Sarah MARCELINO
E-mail: sarahsabino17@gmail.com
Pedro Alexandre HENRIQUE
E-mail: pedroalexandre.df@gmail.com
Pedro Henrique Melo ALBUQUERQUE
E-mail: pedroa@unb.br
Faculty of Economics, Business and Accounting (FACE)
University of Brasilia, Brazil

PORTFOLIO SELECTION WITH SUPPORT VECTOR MACHINES IN LOW ECONOMIC PERSPECTIVES IN EMERGING MARKETS

Abstract. *Some of the innovations are presented here are the use of the ranking probability forclassification of assets and the balance of the cost by the type of forecast error.The study was conducted on two different samples. The first sample consists on assets that are part of the Ibovespa, considering the portfolio valid from January to April 2015 and this sample was compared to the portfolio generated by SVM and the BOVA11 index fund. The second sample consists of the assets comprising the IBrX-100 index, the portfolio valid from January to April 2015 which similarly was compared with the portfolio generated by SVM and the BRAX11 index fund. In order to evaluate the proposed strategy results were also compared to the free return indicator CDI risk.The portfolio cumulative return of the sample selected by SVM was 94.15%, compared to -14.42% of BOVA11 that replicates the Ibovespa. While the portfolio selected on the portfolio of IBrX100 was 38.25% against 13.86% of BRAX11 index fund. For the period rated the CDI presented a return of 53.75%. The average cumulative return of assets in the study period was in the case of portfolio of the Ibovespa 57.1% and 34.4% for IBrX100 portfolio.*

Key words: *Support Vector Machines, Pattern Recognition, Classification Methods, Portfolio Choice, Investment Decisions, Financial Risk.*
Jel Classification:C10, C38, C45, C58 C63, G11

INTRODUCTION

Stock selection is a challenging and crucial part of investor decision-making. Considering the huge amount of available assets in the financial market, according to Fan and Palaniswami (2001), the challenge of selecting stocks is to identify assets with potential to outperform the market next year.

According toTay and Cao (2001), 'predicting financial time series is considered one of the most challenging applications of time series forecasting.' Abu-Mostafa and Atiya (1996) point out that speculators, investors and companies

in their quest to predict market behavior, assume that future events are based at least in part on events and present and past data. However, financial series are permeated by noise, non-stationary and deterministic chaos.

This context leads many economists to adopt the efficient market hypothesis, which considers the changes in stock prices are independent of the past and follow a random pattern (Abu-Mostafa and Atiya 1996). According to Fama (1970), price changes would then be unpredictable and any change in price would represent an immediate reaction to a new event or an unexpected change in supply or demand. If there were any unexpected opportunity to profit, for example, investors would explore it immediately so that the price would back to the level it was when this opportunity did not exist. Also according to this theory, any relevant standards should reflect the current price, but if the market is efficient to the point of all stock prices fully reflect all available public information, we can not expect that this analysis can identify in advance investments with higher returns to the market. While there are several discussions on the efficient market hypothesis, it is difficult to refute it or not (Abu-Mostafa and Atiya1996).

In data mining perspective, future stock returns are considered to some extent predictable. According to Fan and Palaniswami(2001) the prediction problem involves the discovery of patterns of relationships in data and useful application of this information to classify actions. One approach that has shown promise for this problem are the Support Vector Machines (Support Vector Machine - SVM) proposed by Boser*et al*in 1992. Originally, the Support Vector Machines were developed to recognize patterns in a data set. According to Albuquerque (2014), through this recognition is possible to complete an inductive inference process, which would be able to make predictions for a data set later observed the estimation of the model parameters.

Since its creation in 1992, the Support Vector Machines marked the beginning of a new era in artificial intelligence, representing a decrease of thought in Statistical Learning Theory (Soman K.*et al* 2011). With the implementation of the principle of the Structural Risk Minimization, which minimizes the upper bound of the generalization error instead of just minimizing the empirical error, the SVM opens a new perspective for modeling machine learning algorithms with higher generalization hability, surpassing most of the difficulties faced by traditional algorithms such as overfitting and high-dimensional data.

This study aimed to verify whether the use of Support Vector Machines contributes to the portfolio return exceeds industry benchmarks in times of low economic perspective. First, the return of the portfolio selected by SVM was compared to the profitability of ETFs (Exchange Traded Funds) BOVA11 and BRAX11. These ETFs (Exchange Traded Funds) are investment funds that aim to achieve a rate of return similar to the performance of the Brazil Index (IBrX-100) and the Bovespa index, respectively. Subsequently, a comparison with two equally weighted portfolios was made, the first consisting of all actions that make up the Ibovespa index and the second, for all that makes up the Brazil Index. Finally, the portfolio return was also compared with the risk free asset CDI (Interbank Deposit

Certificate). The CDI quantifies the cost of money for banks on a particular day and your average daily rate is often used as a parameter to assess the profitability of investment funds in Brazil, since it is the basis on which most fixed income securities is calculated.

THEORETICAL FRAMEWORK

The first direct application of Support Vector Machines in finance refers to the model of application for stock classification and training portfolio approach proposed by Fan and Palaniswami in 2001. The usefulness of SVM was tested with accounting information of 37 financial figures shares traded on the Australian Stock Exchange for the period from 1992 to 2000. The results were compared with a benchmark model that was determined by the authors as a portfolio of investments equally weighted composed of all the available stocks for classification.

The expected return of the shares was defined as the binary dependent variable, and may take two values: $+1$ representing shares with exceptional return and -1that is considered normal shares. Thus, the stocks that were between the third and fourth empirical quantile of the returns distribution of the Australian Stock Exchange companyswere classified as belonging to Class 1, the best actions. Those who had return between the first and third empirical quantile, were the Class 2, the worst stock class. When the SVM was used to select 25% of the shares of each year, the equally weighted portfolio achieved a total return of 208% over a period of five years, outperforming the benchmark performance that generated a return of 71%. Therefore, the SVM proved to be very useful for selecting actions and this result is confirmed also by other studies.

Recently, Huerta *et al*(2013) developed a similar study of the Fan and Palaniswami (2001) and pointed out that they have chosed SVM to identify stocks with high or low expected return due to its simplicity and effectiveness. Two differential of their approach are the fact that the SVM has been applied monthly to adjust to market changes and the selection of data that was used to train the SVM. They didin't use all available data, but a set of data present in the highest and lowest quantile of the historical distribution, also called tail data. According to the authors, the percentage of shares within the chosen quantile is enough for the SVM learn the correlations between the characteristics of the action and the class to which it belongs. It was given a quantile of 20%, then 20% of the highest return of shares and 20% of the lower stock returns were chosen. According to this approach, 40% of these data are sufficient to train the model and the omission of actions that are in the middle of the distribution leverages the performance, making it possible to train the classifier faster. The collected data are included in the 1981-2010 period, with the removal of a common database CRSP / Compustat. The assets in the sample belonged to different sectors of the economy and how each sector has unique characteristics, the authors built a model for each of them.

The portfolios were formed with the classification of outputs of the SVM and the study reached an annual return of 15% with close to 8% volatility for the portfolio formed.

The study of Emir*et al* (2012) aimed to build a great financial model that would allow the classification of the best stocks of the Turkish market. For this purpose, each year, the shares that had the 10 highest returns were classified as '1' and the other, classified as '0'. According to the authors, the data dimensionality reduction application before processing these for classification improves the final result.

Data were collected for each share they composed the Istanbul Stock Exchange Index (ISE) in the 2002-2010 period and this study was innovative in using both technical and fundamentalists parameters for analysis. Technical data was 13 Index of Istanbul Stock Exchange indicators (ISE) and fundamental analysis was performed using 14 indicators considered essential to represent the companies in the ISE as a whole, despite belonging to different sectors.For comparison purposes, a Neural Network model was applied in the same circumstances and the results showed that the Support Vector Machines were superior in the accuracy of the forecast. Therefore, the empirical results of the study of Emir*et al* (2012) also corroborate the success of SVM as a model to forecast financial time series.

In the same context of approaches for building portfolios, Gupta*et al* (2012) developed a hybrid approach to facilitate investors decision-making. First, using the Support Vector Machines to classify stocks into three predefined classes according to their performance on three financial indicators: liquidity, return and risk.The database consisted of 150 assets listed in the National Stock Exchange (NSE), the main market of financial assets in India. The training set was composed of 60% of all data and the test set for 40%. The second step was the application of a genetic algorithm, more specifically, Real Coded Genetic Algorithm (RCGA) in each of the three classes for formation of optimal portfolios. The portfolio formed from the Class 1 shares showed higher liquidity, but an average risk level. The portfolio formed from the Class 2 had a higher level of return and higher level of risk. Already the Class 3 portfolio had the lowest level of risk compared to other portfolios, and as expected, an average return level. Thus, the authors conclude that investors looking for higher liquidity should invest in Class 1. Those investors looking for higher returns should opt for Class 2 and those looking for safer investments should invest in assets of Class 3. These results indicate that the developed approach is able to classify assets with good accuracy and even more, can generate optimized portfolios for each asset class according to consumer preferences.

The Support Vector Machines can also be applied in the prediction of market direction. Kim (2003) analyzed the applicability of Support Vector Machines in predicting the direction of the daily changes in stock prices compared to two models: BPN (Back-Propagation Neural Network) and CBR (Case-based reasoning). The data used were the daily observations of share prices that make up

the Korean Market Index (KOSPI) and 12 technical indicators for the period January 1989 to December 1998.

The author classified the daily changes of prices in two classes: '0' or'1'. The first class consisted of stock price of the day after which was lower than the previous day. The second class was composed of shares whose index on the following day was higher compared to the previous day. 80% of the data were used for training and estimation of parameters and the remaining 20% were used for model validation.The empirical results show that in the validation set, the SVM achieved a performance in predicting 57.83% against 54.73% and 51.97% of the BPN and CBR models, respectively. It is evident then that the Support Vector Machines exceeded the two models in the prediction accuracy level and this can be attributed to the fact that the SVM implements the principle of the Structural Risk Minimization, allowing better generalization. This study concluded that the SVM is a promising alternative in forecasting financial time series.

Zhang and Zhao (2009) applied the SVM in the foreign exchange market to predict changes in the euro / dollar exchange rates. In this study, the inputs to the model were technical indicators, with data from the Bloomberg system in the range of 10 July 2007 to 9 July 2009.According to the authors, analysis of technical indicators in the exchange market can be described as a general learning problem. First we must recognize that the technical analysis is valid, that is, if the technical indicators and the trend of the exchange rate have some intrinsic connection. If there is such a relationship, the key to the problem is to find a function that minimizes the Expected Risk and that is applicable in a large sample. In this context, the inputs are the technical indicators and outputs that indicate the change in the futures price, derived from the relationship between the indicators and the trend of the exchange rate. However, the choice of indicators is not an easy task.

As in most studies using Support Vector Machines, Zhang and Zhao (2009) classify the model output into two classes: Class 1, made with the remarks in which there was an increase in the price, that is, $y_i = +1$ and Class 2, formed by the observations that had a decrease in the price, that is, $y_i = -1$. The days when there was no change in the price were ignored. Empirical results show that the accuracy of predictability of the SVM is greater than 60%. Therefore, Zhang and Zhao (2009) concluded that with the SVM, you can make independent forecasts the complexity of the financial market.

METHODOLOGY

Methodology of Support Vector Machines

The Support Vector Machines are indigenous to the study of Boser*et al* (1992) and its biggest advantage is the construction of a hyperplane that separates the data into two or more classes, to achieve the maximum separation between them. According Fan and Palaniswami (2001), the application of the principle of minimization of Empirical Error applied in quite widespread methods as Neural Networks, does not guarantee a lower actual error. The SVM solves this issue by

implementing the principle of Structural Risk Minimization, which seeks to minimize the upper bound of generalization error, rather than just minimize the error of the estimation process. This means that the classification of new observations of unknown classes, the chance of an error in the prediction based on classifier learning will be minimal.

Thus, the SVM learning can be understood as the discovery of the central hyperplane that maximizes the margin so that the comments of Class 1 ($y_i = +1$)are as separate as possible from the observations of Class 2 ($y_i = -1$). Figure 1 illustrates the concept of maximum margin for a set of data that can be separated from linear and direct manner.

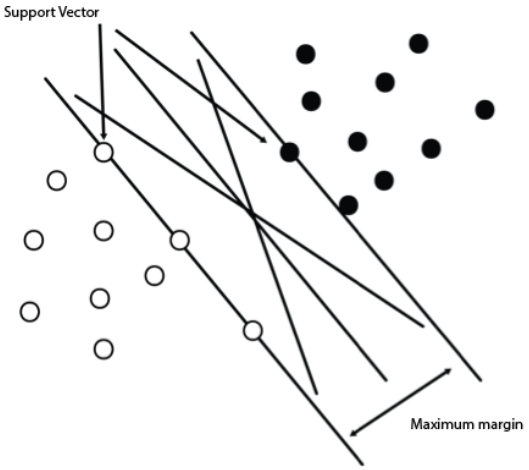

Figure 1–Maximum margin classifier

The goal then is to find the maximum margin hyperplane, written as$w_1x_1 + w_2x_2 - \gamma = 0$,and two other plans that will limit each class, taking the form as$w_1x_1 + w_2x_2 - \gamma \geq +1$ and$w_1x_1 + w_2x_2 - \gamma \leq -1$. In other words, the SVM aims to find two planes such that the points with$d = -1$satisfy the restriction $w_1x_1 + w_2x_2 - \gamma \leq -1$and points with $d = +1$ satisfy$w_1x_1 + w_2x_2 - \gamma \geq +1$.

The distance between these two planes is$\frac{2}{\sqrt{w_1^2+w_2^2}}andwand\gamma$should maximize the distance and satisfy the restrictions at the same time.

Maximizing the margin$\frac{2}{\sqrt{w_1^2+w_2^2}}$is equivalente to minimize its mutual$\frac{w_1^2+w_2^2}{2} = \frac{1}{2}w^Tw$, then the quadratic programming problem may be written in two ways:

Maximize: $\varsigma = \frac{2}{||w||}$ (1)

Subject to:

$\mathbf{D}(\mathbf{A}w - \gamma 1) \geq 1$

For $w \in \mathbb{R}^p, \gamma \in \mathbb{R}$.

or,

$$\text{Minimize} : \varsigma^* = \frac{1}{2} w^T w \quad (2)$$

Subject to:

$$\mathbf{D}(\mathbf{A}w - \gamma 1) \geq 1$$

For $w \in \mathbb{R}^p, \gamma \in \mathbb{R}$.

The primal problem of SVM can also be written in its dual form, given by Dual of Wolfe (1961):

$$Max_{\lambda \geq 0}\left[Min_{w,\gamma} \mathrm{L}(\mathrm{w}, \gamma, \lambda)\right] \quad (3)$$

In the Lagrangian form, the problem Dual is described by:

$$L(\lambda) = -\frac{1}{2}\lambda^T \mathbf{D}\mathbf{A}\mathbf{A}^{\mathbf{T}}\mathbf{D}\lambda + \lambda^T 1 \quad (4)$$

Subject to:

$$1^T \mathbf{D} 1 = 0$$

$$0 \leq \lambda \leq \mathrm{C}1$$

A linear separation problem is only a particular case and therefore it becomes necessary a formulation able to handle more complex problems. When the data are not linearly separable in its original size, SVM strategy is to build a mapping of data in a higher dimensional space, in a 'character space' in which they are linearly separable.

The nonlinear dependence relationship is represented by a matrix **A** and a target variable, represented by vector*y*, where the matrix is the input variable and *y*, is the output. The matrix**A**has dimension*n x p*, where each row represents an observation of a population and each column, a feature, that is, a variable of the population. The vector*y*represents the group where each observation belgongs and its dimenions are*n x 1*, containing only the values $+1$ or-1. Figure 2 illustrates the non-linear mapping process in the characteristic space.

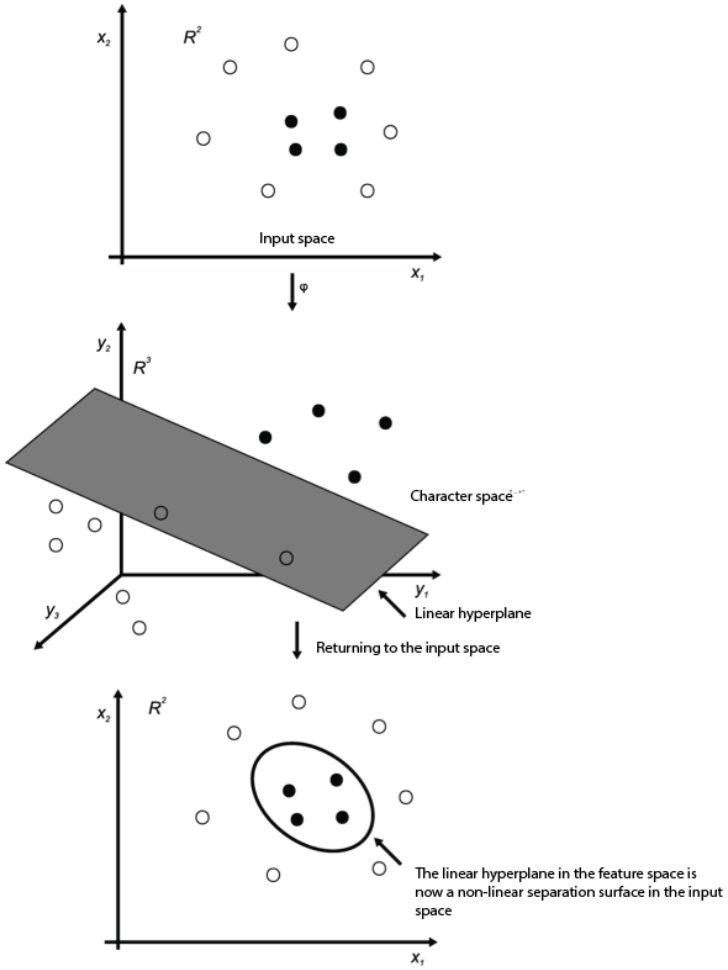

Figure 2–Mapping process

It is therefore evident that the strategy for working with non-linearity data is to create new dimensions through the mapping process and this is described as follows:

$$x \to \phi(x) \qquad (5)$$
$$\mathbb{R}^p \to \mathbb{R}^q \quad \text{tal que } q \gg p.$$

The mapping in question is defined by the function Kernel $\mathrm{K}(x_i, x_j)$which is a measure similarity or proximity between points and is described briefly as:

$$\mathrm{K}(x_i, x_j) = \phi(x_i)^T \phi(x_j) \qquad (6)$$

The kernel takes each point $x$on a map$\phi(x)$. Thus, instead of working with the matrix **A** it works with**F**, constructed from de $\phi(x)$. Since**F**is very large,

the calculation of the matrix $\mathbf{FF^T}$requires an enormous amount of operations, which makes the process very costly. This explosion of dimensionality can be avoided by using the Kernel Trick.The function$\mathrm{K}(x_i, x_j)$, as a function of the input space, does not require the mapping, but the result scalar product $\phi(x_i)^T\phi(x_j)$in the feature space. So just that thematriz$\mathbf{FF^T}$ is known to build an SVM operating in an extremely high dimensional space, even infinite.

In non-linear cases, we want to find the hyperplane with minimum points that contibute to the error. The conditions of maximum margin and the minimization of the number of points that contribute to the error are contradictory, because a higher margin will generate more points with errors. Therefore, the parameter Cis introduced and it represents the value of the error, that is, the weight of making a wrong classification.Considering also that the kernel matrix replaces the matrix$\mathbf{A}$in the formulation of SVM, the non-linear separation problem can be written as

$$\text{Minimize}: \varsigma^* = \frac{1}{2} w^T w + \mathrm{C}1^T \xi \quad (7)$$

Subject to

$$\mathbf{D}[\mathbf{\Phi}(\mathbf{x})w] + \xi \geq 1$$

$$\xi \geq 0$$

Where,

$$\mathbf{\Phi} = \begin{pmatrix} \phi(\mathbf{x}_1)^T \\ \phi(\mathbf{x}_2)^T \\ \vdots \\ \phi(\mathbf{x}_n)^T \end{pmatrix} \text{ e } w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \quad (8)$$

As in the linear case, to solve the problem (7) is interesting to work with the Dual Wolfe (1961), given by:

$$\mathrm{L}(\lambda) = -\frac{1}{2}\lambda^T \mathbf{D}\mathbf{\Phi}(\mathbf{x})\mathbf{\Phi}(\mathbf{x})^{\boldsymbol{T}}\mathbf{D}\lambda + \lambda^T 1 \quad (9)$$

Subject to:

$$0 \leq \lambda \leq \mathrm{C}1.$$

Procedures of collection and data analysis

This study used two samples, the first being formed by the actions that make up the theoretical portfolio of BOVA11 and the second, by the actions that make up the BRAX11. Since the theoretical portfolios of index funds are not fixed and adjusted every four months, this study adhered to the theoretical portfolios of both valid indices from January to April 2015. Data were collected in the earliest date of Economática system available until 2014.

Firstly the data of quarterly financial indicators was collected to be used as inputs to the SVM and the selected indicators were: Earnings per Share, Book Value per Share, Total Liabilities / Total Assets (%), Total Liabilities /

Stockholders Equity (%), Fixed Assets / Stockholders Equity (%),Total Assets / Total Liabilities, Current Ratio, AssetsTurnover, Equity Turnover, Gross Margin (%), Net Margin (%),Return on Assets, ROE (Ending Stock Equity) (%), ROE (Average Stock Equity) (%),ROE (Starting Stock Equity) (%). The next step was the collection of historical quarterly stock prices of the portfolios in order to calculate the return. Observations were taken with incomplete data and any data interpolation was not necessary due to the large size of the base. Then, databases were formed with indicators and quarterly returns. The first for BOVA11 actions with 20 quarters and the second for BRAX11 actions with 19 quarters. The size of the bases of each sample was different due to the availability of data for these actions in Economática. For data analysis we used the free software R, more specifically, the kernlab package.

The shares were then ranked due to the return presented and classified into two classes. TheClasse 1 ($y_i = +1$)was composed of 25% of the shares with higher returns and Class 2 ($y_i = -1$)was composed of the remaining shares representing 75%.In order to prevent overtraining the Cross Validation method was used and the databases were divided into three sets with diferentes time cutouts. The training set represented25.92%of the quarters, the validation set29.47%and the test set, 44.61%. The test set size was greater than the others due to concentration of data in the most recent years of the time frame of the study. This contributed to the model generalization ability test be more reliable.

The next step was the construction of Support Vector Machine for classification of stock performance through learning with historical data in two classes. The kernel function used was the Gaussian kernel. This is the most popular in the applications of SVM in finance, mainly because it maps a point in an infinite dimensional space, allowing more widespread and quick search of the optimal solution. For this reason, was used in this research as well as in studies of Tay and Cao (2001), Huerta*et al* (2013), Emir*et al* (2012) and Kim (2003). Its traditional formulation is given by:

$$k(x, y) = \exp\left(\frac{\|x-y\|^2}{2\sigma^2}\right), \quad (10)$$

whereas $\sigma$is the measure of nonlinearity of SVM.

According to Kim (2003), one of the advantages of SVM is that it depends on a small number of parameters, unlike most forecasting models. It is noted that since the kernel is chosen, the SVM depends on just one parameter, the C. There is no set value for each of the constant and although they are few, the choice of these is essential for the proper performance of the model, since the prediction problem is directly related to the trade-off between generalization ability of the classifier and its complexity.

Note that the data set of this study is unbalanced and therefore there is a bias towards the Class 2, as it is three times larger than Class 1. Veropoulos*et al* (1999) suggest the use of different parameters Cto solve such a problem, which implies a change in the problem formulation 9:

$$L(\lambda) = -\frac{1}{2}\lambda^T \mathbf{D}\boldsymbol{\Phi}(\mathbf{x})\boldsymbol{\Phi}(\mathbf{x})^{\boldsymbol{T}}\mathbf{D}\lambda + \lambda^T 1 \quad (11)$$

Subject to

$0 \leq \lambda \leq C_+,\ y_i = +1$

$0 \leq \lambda \leq C_-,\ y_i = -1.$

Note that for a portfolio selection problem is more costlyto miss the classification of one stock of Class 1 than to misclassificate one stock of Class 2. Therefore, in this study the SVM was built with different error weights for each class, allowing different levels of error and thus a more effective control over the sensitivity of the model.

To define the parametersC e σ, a sequence for the parameterCwas created ranging from 1to 1000 and one for the parameter σ, ranging from0.00000001to4. A grid was constructed with these two sequences to find the optimal combination of parameters, that is, the pair of parameters that generates the smallest error in the validation step. As a measure of accuracy of SVM performance, we used the ratio of the number of times a stock belonging to Class 1 was classified by the SVM as belonging to Class 2 by the total number of Class 1 stocks, that is, how much the SVM misclassificated the stocks of Class 1.

Th results of Lai *et al*(2006)study showed that if an investor selects only good quality stocks, a more voluminous portfolio not necessarily exceeds a portfolio with few actions. Thus, it is wise for investors to select a threshold number of shares and that all are of good quality. In order to improve the way in which the assets of a portfolio are selected, this study addressed the classification of shares through a new perspective.Whereas the distance from one point to the hyperplane, which limits the class in which it is inserted, is related to the classification error probability, geometric information can be used to interpret the SVM outputs as probabilities. Thus, after training the machine with the optimal parameters of the SVM prediction function was programmed to return in the test period, the probability class and not just the scalars$+1$ e -1.

According Platt (1999), to construct a classifier that produces probabilities of class-input relationship is very useful in a practical case of pattern recognition. According to the same author, outputs in the form of probabilities are needed when a classifier directs a small part of a more general decision and outputs have to be combined to make the final decision. Hence, the shares were classified and ranked according to the likelihood of belonging to Class 1 $y_i = +1$, that is,the 25% of shares with more likelihood to have an exceptional return were selected to form the portfolio. Note also that the classification made by means of probabilities allows the control of the portfolio size because it is guaranteed that in each period will be selected a minimum or maximum number of shares, which reduces the overall risk through diversification.

Through the classification of shares, the SVM was used to select portfolios quarterly considering each sample separately. The returns of these

portfolios were compared to the quarterly returns of the following markets benchmarks: the profitability of BOVA11, the profitability of BRAX11, the return of the equally weighted hypothetical portfolio of shares composing the Bovespa Index, the return of the equally weighted hypothetical portfolio of stocks that made up IBrX-110 and the profitability of CDI.

Through an ex-post analysis of the return distribution, the VaRindicated with 5% chance of error, how much could be lost in a quarter considering the worst scenario.

RESULTS AND DISCUSSION

In this research the Support Vector Machine was built having as inputs the quarterly results of 15 financial indicators. It is inferred that these 15 indicators explain the quarterly net return for the following period of each asset, and the return determines the classification of the stock in one of two classes($y_i = +1$ ou $y_i = -1$).

The probabilistic model of ksvm function was used for the SVM interpret the outputs as the probability of the assets are classified as $+1$. Thus, the SVM decision function has not classified assets in Classes 1 and 2 only by the output signal, but by their probability of taking the value$+1$. Class 1 was then composed of 25% of shares with greater odds and Class 2 for the remainder of the assets.

The search for optimal parameters was made in the same grid for the two samples, however, the optimized pairs found were different. This shows that different groups of assets are related in various forms with their ratios, which confirms the hypothesis that different characteristics of an asset, as its economic sector and its relationship with other sectors, are important information to compare their indicators and can help to predict their level of return. Thus, the separation in small groups for training of assets can contribute to better classification and prediction.In order to avoid the bias created by the unbalanced database was used to Class 1 $C_+ = 80\%$and for Class2,$C_- = 20\%$.

Whereas the first sample of the ETF BOVA11 the optimal values found were $C = 785.93$and$\sigma = 0.402$with a0.471 error. For the second sample, the ETF BRAX11, the optimal pair of parameters found wasfoi$C = 714.57$and$\sigma = 1.99$with0.862 error.It is observed that the accuracy of SVM in the sample of BRAX11, that is, the 100 most liquid shares, was not satisfactory, possibly because it is a more general set of actions and the machine requires more variables for prediction or a more robust model of accuracy.

To test the applicability of the SVM training portfolios, the returns generated by the machine were compared with those generated by ETF returns in the same period. In the first benchmarking, SVM classified the actions that comprised the BOVA11 theoretical portfolio valid from January to April 2015. With the classification and selection of the best actions of this set, an equally weighted portfolio was formed which had a cumulative return in 20 quarters of 94.15%. In the same period, the BOVA11 had a cumulative return of -14.42%.

Regarding the risk of investments, the BOVA11 showed a VaR of −14.10% and the SVM portfolio, −6.05%.

In the second benchmarking, SVM classified the stocks that comprised the theoretical portfolio for BRAX11 valid from January to April 2015. The return of the equally weighted portfolio made up of stocks selected by SVM within that second set presented a cumulative return in 19 quarters of 38.25%, while the ETF BRAX11 presented a cumulative return of13.86%. The VaR of this second portfolio SVM was −8.70%and for BRAX11 was−10.50%.

The economic conditions in Brazil along the defined time frame of this research was marked by strong economic slowdown, possibly has increased the discrepancy between the returns. For this reason, the returns of two other market benchmarks were also estimated. The first consists of all67 shares of the Ibovespa index used in this study and the second of every 100 shares of theoretical portfolio of IBrX-100. That is, the cumulative returns were calculated for portfolios composed of all actions of the theoretical portfolio and not just those that were classified as good and selected by SVM. The cumulative return on the 20 quarters of the third benchmark was 55.81%with risk −7.60%. The accumulated return of the fourth benchmark was 34.41% with VaR −9.50%. So again,the return of the portfolio chosen by the SVM was higher than the market benchmarks and the risk of SVM portfolio was also lower.

Finally, the returns of SVM portfolios were also compared to the return of a risk-free asset for the same period. In the 20 quarters analyzed in the first sample, the CDI showed a return of 57.10%and 19 quarters of the second sample, 53.75%. So, for all of the assets of theoretical BOVA11 portfolio, the SVM presented a portfolio 69.89% higher than the CDI. As for the theoretical portfolio of BRAX11, the portfolio return SVM was 28.84% lower than that presented by the CDI. Table 1 brings together all the results found in the research.

Opção de Investimento	Retorno Acumulado	VaR
Portfolio SVM 1	94.15%	-6,50%
Portfolio SVM 2	38.25%%	- 8.70%
BOVA11	-14.42%	- 14.10%
BRAX11	13.86%	- 10.50%
Equally weighted portfolio	55.81%	- 7.60%
Equally weighted portfolio	34.10%	- 9.50%
CDI (20 quarters)	57.10%	-
CDI (19 quarters)	53.75%	-

Caption: Table 1 –Results of the research

Figure 3 compares the cumulative returns in 20 quarters of the CDI, the BOVA11 the first portfolio selected by SVM and also weighted theoretical portfolio of shares composing the Bovespa index, also called market benchmark. Figure 4in turn, compares the cumulative returns in 19 quarters of the CDI, the BRAX11, the second portfolio selected by SVM and also weighted theoretical portfolio of shares that made up the IBrX-100, also called market benchmark.

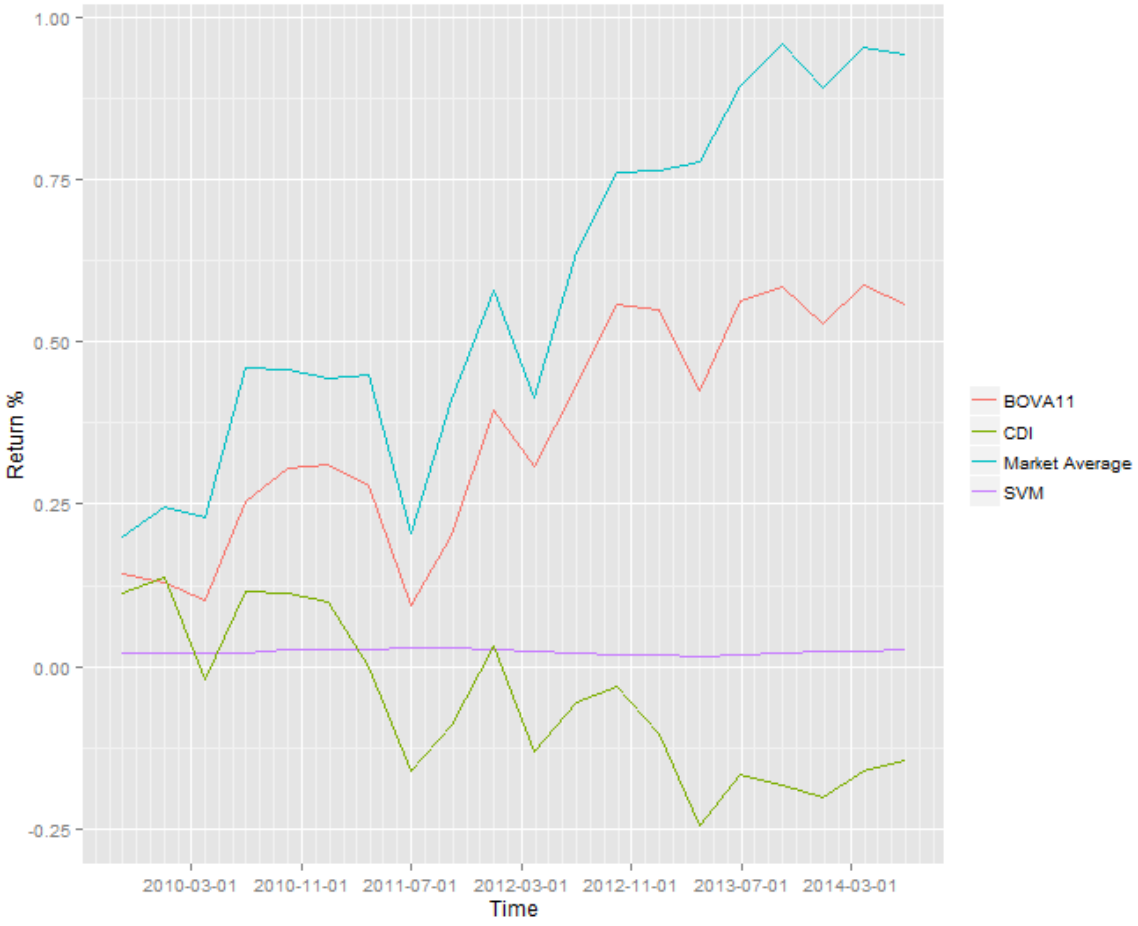

Figure 3 – Comparison of cumulative returns for BOVA11

Figure 4– Comparison of cumulative returns for BRAX11

CONCLUSIONS AND RECOMMENDATIONS

This research aimed to examine in the selection of portfolios whether the use of Support Vector Machines really contributes to a greater returnthan market benchmarks in the Brazilian context.

Two sets of assets were used in this research. The first was formed by 67 stocks that make up the theoretical portfolio of BOVA11 and the second, the 100 stocks that make up the BRAX11, both considering the theoretical portfolios of ETFs valid from January to April 2015. Financial indicators data and prices of each asset were collected at the Economática system from the earliest date available until 2014.

Trough the ranking of returns, the shares were classified into two classes:Class 1 ($y_i = +1$)composed of 25% of the shares with higher returns and Class 2 ($y_i = -1$)composed of the remaining shares representing 75%.The

Support Vector Machine was then built with the Gaussian kernel for performance classification of actions through learning with historical data in two classes. However, the classification of stocks by the machine was approached in a new perspective. The outputs of SVM were converted into probabilities of assets been classified as $+1$nd through the ranking of these probabilities, the actions that were within the 25% more likely to have been classified as a Class 1 and the others, constituted a Class 2. This research also innovated by building a Machine of Support Vector with different error weights for each class, allowing different levels of error and therefore a more effective control over the sensitivity of the model.

With the optimal parameters $\mathrm{C} = 785.93$ e $\sigma = 0.402$for the first sample, the SVM showed an accuracy of 59.80%, which is a very good result. However, in the second sample, with the optimal parameters$\mathrm{C} = 714.57$ e $\sigma = 1.99$, the accuracy of the model was13.80%. It is observed that the accuracy level fell significantly from one sample to another, possibly because the nature of the theoretical portfolios of the samples are very different. While the BOVA11 reflects the profitability of the Bovespa index, that is, the average yield of the shares with the largest representation in the Bovespa, the BRAX11, reflects the profitability of IBrX-100, which consists of a much more extensive portfolio with only marketable stock. These differences in results contribute to the hypothesis that different sets of assets are related in various ways with your ratios and perhaps the segregation of them by the economic sector in which they live will contribute to better accuracy, selection and prediction.

With stocks selected by SVM, quarterly portfolios were formed for each sample and their returns were compared with market benchmarks returns. The portfolio cumulative return selected by SVM with the actions of the first sample was 94.15%in 20 quarters with VaRof-6.50%. In this same period, the ETF BOVA11 showed a cumulative return of -14.42%andVaRequal to-14.10%.The accumulated return of the portfolio selected by SVM with the actions of the second sample was $38{,}25\%$in 19 quarters withVaR de -8.70%. The ETF BRAX11 presented in the same period a return of13.86%andVaRof-10.50%.

The returns of SVM portfolios were higher than both benchmarks, however, the period of the test set was marked by strong economic slowdown, a consequence of the slowdown of GDP (Gross Domestic Product), increased interest rates and high inflation. This economic context in Brazil was largely the result of government policies and economic decisions, but also a rate of less favorable international growth. Due to the possible increase of the discrepancy between the ETFs returns and portfolios selected by SVM by economic conditions, other two market benchmarks were calculated. One formed by all 67 shares of theoretical portfolio of BOVA11 used in this study and the other, by all 100 shares of the theoretical portfolio of BRAX11. In other words, the cumulative returns were calculated for portfolios composed of all actions of the theoretical portfolio and not just those that were classified as good and selected by SVM. Thus, the cumulative return on the 20 quarters of the third benchmark was $55.81\%$with

VaR−7.60%. The accumulated return of the fourth benchmark was 34.41%with VaR−9.50%. As with the comparison made with the cumulative return of ETFs compared to the cumulative returns of also weighted portfolios, the returns of portfolios selected by SVM were superior in both samples.

Compared to a risk-free asset CDI, the first portfolio selected by SVM, which considered the actions of the theoretical BOVA11 portfolio, presented a return 64.87% higher. But the other, the second portfolio selected by SVM, which considered the actions of the theoretical BRAX11 portfolio, presented a cumulative return 28.84%lower. However, these results were not considered as significant because the CDI be ballasted by the Selic rate and thus provide much higher returns in situations of high inflation and rising interest rates.

It is evident then that the portfolios formed by SVM outperformed the equities market in times of low economic perspective. It is also known that in difficult economic situations, fixed income has a very great attractiveness gain, usually motivated by monetary policy, which indicates that at a time of economic growth, the results of SVM are even more attractive.

Thus, the results of this study corroborate the hypothesis of superiority innovative method of Support Vector Machines in the formation of portfolios, characterized by the construction of a hyperplane that separates the data into two or more classes, to achieve the maximum separation between them and the implementation of the principle of the Structural Risk Minimization, which seeks to minimize the upper limit of generalization error, rather than just minimize the error of the estimation process.

Many gaps in portfolios training approach through Support Vector Machines can be explored. For future study, it is suggested improving the way of defining the optimal parameters, construction machine with different types of Kernel, determining the most appropriate inputs to the model and adequacy of the model used accuracy, so that it is consistent with the event studied and the results to be reached. Insights like these could amount to a large extent the accuracy of classification and the return generated by the portfolio, thus contributing to the improvement of the method.

6. REFERENCES

[1] **Abu-Mostafa, Y. S.; Atiya, A. F.(1996),*Introduction to Financial Forecasting*.** *Applied Intelligence, Springer,* v. 6, n. 3, pp. 205–213;

[2] **Albuquerque, Pedro H. M.(2014),*Previsão de séries temporais financeiras por meio de máquinas de suporte vetorial e ondaletas*.** Working Paper, São Paulo, Brazil/ Universidade de São Paulo, Instituto de Matemática e Estatística;

[3] **Boser, Bernhard E.; Guyon, Isabelle M.; Vapnik, Vladimir N.A.(1992),*Training Algorithm for Optimal Margin Classifiers*.** *In: Annual*

Workshop on Computacional Learning, 5, 1992, Pittsburgh. ACM Press. Pittsburgh: Haussler D, Jul pp.144-152;
[4**] Emir, Senol; Dinçer, Hasan; Timor, Mehpare (2012),*A Stock Selection Model Based on Fundamental and Technical Analysis Variables by Using Artificial Neural Networks and Support Vector Machines*. *Review of Economics & Finance*, Mar. Istanbul, pp. 106-122;
[5] **Fama, E. F (1970),*Efficient Capital Markets: A Review of Theory and Empirical Work***. *The Journal of Finance*, Chicago: University of Chicago, v. 25, n. 2, pp.383-417;
[6] **Fan, Alan; Palaniswami, Marimuthu (2001*), Stock Selection Using Support Vector Machines*. *Neural Networks, 2001. Proceedings. IJCNN 2001. (IEEE World Congress on Computational Intelligence).IEEE International Joint Conference on.* [S.l.], 2001. v. 3, pp. 1793–1798;
[7] **Gupta, Pankaj; Mehlawat, Kumar M.; Mittal, Garima(2012),*Asset Portfolio Optimization Using Support Vector Machines and Real-Coded Genetic Algorithm***. *Journal of Global Optimization, Springer*, v. 53, n. 2, pp. 297–315;
[8] **Huerta, Ramon; Corbacho, Fernanda; Elkan, Charles(2013),*Nonlinear Support Vector Machines Can Systematically Identify Stocks with High and Low Future Returns***. *Algorithmic Finance,* IOS Press, v.2, n. 1,pp. 45–58;
[9] **K. Veropoulos; C.Campbell; N. Cristianini (1999),*Controlling the Sensitivity of Support Vector Machine***.*In: Proceedings of the International Joint Conference on Artificial Intelligence*;
[10] **Kim, Kyoung-jae (2003),*Financial Time Series Forecasting Using Support Vector Machines***.*Neurocomputing*, Elsevier, v. 55, n. 1, pp. 307–319;
[11] **Lai, Kin K. et al.(2006),*A Double-Stage Genetic Optimization Algorithm for Portfolio Selection***.*In: International Conference, 13, ICONIP*, Hong Kong, China. *Springer-Verlag Berlin Heidelberg*, pp. 929-937;
[12] **Platt, John C.(1999),*Probabilistic Outputs for Support Vector Machines and Comparsons to Regularized Likelihood Methods***. *Microsoft Research*;
[13] **Soman, K. P.; Loganathan, R.; Ajay,V.(2011),*Machine Learning with SVM and Other Kernel Methods*;** (PHI Learning Private Limited: New Delhi);
[14] **Tay, Francis. E.H.; Cao, Lijuan (2001),*Application of Support Vector Machines in Financial Time Series Forecasting***.*Omega*, Elsevier, v. 29, n. 4, pp. 309–317;
[15] **Wolfe, P. (1961),*A Duality Theorem for Non-linear Programming***.*Quarterly of Applied Mathematics*, n. 19, pp. 239–244;
[16] **Zhang, Zuoquan; Zhao, Qin (2009), *The Application of svms Method on Exchange Rates Fluctuation***. *Discrete Dynamics in Nature and Society*, Hindawi Publishing Corporation, v. 2009.