NEURAL NETWORKS AND INVESTOR SENTIMENT MEASURES FOR STOCK MARKET TREND PREDICTION

SALIM LAHMIRI

Department of Computer Science, UQAM, Montreal, Canada

ABSTRACT

Soft computing methods and various sentiment indicators are employed to conduct out-of-sample predictions of the future sign of the stock market returns. In particular, we assess the performance of the probabilistic neural network (PNN) against the back-propagation neural network (BPNN) in predicting technology stocks and NYSE up and down moves. Genetic algorithms (GA) are employed to optimize the topologies of the BPNN. Our results from Granger causality tests show strong evidence that all stock returns are strongly related to at least one of the sentiment variables. In addition, the results from simulations show that the GA-BPNN is more capable of distinguishing between market ups and downs than the PNN. Finally, the simulations show that trading given decision rules (for example; buy stock if predicted return is higher than a given threshold) yields to higher accuracy than predicting the stock market ups and downs.

Keywords: *Artificial Intelligence, Classification, Stock Market*

1. INTRODUCTION

Large profits can be earned by trading in stock markets. Therefore, investors are highly interested in forecasting the future trend of stock market indices and stock prices. The purpose of prediction is to reduce uncertainty associated to investment decision making. However, forecasting stock markets is a challenging task since the dynamics of the market are very complex and non-linear. For instance, many factors affect the stock market such as business and economic conditions, political events and investor's expectations. There is an abundant theoretical and empirical literature exploring the economics and the behaviour of stock markets. For instance, empirical finance has documented that traditional asset-pricing which are based on statistical methods such as the capital asset pricing model [1][2], the asset pricing model [3], and inter-temporal capital asset pricing model [4] all fail to explain and predict future stock returns. On the other hand, behavioural finance provides an alternative theory regarding financial markets. Based on experimental psychology literature, behavioural finance considers that cognitive biases could affect asset prices. Indeed, investor sentiment and limited arbitrage are the main arguments on which the theory of behavioural finance relies. In particular, the theory of investor's sentiment states that investors make investment decisions according to their sentiments (emotions) instead of following a fully rational process. Then, stock prices could be affected by sentiment (irrational behaviour). Many papers in the field of behavioural finance document the effect of sentiment on stock markets [5-8]. Other studies investigate the role of sentiment variables in the prediction of stock returns and financial futures. For example, [9-11] find that sentiment measures help predict returns on futures. On the other hand, [10] concludes that investor sentiment may have significant effects on the cross-section of stock prices. In addition, Baker and Wang [11] show that the forecasting power of sentiment measures is extremely limited once past returns are included as predictors. Based on the investor psychology, the behavioural finance literature has proposed many proxies of investors' sentiment, including investors surveys [12-14], investor mood [15][16], mutual fund flows [16][17], trading volume [17][19], retail investor trades [20][21], and closed-end fund discount [22-24] among others. In the previous works [5-24] linear statistical regressions were used to model and predict stock market returns with investor sentiment using in-sample data. In this study we consider the problem of stock market trend prediction. Indeed, predicting stock market trends is a classification problem that categorizes markets returns as up and down moves, which is easier than the price variation prediction as in [5-24].

There have already been studies looking at the direction or trend of movements of stock markets using BPNN [25-27] and PNN [28-33]. However, none of these studies provide a comparative evaluation of different intelligent classification techniques regarding the ability to predict the sign of stocks and index returns using sentiment measures as inputs. Our contribution is to use sentiment measures in the prediction of daily trend of individual stocks in out-of-sample data from the US technology sector and stock market index using neural networks with sentiment indicators. We rely on the technology sector because investors are strongly interested in investing in high-technology companies in the US and Europe since the late 1999s and the literature does not contain work that explores these companies with soft computing techniques and using indicators related to investor psychology. The well known back-propagation neural network (BPNN) is genetically optimized to predict the up and down moves of stocks, and it performance is compared to those of the probabilistic neural network (PNN).

The BPNN is a feed-forward network introduced by Rumelhart [33]. Given input–output pairs, the system is trained using back propagation gradient descent with momentum, and consequently adjusted to approximate any non-linear function, which make the system powerful in classification problems. On the other hand, the PNN was proposed by Specht [34]. It is built based upon the Bayesian method of classification. Indeed, the PNN employs Bayesian decision-making theory based on an estimate of the probability density of the data. The main advantage of the Bayesian method is to be able to classify a new sample with the maximum probability of success given a large training set using prior knowledge [35]. The PNN combines the simplicity, speed and transparency of traditional statistical classification models and the computational power and flexibility of back-propagated neural networks [36]. According to Kim and Chun [37], PNN outperforms back-propagation in discovering local patterns in time series, particularly in the absence of noise.

The rest of the paper is organized as follows. Section 2 describes the methodology. Section 3 outlines the simulation results. Section 4 concludes the paper.

2. METHODOLOGY
In this study we utilize US daily time series for the returns of three companies (Apple, Cisco, and General Electric) from the technology sector and one equity index (NYSE) from January 3rd 2000 to December 31st 2008. The first 80% observations of the data are used for training and the remaining 20% is used for testing. For each company and the equity index i, the return time series are computed according to:

$$R_{i,t} = \log(p_{i,t}) - \log(p_{i,t-1})$$

where p is the closing price and t is time script. Figure 1 shows the return series.
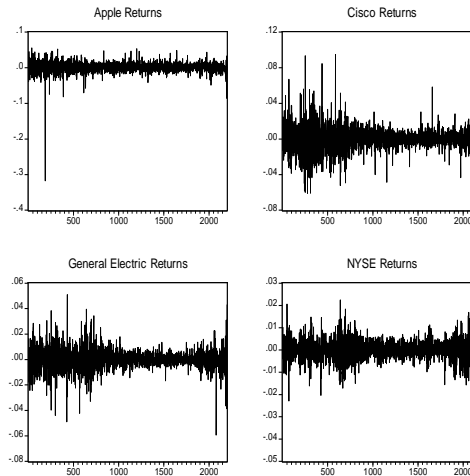

Figure 1: Return series

In this study, four measures of investor sentiment are used to predict future stock market returns. The first measure is the Volatility Index (VIX) of the Chicago Board Options Exchange [38] which is an estimate of the implied volatility of S&P 500 index options. The VIX is viewed as a fear index; that is high (low) levels indicate bearish (bullish) sentiment [10]. The second measure is the State Street's Investor Confidence Index (ICI) [39] that measures the attitude of investors to risk. Figure 2 exhibits the VIX and the ICI time series.
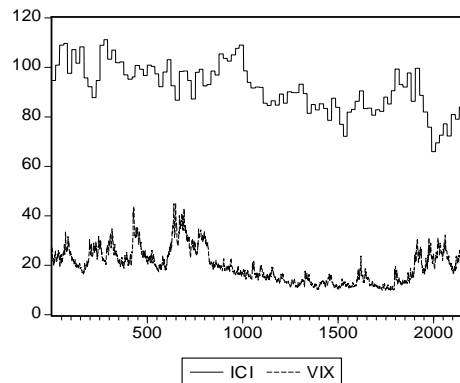

Figure 2: ICI, and VIX series

According to Hirshleifer [40], a lack of accurate information and greater uncertainty about stocks leads to psychological biases. Moreover, Greater information uncertainty is highly related to future stock returns [41]. Harris [42] and Godek [43] suggest that the uncertainty about the stock price should be estimated using stock returns volatility and trading activity. Indeed, sentiment is related to high volatility [8]. Consequently, the log of volume series (Figure 3) and measures of the volatility of return series are considered in our study as the third and the fourth sentiment indicators respectively. In the next step, returns series are modeled by ARMA processes and APARCH models to estimate and extract volatility series.
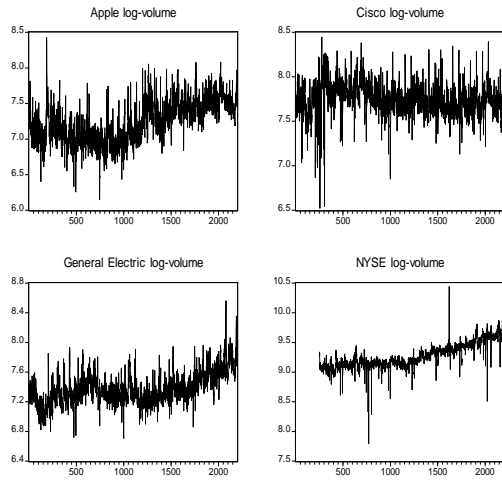

Figure 3: Log-volume series

2.1 Volatility modeling and extraction

To estimate the volatility of stocks and market returns, the following methodology is employed. Assuming that $R_{i,t}$ follows an ARMA(p,q) process, the conditional variance is modelled using the asymmetric power GARCH model APARCH (m,n) introduced by Granger and Engle [44]. First, the mean equation is estimated:

$$R_{i,t} = \alpha_s + \sum_{t=1}^{p} \rho_{i,t-1} R_{i,t-1} + \sum_{t=1}^{q} \phi_{i,t-1} \varepsilon_{i,t-1} + \varepsilon_{i,t}$$

$$\varepsilon_{i,t} \sim \left(0, \sigma_{i,t}^2\right)$$

To identify[1] the degrees p and $q,$ we make use of the Akaike information criterion (*AIC*) and Schwarz criterion (*SC*) computed as follows [45]:

$$AIC = -2\ell/T + 2k/T$$

$$SC = -2\ell/T + (k \log T)/T$$

$$\ell = \left(-T/2\right)\left(1 + \log(2\pi) + \log(\hat{\varepsilon}_S' \hat{\varepsilon}_S / T)\right)$$

where k and T are respectively the number of coefficients and sample size used for estimation and ε is the error term from the mean equation. The APARCH (m,n) model is given by the following variance equation:

$$\sigma_{s,t}^{\delta} = \omega_s + \sum_{j=1}^{m} \beta_{s,j} \sigma_{s,t-j}^{\delta} + \sum_{i=1}^{n} \theta_{s,i} \left(\left|\eta_{s,t-i}\right| - \gamma_{s,i} \eta_{s,t-i}\right)^{\delta}$$

where δ is the power parameter of the standard deviation σ and γ is a parameter that captures asymmetry effect up to a given order. In this study, γ is set to 1. The model APARCH is estimated with errors ε following a generalized exponential distribution (GED) which is introduced by Subbotin [46]. For instance, [47][48] found that the out-of-sample performance for the GARCH family models is worse with normal distribution. The orders m and n were arbitrarily set to 1 and the obtained parameters ($\omega, \beta, \theta, \gamma$) are all statistically and highly significant[2]. The APARCH (m,n) model provides three interesting advantages. The power parameter δ of the standard deviation can be estimated within the variance equation rather than imposed. Squared power transformation may lead to a sub-optimal model when the data is non-normally distributed [49]. Moreover, the power term is suitable to model volatility clustering -low volatility periods followed by high volatility periods- by changing the importance of the outliers [50]. Finally, the volatility series used are those extracted from the APARCH (m,n) equation. They are shown in Figure 4.
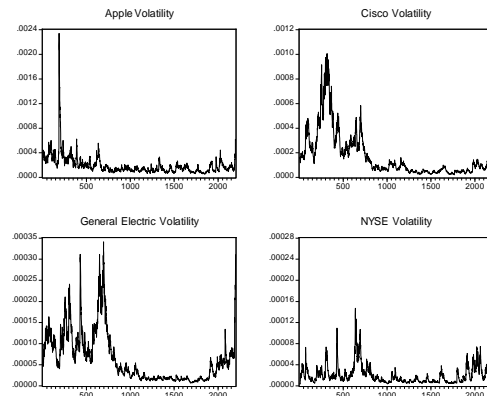

Figure 4: Volatility series

---

[1] The parameters (p,q,α,ρ,ϕ) are not reported, but are available upon request.

[2] They are not reported, but are available upon request.

The five inputs (sentiment indicators) are first selected after running Granger causality tests [51]. The Granger causality test allows considering only inputs which have a highly statistical causal effect on future stock returns. The test is based on bivariate regressions of the form:

$$y_t = \beta_{y,0} + \beta_{y,1} y_{t-1} + ... + \beta_{y,k} y_{t-k} + \delta_{x,1} x_{t-1} + ... + \delta_{x,k} x_{t-k} + \eta_t$$

$$x_t = \beta_{x,0} + \beta_{x,1} x_{t-1} + ... + \beta_{x,k} x_{t-k} + \delta_{y,1} y_{t-1} + ... + \delta_{y,k} y_{t-k} + \upsilon_t$$

where η and υ represent Gaussian disturbances. Then, F-statistics are computed as the Wald statistics for the joint hypothesis:

$$\delta_1 = \delta_2 = ... = \delta_k = 0$$

The F-statistics allows testing whether the coefficients on the lagged x's are statistically significant in explaining the dependant y. In this study, the number of lags, k, was arbitrarily set to 5. Table 1 provides the obtained results for the Granger causality test for each stock. Only statistically significant inputs are reported.

Table 1
Summary of pairwise Granger causality tests

Null Hypothesis	F-Stat	Probability
Apple		
Volatility does not Granger Cause returns	3.98117	0.00134
VIX does not Granger Cause returns	2.07191	0.05609
Cisco		
Volume does not Granger Cause returns	3.4265	0.00435
VIX does not Granger Cause returns	3.29944	0.00568
General Electric		
VIX does not Granger Cause returns	6.65516	3.70E-06
NYSE		
ICI does not Granger Cause returns	2.42607	0.03337
Volatility does not Granger Cause returns	9.18499	1.20E-08
VIX does not Granger Cause returns	5.01435	0.00014

2.2 The BPNN

The BPNN introduced by [33] has feed forward connections and uses the back-propagation algorithm optimized based on the gradients method. The topology of the network consists of three layers: an input layer, a hidden layer, and an output layer. All nodes of a layer are connected to all the nodes in the next higher layer. On the other hand, there are no connections among neurons in the same layer. Activation functions are used in the hidden layers to introduce non-linearity into the network to approximate non-linear functions. Indeed, it is the non-linearity that makes the BPNN powerful. The sigmoid functions such as the logistic and hyperbolic tangent functions, and the Gaussian function are the standard choices. The training of a BPNN involves three stages: the feed-forward of the input training pattern, the calculation and back-propagation of the error, and the adjustment of the weights. In particular, the adaptation of the weights is derived based on the gradient descent method and error back-propagation to minimize the error function E given by:

$$E = 0.5 \sum_{j=1}^{k_j} (d_j - y_j)^2$$

Here, y_j and d_j are respectively the actual and the desired output in each node j and k is the number of output neurons. The error is then back-propagated by the gradient descent through the network by adjusting the new weights according to this equation:

$$\Delta W(t) = -\gamma \left(\partial E / \partial W \right) + \alpha \Delta W(t-1)$$

Where ΔW is the weight change at time t and the parameters γ and α are respectively the learning rate and the momentum coefficient. This latter, makes the convergence faster and the training more stable.

2.3 The PNN

Unlike the back-propagation neural networks, the probabilistic neural network [34] requires only a single presentation of each pattern. The PNN employs an exponential activation function rather than the sigmoid function that is commonly used in the MLP. Then, a PNN can identify nonlinear decision boundaries that approach the Bayes optimal [52]. The basic network topology consists of four layers. The first layer is the inputs layer. In the second layer, the probability density function (PDF) of each group of patterns is directly estimated from the set of training samples using [53] window approximation method. The third layer performs the summation of all PDFs. Finally, the Bayesian decision is made in the fourth layer. In sum, the network structure of PNN is similar to back-propagation neural network; but the main difference is that the transfer function is replaced by exponential function and all training samples are stored as weight vectors. For instance, the PDF is assumed to follows a Gaussian distribution. Then,

the PDF for a feature vector X to be of a certain category A is given by:

$$f_A(X) = \left(1/(2\pi)^{0.5p}\sigma^p\right)m^{-1}\sum_{i=1}^{m}\exp\left((X-X_{Ai})'(X-X_{Ai})/2\sigma^2\right)$$

Where, p is the number of patterns in X, m is the number of the training patterns of category A, i is the pattern number, and σ is the smoothing factor of the Gaussian curves used to construct the PDF. The value of σ is optimized during training based on the clearest separation of classes with the highest classification rate [54-58].

2.4 Genetic algorithms

A genetic algorithm [59-62] was used to automatically optimize the architectures of the networks. GA was used for three reasons. First, the architectural design is crucial to the success of a network's information processing capabilities [62]. Second, genetic search provides an advantage over expert experience in building neural networks and also over constructive and destructive algorithms [63-64]. Finally, genetic algorithms allow the convergence speed of artificial neural networks to be faster because of the search multiple initial states and the effect of mutation operations [64]. The process of a genetic algorithm is iterative and consists of the following steps[3]:

1. Create an initial population of "genotypes", it is a genetic representation of neural networks, and network architectures are randomly selected.
2. Train and test the neural networks to determine how fit they are by calculating the fitness measure of each trained network i. The fitness function is calculated as: $f_i = 1/MSE_i$, where MSE is the mean squared error.
3. Compare the fitness of the networks and keep the best top 10 for future use.
4. Select better networks from each completed population by applying the selection operator.
5. Refill the population back to the defined size.
6. Mate the genotypes by exchanging genes (features) of the networks.
7. Randomly mutate the genotypes according to a given probability.
8. Return back to step 2 and continue this process until stopping criteria (RMSE<ε) is reached.

The initial parameters of the BPNN to be optimized are described in Table 2.

---

[3] Detailed comprehension of genetic algorithms is given in [62].

Table 2
Initial parameter space of the genetic algorithm

Hidden layers	
Neurons by each hidden layer	Maximum 8
Activation functions	Linear, sigmoid, hypertang
Size of initial population	30
Selection	0.50%
Refill	Cloning
Mating	Tail swap
Mutations	Random exchange at 0.25%
Number of passes	20
Learning rate parameter range	0.4
Momentum rate parameter range	0.3
Stopping criterion	10 generations

Notes: sigmoid function is defined as $sigmoid(t)=1/\left(1+e^{-t}\right)$, hypertang function is defined as $\tanh(t)=\left(e^t-e^{-t}\right)/\left(e^t+e^{-t}\right)$

In order to predict the future trends of stocks and the NYSE, the following model is approximated using neural networks:

$$y_{i,t} = f(R_{i,t-1}, R_{i,t-2}, S_{i,t-1})$$

where R^4 is returns, S is a matrix of sentiment indicators selected following the Granger causality test, f is an unknown function to be approximated, y is the future trend, t is time script and i is the series to be modeled. All inputs (R and S) are standardized to help the intelligent systems to converge efficiently according to the following the transformation given by:

$$x' = \left(2x - (Max(x)+Min(x))\right)/\left(Max(x)-Min(x)\right)$$

The output is defined as follows:

$$y_{i,t} = \left\{0 \quad if \quad R_{i,t}<0; 1 \quad if \quad R_{i,t}>0\right\}$$

The previous definition means that the trading strategy is buying stock if its predicted future return is positive (future stock price is expected to increase) and selling stock if its predicted future return is negative (future stock price is expected to fall). The out-of-sample daily predictions were conducted and the prediction accuracy was measured. The accuracy is calculated based on the number of correct classifications (Hit Ratio). The highest neural output of the network indicates the category or class the data record falls into. For example, 100% accuracy is where all records are properly classified and 0% accuracy is where none are properly classified.

---

[4] Future trends are predicted with past return at *t-1* and *t-2* since all return series follows an ARMA(2,2) process. Indeed, all autoregressive coefficients up to order two were found to be statistically significant.

3. RESULTS AND DISCUSSION

First of all, the Granger causality tests shows strong evidence that sentiment variables statistically cause shifts in stock returns. This is consistent with [12][13] findings. Table 3 presents the correct classification rates from the two systems. In terms of prediction accuracy, the GA-BPNN has the highest hit ratio (55.75%) in predicting the sign of future return, while the best hit ratio obtained by PNN is 52.83%. On the other hand, the lowest hit ratios were 48.14% and 53.32% obtained by PNN and GA-BPNN respectively. As a result, the GA-BPNN outperforms the PNN in all stocks including the NYSE. Although the prediction accuracy is low, the overall results are interesting since some previous studies reported that stock prices are approximately close to the random walk process and; consequently; an accuracy of 56% in the predictions is a satisfying result for stock forecasting [65-68].

Table 3:
 Out-of-sample classification accuracy in %

	PNN	GA-BPNN
Apple	51.95	55.75
Cisco	48.14	54.34
GE	52.30	54.69
NYSE	49.96	53.32

According to empirical behavioural finance, high impulsivity investors and investors with high confidence in the future are more likely to take financial risks [69][70]. Among many possible reasons for future risk taking is fully substitutable stocks to purchase. Indeed, in many situations investors act as if they extrapolate a positive price trend by overbuying winners [71]. In addition, investors who extrapolate trends in stock prices are likely to follow a momentum investment strategy and buy winners [72][73]. In particular, if momentum investment strategy is adopted then investors are more likely to buy shares that have a higher value [74]. Then, in order to improve the prediction of the future trend in technology stocks and the NYSE, a trading rule is defined to buy stocks. For instance, the output is defined according to three strategies as follows:

Strategy.1:
$$y_{i,t} = \left\{ 0 \quad if \quad R_{i,t} < 1\%; 1 \quad if \quad R_{i,t} > 1\% \right\}$$

Strategy.2:
$$y_{i,t} = \left\{ 0 \quad if \quad R_{i,t} < 1.5\%; 1 \quad if \quad R_{i,t} > 1.5\% \right\}$$

Strategy.3:
$$y_{i,t} = \left\{ 0 \quad if \quad R_{i,t} < 2\%; 1 \quad if \quad R_{i,t} > 2\% \right\}$$

The previous output definitions allow investigating the relationship between trading rules (buying the stock if the expected trend is up more than *s %*) and the performance of the neural network systems. The forecasting performance of each trained system was compared and analyzed depending on the trading strategy. The results obtained from simulations are given in Table 4.

Table 4
Classification accuracy (in %) given trading rules

Systems	GB	GB	GB	PNN	PNN	PNN
Rules	1%	1.50%	2%	1%	1.50%	2%
Apple	82.65	86.36	94.60	80.97	86.27	94.51
Cisco	84.60	91.33	95.31	83.36	90.09	94.16
GE	92.39	96.55	98.67	91.59	96.02	98.41
NYSE	97.79	98.67	99.47	97.52	98.58	99.38

GB is GABPNN.

The simulations show that the performance of the systems improves when the trading rules are considered. The lowest and the highest hit ratios for the PNN are 82.65% and 99.47%. On the other hand, lowest and the highest hit ratios for the GA-BPNN are 80.97% and 99.38%. The lowest hit ratios are obtained with Apple given 1% decision rule and the highest hit ratio is obtained with NYSE given 2% decision rule. Recall that the lowest and highest accuracy rates obtained when the trading rule is 0% are respectively 48.14% for the PNN and 53.32% for the GA-BPNN. Then, a predictive system based on defined decision rules (buy stock if predicted up is more than *s %*) performs much better than a predictive system that predicts both future directions: up and down. Moreover, the performance of the systems increases with the trading rules: an increase in the decision rule leads to an improvement in the performance of the systems. Finally, the simulations confirm that the GA-BPNN is suitable for stock market trend prediction than the PNN since it achieves higher accuracy in all stocks and all strategies.

Previous studies have shown that PNN provide good classification rates than traditional BPNN in many different applications [75][76]. However, our findings show the superiority of GA-BPNN over PNN. Indeed, the PNN has a fixed topology and; on the other hand; the topology of the BPNN is optimized using genetic algorithms in this study. This could explain why BPNN outperforms the PNN. Thus, the role of genetic algorithm is

important in optimizing the BPNN architecture and achieving higher performance. This is consistent with [62]. Also, the reason that GA-BPNN outperforms the PNN in this study is that PNN is sensitive to noisy data such as financial time series. Finally, it is important to mention that faster than GA-BPNN with similar classification accuracy for up trend detection, the PNN would be more appropriate in real time applications when trading strategies are considered.

4. CONCLUSION

Predicting stock market trends is a classification problem that categorizes markets returns as up and down moves, which is easier than price prediction. There are several neural network architectures and statistical methods to perform classification tasks. In this paper, two different neural network architectures are used to predict future trends of the NYSE and three stocks from the technology sector. The neural networks are the Back-Propagation neural networks (BPNN) and the Probabilistic Neural Network (PNN). Genetic algorithms (GA) are used to optimize the topology of the BPNN. Investor sentiment measures are used as inputs to the neural networks. First, Granger causality tests are applied to identify which measures are statistically related to each company and equity market index. Similar to prior studies, our findings show that individual stock returns are highly related to the sentiment of the investor according to Granger causality tests. Second, soft computing techniques – artificial neural networks - are employed to model and predict the future sign of stock returns and the market. The simulations show that the GA-BPNN is suitable for stock market trend prediction than the PNN since it achieves higher accuracy in all stocks and in all strategies.

According to financial theory, investors seek to maximize their final wealth. And, according to empirical behavioural finance; when a momentum investment strategy is adopted investors are more likely to buy shares that have a higher value. Therefore, it is interesting to design intelligent systems to predict future up trends based on suitable trading rules. The results show that trading given decision rules (buy a stock if predicted return is higher than a given threshold) yields to higher accuracy than trading on the basis of predicted trend up or less than 0%.

For future research it is suggested to compare the prediction accuracy of neural networks systems and support vector machines (SVM). More importantly, it is suggested to design and implement an ensemble system. In addition, it would be interesting to take into account both economic variables and technical indicators along with sentiment measures to predict stock market future trends.

REFRENCES:

[1] W.F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. Journal of Finance **19** (1964), 425-442.

[2] J. Lintner . The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Review of Economics and Statistics **47** (1965), 13-37.

[3] S. Ross. The Arbitrage Theory of Capital Asset Pricing. Journal of Economic Theory **13** (1976), 341-360.

[4] R.C. Merton. An Intertemporal Capital Asset Pricing Model. Econometrica **41** (1973), 867-887.

[5] J.B. De Long, A. Shleifer, L.H Summers and R.J Waldmann. Noise trader risk in financial markets, Journal of Political Economy **98** (1990), 703-738.

[6] K.L. Fisher and M. Statman. Investor Sentiment and Stock Returns. Financial Analysts Journal **56** (2000), 16-23.

[7] M. Baker and J. Wurgler. Investor Sentiment and the Cross-Section of Stock Returns. National Bureau of Economic Research,Working Papers 10449, 2004.

[8] M. Baker and J. Wurgler. Investor Sentiment and the Cross-Section of Stock Returns. Journal of Finance **61** (2006), 1645-1680.

[9] R.S. Neal and S.M. Wheatley. Adverse selection and bid-ask spreads: Evidence from closed-end funds. Journal of Financial Markets **1** (1998), 121-149.

[10] D.P. Simon and A. Roy. Wiggins III. S&P Futures Returns and Contrary Sentiment Indicators. The Journal of Futures Markets **21** (2001), 447–462.

[11] C. Wang. Investor Sentiment and Return Predictability in Agricultural Futures Markets. The journal of Futures Markets **21** (2001), 929-952.

[12] G.W. Brown and M.T Cliff. Investor Sentiment and Asset Valuation. Journal of Business **78** (2005), 405–40.

[13] M. Lemmon and E. Portniaguina. Consumer Confidence and Asset Prices: Some Empirical Evidence. Review of Financial Studies **19** (2006), 1499–1529.

[14] L. Xiaoli Qiu and I. Welch. Investor Sentiment Measures. Social science research network Working Paper Series, 2006. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=589641

[15] M.J. Kamstra, L.A. Kramer and M.D. Levi. Winter Blues: A SAD Stock Market Cycle. American Economic Review **93** (2003), 1257–63.

[16] A. Edmans, D. Garcia and Øyvind Norli. Sports Sentiment and Stock Returns. Journal of Finance **62** (2007), 1967-1998.

[17] M. Baker and J. Stein. Market Liquidity as a Sentiment Indicator. Journal of Financial Markets **7** (2004), 271-299.

[18] J. Scheinkman and W. Xiong. Overconfidence and Speculative Bubbles. Journal of Political Economy **11** (2003), 1183–1219.

[19] R. Greenwood and S. Nagel. Inexperienced Investors and Bubbles. 2006. http://www.people.hbs.edu/rgreenwood/Mfage8.pdf.

[20] A. Kumar and C. Lee. Retail Investor Sentiment and Return Comovement. Journal of Finance **61** (2006), 2451 – 2486.

[21] M.E. Zweig. An Investor Expectations Stock Price Predictive Model Using Closed-End Fund Premiums. Journal of Finance **28** (1973), 67–87.

[22] C. Lee, A. Shleifer and R.H. Thaler. Investor Sentiment and the Closed-End Fund Puzzle. Journal of Finance **46** (1991), 75–109.

[23] N. Robert and S.M. Wheatley. Do Measures of Investor Sentiment Predict Returns?. Journal of Financial & Quantitative Analysis **33** (1998), 523–48.

[24] K. Kohara, Y. Fukuhara and Y. Nakamura. Selective presentation learning for neural network forecasting of stock markets. Neural Computing & Applications **4** (1996), 143–148.

[25] J.C. Yao, L. Tan and H.L. Poh. Neural networks for technical analysis: a study on KLCI. International Journal of Theoretical and Applied Finance **2** (1999) 221–241.

[26] M. Qi and G.S. Maddala. Economic factors and the stock market: a new perspective. Journal of Forecasting **18** (1999) 151–166.

[27] H. Tan, D. Prokhorov and D. Wunsch. Conservative thirty calendar day stock prediction using a probabilistic neural network. Proceedings of the IEEE Conference Computational Intelligence for Financial Engineering, 1995, 113–117.

[28] K. Schierholt and C.H. Dagli. Stock Market Prediction Using Different Neural Network Classification Architectures. Conference on computational intelligence for financial engineering IEEE/IAFE, 1996.

[29] W. Emad Saad, D.V. Prokhorov and D.C. Wunsch. Comparative Study of Stock Trend Prediction Using Time Delay, Recurrent and Probabilistic Neural Networks. IEEE Transaction on Neural Networks, 1998, 1456-1470.

[30] S.H. Kim and S.H. Chun. Graded forecasting using an array of bipolar predictions: application of probabilistic neural networks to a stock market index," International Journal of Forecasting **14** (1998), 323–337.

[31] L.M.Y Leung, H. Daouk and A.-S. Chen. Forecasting stock indices: a comparison of classification and level estimation models. International Journal of Forecasting **16** (2000), 173–190.

[32] A.S. Chen, M. T Leung and H. Daouk. Application of neural networks to an emerging financial market: Forecasting and trading the Taiwan Stock Index. Computers and Operations Research **30** (2003), 901–923.

[33] D.E Rumelhart, G.E. Hinton and R.J. Williams. Learning internal representations by error propagation. In: Rumelhart, D.E., McClelland, J.L. (Eds.), Parallel Distributed Processing, Vol. 1, MIT Press, Cambridge, MA, 1986, 318-362.

[34] D. Specht. Probabilistic Neural Networks for Classification, Mapping, or Associative Memory. IEEE International Conference on Neural Networks, 1988.

[35] P. Wasserman. Advanced Methods in Neural Computing. New York: Van No strand Reinhold, 1993.

[36] D. Specht. Probabilistic Neural Networks. Neural Networks **3** (1990), 109-118.

[37] S.H. Kim and S.H Chun. Graded forecasting using an array of bipolar predictions: application of probabilistic neural networks to a stock market index. International Journal of Forecasting **14** (1998), 323–337.

[38] The Chicago Board Options Exchange at www.cboe.com.

[39] The State Street's Investor Confidence Index at www.statestreet.com.

[40] D. Hirshleifer. Investor psychology and asset pricing. Journal of Finance **56** (2001), 1533–1596.

[41] F.X. Zhang. Information Uncertainty and Stock Returns. The journal of Finance (2006), 105-136.

[42] L.E Harris. Stock price clustering and discreteness. Review of Financial Studies **4** (1991), 389−415.

[43] P.E. Godek. Why NASDAQ market makers avoid odd-eighth quotes. Journal of Business **65** (1996), 509−528.

[44] Z. Ding, C.W.J Granger and R.F Engle. Long memory property of stock market returns and a new model. Journal of Empirical Finance **1** (1993), 83–106.

[45] C. Gourieroux. ARCH Models and Financial Applications. Springer-Verlag, 1997.

[46] M.T. Subbotin. On the law of frequency error. Matematicheskii Sbornik **31** (1923), 296–301.

[47] A. Wilhelmsson. GARCH Forecasting Performance under Different Distribution Assumptions. Journal of Forecasting **25** (2006), 561–578.

[48] F.P Ghijsels. Additive outliers, GARCH and forecasting variance. International Journal of Forecasting **15** (1999), 1–9.

[49] M. McKenzie and H. Mitchell. Generalised asymmetric power arch modeling of exchange rate volatility. Applied Financial Econnomics **12** (1999), 555–564.

[50] M. McKenzie, H. Mitchell, R. Brooks and R. Faff. Power arch modelling of commodity futures data on the London metal exchange. European Journal of Finance **7** (2001), 22–38.

[51] C.W.J Granger. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. Econometrica **37** (1969), 424-438.

[52] E. Parzen. On Estimation of a Probability Density Function and Mode. Annals of Mathematical Statistics **33** (1962), 1065-1076.

[53] Z.R. Yang, M.B. Platt and H.D Platt. Probabilistic Neural Networks in Bankruptcy Prediction. Journal of Business Research **44** (1999), 67–74.

[54] C. Gaganis and F. Pasiouras, M. Doumpos. Probabilistic neural networks for the identification of qualified audit opinions. Expert Systems with Applications **32** (2007), 114–124.

[55] Y. Chen, X. Zhang, M.Q. Yang and J.Y. Yang. Ensemble of Probabilistic Neural Networks for Protein Fold Recognition. Proceedings of the 7th IEEE International Conference on Bioinofmratics and Bioengineering, 2007, 66-70.

[56] L. Wang and D. Wen. Recognition of Arterial Pulse Wave of Pregnant Woman Based on Probabilistic Neural Network. Proceedings of the 8th IEEE International Conference on Bioinofmratics and Bioengineering, 2008, 1955 - 1958.

[57] S.-Y. Ooi, A.B.J Teoh and Thian-Song Ong. Compatibility of biometric strengthening with probabilistic neural network. International Symposium on Biometrics and Security Technologies, 2008, 1 – 6.

[58] S. B. Sert and Ö. Kalenderli. Determination of source voltage from audible corona noise by neural networks. IEEE Transactions on Dielectrics and Electrical Insulation **16** (2009), 224-231.

[59] J.H. Holland. Outline for a logical theory of adaptive systems. Journal of the ACM 3 (1962) 297-314.

[60] J.H. Holland. Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor, 1975.

[61] J.H. Holland. Adaptation. In: Rosen, R., Snell, F.M. (Eds.), Progress in Theoretical Biology, Academic Press, New York, **4** (1976), 263-293.

[62] X. Yao. 1999. Evolving Artificial Neural Networks. Proceedings of the IEEE, **87** (1999).

[63] G.F. Miller, P.M. Todd and S.U. Hegde. Designing neural networks using genetic algorithms. Proc. 3rd Int. Conf.Genetic Algorithms and Their Applications, J. D. Schaffer, Ed. San Mateo, CA: Morgan Kaufmann, 1989, 379–384.

[64] D.E. Goldberg. Genetic Algorithm in Search, Optimization, and Machine Learning, Addison Wesley, 1989.

[65] B. Qian and K. Rasheed. Stock market prediction with multiple classifiers. Applied Intelligence **26** (2007), 26:25–33.

[66] S. Walczak. An empirical analysis of data requirements for financial forecasting with neural networks. Journal of Management Information Systems **17** (2001) 203–222.

[67] G. Tsibouris and M. Zeidenberg. Testing the efficient markets hypothesis with gradient descent algorithms, In: Refenes AP (ed) Neural networks in the capital markets, JohnWiley & Sons, Chichester, England, Chap 8, 1995, 127–136.

[68] D.J.E. Baestaens, W.M. Van Den Bergh and H. Vaudrey. Market inefficiencies, technical trading and neural networks. In: Dunis C (ed) forecasting financial markets, financial economics and quantitative analysis. John Wiley & Sons, Chichester, England, 1996, 254–260.

[69] A. Shleifer and R. Vishny. The limits of arbitrage. Journal of Finance **52** (1997), 35-55.

[70] W.F.M. Debont. Stock prices: Insights from behavioral finance. In A. Lewis (Ed.), The Cambridge Handbook of Psychology and Economic Behavior, Cambridge, Cambridge University Press, 2008, 64-104.

[71] T. Odean. Do investors trade too much? American Economic Review **89** (1999) 103-124.

[72] P.A. Andreasson. Explaining the price-volume relationship: The difference between price changes and changing prices. Organizational Behavior and Human Decision Processes **41** (1988), 371-389.

[73] W.F.M. DeBondt. Betting on trends: Intuitive forecasts of financial risk and return. International Journal of Forecasting **9** (1993), 355-371.

[74] H. Hong and J.C. Stein. A unified theory of underreaction, momentum trading, and overreaction in asset markets. Journal of Finance **54** (1999), 2143-2184.

[75] S. Ramakrishnan and I.M.M. El Emary. Comparative study between traditional and modified probabilistic neural networks. Telecommunication Systems 40 (2009), 67–74.

[76] Y.L. Chong and K. Sundaraj. A study of back-propagation and radial basis neural network on EMG signal classification. Proceeding of the 6th International Symposium on Mechatronics and its Applications (ISMA09), Sharjah, UAE, 2009, 24-26.