100:

$$RSI_t = 100 \frac{EMA_n^U}{EMA_n^U + EMA_n^D} \quad (1)$$

where RSI_t is the Relative Strength Index at time t .

The 14-day RSI, a popular length of time utilized by traders, is also applied in this study. The RSI ranges from 0 to 100 however the range has been normalized between -1 and +1 in order to place it in the SVM.

- Moving Average Convergence Divergence (MACD)

The MACD is designed mainly to identify trend changes. It is constructed based on moving averages and is calculated by subtracting a longer exponential moving average (EMA) from a shorter EMA. The MACD is shown in equation 2:

$$MACD(n) = EMA_k(i) - EMA_d(i) \quad (2)$$

where:

$$EMA_n(i) = \alpha \times S_i + (1 - \alpha) \times EMA_n(i - 1)$$

$$\alpha = \frac{2}{1 + n}$$

being n the number of days in the exponential average, and S_i is the asset price on it day.

In this article, $k=12$ and $d=26$ day EMA's are selected, which are commonly used time spans in order to calculate MACD [12].

The range of MACD has been normalised between -1 and +1 in order to use it in the SVM.

C. Design of Trading Rule

The inputs of the SVM are RSI and MACD. The outputs of the SVM are the movement of the market and the degree of set membership.

The trading rule is generated by the SVM Classifier

(SVMC). We explain the design of the trading rule below.

First, the SVM analyses the inputs classified in buy situations or sell situations.

Second, the SVM tries to separate the different prices of the simulated markets in two classes: buying and selling situation, with the inputs mentioned earlier.

Third, the SVM uses the kernel function Heavy Tailed Radial Basis Function and the C parameter value is 10 in order to make the forecasting.

Fourth, the hit ratio is calculated for the different testing periods.

Finally, given a value of the RSI, MACD, the SVMC predicts the upward or downward movement for the following week and the intensity of that movement.

D. The Outputs of the SVM

The outputs of the SVM are the market movement (bullish or bearish), and the degree of set membership.

IV. RESULTS

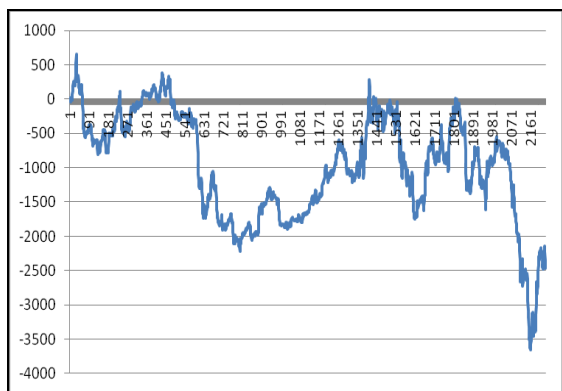


Fig. 6. Bullish trend results.

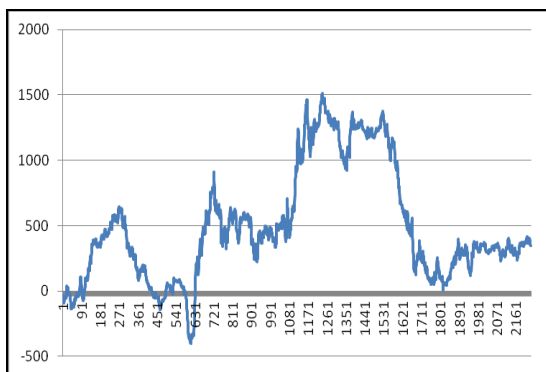


Fig. 7. Bearish trend results.

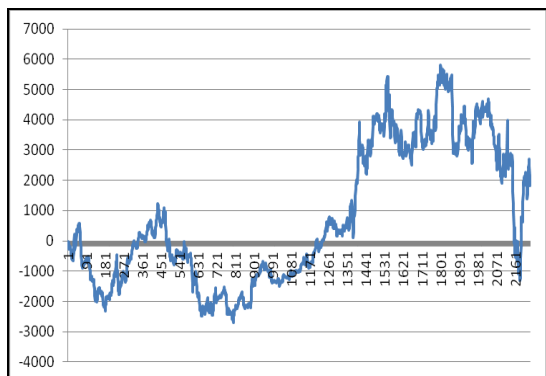


Fig. 8. High volatility results.

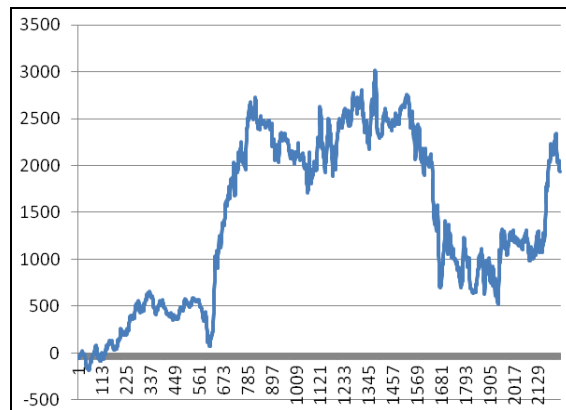


Fig. 9. Low volatility results.

We present the four charts (see Fig. 6, Fig. 7, Fig. 8, Fig. 9) which are created by the SVM results. The charts show the achieved points by the trading rule in each simulated market.

V. CONCLUSIONS

As it is shown in results section, this trading strategy or parameters SVM configuration achieve better results in high volatility markets or low volatility markets than trend markets.

REFERENCES

- [1] C. L. Dunis, R. Rosillo, D. D. L. Fuente, and R. Pino, "Forecasting IBEX-35 moves using support vector machines," in *Neural Computing and Applications*, vol. 21, pp. 3-10, 2012.
- [2] V. N. Vapnik, *Statistical Learning Theory*, New York: Wiley, 1998.
- [3] C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, pp. 121-167, 1998.
- [4] T. Evgeniou, M. Pontil, and T. Poggio, "Regularization networks and support vector machines," *Advances in Computational Mathematics*, vol. 13, pp. 1-50, 2000.
- [5] X. Xu, C. Zhou, and Z. Wang, "Credit scoring algorithm based on link analysis ranking with support vector machine," *Expert Systems with Applications*, vol. 36, pp. 2625-2632, 2009.
- [6] L. Cao and F. Tay, "Financial forecasting using support vector machines," *Neural Computing and Applications*, vol. 10, pp. 184-192, 2001.
- [7] L. Cao and F. Tay, "Support vector machine with adaptive parameters in financial time series forecasting," *IEEE Trans. on Neural Networks*, vol. 14, pp. 1506-1518, 2003.
- [8] K. Kim, "Financial time series forecasting using support vector machines," *Neurocomputing*, vol. 55, pp. 307-319, 2003.
- [9] W. Huang, Y. Nakamori, and S. Y. Wang, "Forecasting stock market movement direction with support vector machine," *Computers & Operations Research*, vol. 32, pp. 2513-2522, 2005.
- [10] R. Rosillo, D. D. L. Fuente and J. A. L. Brugos, "Technical analysis and the Spanish stock exchange: testing the RSI, MACD, momentum and stochastic rules using Spanish market companies," *Applied Economics*, vol. 45, pp. 1541-1550, 2013.
- [11] J. W. J. Wilder, *New Concepts in Technical Trading Systems, Trend Research*, Greensboro, NC: Hunter Publishing Company, 1978.
- [12] J. J. Murphy, *Technical Analysis of the Financial Markets*, Institute of Finance, New York, 1999.



Rafael Rosillo is a lecturer in Business Management at University of Oviedo since 2011. He got the Ph.D. in Business Management. He has also a MBA and a degree in Computer Engineering. His researches are focused on artificial intelligence, stock markets, forecasting and supply chain. He has collaborated like visiting researcher in John Moores University (Liverpool) and University College of Dublin. He is author of several publications in different national and international journals and chapters of books. He has participated in several international scientific

conferences and several research projects.



Javier Giner is a Ph.D. in Physics. He is also a senior lecturer and researcher of the Financial Economy and Accounting Department, University of La Laguna, Tenerife, Spain since 2004. His research is focused in option models valuation and financial simulation process. Several national and international publications. Participation in 10 research projects. He is a visiting lecturer at the University of Salerno, Naples on seminars and courses on finance, option models, financial management and quantitative finance.

journals and has taken part of the scientific committee in some international conferences.



Borja Ponte is a Ph.D. student at the Department of Business Administration of the University of Oviedo. His master's Thesis is "The Bullwhip Effect in Supply Chains: An Approach based on Artificial Intelligence Techniques", qualified with honors, represents its introduction into the world of research, trying to combine the fields of Logistics and Multiagent Systems. He has presented papers at two international conferences, and he has participated in the work presented in two others.



Javier Puente is an associate professor of operations management at the University of Oviedo, Spain. His key research topics include: applied artificial intelligence, production planning and control, quality of service and supply chain management. He has published a great number of research papers in a number of leading journals such as: International Journal of Production Economics, Applied Artificial Intelligence, Computers & Industrial Engineering, Applied Soft Computing or Artificial Intelligence in Medicine among others. He also has published in several international conferences: IC-AI, GECCO, POMS, EUSFLAT. He has been usual referee of several international