Designing Financial Strategies based on Artificial Neural Networks Ensembles for Stock Markets

Julia de Mello Assis (In Memorian)
PPGMMC - CEFET-MG
Belo Horizonte – MG, Brazil
Email: juliamello.assis@gmail.com

Adriano C. M. Pereira
DCC - UFMG / PPGMMC - CEFET-MG
Belo Horizonte – MG, Brazil
Email: adrianoc@dcc.ufmg.br

Rodrigo Couto e Silva
PPGMMC - CEFET-MG
Belo Horizonte – MG, Brazil
Email: rodcouto@gmail.com

***Abstract*—Before the advent of computers and Internet, the stock market investors perform their operations based mainly on intuition. With the growth of investments and online stock trading, a continued search for better tools to improve the prediction of stock market trends has become necessary in order to increase profits and reduce risks. In this work we propose and evaluate some algorithmic trading (algotrading) strategies, based on an Ensemble of artificial neural networks (ANN), to support the decision of stock market's investors. Thirty different ANN models, using different input sets of price, volume and technical indicators, were analyzed for different stock symbols (i.e., companies from different economy sectors) of the main Brazilian Stock Market – BM&FBovespa. Moreover, ensembles that combine the best ANN models were modeled and validated through different experiments. The results confirm the potential of the proposed strategies for algotrading.**

***Index Terms*—Artificial Neural Networks. Stock Markets. Financial Computing. Machine Learning. Data Science.**

I. Introduction

Investors seek to base their decisions with tools and techniques that can offer them a possibility of financial return or risk management higher than one from a currency, reducing future uncertainty. Yao, Tan and Poh [1] explains that the market is sensitive to economic, political and psychological factors, such as fear, euphoria, pessimism or optimism; which makes the movement of historical series of prices volatile, speculative and difficult to predict.

Artificial Neural Networks (ANNs) have been used to perform time-series forecasts for a long time, according to the state-of-the-art study in Forecast with ANNs performed by [2]. Moreover, this was also verified by [3] and [4].

According to [5], even though there are many studies in the field there are still many possibilities of applications using combinations of techniques not yet analyzed and there is also a shortage in studies and experiments in the area, mainly in emergent countries, such as Brazil.

According to [6] ANNs' ability to extract nonlinear data patterns makes it one of the most useful forecast tools in this context. Classification using multiple ANN topologies such as in [7], as well as the use of Ensembles as [8], becomes relevant to improve the quality and accuracy of the forecast.

The main objective of this paper is to support the decision of an investor in the purchase and sale of stocks of the stock exchange, using a negotiation strategy based on the forecasts of an ANN Ensemble. For this, some Artificial Neural Networks will be modeled and evaluated. Also, Trading strategies will be elaborated from the results of the ANNs, a computational technique will be implemented that will merge the machine learning algorithms that obtained the best results in different datasets and finally, statistically validating the trading strategies.

As contributions from this work to the community we can mention: a new investment method using a combination of ANNs and the validation of the model using real data from the Brazilian financial market.

The remainder of this paper is organized, as follows. The related works are presented in Section II. Some explanations about Stock Market and ANNs are showed in Section III. Section IV describes the methodology used in this work. The experiments and analyses are showed in Section V, followed by the conclusions, which are presented in Section VI.

II. Related Work

This section presents some work related to the application of artificial neural networks (ANN) [9] for forecasting temporal series in the financial market.

In [7], a model for buying and selling shares of the Brazilian market with triggers is built, based on predictions of artificial neural networks. Fifty different network arrays were structured, where the number of neurons and the input database were changed. This database contains information obtained with technical indicators. Finally, in one of the test simulations, the model reached a profitability higher than 70%.

James [10] develops a Stock Trading System to predict trading signals based on KABAN Cell Neuron Network. This work evaluates 49 trading days, achieving up to 67% in terms of annualized investment return.

Although the ANNs have been used for a long time in order to achieve better time-series predictions according to [3] and [11], recent work has explored the ANN Ensemble for its ability to improve forecast accuracy, as evidenced in the works of Hayashi et. al [8] - Three-MLP Ensemble Re-RX algorithm and recent classifiers for credit-risk evaluation, Pulito et. al [12] - Genetic optimization of ensemble neural networks for complex time series prediction, Barrow and Crone [13] - Crogging (cross-validation aggregation) for forecasting: a novel algorithm of neural network ensembles on time series subsamples, Escovedo et. al [14] - Using ensembles

978-1-5090-6014-6/18/$31.00 ©2018 IEEE

for adaptive learning: A comparative approach, and Dong et. al - [15] Neural networks and AdaBoost algorithm based ensemble models for enhanced forecasting of nonlinear time series, thus generating powerful decision support systems.

The Three-MLP Ensemble Re-RX Algorithm proposed by Hayashi et. al [8] presented results in which the ensemble achieved better accuracy in the prediction than other techniques, such as the fuzzy SVM with a Radial Basis Function (RBF) proposed by Zhang. The ensemble achieved improvements from 72.07% to 98.14% in terms of the accuracy in forecasting.

All these works are similar to what will be presented in this paper, since they use ANNs to predict financial time series. However, none of these studies has developed a technique based on the combination of several different ANN models with learning algorithms to create a new model based on a ranking of the best accurate ANN models. Therefore, this work is innovative in two main points: 1. In the proposal of the method to select which network models will be part of the ensemble; 2. In the application of a new approach of ANN ensembles to support investment strategy decisions.

III. Fundamentals

Artificial Neural Networks (ANN) are based on biological neurons and composed of many parallel computing devices, called artificial neurons, connected one to another [16].To do forecasting, an ANN must be trained using samples of a time series that represents a certain phenomenon at a given scenario and after it can be used to predict (forecast) the behavior at subsequent times for other scenarios [17]

The stock exchange is a place where stocks issued by companies that open their capital are traded, in order to have a new source of financial resources. Brazil's main stock exchange is the BM&FBOVESPA, where shares of different classes are traded: shares, contracts index, commodities, forward transactions, among several others. This paper attempts to predict the trends of some assets present at BM&FBOVESPA.

There are several strategies of investment analysis that allow to identify the best opportunity to buy and sell shares, besides facilitating the evaluation of the assets that have better prospects towards valuation and payment of dividends. One such tactic is technical analysis, which is the study of past price movements and asset trading volumes, in order to predict the future price movement [1].

The work in [5] describes that the indicators from technical analysis are based on mathematical functions that impose consistency and discipline on researchers in pricing. The authors in [18] explain that these mathematical methods are commonly used by financial traders to predict transitions between different market regimes. This project will use technical analysis indicators combined with some stock information to try to strengthen the prediction power of trends.

IV. Model and Methodology

The objective of this research is to create an intelligent computational tool capable of generating triggers of buying and selling of financial assets, in order to support the decision of investors. The intelligence contained in this tool is based on a set of ANNs that learn and recognize patterns of behavior in asset historical data as well as in their technical indicators.

An overview of the experimental methodology proposed and applied in this work is illustrated in Figure 1.

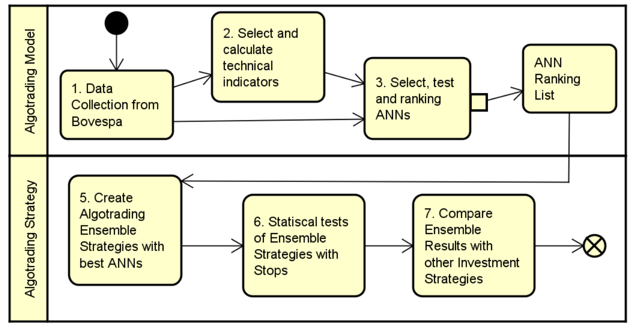

Figure 1. Methodology steps - Algorithmic Trading Model and Strategy

In the initial phase of the research the database is collected: the historical series of price, quantity of trades and traded volume of the stocks CIEL3, ITUB4, VALE5 and the BOVA11 index are selected. The BOVA11 Index Fund is chosen among assets because it behaves similarly to BM&FBovespa's main index, the IBovespa. The remaining assets are selected in order to diversify the validation of the model and identify a possible bias in the accuracy of the networks in different market sectors of the Companies. Therefore, ITUB4 is a stock of the Brazilian bank Itau Unibanco Holding S/A, CIEL3 is a stock of Cielo, a Brazilian company of credit transactions, i.e, working in the financial sector, finally VALE5 is a stock of Vale S/A, known as a Brazilian multinational mining company.

The period from January 1, 2014 to December 31, 2014 was selected to be used, as they were at the initial stage of the survey the most recent data available that included enough to determine a possible annual seasonality. The 15-minute data granularity or periodicity is suitable for simulation of Day Trade and Swing Trade (short duration) strategies. However, they would need to be revalidated for HFT-type strategies, differing at this point from the research of Silva [19], or even Long Positions, as in the work [20].

The price data from *candles*, including the trading volume and the number of shares traded, are normalized in this phase, before being informed to the ANNs, according to Equation 1.

$$X(i) = \frac{X(i) - Min(X)}{Max(X) - Min(X)} \tag{1}$$

In the second phase of this work, the financial technical indicators of each selected asset are calculated. Indicators that have already been normalized are used as input to ANNs [21]:

- On-balance Volume (OBV) - Volume Balance that relates volume to price changes;
- Relative Strength Index (RSI) Measures the evolution of the relationship of forces between buyers and sellers over time; Gives overbought and oversold signals;

- Simple Moving Average (SMA) is formed by the average price of a security over a specific number of periods;
- Exponential Moving Average (EMA), an extension of SMA, reduce the lag by applying more weight to recent prices;
- Money Flow Index (MFI), is an oscillator that uses both price and volume to measure buying and selling pressure. Is a tool to find conditions of overbought and oversold.

Those informations forms the database of the experiment or the inputs for the ANNs. We divided the whole data into three smaller packages: a large part to train and two smaller ones to validate and test. The packages do not contain repeated data and it prevents overfitting of ANNs.

During the third phase, three ANNs were chosen and modeled as described on the Table I, each one with the architecture of two intermediate layers and ten neurons in each layer.

Table I
ANNs

Code	ANN
N1	Feedforward
N2	Cascade Forward
N3	Elman

For the selected ANNs, each of the ten training algorithms was applied, as described in Table II.

Table II
LEARNING ALGORITHM

Code	Learning Algorithm
LA1	Levenberg-Marquardt
LA2	Bayesian Regulation
LA3	Scaled Conjugate Gradiente
LA4	Resilient backpropagation
LA5	Conjugate gradient backpropagation with Powell-Beale restarts
LA6	Gradient descent with adaptive learning rate backpropagation
LA7	Gradient descent with momentum backpropagation
LA8	One-step secant backpropagation
LA9	BFGS quasi-Newton backpropagation
LA10	Conjugate gradient backpropagation with Fletcher-Reeves updates

Thus, thirty different combinations of the network-training pair were obtained. To each of them, the price, volume, quantity traded and each of the calculated technical indicators were applied as inputs, and the forecast of the trend in the next 15-minute granular *candles* period is expected to be output.

The network training is performed using 60% of the database. Here was used hyperbolic tangent sigmoid transfer function as a transfer function of its layer for hidden layers and output layer and as a performance function was used mean squared normalized error. Holdout cross validation is still performed. After this, the network model capable of classifying is obtained. After that there's no more training.

With 20% of the sequential data the network validation process is carried out, in order to filter out the best networks. Thus, trained networks are simulated here using a completely new and unknown data for them. After this process, the classification power of each model is calculated, only for the uptrend, from a precision metric and then the five best ANNs are selected, considering the network-training pair.

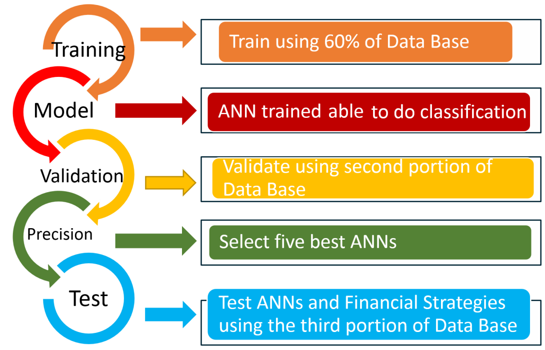

Figure 2. Model Mechanic

With these ANNs a further step is applied: the test (generalization) of each, utilizing the remaining 20% of the database and again, this is another simulation using trained and validated networks with a completely new and unknown data for them. Here the model is already trained and the best ANNs have already been selected by the validation stage, so this is where we test the modeled financial strategies. If the ANNs were able to make a good prediction in the validation set, it is expected to work well with other new data, that is, we hope that the strategies can make good buying and selling decisions. All of this explanation can be seen more clearly on Figure 2.

This process is performed month by month for each asset and repeated 30 times. As it is a stochastic process, there is variation in the results at each execution and its magnitude can be large enough to take important conclusions from our experiment, besides enabling a better statistical analysis of the experiment.

The following phases deal with the assembly of the strategies of operation and simulation of the negotiations. In the fourth phase, three Algotrading Strategies were created based on the combination of the top five ANNs, that is, the five classifiers that obtained the highest average values out of the 30 uptrend precisions obtained in the 30 executions. The ensemble technique was modeled because of the difficulty in selecting for a month a single classifier that was exceptionally better than the others, since many had rates with very close values, and also because it was not possible to define which ANN in specific should be selected for a given stock, since the months did not present a repetition of the best classifiers, that is, each month contained a set of ANNs with good results that differed.

Strategies work like Robot Advisors who consider each ANN's trend signals to assess whether they will indicate the stock purchase at that time. The purchase indication is given when it is verified that the next candle has a raising trend. The purchase signal is generated in each strategy when:

- Strategy 1 - 3 or more ANN agree with buy signal;
- Strategy 2 - 4 or more ANN agree with buy signal;
- Strategy 3 - All 5 ANN must agree with buy signal;

As a form of financial validation of the strategies, for each purchase signal, a simulation of order of entry in the market is generated in the amount corresponding to 1 lot of 100 shares. A way to control the risk involved in trades is to establish exit rules. The stop-loss strategy will cause the exit when a predefined price is reached. A low stop-loss can avoid profitable operations, while a high one can cause a higher loss. The strategies were executed using each stop configuration, in order to verify which would result in a better financial return. The Stop configuration settings used are:

- Stop Loss = 1.0%; Stop Gain= 1.0%;
- Stop Loss = 1.0%; Stop Gain= 0.5%;
- Stop Loss = 0.5%; Stop Gain= 0.5%;

Thus, after validating and finding the top five ANNs, these are combined and the trading strategies are created and performed using the ensemble selected.

V. Results

This section presents reached results, organized by two subsections: results of the performance of the ANNs on the historical series of the selected stocks and evaluating of trading strategies results. Due to paper space constraints, we decided to present the main results for some of the stocks. We plan to extend this paper later in order to present all results and different strategies in a scientific journal or magazine.

A. ANNs Prediction Performance Results

To classify and select the five best neural networks the Precision metric was calculated. With its value it is possible to assess how many times that each network gave a true prediction. This is calculated for each month and network.

$$Precision(High) = \frac{TP}{TP + FP} \tag{2}$$

As this is a rate computed from a confusion matrix for a binary classifier, it is understood that True Positive(TP) is the number of true predictions (Type II error) e False Positive(FP) is the number of false predictions (Type I error). The aim of this work is to find good classifiers for high movement trends, so the values of true positives show the correct predictions and the false positive evidence the wrong predictions.

$$Recall(High) = \frac{TP}{TP + FN} \tag{3}$$

Recall measures how many of the positives does the model return. Recall (see Equation 3) displays the calculation for ratings of bullish.

F1-Score (see Equation 4) is the harmonic average between Precision and Recall.

$$F1Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

The selected ANNs for the Ensemble Strategy are the five ones that achieved the highest values of the average of the 30 precision values (a value for each model execution) for prediction of high trend in the validation set of the networks. Table III presents the best ANNs according to this indicator, for VALE5 stock, in January, CIEL3 in April and BOVA11 in October of 2014. This table also shows the average results for Recall and F1-Score also for High movement trends. All these values are results from the ANNs validation set. The results were also generated for the other months of the year and for the ITUB4 stock, but will not be displayed in this paper, due to space constraints. Those months were chosen because they show more variety and variance (in the financial strategies), so we can evidence better our results. Moreover, we can exhibit more results from more months of the year.

Table III
Top Five ANN for VALE5, CIEL3 and BOVA11 Stocks in 2014

Month	**ANN**	**Precision**	**Std. Dev. (Precision)**	**Recall**	**F1 Score**
			VALE5		
	N2LA4	0.58	0.23	0.36	0.35
	N2LA1	0.53	0.32	0.24	0.25
JAN	N1LA4	0.53	0.19	0.40	0.38
	N1LA2	0.51	0.06	0.62	0.51
	N2LA6	0.50	0.29	0.40	0.35
			CIEL3		
	N2LA2	0.67	0.06	0.10	0.18
	N3LA2	0.59	0.10	0.53	0.50
APR	N1LA2	0.57	0.09	0.53	0.49
	N1LA10	0.55	0.11	0.72	0.58
	N3LA10	0.54	0.24	0.60	0.48
			BOVA11		
	N1LA9	0.57	0.10	0.73	0.57
	N2LA5	0.56	0.10	0.71	0.56
OCT	N2LA1	0.56	0.09	0.75	0.59
	N3LA1	0.56	0.16	0.64	0.52
	N3LA4	0.54	0.03	0.73	0.59

For VALE5, it was verified that in January the precisions of the top five ANNs were above 50%, while recall and F1-Score showed rates below this percentage. For CIEL3, in April it is verified for the majority precisions higher than 55%. Recall values are above 50% and F1-Score above 45%. Finally, for BOVA11 in the month of October we can see precisions above 55%, while recall presented values above 70% and F1-score above 50%.

We have tested some classification algorithms, such as SVM and Random Forest, however the results were similar or a little bit worse than our best ANN models. Therefore we focus on evaluating our ANN models and their ensembles and not to compare to others. For problems like stock market forecasting, we do not expect high values of accuracy, since it is extremely difficult to predict the prices movement.

These five ANNs are combined to create signals for a financial operating strategy in the stock market. This process are repeated for each month of the year.

B. Algorithmic Trading Strategy Results

The financial strategy was executed on the test data of the experiment (portion with 20% finals of the database of *candles*, a completely new and unknown data for the selected ANNs). As in the previous section, this one will shows the results for the financial strategy using VALE5 stock in January, CIEL3 in April and BOVA11 in October of 2014. The days on which the operations were simulated are as follows:

- VALE5 - JANUARY

  Mon Jan 27 14:45; Open Price: R$ 26.51

Fri Jan 31 17:45 2014; Close Price: R$ 27.95

- CIEL3 - APRIL
  Fri Apr 25 10:00 2014; Open Price: R$ 31.26
  Wed Apr 30 17:45 2014; Close Price: R$ 32.02
- BOVA11 - OCTOBER
  Mon Oct 27 13:00 2014; Open Price: R$ 48.45
  Fri Oct 31 17:15 2014; Close Price: R$ 52.98

A statistical tool will be used to identify which Stop configuration (of the three modeled) that increases the investment return (Equation 5). In order to know which statistical test to use, we tested the normality premises of the observations with Jarque-Bera, the randomness and homogeneity were verified with Runs Test and Bartlett (if the observations are normal, or Levene otherwise), respectively, as explained in [22].

It was verified that the premises of randomness and independence of the variances were satisfied. Therefore, for the strategies that presented normality in the distribution, the statistical test One-way analysis of variance were performed and when there is no normality, Kruskall Wallis was used. For both, a confidence level of 90% was defined. Thus, the hypotheses modeled were:

- H0: Strategies are statistically the same;
- H1: There are statistical differences.

In cases where the null hypothesis was rejected, it was possible to claim which of the strategies presented the best performance using the Tukey multi comparison test.

For each strategy there are a total of 30 observations. As in each execution of the model there is a differences in ANN output, therefore each observation was acquired by performing the financial strategy with the 30 different answers of the defined top five ANNs. The observations used for statistical validation were the Investment Returns (IR), Equation 5. Gain or loss is the subtraction of the value of the purchase of the share with its sale value. If positive, it is a gain; otherwise, a loss.

$$InvestmentReturn(IR) = Profit - Loss \qquad (5)$$

After the statistical validation, it was possible to infer which Stop configuration was more often better than the others in the year for a stock. Coincidentally, for VALE5, CIEL3 and BOVA11 the best Stop strategy was the 1.0% Stop Loss and 1.0% Stop Gain.

Figures 3, 4, 5 display boxplots with the investment returns in R$, for the best financial strategies for the stocks already mentioned. For VALE5 it shows a set of results for the first quarter, CIEL3 has the second quarter and for BOVA11 the fourth quarter of 2014 is shown. In this way, we can show some results of almost all the year, lacking only the third quarter.

As the strategy where the five must agree is more rigid, it performs fewer operations than the others and often decides not to operate, cause of their investment returns being mostly zero and their box being smaller than those of other strategies.

It is seen that to VALE5 (Figure 3) most of the strategies have positive values. The variance is high for strategies where three or more ANNs must agree to the buying operation in the months of January and March, this is also identified for the strategy where four or more must agree on the month of February.

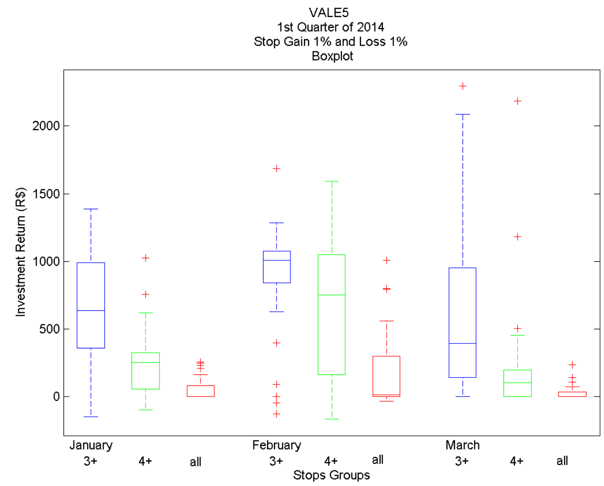

Figure 3. Box-Plots with Investment Returns of Ensemble Strategies with the bests Strategies for VALE5 in 1st Quarter of 2014

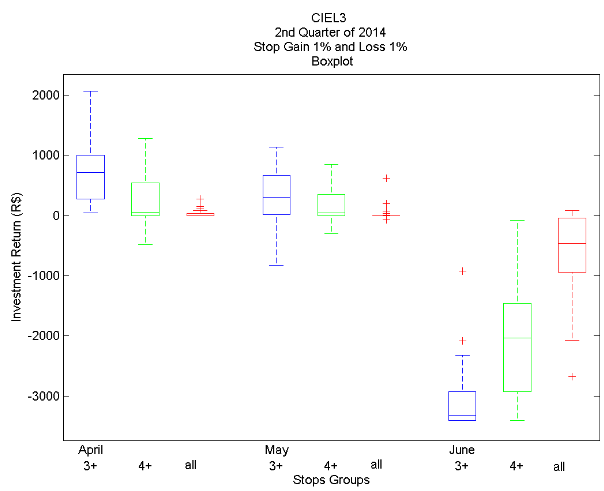

Figure 4. Box-Plots with Investment Returns of Ensemble Strategies with the bests Strategies for CIEL3 in 2nd Quarter of 2014

For CIEL3, it (Figure 4) is verified that the variances are not very large and the majority of the results are above the median value. With the exception of June, all strategies obtained values above zero. The 3+ agree strategy obtained the highest IR values, the 4+ agree strategy had a lower variance compared to the previous one and also was higher than the median.

The strategies for the fourth quarter of 2014 for BOVA11 (Figure 5), with the exception of November, which obtained mostly bad results, were positive, in addition to the majority of their values being above the median. Moreover, the 3+ agree strategy obtained the smaller variance than the others. Tables

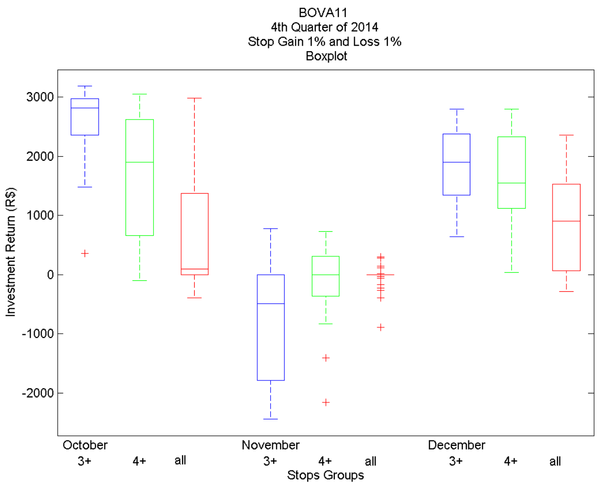

Figure 5. Box-Plots with Investment Returns of Ensemble Strategies with the bests Strategies for BOVA11 in 4th Quarter of 2014

IV, V, VI show in more detail the results of the financial strategies in Figures 3, 4, 5. However, only one month will be shown for each one, the ones already mentioned at the beginning of this session. For each strategy was selected: the best result (the one at the top of the box), the median (the one in the box median line) and the worst result (the one at the bottom of the box).

For all techniques, in all stocks, in the evaluated months, it can be seen that most of the techniques obtained positive investment returns, thus obtaining profits and also seeing high success rates (%Triggers with Profit). In addition, the 3+ agree strategy carried out, in most cases, a greater number of orders and the strategy that the five ANN must agree was the one that least carried out orders due to its rigidity of operation.

Table IV
RESULTS OF THE ENSEMBLES STRATEGIES WITH THE BEST STOP CONFIGURATION FOR VALE5 IN JANUARY OF 2014

VALE5 - January					
Box Plot Result	**Total of Order Triggers**	**%Triggers with Profit**	**IR in Percent**	**IR Mean in R$**	**Total of Stops**
3+ Agree					
Best	90	73.33%	51.51%	R$ 15.40	86
Median	20	95.00%	23.74%	R$ 31.80	20
Worse	14	42.86%	-5.05%	R$ -10.50	14
4+ Agree					
Best	44	81.82%	38.21%	R$ 23.32	44
Median	6	100.00%	9.75%	R$ 43.00	6
Worse	8	25.00%	-3.49%	R$ -12.00	7
All must agree					
Best	6	100.00%	9.75%	R$ 43.00	6

Through Table IV it is possible to verify for the strategy where 3+ agree, except for the worst result, that order triggers with profit are greater than 70%. Here the return investment ranged from R$-147.00 to R$1386.00. The other two months were also analyzed. In February the Investment Return ranged from R$126.00 to R$1685.00 and in March from R$433.00 to R$294.00.

For the strategy where 4+ agree, the success rates, with the exception of the worst case, are higher than 80%. Here the investment return ranged from R$-96.00 to R$ 1026.00. In February the Investment Return ranged from R$-163.00 to R$1591.00 and in March from R$106.00 to R$2185.00.

Finally, the strategy where the five ANNs must agree only decided to operate in the best case. The cases in the box median and floor decided not to operate at any time, leaving with zeroed results. The best case was 100% success in the few 6 trades it made and made a profit of R$258.00. In February the Investment Return ranged from R$163.00 to R$1591.00 and in March from R$106.00 to R$2185.00.

Table V
RESULTS OF THE ENSEMBLES STRATEGIES WITH THE BEST STOP CONFIGURATION FOR CIEL3 IN APRIL OF 2014

CIEL3 - April					
Box Plot Result	**Total of Order Triggers**	**%Triggers with Profit**	**IR in Percent**	**IR Mean in R$**	**Total of Stops**
3+ Agree					
Best	69	86.96%	66.74%	R$ 29.93	69
Median	109	56.88%	23.66%	R$ 6.57	108
Worse	14	50.00%	1.70%	R$ 3.29	13
4+ Agree					
Best	46	86.96%	41.40%	R$ 27.78	46
Median	1	100.00%	1.36%	R$ 42.00	1
Worse	42	33.33%	-15.49%	R$ -11.55	42
All must agree					
Best	7	100.00%	8.92%	R$ 39.29	7

From Table V it is verified for the CIEL3 stock in April that the order triggers with profits are higher than 55%, with the exception of the worse case that was 50%, for the 3+ agree strategy. Here the investment return ranged from R$46.00 to R$2065.00. The other two months of the quarter were also analyzed, in May the Investment Return ranged from R$-822.00 to R$1140.00 and in June from R$-3404.00 to R$-924.00.

For the strategy where 4+ agree, success rates, with the exception of the worst case, are higher than 86%. Here the investment return ranged from R$-485.00 to R$1278.00. In May the investment return ranged from R$-302.00 to R$849.00 and in June from R$-3404.00 to R$-82.00.

Finally, the strategy where the 5 must agree, decided to operate only in the best case. The cases that are in the boxplot median and floor decided not to operate at any time, getting results zeroed. So, the best case was 100% success in the few 7 trades he made and got a profit of R$275.00. In May the Investment Return ranged from R$-71.00 to R$621.00 and in June from R$-2672.00 to R$81.00.

Through the Table VI one can ascertain for the BOVA11 stock in October that all cases of the strategy where 3+ agree, have order trigger with profit rates above 60%. Here the investment return ranged from R$362.00 to R$3184.00. The other two months of the quarters were also analyzed, in November the Investment Return ranged from R$-2439.00 to R$779.00 and in December from R$641.00 to R$2797.00.

For the strategy where 4+ agree, profit rates, with the exception of the worst case, are greater than 80%. Here the

Table VI
RESULTS OF THE ENSEMBLES STRATEGIES WITH THE BEST STOP CONFIGURATION FOR BOVA11 IN OCTOBER OF 2014

BOVA11 - October					
Box Plot Result	**Total of Order Triggers**	**%Triggers with Profit**	**IR in Percent**	**IR Mean in R$**	**Total of Stops**
3+ Agree					
Best	67	86.57%	63.86%	R$ 47.52	67
Median	96	72.92%	56.66%	R$ 29.51	96
Worse	17	64.71%	7.58%	R$ 21.29	17
4+ Agree					
Best	74	82.43%	61.19%	R$ 41.28	74
Median	25	100.00%	37.28%	R$ 74.00	25
Worse	31	45.16%	-1.75%	R$ -3.35	31
All must agree					
Best	75	81.33%	59.73%	R$ 39.73	75
Median	1	100.00%	1.86%	R$ 92.00	1
Worse	9	22.22%	-7.79%	R$ -43.44	9

Table VII
BASELINES AND BENCHMARKS RESULTS

Stock Month/Year	Strategy	Total of Order Triggers	% Triggers with Profit	IR in Percent
VALE5 JAN/14	**Benchmarks**			
	B&H	1	100%	5.20%
	CDI	1	100%	0.16%
	Baselines			
	Random	66	55%	11.15%
	Trivial	72	50%	8.10%
CIEL3 APR/14	**Benchmarks**			
	B&H	1	100%	1.65%
	CDI	1	100%	0.12%
	Baselines			
	Random	64	45%	-1.90%
	Trivial	50	32%	-5.90%
BOVA11 OCT/14	**Benchmarks**			
	B&H	1	100%	1.66%
	CDI	1	100%	0.16%
	Baselines			
	Random	62	48%	11.80%
	Trivial	66	45%	11.20%

investment return ranged from R$-104.00 to R$3055.00. In November the Investment Return ranged from R$-2158.00 to R$724.00 and in December from R$33.00 to R$2800.00.

Finally, the strategy where the 5 must agree got a success rate higher than 80%, with the exception of the worst case, which was 22%. The middle case won a 100% hit rate, but only operated once. For this strategy IR ranged from R$-391.00 to R$2980.00. In November it ranged from R$-894.00 to R$295.00 and in December from R$-283.00 to R$2354.00.

In order to be able to check if these results are reasonable, they will be compared with the following benchmarks and baselines of investments: Buy & Hold (B&H) and CDI, Trivial and Random. All operations of baselines and benchmarks were carried out during the same period and for the same actions already mentioned in that session.

Buy & Hold (B&H) performs only one purchase of the asset at opening value of the initial period and another sale operation at the closing price of the final period. The CDI (or DI Over) is how the Interbank Deposit Rate is known in Portuguese and it is similar to LIBOR interest rates on London too. In this Benchmark, the accumulated rate in the period is used to give the idea of what it would be like to apply the same capital to fixed income. Investment strategies with profitability indexed by the CDI are suitable for investors with high risk aversion.

The Trivial and Random baselines operate as the Ensemble techniques created in this work, differing only in the purchase trigger. In Trivial, the purchase is made when the current candle has high movement trend. In the Random strategy, the buy trigger is a random sign.

The Investment Return (IR in percent) shows the percentage evolution of the capital in the period of operation, adding up all profits and losses of each operation in both ensemble strategies and baselines.

For VALE5, the yield on fixed income CDI accumulated in the period was 0.16% and the Buy & Hold method obtained a profitability of 5.20%. All modeled strategies (Table IV) are more profitable than the CDI and Buy & Hold, except for the worst case scenario.

The Random strategy earned a profit of 11.15%. Here, the strategy where 3+ agree was less profitable than Random just in the worst case. The 4+ agree strategy was more profitable only in its best case and the 5 must agree was less profitable, but more accurate.

The Trivial method earned 8.10% profit. Here the ensemble strategies are better than Trivial in the same way as was verified for Buy & Hold and CDI.

For CIEL3, the CDI yielded 0.12% profit. It can be seen here that the modeled strategies (Table V) are more profitable than the CDI, except for the worst case of the 4+ agree strategy.

Buy & Hold achieved a profitability of 1.65%. The 3+ agree and All 5 must agree strategies are more profitable than Buy & Hold in all cases, but the 4+ agree is only more profitable in the best case.

Random and Trivial yielded negative earnings of -1.90% and -5.90% respectively. All ensemble strategies created are more profitable than the two baselines, with the exception of the worst case of the 4+ agree strategy.

For the BOVA11, the CDI and Buy & Hold models earned profits of 0.16% and 1.66%, respectively. It can be seen here that the modeled ensemble strategies (Table VI) are more profitable than the two methods, except for the worst case of the 4+ agree and the all must agree strategies.

Random and Trivial obtained yields (investment return) of 11.80% and 11.20%, respectively. The ensemble strategies are more profitable than both methods, with the exception of their worst cases and the median case of the all must agree technique.

VI. CONCLUSION

The financial market is complicated, as asset values change rapidly over time through the influence of several factors, making difficult to predict the price series. Several studies agree that ANNs are capable of predicting these series. Thus, this paper main objective is to support the decision of an investor in the stock exchange, using a negotiation strategy based on the forecasts of an ANN Ensemble.

In this work thirty different Artificial Neural Networks (ANN) were modeled, using different input sets of price, volume and technical indicators. From the results of the five best ANNs, in the validation stage of the models, for the VALE5 shares, in January, CIEL3, in April, and BOVA11 in October of 2014, it was possible to verify that all five best ANNs obtained precisions above 50%. The ANNs for BOVA11 share was the ones with the lowest precision variation and ANNs for CIEL3 showed the highest variation. In most cases, the standard deviation is small and with the exception of VALE5 case, the ANNs also presented recall values above 50%.

The financial strategies were statistically evaluated and from this it can be seen which Stop-loss configuration increases the investment return when applied alongside one of the algorithmic trading strategies. Coincidentally for the three actions, was the configuration Stop Loss of 1.0% and Stop Gain of 1.0%. Also by this validation, it is possible to infer that the strategy where the five ANNs should agree is more rigid, because it performs fewer operations and often decides not to operate, which causes the large amount of zeroed values for the investment returns and small variance.

For all strategies, for the quarters analyzed, it can be verified that the majority obtained Positive Investment Returns, that is, profits and success rates (%Triggers with Profit) over 70%. In addition, the 3+ agree strategy carried out a greater number of orders and the strategy that the five ANN must agree was the one that made the least orders, although it obtained good IR Percentage values.

Thus, Ensemble 3+ agree strategies were in general more profitable than other ensemble like strategies 4+ agree and all must agree. It is evidenced in the results how the operations with the Ensemble of ANNs strategies can be more profitable than the presented baselines and benchmarks.

ANN ensembles strategies show in their median results for CIEL3 and VALE5 that the investment returns are smaller when compared to the Trivial and Random baselines. Although it is found that the median and best results of Ensemble 3+ agree, they were much more profitable than all the benchmarks (CDI and B&H) and baselines (Random and Trivial) studied.

It was concluded that the more ANNs we have agreeing in the ensemble, more accurate or assertive the prediction is. On the other hand, smallest will be the number of operations, and, consequently, the profits may be smaller in the more accurate strategies. The more assertive techniques can better meet the needs of investors with a higher risk aversion profile.

Acknowledgments

This research was supported by the Brazilian National Institute of Science and Technology for the Web (CNPq grant numbers 573871/2008-6 and 477709/2012-5), MASWeb (grant FAPEMIG/PRONEX APQ-01400-14), EUBra-BIGSEA (H2020-EU.2.1.1 690116, Brazil/MCTI/RNP GA-000650/04), CEFET-MG, CAPES, CNPq, and Fapemig.

References

[1] J. Yao, C. L. Tan, and H.-L. Poh, "Neural networks for technical analysis: a study on klci," *International journal of theoretical and applied finance*, vol. 2, no. 02, pp. 221–241, 1999.

[2] G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks:: The state of the art," *International journal of forecasting*, vol. 14, no. 1, pp. 35–62, 1998.

[3] E. M. Azoff, *Neural network time series forecasting of financial markets*. John Wiley & Sons, Inc., 1994.

[4] G. S. Atsalakis and K. P. Valavanis, "Surveying stock market forecasting techniques–part ii: Soft computing methods," *Expert Systems with Applications*, vol. 36, no. 3, pp. 5932–5941, 2009.

[5] L. C. Martinez, D. N. da Hora, J. R. d. M. Palotti, W. Meira, and G. L. Pappa, "From an artificial neural network to a stock market day-trading system: A case study on the bm&f bovespa," in *2009 International Joint Conference on Neural Networks*, pp. 2006–2013, IEEE, 2009.

[6] X. Wang and M. Han, "Multivariate time series prediction based on multiple kernel extreme learning machine," in *2014 International Joint Conference on Neural Networks (IJCNN)*, pp. 198–201, IEEE, 2014.

[7] F. D. Paiva, *Redes neurais para decisões no mercado de ações brasileiro*. PhD thesis, 2014.

[8] Y. Hayashi, Y. Tanaka, S. Yukita, S. Nakano, and G. Bologna, "Three-MLP Ensemble Re-RX algorithm and recent classifiers for credit-risk evaluation," pp. 1–8, IEEE, July 2015.

[9] X. Yao, "Evolving artificial neural networks," *Proceedings of the IEEE*, vol. 87, no. 9, pp. 1423–1447, 1999.

[10] C. James, "Kanban Cell Neuron network Stock Trading System (KCN-STS)," pp. 1–8, IEEE, July 2015.

[11] G. P. Zhang, "Time series forecasting using a hybrid arima and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.

[12] M. Pulido, P. Melin, and O. Castillo, "Genetic optimization of ensemble neural networks for complex time series prediction," in *The 2011 International Joint Conference on Neural Networks*, 2011.

[13] D. K. Barrow and S. F. Crone, "Crogging (cross-validation aggregation) for forecastinga novel algorithm of neural network ensembles on time series subsamples," in *Neural Networks (IJCNN), The 2013 International Joint Conference on*, pp. 1–8, IEEE, 2013.

[14] T. Escovedo, A. V. A. da Cruz, M. Vellasco, and A. S. Koshiyama, "Using ensembles for adaptive learning: A comparative approach," in *Neural Networks (IJCNN), The 2013 International Joint Conference on*, pp. 1–7, IEEE, 2013.

[15] Y. Dong, J. Zhang, and J. M. Garibaldi, "Neural networks and adaboost algorithm based ensemble models for enhanced forecasting of nonlinear time series," in *2014 International Joint Conference on Neural Networks (IJCNN)*, pp. 149–156, IEEE, 2014.

[16] R. Raposo and A. D. O. Cruz, "Stock market prediction based on fundamentalist analysis with fuzzy-neural networks," in *Proceedings of 3rd WSES International Conference on Fuzzy Sets*, Citeseer, 2002.

[17] I. Basheer and M. Hajmeer, "Artificial neural networks: fundamentals, computing, design, and application," *Journal of microbiological methods*, vol. 43, no. 1, pp. 3–31, 2000.

[18] G. Armano, M. Marchesi, and A. Murru, "A hybrid genetic-neural architecture for stock indexes forecasting," *Information Sciences*, vol. 170, no. 1, pp. 3–33, 2005.

[19] E. Silva, H. Brandão, D. Castilho, and A. C. Pereira, "A binary ensemble classifier for high-frequency trading," in *2015 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, 2015.

[20] F. A. de Oliveira, C. N. Nobre, and L. E. Zarate, "Applying Artificial Neural Networks to prediction of stock price and improvement of the directional prediction indexCase study of PETR4, Petrobras, Brazil," *Expert Systems with Applications*, vol. 40, no. 18, pp. 7596–7606, 2013.

[21] C. Coello, V. Cutello, K. Deb, S. Forrest, G. Nicosia, and M. Pavone, *Parallel Problem Solving from Nature - PPSN XII: 12th International Conference, Taormina, Italy, September 1-5, 2012, Proceedings*. No. pt. 2 in Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2012.

[22] D. C. Montgomery, G. C. Runger, and N. F. Hubele, *Engineering statistics*. John Wiley & Sons, 2009.